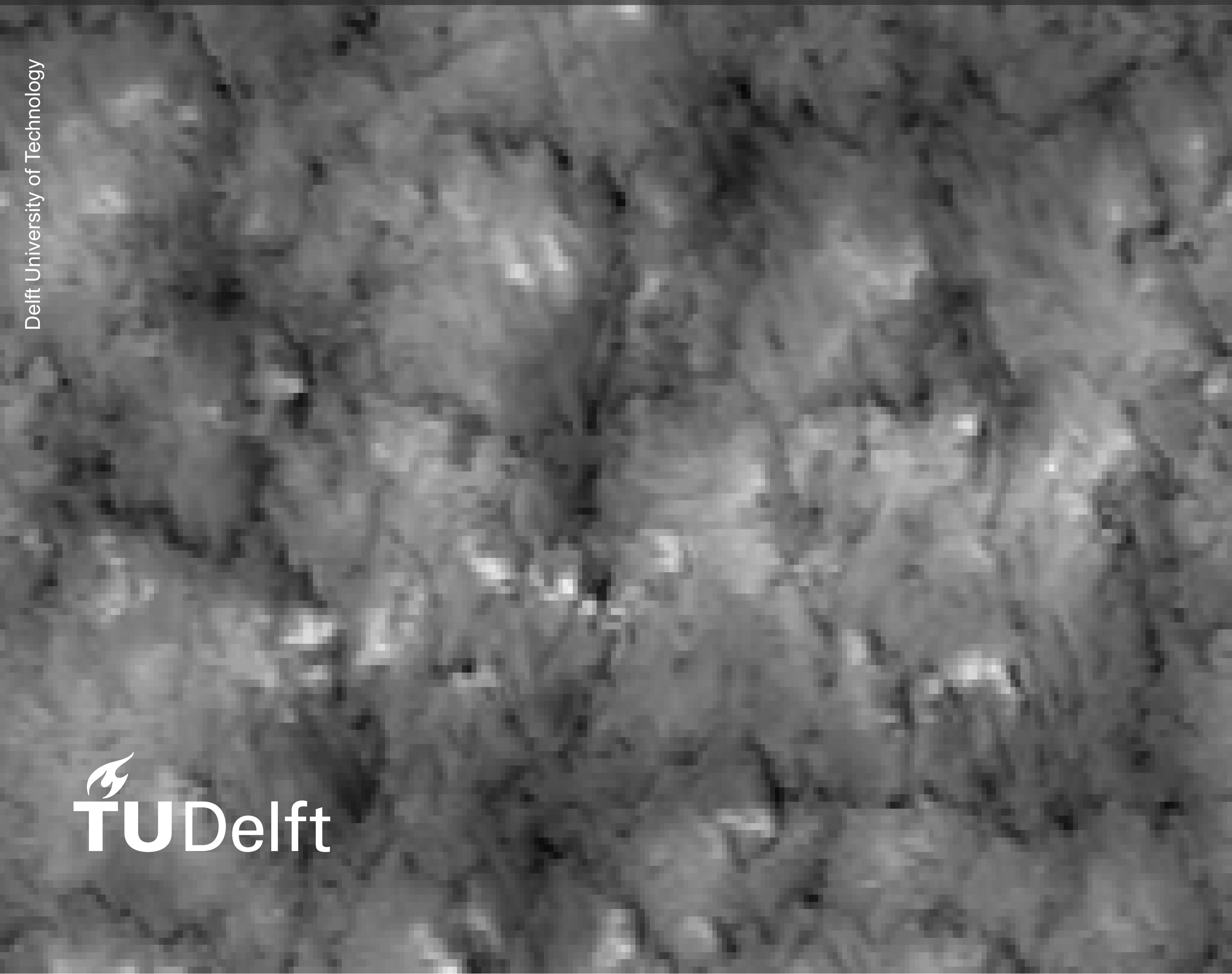
Data assimilation of observed cloud fields in LES model

Applying a three-dimensional nudging tendency to thermodynamic properties during LES model spin-up for increased agreement with observations

M. L. Steerneman



Delft University of Technology

Data assimilation of observed cloud fields in LES model

Applying a three-dimensional nudging tendency to thermodynamic properties during LES model spin-up for increased agreement with observations

by

M. L. Steerneman

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday September 19, 2023 at 3:00 PM.

Student number:	4670116	
Project duration:	January 16, 2023 – September 19, 2023	
Thesis committee:	Dr. S. R. de Roode,	TU Delft, supervisor
	Prof. dr. H. J. J. Jonker,	TU Delft
	Dr. ir. M. Rohde,	TU Delft

An electronic version of this thesis is available at http://repository.tudelft.nl/.
Cover image: the liquid water path in the atmosphere as found by a simulation of a stratocumulus-topped boundary layer.

TU Delft

Preface

As a (future) physicist, I am quite fond of numbers. Therefore, I thought I would give a short overview of my thesis project in statistics. The thesis that lies before you consists of 22636 words spanning 78 pages. It is the product of 9 months of hard labor, during which I have considered tossing my computer out of the window at least 3 times, and have decorated my notebooks with 76 doodles. 48 different Python scripts have helped me convert the data of 198 DALES runs into useful results, only a fraction of which are presented in this work.

In around 30 meetings, my supervisor Stephan showed and shared his extensive knowledge and enthusiasm for not just the 1 topic of stratocumulus, but many other unrelated subjects as well. Stephan, yours is an enthusiasm that cannot help to rub off on the people around you and I thank you for the way it has motivated me time and time again. Furthermore, I would like to express my gratitude to Martin Rohde and Harm Jonker, who along with Stephan will allow me to defend my thesis (hopefully just the 1 time). I am also indebted to my fellow inhabitants of master thesis rooms 3.36 and 3.34, for providing much-needed distractions during the 3 breaks which we stuck to so rigorously each day. Finally, I want to thank my family, friends, and Celine for the infinite amount of love and support they have provided me in my life.

M. L. Steerneman
Delft, September 2023

Abstract

This thesis investigates the implementation of three-dimensional nudging into large-eddy simulation (LES) to assimilate observed atmospheric data into an LES model. 3D-nudging 'pushes' the thermodynamic fields in a simulation towards the desired observed fields. The aim is to test if such a method is useful in improving solar forecasts of stratocumulus-topped boundary layers. For this purpose 3D-nudging LES solar forecasts are compared to persistence forecasts and conventional LES-based forecasts. As a proxy for observations, exact thermodynamic fields from LES were used in this research. Using LES fields is advantageous as it provides full 3D thermodynamic fields but also dynamic fields for checking the turbulence in the different methods. Results show that 3D-nudging is quite capable of replicating the desired thermodynamic fields. Unfortunately, nudging comes with a penalty as it causes the turbulence built up in a simulation to be flawed. This effect is mitigated by the design of variations on the nudging technique, the most promising of which is multiple time fields nudging, which nudges the thermodynamic fields in a simulation to subsequent desired fields every 10 minutes during the nudging period. Solar forecasts found by this method are found to be more accurate than the persistence and regular LES methods on forecast horizons of 30 minutes and larger. Approaches proposed in this study to approximate thermodynamic fields from observational data estimate thermodynamic fields to a reasonable accuracy but are far from perfect, and thus it should be noted that solar forecast accuracy of the discussed methods will be less accurate when applied to real observations. Further research is recommended to focus on the use of the 3D-nudging methods in more LES case studies, and on devising better methods for the estimation of thermodynamic fields from observations.

Contents

Nomenclature

Abbreviations

Abbreviation	Definition
ABL	Atmospheric boundary layer
ASTEX	The Atlantic Stratocumulus Transition Experiment
CBL	Convective boundary layer
DALES	Dutch Atmospheric Large-Eddy Simulation
FS	Forecast skill
GCM	Global circulation model
KNMI	Royal Netherlands Meteorological Institute
LCL	Lifting condensation level
LES	Large-eddy simulation
LWP	Liquid water path
NWP	Numerical weather prediction
RMSE	Root mean square error
SFS	Subfilter-scale
STBL	Stratocumulus-topped boundary layer
(V)TKE	(Vertically integrated) turbulent kinetic energy

Symbols

Symbol	Definition	Unit
c_f	LES filter width	-
c_h, c_m, c_ϵ	Proportionality constants	-
COT	Cloud optical thickness	-
c_p	Specific heat at constant pressure	J kg^{-1} K^{-1}
c_{qT}	Constant used in the relationship between T' and q_t'	K (kg kg^{-1})$^{-1}$
c_sfc	Constant relating the surface SHF and LHF	K (kg kg^{-1})$^{-1}$
e	Subfilter-scale turbulence kinetic energy	m^2 s^{-2}
e	Water vapor pressure	Pa
e_sat	Saturation vapor pressure	Pa
f	Coriolis parameter	s^{-1}
F	Energy flux	W m^{-2}
$\mathcal{F}_\mathrm{i}$	Forcing	m s^{-2}
$F_\mathrm{LW,net}$	Net longwave radiation flux	W m^{-2}
FS	Forecast skill	-
g	Earth's gravitational acceleration	m s^{-2}
H	Cloud thickness	m
K_h	Eddy diffusivity for thermodynamic scalars	m^2 s^{-1}
K_m	Eddy viscosity for momentum	m^2 s^{-1}
LHF	Latent heat flux	W m^{-2}
L_v	Latent heat release due to condensation of water vapor	J kg^{-1}
LWP	Liquid water path	kg m^{-2}
l_Δ	Geometric mean of LES mesh sizes	m
m	Mass	kg
N_d	Cloud droplet number concentration	m^{-3}

Symbol	Definition	Unit
p	Pressure	Pa
$\mathrm{Pr_T}$	Turbulent SFS Prandtl number	-
q_l	Liquid water specific humidity	kg kg^{-1}
q_sat	Saturated water vapor specific humidity	kg kg^{-1}
q_t	Total water specific humidity	kg kg^{-1}
q_v	Water vapor specific humidity	kg kg^{-1}
R_d	Specific gas constant for dry air	J kg^{-1} K^{-1}
r_e	Effective cloud droplet radius	m
RH	Relative humidity	-
RMSE	Root mean square error (in LWP)	kg m^{-2}
R_v	Specific gas constant for water vapor	J kg^{-1} K^{-1}
SHF	Sensible heat flux	W m^{-2}
$S_{q\mathrm{t}}$	Sources of moisture from melting or formation of ice	kg kg^{-1} s^{-1}
S_{θ_l}	Heat source	K s^{-1}
$S_\phi^{\mathrm{n},1D}$	1D-nudging source term for thermodynamic variable ϕ	[ϕ] s^{-1}
$S_\phi^{\mathrm{n},3D}$	3D-nudging source term for thermodynamic variable ϕ	[ϕ] s^{-1}
T	Temperature	K
t	Time	s
T_v	Virtual temperature	K
u	East-west wind speed	m s^{-1}
v	South-north wind speed	m s^{-1}
VTKE	Vertically integrated turbulent kinetic energy	m^3 s^{-2}
w	Vertical wind speed	m s^{-1}
w_e	Entrainment velocity	m s^{-1}
x	East-west location	m
y	South-north location	m
z	Height	m
z_b	Cloud base height	m
z_i	Inversion layer height	m
z_t	Cloud top height	m
α	Change of q_l with height	kg kg^{-1} m^{-1}
α_K	Kolmogorov constant	-
β	Factor introduced in the relation between q_t' and q_l'	-
γ	Change of saturation specific humidity with temperature	kg kg^{-1} K^{-1}
ϵ, ϵ_I	Dimensionless constants relating R_d and R_v	-
η	Dimensionless constant for relations between fluctuation in limiting cases	-
θ	Potential temperature	K
θ_l	Liquid potential temperature	K
θ_v	Virtual potential temperature	K
ϑ_0	Solar zenith angle	°
λ	SFS turbulent eddy characteristic length scale	m
ν	Deviatoric sub-grid momentum flux	m^2 s^{-2}
Π	Exner function	-
π	Modified pressure	m^2 s^{-2}
ρ	Density	kg m^{-3}
σ_r	Reference standard deviation in LWP field	kg m^{-2}
τ	Nudging time scale	s
τ_{1D}	One-dimensional nudging time scale	s
τ_{3D}	Three-dimensional nudging time scale	s
φ	Latitude	°
ω	Angular velocity of the Earth	s^{-1}

List of Figures

1 Introduction

1.1. Importance of solar forecasting

Climate change is one of the biggest problems our world is facing right now. Greenhouse gas emissions caused by mankind are leading to a significant rise in the temperature of the Earth (Masson-Delmotte et al., 2021), resulting in a variety of problems that we are already starting to encounter and that will only escalate in the future (Patz et al., 2005, Pörtner et al., 2022). With a global share of 73.2% (Ritchie et al., 2020), the energy sector is the largest contributor to greenhouse gas emissions by far. To curb climate change, emissions in the energy sector will need to be reduced drastically, and should eventually be brought to zero. Achieving this requires a transition from fossil fuels to renewable energy sources. About 12% of the renewable energy that is currently generated has the sun as its source, and this fraction is growing rapidly (IEA, 2022b). Projections based on the Net Zero Emissions by 2050 Scenario indicate that solar energy production will increase by 640% to generate a third of all renewable energy by 2030 (IEA, 2022a). This enormous increase in solar energy poses its own set of challenges. Large variability in the output of solar farms due to changes in the weather leads to great challenges in reliably meeting energy demand and maintaining a stable and fail-safe energy grid (Das et al., 2018). Solar energy generation can be predicted by using forecasts of solar radiation. Therefore, solar forecasts are crucial for the integration of solar energy into the energy grid.

1.2. Conventional solar forecasting for stratocumulus

The field of atmospheric physics helps to create solar forecasts by developing computational models to predict the weather. For weather forecasting, general circulation models (GCMs) or numerical weather prediction (NWP) models are typically used. Accurate representation of clouds in these models is particularly important for solar forecasting, as clouds reflect a large part of the incoming sunlight. Stratocumulus is a cloud type that is particularly difficult to represent in forecasting models. Stratocumulus clouds have an average albedo of 0.6 (Barry and Chorley, 2003), meaning they reflect 60% of incident solar radiation. As such, they have a large effect on the available solar radiation at the Earth's surface. Stratocumulus clouds occur at low altitudes (heights ranging from 500-2000 m (Wood, 2015)) in a shallow layer called the atmospheric boundary layer (ABL). The ABL is found in the lowest part of the troposphere and is a layer of turbulent air that is coupled with the Earth's surface. It is divided from the laminar air above it by a thermal inversion layer. Within this thermal inversion layer, which is only a few tens of meters thick (Wood, 2012), the temperature rises significantly and the humidity of air decreases drastically with height.

Stratocumulus cloud cover is greatly underestimated in NWP and GCMs, as their large grid size renders these models unable to represent the sharp gradients in the inversion layer properly (Ma et al., 1996, Duynkerke and Teixeira, 2001, Siebesma et al., 2004, Mathiesen et al., 2013). This leads to erroneous forecasts of the downwelling solar radiation. To bypass this inaccuracy, techniques have been developed where information on cloud coverage and cloud motion is extracted from atmospheric images and used to propagate the cloud structures in time (Perez et al., 2010, Marquez et al., 2013, Chow et al., 2011). Furthermore, Wang et al. (2019) have developed an algorithm based on this method which outperforms the NWP model HARMONIE, developed by the Royal Netherlands Meteorological Institute (KNMI), for a forecast horizon of 2-3 hours. The approach described here assumes that no other physical processes other than advection work on the clouds. Such a method is more commonly known as persistence forecasting.

The persistence technique as described above leaves out all other physical processes besides advection, like turbulence and microphysics. However, especially in a relatively thin cloud like stratocumulus, ignoring these physical processes negatively affects stratocumulus prediction. Incorporating these processes into forecasting models is therefore expected to yield even more accurate results compared to persistence models. Unfortunately, the large mesh sizes used in GCMs (standard use of 5-11 vertical grid levels below 1 km (Wyant et al., 2007)) and NWP (lowest levels divided by pressures of 10 hPa, equivalent to roughly 75 meters (Chou, 2011)) models do not allow analytic solving of these physical processes, which typically occur on a smaller scale. Processes like turbulent transport and microphysics are accounted for by parameterizations that can introduce serious inaccuracies.

1.3. LES models as solar forecasting tool

For this reason, studies on the ABL, where stratocumulus occurs, are typically conducted with atmospheric models based on large-eddy simulation (LES). LES is a computational method that simulates turbulence in fluids by simplifying the mathematical approach to solving turbulence. It assumes that the turbulent eddies on small length scales (smaller than the grid size) have little to no effect on the behavior of the entire system and so solves all equations only on the scale of the largest turbulent eddies present in the atmosphere. Physical processes on the smaller turbulent scales are approximated using a parameterization. The larger scales can hold up to 90% of all the turbulent kinetic energy (TKE) in the system (Heus et al., 2010) and are resolved by numerical solving of the Navier-Stokes equations. So, the large-scale transports are resolved, and all transport that is smaller than the grid size is parameterized. The grid size of LES is at least one order of magnitude smaller than the largest turbulent eddy size. Generally, this translates to a grid size of 5-20m for a cloud-topped ABL. This grid size is significantly smaller than the grid size in GCM and NWP. A smaller grid size means that less of the atmospheric transport is parameterized. So, LES is less reliant on its parameterization than GCM or NWP and this leads to a more accurate representation of stratocumulus (Duynkerke et al., 2004).

1.3.1. Statement of the problem: the LES spin-up phase

Unfortunately, a regular LES run cannot be relied upon to produce an accurate forecast for short time periods. At the start of an LES run, no turbulence is present in the system. The turbulence has to be built up during a so-called 'spin-up' period, which lasts approximately two hours. LES induces turbulence by initially setting the horizontal thermodynamic fields to be homogeneous and equal to a mean vertical profile given as model input. Next, pseudo-random perturbations are assigned to the thermodynamic variables at each grid point. The differences between horizontal values of thermodynamic quantities like temperature and humidity allow turbulence to develop. The spin-up phase ends when the turbulence has evolved to a quasi-steady state. In figure 1.1, the spin-up phase is visualized. It gives a plot of the vertically integrated TKE found from a run using the Dutch Atmospheric Large-Eddy Simulation (DALES). The TKE is the mean kinetic energy per unit mass of turbulent eddies. The figure clearly shows how the TKE varies wildly before arriving at a quasi-steady state after approximately 2 hours.

The physically unrealistic turbulence during spin-up makes results found during this period unreliable. Moreover, it can lead to an unrealistic evolution of the simulation, causing it to drift from the situation as observed. Research by de Roode et al. (2019) shows that this can give significant differences between simulations. Besides, the perturbations that are assigned to the initial fields will result in heterogeneous thermodynamic fields that do not align with the observed conditions. For high-resolution solar forecasting, this can be problematic, as both observations (Albrecht et al., 1990, Platt, 1976) and simulations (de Roode and Los, 2008) of the stratocumulus cloud layer indicate that whereas the cloud top has a height that is relatively homogeneous horizontally, the horizontal cloud base height can differ significantly. This is caused by heterogeneous fields of thermodynamic variables such as temperature and humidity. Such a varied cloud base field impacts the albedo of the cloud (McKee and Cox, 1974, Harshvardhan and Randall, 1985, Pincus et al., 1999), and hence directly affects the amount of sunlight received at the surface. Poorly represented cloud fields can thus lead to unreliable solar forecasts.

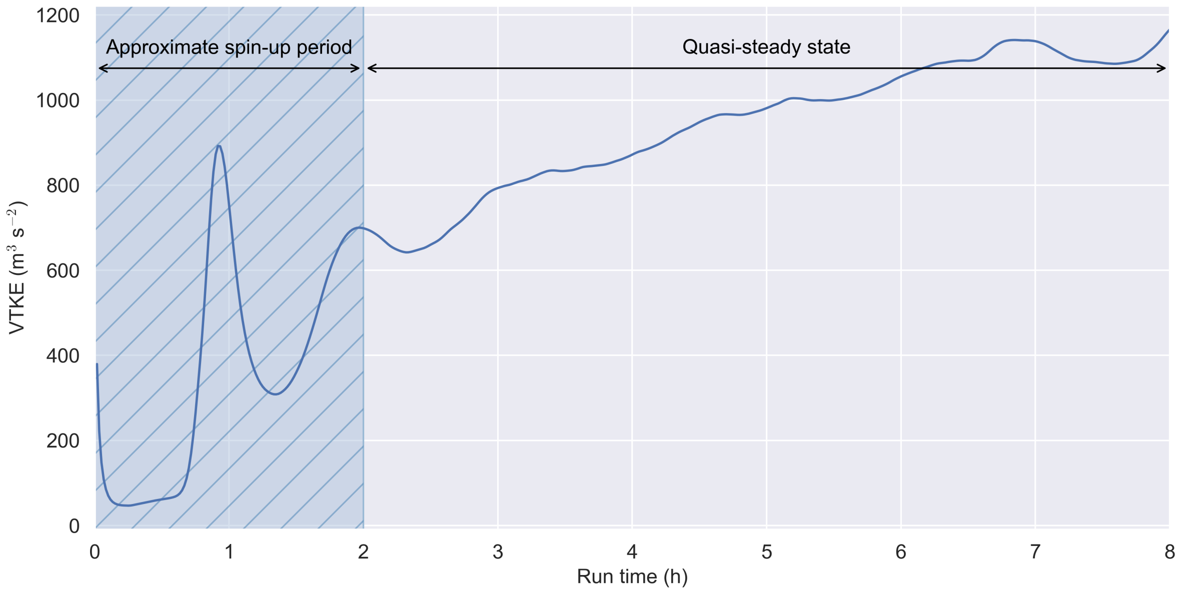


Figure 1.1: The vertically integrated turbulent kinetic energy (VTKE) as a function of simulation time. The first two hours are marked by an arced area in blue to indicate the approximate spin-up period. Results generated by an LES run simulating a stratocumulus-topped boundary layer (STBL) (set-up explained in sec. 5.1 and 5.2)

Additionally, the structure of the fields affects the evolution of the thermodynamics. Inaccurate fields can thus also cause the simulation to diverge further from observations. These issues can be partially avoided by initializing the LES run with the thermodynamic fields as found from observations. In such a way, no pseudo-random perturbations have to be assigned, and the field will bear a stronger resemblance to the observations. However, such a simulation will still have zero initial turbulence, making it evolve unrealistically (fig. 1.1).

To address the unrealistic development of thermodynamics during LES spin-up, the concept of nudging has been introduced. This method adds an extra term to the tendency of the temperature and humidity, making the values advance to a desired state. In other words, the simulation is 'pushed' to the values found in observations. One-dimensional nudging of the temperature and humidity in LES has been implemented broadly already, either to nudge to mesoscale model results (Schalkwijk et al., 2015, Draxl et al., 2021) or observations (Blossey et al., 2013, Atlas et al., 2020). In one-dimensional nudging, all values in a horizontal field are pushed towards a mean value as given by a vertical profile. This results in an improved agreement of the horizontal mean state between the simulation and observations, as the initial wild evolution of turbulence (fig. 1.1) has less effect on the mean state. However, this one-dimensional form of nudging does not take the horizontal heterogeneity of thermodynamic fields into account. As mentioned before, the structure of the field affects the albedo and temporal evolution of a cloud. So, to receive an accurate solar forecast, the observed thermodynamic fields should be assimilated into the LES model.

1.4. Research aim

For this purpose, this research explores a novel nudging technique. It aims to capture an observed LWP field in the simulation whilst maintaining the realistic turbulence generated during the spin-up period (fig. 1.1). This is tested using a 3D nudging procedure, which nudges the simulation towards three-dimensional thermodynamic fields, during the spin-up period. As only two-dimensional, LWP fields can be derived from satellite retrievals, this model uses 3D temperature and humidity fields generated by LES. The goal is to create a method that makes a large-eddy simulation start with physical turbulence as well as the observed fields, ultimately aiming to improve the ability of the LES model to correctly predict the occurrence and development of stratocumulus in the atmosphere. If achieved, this could lead to an improved solar forecast that can perhaps rival the standard persistence solar forecasts.

1.5. Research questions

The novel nudging technique is central to the research question posed in this thesis:

Does the implementation of three-dimensional nudging of the thermodynamic fields during the spin-up phase of an LES model give a better solar forecast for a stratocumulus-topped boundary layer than those generated by conventional methods?

To help formulate an answer to this question, three sub-questions were formulated:

- Is the three-dimensional nudging method capable of including the desired three-dimensional thermodynamic fields in LES, and what nudging time scales are required to do so?
- Does the implementation of the three-dimensional nudging term have an effect on the build-up of turbulence during the spin-up period, and how does this affect the evolution of the thermodynamic fields in time?
- Can thermodynamic fields for use in LES methods accurately be determined from ground-based and satellite observations?

1.6. Thesis stucture

After this introduction, this thesis continues with a description of the thermodynamic and dynamic processes relevant to a study of stratocumulus and a characterization of the stratocumulus cloud in chapter 2. Subsequently, chapter 3 describes a method of determining thermodynamic fields from observations. In chapter 4, DALES is introduced and its governing equations are given. Chapter 5 discusses the set-up of the experiments as well as the implementation of the 3D-nudging and persistence methods into DALES. Next, chapter 6 shows the results generated during this research, which are discussed in chapter 7. Finally, conclusions and recommendations that arise from this study are given in chapter 8.

2

Thermodynamics and dynamics of stratocumulus

In this chapter, thermodynamic properties of the atmosphere are introduced that are essential for large-eddy simulation of stratocumulus. Additionally, the equations governing the dynamics of stratocumulus in the atmosphere are outlined. Finally, a description of stratocumulus clouds is given, including their formation and evolution over time, drawing on the theory from the earlier sections.

2.1. Atmospheric thermodynamics

A fundamental aspect of studying cloud behavior is understanding the dynamics of water in the atmosphere, as clouds are formed by condensation of water vapor. Before describing all physical properties relevant to thermodynamic transport in the atmosphere, this section discusses the convective boundary layer (CBL), where stratocumulus occurs. The description given here is based on the contents of the syllabus *Atmospheric Physics* by de Roode (2021) and the textbook *Atmospheric Science* by Wallace and Hobbs (2006) unless explicitly stated otherwise. For a full derivation of the expressions below, the reader is referred to these texts.

2.1.1. Convective boundary layer

An ABL is referred to as a CBL or a mixed layer when the air inside it becomes vertically well-mixed. This usually happens during daytime hours. Vertical mixing of the air is caused by the generation of turbulence in the ABL. Turbulence, the irregular and chaotic motion of a gas or liquid, is composed of eddies of different sizes. Eddies are defined as areas where the flow of a fluid is different than the mean direction of the flow. The turbulence in the mixed layer is driven by two processes. The first is induced by solar heating of the surface, which causes the air just above the surface to be warmer than that higher up. Because of this, turbulent convective plumes of warm, ascending air and cold, descending air are generated. The second driver of turbulence is surface friction of horizontal wind, which causes the horizontal wind speeds close to the surface to be smaller than the speeds higher up (de Roode, 2021). Different wind speeds between vertical layers promote the transport of air between the layers. Because of the turbulence, quantities like temperature and humidity become well mixed throughout the CBL. However, the CBL is capped by the aforementioned thermal inversion layer, which acts as a sort of lid, not allowing thermal plumes to escape the boundary layer.

2.1.2. Equations for humidity

Water is always present in air, either in vapor, liquid, or solid form. Moreover, it is constantly changing phases. To quantify the amount of water in a specific phase in the atmosphere, the specific humidity is introduced:

$$q_{\mathrm{p}} = \frac{m_{\mathrm{p}}}{m_{\mathrm{tot}}}, \tag{2.1}$$

where the phase $\mathrm{p} \in \{\mathrm{v}, \mathrm{l}\}$ of the water is either water vapor or liquid water. For simplicity, the liquid water specific humidity also includes moisture in the form of ice or rain. The mass of water in a certain phase is given by m_p, and the total mass $m_\mathrm{tot} = m_\mathrm{v} + m_\mathrm{l} + m_\mathrm{d}$ is the sum of the mass of the water in all phases and the mass of dry air m_d. In the absence of any sources or losses (like precipitation) of moisture in the atmosphere, the total water specific humidity q_t is a conserved variable and defined as:

$$q_\mathrm{t} = q_\mathrm{v} + q_\mathrm{l}. \tag{2.2}$$

The specific humidity is a key parameter when dealing with clouds, as they consist of water in either liquid or solid form. A cloud can grow when water vapor condensates into a liquid or even solid state, and will shrink or dissipate when liquid or solid water evaporates to water vapor. Condensation and evaporation are continually occurring in air. One speaks of a net condensation when more water molecules arrive at a liquid surface than leave, and a net evaporation for the opposite situation.

The rate at which the two processes occur depends on many factors, one of these being the temperature of the air. Warmer air has more energetic molecules, and molecules with a high energy evaporate more readily. Evaporation decreases more heavily with temperature than condensation. As such, when air cools down enough, there can be a net condensation, allowing cloud droplets to form. This temperature is known as the dew point temperature.

When the amount of water vapor in an air parcel is saturated, water droplets start forming. This occurs when the water vapor pressure e is equal to its saturated value e_sat. The saturation vapor pressure is used to determine a widely used quantity in meteorology, the relative humidity $\mathrm{RH} = e/e_\mathrm{sat}$. For typical temperatures in the ABL, an empirical relation for the saturation vapor pressure is given by (Stull, 1988)

$$e_\mathrm{sat} = 610.78 \exp\left[\frac{17.2694(T - 273.16)}{T - 35.86}\right], \tag{2.3}$$

where T denotes the absolute temperature in Kelvin.

It is more useful to express the water vapor pressure and saturation vapor pressure in terms of specific humidity. When air is saturated, it cannot hold any more water vapor, and therefore $q_\mathrm{v} = q_\mathrm{sat}$. The saturated water vapor specific humidity q_sat can be derived from Dalton's law of partial pressures (Dalton, 1801) and reads

$$q_\mathrm{sat} = \frac{\epsilon e_\mathrm{sat}}{p + e_\mathrm{sat}(\epsilon - 1)}, \tag{2.4}$$

where the dimensionless constant $\epsilon = R_\mathrm{d}/R_\mathrm{v}$ is the ratio of the specific gas constant for dry air R_d and the specific gas constant for water vapor R_v. In the troposphere, the total pressure p is much larger than the saturated water vapor pressure e_sat, and equation 2.4 can be simplified to

$$q_\mathrm{sat} \approx \epsilon \frac{e_\mathrm{sat}}{p}. \tag{2.5}$$

2.1.3. Liquid potential temperature

The potential temperature is denoted by θ and it represents the temperature that a parcel of air would have if it was brought adiabatically from its own temperature and pressure state to a reference pressure state p_0. An expression for the potential temperature of an air parcel in terms of its actual temperature T and pressure p and the reference pressure p_0 is found by using the concept of energy conservation and the ideal gas law ($p = \rho R T$, with ρ the density). The derivation finally finds

$$\theta = T\left(\frac{p_0}{p}\right)^{\frac{R_\mathrm{d}}{c_p}} = \frac{T}{\Pi}, \tag{2.6}$$

which has the specific heat of dry air at constant pressure as c_p, and introduces the Exner function Π:

$$\Pi = \left(\frac{p}{p_0}\right)^{\frac{R_\mathrm{d}}{c_p}}. \tag{2.7}$$

Typically the reference pressure p_0 is taken to be 1000 hPa. The potential temperature of a dry adiabatic parcel is constant with height, making it useful when studying dry atmospheres. In the presence of liquid cloud water, an even more useful quantity is the liquid potential temperature

$$\theta_\mathrm{l} \approx \theta - \frac{L_\mathrm{v}}{c_p \Pi} q_\mathrm{l}, \tag{2.8}$$

which also includes the effect on heat of phase changes of water using L_v, the latent heat release due to condensation of water vapor. As with the total specific humidity, the liquid potential temperature is a conserved value in the absence of loss or gain of humidity in the atmosphere.

The variables q_t and θ_l are conserved in the ABL for vertical adiabatic motions. Moreover, the vigorous turbulent mixing makes these variables more or less constant throughout the boundary layer. The typical vertical profiles of these variables in a stratocumulus-topped boundary layer (STBL) shown in figure 2.1 illustrate this. In the thin inversion layer, θ_l experiences a sharp increase, whilst q_t drastically decreases. In the free troposphere above the inversion layer, θ_l increases constantly with height, and the q_t exhibits a constant decrease with height. The figure also illustrates the vertical profiles for some of the other variables introduced throughout this section. Clearly, these are not conserved within the entire boundary layer, highlighting the practicality of using θ_l and q_t.

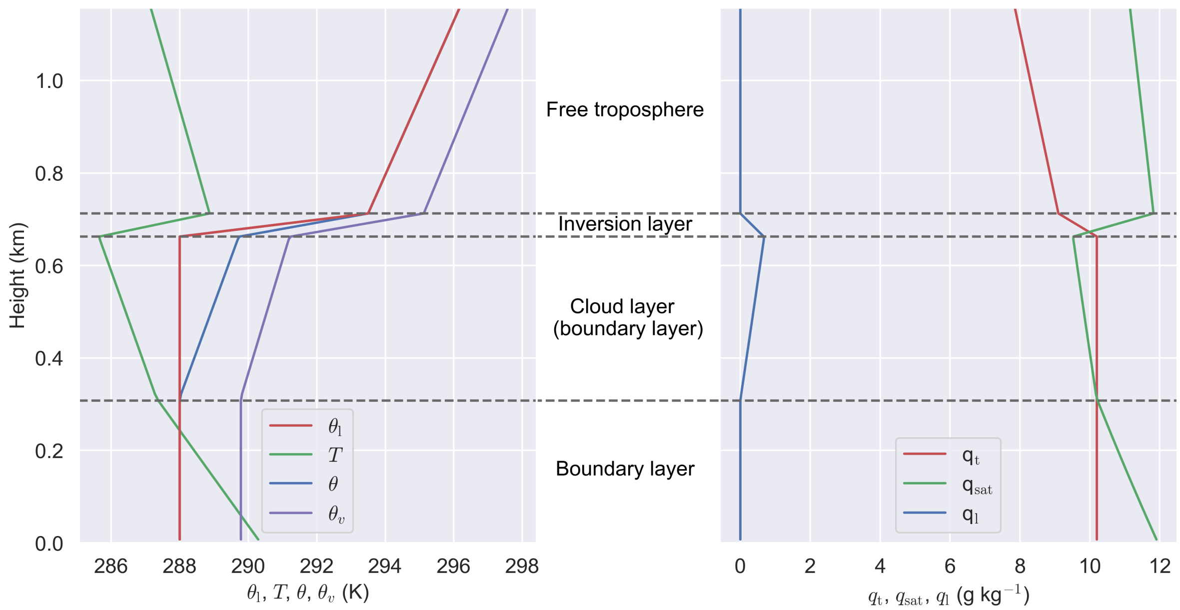


Figure 2.1: The vertical input profiles of θ_l and q_t for the Atlantic Stratocumulus Transition Experiment (ASTEX) case, simulating an STBL, as calculated using the method described in section 5.2. Profiles are only shown for the lowest 1.15 km of the atmosphere but actually extend to approximately 3 km. Vertical profiles for T, θ, θ_v, q_sat and q_l are calculated from the input profiles. The profiles shown are typical for the STBL.

2.1.4. Virtual potential temperature

Buoyancy is one of the main sources of turbulence in fluids. For instance, warm parcels of air are positively buoyant: they rise as their density is lower than that of the surrounding air. The buoyancy of an air parcel can be expressed using the virtual temperature T_v. The virtual temperature is the temperature at which a dry air parcel would have the same density as a parcel containing moisture with temperature T. It is derived using the ideal gas law and Dalton's law of partial pressures and can be denoted as

$$T_v = T(1 + \epsilon_\mathrm{I} q_\mathrm{v} - q_\mathrm{l}), \tag{2.9}$$

where another dimensionless constant $\epsilon_\mathrm{I} = R_\mathrm{v}/R_\mathrm{d} - 1$ is introduced. The virtual potential temperature θ_v is found by dividing T_v by the Exner function (eq. 2.7):

$$\theta_v = \frac{T_v}{\Pi} = \theta(1 + \epsilon_\mathrm{I} q_\mathrm{v} - q_\mathrm{l}). \tag{2.10}$$

Beyond its use for expressing buoyancy, θ_v is useful as it is conserved for unsaturated moist air parcels. This is illustrated in figure 2.1, where the typical profile of θ_v in the cloud-topped ABL has a constant value for heights below the cloud layer.

2.1.5. Cloud albedo

The albedo of a cloud marks the ratio of incident solar radiation a cloud reflects to the total amount of radiation it receives. Solar radiation not reflected by the cloud is either absorbed by it or transmitted towards the Earth. A cloud with a large albedo thus greatly diminishes the amount of available sunlight at the surface, making cloud albedo an important factor in solar forecasting. For shortwave radiation like that from the sun, cloud albedo depends mostly on the cloud optical thickness COT (sometimes τ in literature) and the solar zenith angle ϑ_0 (Wood, 2012). The solar zenith is defined as the angle between the Sun's rays and the vertical direction. For higher solar zenith angles, radiation travels through a larger part of the cloud, making reflection more likely. The cloud optical thickness qualifies the extent to which a cloud prevents radiation from passing through it. In stratocumulus clouds COT is found from the vertical integral of the ratio of cloud liquid water ρq_l to the effective cloud droplet radius r_e (Wood, 2012):

$$\mathrm{COT} = \frac{3}{2\rho_\mathrm{l}} \int_{z_\mathrm{b}}^{z_\mathrm{t}} \frac{\rho q_\mathrm{l}}{r_\mathrm{e}} \mathrm{d}z, \tag{2.11}$$

where ρ gives the density of air and ρ_l gives the density of liquid water. The integration bounds for COT are the cloud base height z_b and cloud top height z_t.

2.1.6. Liquid water path and cloud thickness

The total amount of liquid water in an atmospheric column is given by the liquid water path (LWP):

$$\mathrm{LWP} = \int_{z_\mathrm{b}}^{z_\mathrm{t}} \rho q_\mathrm{l} \mathrm{d}z. \tag{2.12}$$

The expression for the LWP is very similar to the expression of the cloud optical thickness (eq. 2.11), and thus the LWP of a cloud is closely related to its albedo. In this research, the LWP will be taken as a proxy for the cloud albedo.

Satellite observations are frequently used to infer the LWP. In stratocumulus, the fields of θ_l and especially q_t are strongly correlated to the LWP field (de Roode and Los, 2008), as illustrated in figure 2.2. This shows how large the effect of the thermodynamics of the atmosphere on the LWP is.

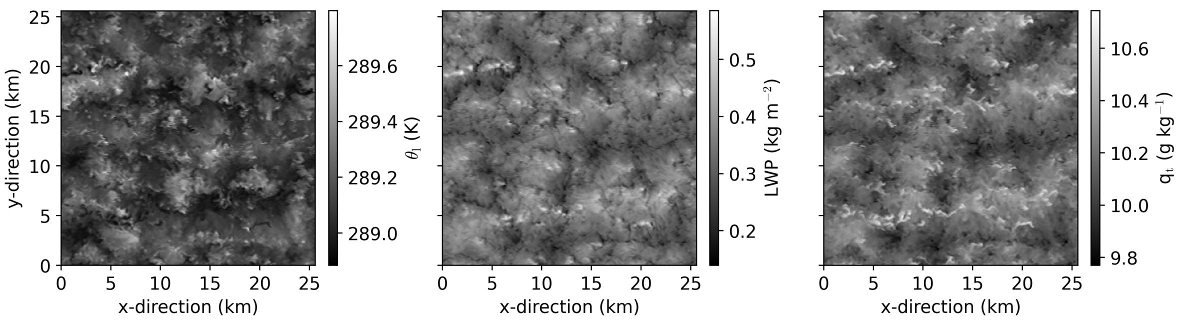


Figure 2.2: Snapshot of the fields of θ_l, LWP and q_t for a DALES run of a stratocumulus case. Details of the DALES run are given in sections 5.1 and 5.2. Fields extracted after 8 hours, and the fields of θ_l and q_t are extracted at a height in the middle of the cloud layer.

2.2. Dynamics

The atmospheric state is determined by processes like radiation, precipitation, large-scale advection, and turbulent motion. In this section, the basic set of equations that govern the dynamics in the atmosphere are given. Before presenting these expressions, it is helpful to introduce Reynolds decomposition, which can be used to decompose turbulent motion from large-scale motion.

Reynolds decomposition splits the temporal or spatial series of a scalar ϕ into two parts: $\phi = \overline{\phi} + \phi'$. Here, $\overline{\phi}$ denotes the mean of the scalar over a certain span, and ϕ' gives the turbulent fluctuation from this mean. The scalar ϕ can refer to the thermodynamic variables presented in the last section, but also to the east-west, south-north, or vertical wind speeds (u, v, and w respectively) in the atmosphere.

Reynolds decomposition and averaging are applied to the ideal gas law and the conservation equations for momentum, mass, heat, and humidity to convert these into equations describing large-scale advection and turbulent flow. In atmospheric flow, molecular processes are negligible with respect to the other terms and are thus omitted from the equations. A more detailed derivation of the expressions in this section can be found in textbooks like Stull (1988).

2.2.1. Mass conservation

In the boundary layer, the incompressibility approximation holds (de Roode, 2021). Consequently, Reynolds decomposition and averaging of the conservation of mass gives

$$\frac{\partial \overline{u_\mathrm{i}}}{\partial x_\mathrm{i}} = 0, \quad \frac{\partial u_\mathrm{i}'}{\partial x_\mathrm{i}} = 0, \tag{2.13}$$

where $u_\mathrm{i} = (u, v, w)$ gives the wind speed, and $x_\mathrm{i} = (x, y, z)$ gives the position in the east-west, south-north, and vertical direction.

Advection terms which usually arise when using Reynolds techniques are given by

$$u_\mathrm{i}' \frac{\partial \phi'}{\partial x_\mathrm{i}}, \tag{2.14}$$

where $\phi \in \{\theta_\mathrm{l}, q_\mathrm{t}, \theta_v, u, v, w\}$. Adding the term $\phi' \partial u_\mathrm{i}' / \partial x_\mathrm{i}$ (which is zero as a result of eq. 2.13) to the above allows one to write:

$$u_\mathrm{i}' \frac{\partial \phi'}{\partial x_\mathrm{i}} + \frac{\phi' \partial u_\mathrm{i}'}{\partial x_\mathrm{i}} = \frac{\partial u_\mathrm{i}' \phi'}{\partial x_\mathrm{i}}. \tag{2.15}$$

This equation is convenient, as turbulent effects are usually represented using the notation on the right-hand side. For example, the bouyancy turbulent flux is expressed as $\overline{w'\theta_v'}$, the total specific humidity turbulent flux as $\overline{w'q_\mathrm{t}'}$ and the vertical velocity variance as $\overline{w'^2}$. Each indicates the change of the specific variable due to turbulent transport.

2.2.2. Momentum conservation

Applying Reynolds decomposition and averaging to the momentum equations, and rewriting using mass conservation, gives

$$\frac{\partial \overline{u_\mathrm{i}}}{\partial t} + \overline{u_\mathrm{j}} \frac{\partial \overline{u_\mathrm{i}}}{\partial x_\mathrm{j}} = -\delta_{\mathrm{i}3} g + f \epsilon_{\mathrm{ij}3} \overline{u_\mathrm{j}} - \frac{1}{\overline{\rho}} \frac{\partial \overline{p}}{\partial x_\mathrm{i}} - \frac{\partial \overline{u_\mathrm{i}' u_\mathrm{j}'}}{\partial x_\mathrm{j}}, \tag{2.16}$$

where the Kronecker-delta $\delta_{\mathrm{i}3}$ ensures the gravitational acceleration g only works in the vertical direction, and the Levi-Civita symbol $\epsilon_{\mathrm{ij}3}$ works on the Coriolis parameter $f = 2\omega \sin\varphi$ with $\omega = 7.27 \times 10^{-5}$ s^{-1} giving the angular velocity of the Earth and φ giving the latitude. When turbulence is horizontally homogeneous, terms containing the derivative of the wind fluctuation with respect to horizontal position are negligibly small in comparison to the other terms in the equations and can be omitted.

Equations can also be found for the wind speed fluctuation. This is especially relevant for the fluctuation in vertical wind speed w', as fluctuations in vertical wind speed determine the magnitude of the turbulent fluxes. In turn, turbulent fluxes are key in the distribution of heat and humidity across the STBL. The fluctuation in vertical wind speed can be determined using

$$\frac{\partial w'}{\partial t} + u_\mathrm{i} \frac{\partial w'}{\partial x_\mathrm{i}} = \frac{\theta_v'}{\overline{\theta_v}} g - \frac{1}{\overline{\rho}} \frac{\partial p'}{\partial z}. \tag{2.17}$$

This equation shows that parcels of air that are warmer than their surroundings (positive θ_v', eq. 2.10) tend to rise through the air, whereas colder air parcels (negative θ_v') tend to sink. This process is crucial to the turbulent mixing in the stratocumulus cloud layer and the ABL as a whole.

2.2.3. Heat conservation

The conservation of energy in terms of θ_l is given by

$$\frac{\partial \overline{\theta_l}}{\partial t} + \overline{u_i}\frac{\partial \overline{\theta_l}}{\partial x_i} = -\frac{1}{\overline{\rho} c_p \Pi}\frac{\partial \overline{F_i}}{\partial x_i} - \frac{\partial \overline{w'\theta_l'}}{\partial z} + S_{\theta_l}, \quad (2.18)$$

where the flux term $\overline{F_i}$ gives heat added to air by processes like radiation, and S_{θ_l} gives source or sink terms like the evaporation of rain. Turbulence is assumed to be horizontally homogeneous.

2.2.4. Water conservation

For q_t, the equation reads

$$\frac{\partial \overline{q_t}}{\partial t} + \overline{u_i}\frac{\partial \overline{q_t}}{\partial x_i} = -\frac{\partial \overline{w'q_t'}}{\partial z} + S_{q_t}, \quad (2.19)$$

which also assumes horizontally homogeneous turbulence. The source term S_{q_t} gives the contribution of ice or precipitation, as ice and rain were not included in the definition of the total specific humidity (eq. 2.2).

2.3. Stratocumulus clouds

Of all cloud types, stratocumulus (fig. 2.3) covers the largest part of the Earth (an annual average of approximately one-fifth of the Earth's surface, (Wood, 2012)). Moreover, it is the cloud type that occurs most often over Western Europe (Warren et al., 1986). Stratocumulus clouds exert a large effect on the Earth's radiative balance (Hartmann et al., 1992). Strikingly, Randall et al. (1984) noted that a 4% increase in global stratocumulus coverage could completely offset the global warming effect. Unfortunately, global climate models suggest that warming of the atmosphere leads to decreased stratocumulus cloud coverage (Dufresne and Bony, 2008).

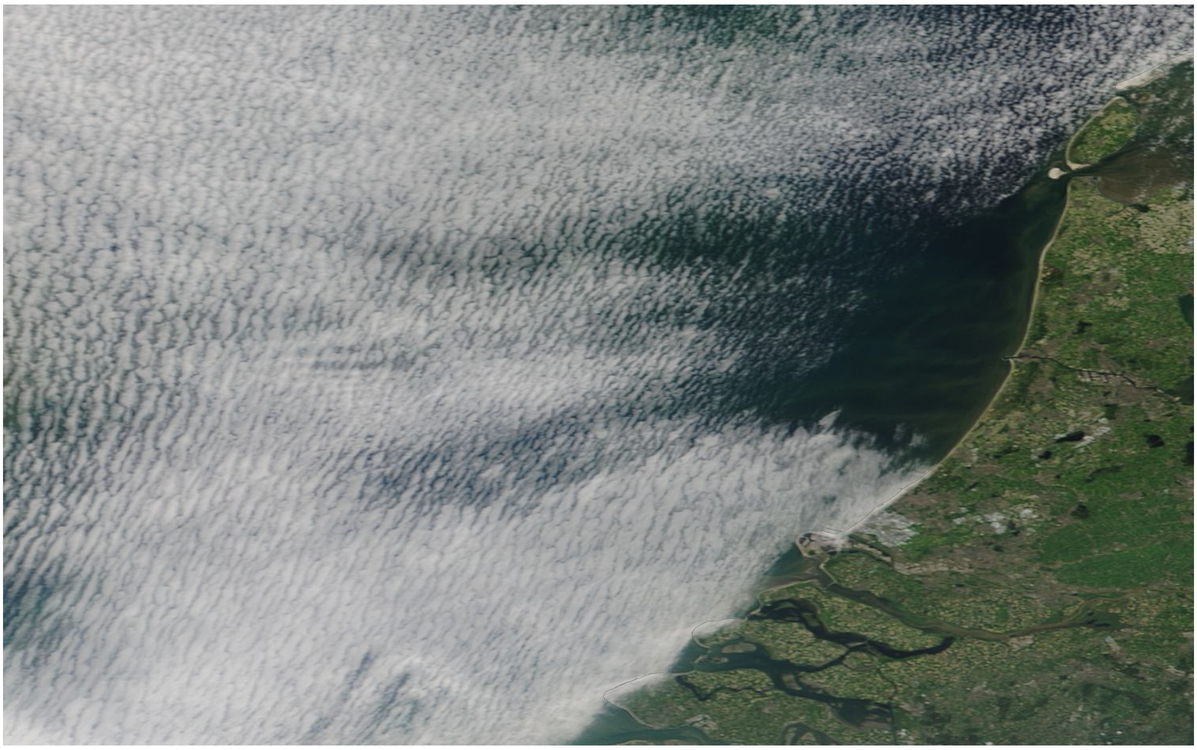


Figure 2.3: Snapshot of a stratocumulus cloud deck off the west coast of the Netherlands. Captured by NASA (2013), on June 6, 2015.

Stratocumulus occurs at the top of the CBL, which is generally found at heights ranging from 500-2000 m (Wood, 2015), where it is capped by the inversion layer. Furthermore, stratocumulus clouds are usually only a few hundred meters thick (Wood, 2012), but can have a cloud fraction of close to one (de Roode, 2021), indicating complete sky coverage. As explained in the introduction (ch. 1), stratocumulus clouds significantly impact solar forecasts due to their high albedo. Moreover, the thin depth of the thermal inversion layer makes stratocumulus difficult to resolve in typical atmospheric models, consequently making them a crucial part of research into the enhancement of solar forecasts.

2.3.1. Occurence

The formation of stratocumulus clouds depends on the presence of two elements. First of all, a strong thermal inversion layer should cap the CBL. A thermal inversion layer can be formed as a result of the downward motion of air in the Hadley circulation (van der Dussen, 2015) or under a high-pressure system. The second necessary element is a cold and moist surface. This is why stratocumulus forms especially often over cool ocean currents. The high humidity of the ocean provides a potent source of moisture in the atmosphere. Due to the cold surface, the thermal inversion layer is strong. Because the boundary layer is well mixed, the humidity is transported upwards. However, the inversion layer traps the moisture inside the boundary layer, causing it to accumulate, eventually saturating the air and forming a stratocumulus cloud which has its top at the inversion layer.

2.3.2. Evolution and advection

This subsection presents the key processes affecting the evolution of stratocumulus, including advection, radiation, precipitation, and turbulence. A schematic of all the key processes is given in figure 2.4.

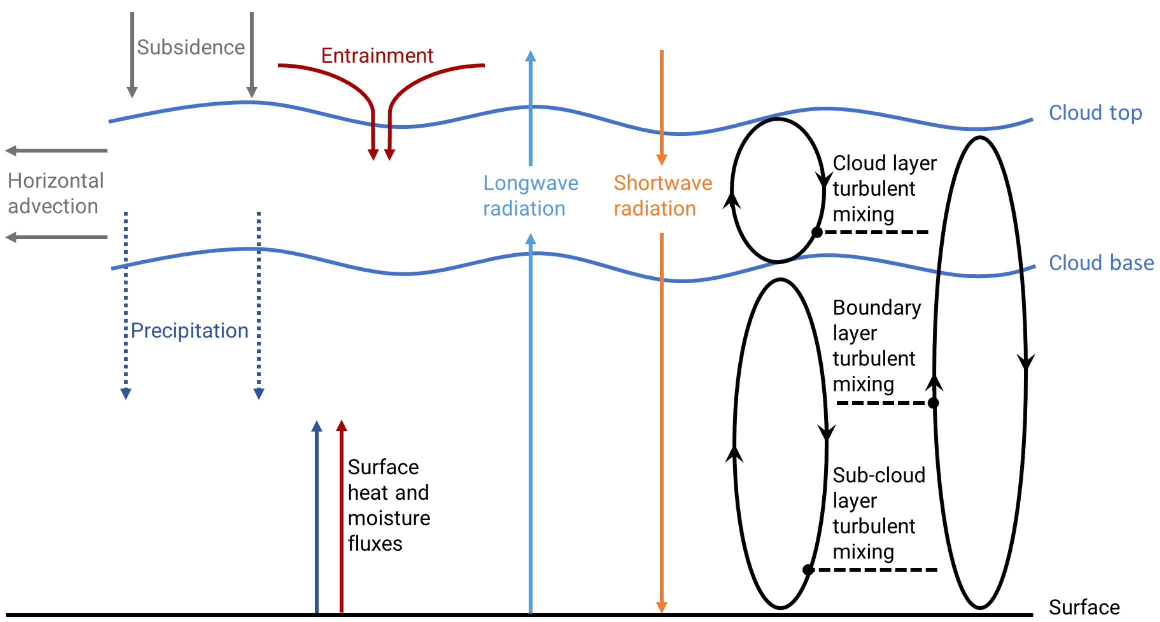


Figure 2.4: Schematic of an STBL and all the processes relevant for its evolution.

Large-scale advection

Large-scale advection is a very relevant process for stratocumulus. For instance, stratocumulus clouds found above the Netherlands are typically formed over the North Sea and then transported over land by horizontal advection. Mean wind speeds $\overline{u}$, $\overline{v}$ and $\overline{w}$ transport clouds through the atmosphere. Typically, the mean vertical wind speed $\overline{w}$ is small in a boundary layer, but not non-zero. A non-zero mean vertical wind speed can cause the boundary layer to either grow or shrink, and this is referred to as large-scale subsidence, often expressed as $\overline{w}_{\mathrm{h}}$. In stratocumulus, subsidence pushes the inversion layer and so the cloud top down, causing the cloud to thin. Horizontal mean wind speeds $\overline{u}$ and $\overline{v}$ generally are much larger than $\overline{w}_{\mathrm{h}}$. Large-scale advection is represented by the second term on the left-hand side in equations 2.16, 2.18 and 2.19.

Radiation
Stratocumulus clouds emit longwave radiation isotropically, in a way similar to that of black bodies (van der Dussen, 2015). At the top of the cloud, the cloud emits more radiation than it receives, causing a radiative cooling. A direct effect of this radiative cooling is that as the cloud layer temperature decreases, q_{sat} decreases (eq. 2.4), leading to more saturated air and causing a thickening of the cloud. Furthermore, the loss of longwave radiation generates cold downdrafts (eq. 2.17). As such, radiative cooling generates turbulence within the cloud that causes the air inside it to be well-mixed. Moisture that evaporates at the surface and enters the cloud through updrafts is thus well distributed across the cloud layer, further causing the cloud layer to thicken.

Shortwave solar radiation also plays a role in the evolution of stratocumulus clouds. A significant part of the solar radiation incident on a stratocumulus cloud deck is reflected (60% on average (Barry and Chorley, 2003), with maxima up to 80% (van der Dussen 2015)). Of the remaining solar radiation, most travels on towards the Earth's surface, but a small part is absorbed by the cloud. This causes a warming tendency which partly counteracts the effect caused by radiative cooling, so that during the day stratocumulus cools and thus thickens less strongly than during the night. Consequently, observations indicating that stratocumulus clouds tend to have greater thickness during the night compared to the daytime (Wood et al., 2002) can be explained.

Entrainment
Turbulence inside the cloud layer can cause warm thermals to penetrate the inversion layer and rise into the free troposphere above. Here, their buoyancy is damped due to the relatively warm surrounding air, and the thermals descend downwards into the cloud layer again. However, during this descent, warm, dry air from the free troposphere is dragged along into the cloud layer, where the air is more cold and moist. This mixing of air from the free troposphere into the boundary layer is known as entrainment. The entrainment velocity is denoted as w_{e}. Entrainment of air into the boundary layer acts to deepen the boundary layer, which would usually contribute to a deepening of the cloud layer. However, the air from the free troposphere works to warm and dry the stratocumulus, and the subsequent evaporation of cloud liquid water leads to a thinner cloud. As cloud thinning by the latter effect usually dominates cloud thickening by a deepening of the boundary layer, entrainment generally causes a net thinning of stratocumulus (Randall, 1984). As the air is descending into the cloud layer ($w' < 0$), and it is warmer ($\theta_{\mathrm{l}}' > 0$) and dryer ($q_{\mathrm{t}}' < 0$), entrainment is usually marked by large, negative liquid potential temperature turbulent flux $\overline{w'\theta_{\mathrm{l}}'}$ and a large, positive $\overline{w'q_{\mathrm{t}}'}$.

Precipitation
Precipitation can sometimes occur in stratocumulus clouds, particularly in relatively thick stratocumulus. The amount of cloud droplets in a unit volume of cloud is known as the cloud droplet number concentration N_{d}. Clouds with a small N_{d} on average have smaller cloud droplets than clouds with large N_{d}. For precipitation to occur, cloud droplets have to grow large so they are heavy enough to fall to the Earth. Stratocumulus clouds have a large N_{d} (~ 150 cm^{-3} over the ocean, Martin et al., 1994), so precipitation does not occur so often, and if it does, the precipitation is usually in the form of drizzle (Comstock et al., 2005, van Zanten et al., 2005). Precipitation removes moisture from the atmosphere (a negative $S_{q_{\mathrm{t}}}$ in eq. 2.19), causing a decrease in total humidity. This causes the cloud layer to thin.

Breakup
On one hand, longwave radiative cooling and evaporation of surface moisture cause the stratocumulus cloud to thicken. On the other hand, large-scale subsidence, entrainment, solar heating, and precipitation work to thin the stratocumulus cloud. Whether a stratocumulus cloud as a whole thins or thickens depends on the balance between all of these processes. Van der Dussen et al. (2014) show how an LWP budget analysis can be done to determine if the cloud thickens or thins. If one of the thinning processes is strong, it can cause the stratocumulus cloud to break up.

Dissipation can also occur when entrainment has deepened the boundary layer to a certain extent. In these situations, heating of the cloud layer with solar radiation during the day can cause the cloud layer to be decoupled from the sub-cloud layer. This cuts the cloud layer off from its moisture source at the surface and causes stratocumulus to dissipate (de Roode et al., 2016). It is then usually replaced by cumulus clouds or a broken stratocumulus cloud with cumulus clouds rising into it (Wood, 2015).

3

Estimating thermodynamic fields from satellite images

The nudging and persistence methods developed in this research make use of fields of θ_l and q_t. This chapter shows how it is possible to use satellite and ground-based observations to estimate these fields. First, it is demonstrated how cloud top and base heights can be approximated from observations. These are used in the approach described next, estimating the liquid water specific humidity fields from an LWP field. Finally, a technique to approximate the total specific humidity and liquid potential temperature fields from the liquid specific water humidity is denoted, which was derived especially for this thesis. Additionally, two 'limiting' cases to this approach are identified.

3.1. Cloud base and top height

By assuming the vertical profile of q_l to be linear, a good approximation within shallow stratocumulus (Wood and Taylor, 2001), the cloud depth can be approximated from the LWP. To find the cloud depth H, first, the cloud base height z_b has to be determined. Saturated water vapor specific humidity depends on the pressure and on the temperature (eq. 2.4), and so it generally decreases as an adiabatic parcel rises. At a certain height, the water vapor specific humidity becomes equal to the saturated water vapor specific humidity ($q_\mathrm{v} = q_\mathrm{sat}(T, p)$), indicating saturation of the parcel. This height is termed the lifting condensation level (LCL), and its value can often be used as a reasonable approximation of the cloud base height (de Roode, 2021). As the water vapor specific humidity q_v is conserved for an unsaturated parcel, the LCL can be evaluated by finding the vertical profile of $q_\mathrm{sat}(T, p)$. To this end, one needs to compute vertical profiles for T and p. In the lower part of the atmosphere, there is no liquid water. When the surface is warmer than the air just above it, the temperature has a constant decrease with height, known as the dry adiabatic lapse rate (Muralikrishna and Manickam, 2017). Combining the dry adiabatic lapse rate with observations of the surface temperature gives a vertical temperature profile. Similarly, a vertical pressure profile can be obtained by using surface pressure observations and the assumption of hydrostatic equilibrium (de Roode, 2021). Using these profiles and equation 2.4, the vertical profile of q_sat can be calculated, allowing for the estimation of the LCL. Exact expressions for the LCL height, temperature, and pressure have been formulated in literature, for example by Romps (2017).

Using the cloud base height, the cloud depth and thus the cloud top height z_t can be determined. Due to the turbulence inside a stratocumulus cloud, it is vertically well mixed. This allows the approximation of a constant water content lapse rate $\mathrm{d}q_\mathrm{l}/\mathrm{d}z$ in the cloud layer. Additionally, the amount of liquid water in the cloud is assumed to be zero at its base and maximum at the top, allowing one to write:

$$q_\mathrm{l}(z) = \alpha(z - z_\mathrm{b}), \quad \text{for } z_\mathrm{b} \leq z \leq z_\mathrm{t}, \tag{3.1}$$

where $\mathrm{d}q_\mathrm{l}/\mathrm{d}z$ is denoted by α. This simple expression of the liquid water specific humidity can be used in the definition of the LWP in equation 2.12. Integration of this expression (assuming a constant total density for air ρ_0) and subsequent introduction of the cloud thickness $H = z_\mathrm{t} - z_\mathrm{b}$ gives

$$\mathrm{LWP} = \frac{1}{2}\alpha\rho_0 H^2, \tag{3.2}$$

which can be rewritten in terms of the unknown H as

$$H = \sqrt{\frac{2\mathrm{LWP}}{\rho_0\alpha}}. \tag{3.3}$$

The last undetermined term left in the expression is the derivative of the liquid water specific humidity with respect to height α. From the definition of total specific humidity, it is derived that

$$\alpha = \frac{dq_\mathrm{t}}{dz} - \frac{dq_\mathrm{v}}{dz} = -\frac{dq_\mathrm{sat}}{dz}, \tag{3.4}$$

where the expression on the far right is found by remembering that q_t is a conserved variable in the atmospheric boundary layer and q_v at cloud level is equal to q_sat. The vertical gradient of q_sat can be approximated by its value at the LCL and a height Δz above the LCL:

$$-\frac{dq_\mathrm{sat}}{dz} = -\frac{q_\mathrm{sat}(T(z_\mathrm{LCL} + \Delta z), p(z_\mathrm{LCL} + \Delta z)) - q_\mathrm{sat}(T(z_\mathrm{LCL}), p(z_\mathrm{LCL}))}{\Delta z}, \tag{3.5}$$

where z_LCL denotes the LCL. Using the moist adiabatic lapse rate (Muralikrishna and Manickam, 2017) and the temperature at the LCL, the temperature at a height above the LCL can be determined. The calculated temperature can then be used to determine q_sat at this height, allowing the calculation of the cloud thickness using equations 3.3-3.5. From there, the calculation of z_t using z_b as given by the LCL is arbitrary. Naturally, α can also be used to approximate the q_l vertical profile throughout the cloud.

3.2. Estimation of the liquid water specific humidity fields

Equation 3.2 gives an expression for the LWP in terms of the cloud base height and the vertical gradient of the liquid water specific humidity. This gradient is assumed to be constant everywhere in the cloud layer. By using Reynolds averaging, one can then find

$$\overline{\mathrm{LWP}} = \frac{1}{2}\alpha\rho_0\overline{H}^2. \tag{3.6}$$

From this assumption, it also follows that q_l has its maximum at the cloud top, and so the maximum mean value of the liquid water specific humidity occurs at the mean cloud top height, or, using the definition in equation 3.1,

$$\overline{q_\mathrm{l}}_\mathrm{max} = \alpha\overline{H}. \tag{3.7}$$

The fluctuation in the cloud top height is assumed to be negligibly small, which means that the maximum liquid water specific humidity for each column can be written as (de Roode and Los, 2008)

$$q_\mathrm{l,max} = \overline{q_\mathrm{l}}_\mathrm{max} + q_\mathrm{l}'. \tag{3.8}$$

It is assumed that q_l' is constant for every height in the column. For every column, it also holds that $q_\mathrm{l,max} = H\alpha$ (eq. 3.1). Inserting this and equation 3.7 into equation 3.8 gives

$$H' = \frac{q_\mathrm{l}'}{\alpha}. \tag{3.9}$$

The expression for LWP can be reformulated as

$$\mathrm{LWP} = \frac{1}{2}\alpha\rho_0(\overline{H} + H')^2, \tag{3.10}$$

written in terms of the mean cloud thickness $\overline{H}$ and the fluctuations of the cloud thickness H'. The fluctuation of the LWP can be found by subtracting the mean LWP from the total LWP, and after rewriting using equation 3.9 it is given by

$$\mathrm{LWP}' = \rho_0\overline{H}q_\mathrm{l}' + \frac{1}{2}\frac{\rho_0}{\alpha}q_\mathrm{l}'^2. \tag{3.11}$$

With the use of standard values for the stratocumulus-topped atmosphere, $\overline{H} = 200$ m, $\alpha = 1.9 \times 10^{-6}$ kg kg^{-1} m^{-1}, $q_1' = 10^{-4}$ kg kg^{-}1, and $\rho_0 = 1.2$ kg m^{-3}, it can be found that the second term on the right-hand side constitutes only about 13% of the first term on the right-hand side. Therefore, it is approximated as

$$\mathrm{LWP}' \approx \rho_0 \overline{H} q_1', \tag{3.12}$$

for simplicity in the calculations. The moist adiabatic lapse rate can be used to determine α (eqs. 3.4 and 3.5), which can in turn be used to determine $\overline{H}$ (eq. 3.6). Then, using the above equation, the q_1' field can be estimated. Combining this field with the vertical profile of $\overline{q_1} = \alpha(z - z_\mathrm{b})$ (eq. 3.1), one finds the estimated field of q_1 for all heights (assuming $q_1 = 0$ outside of the cloud layer). Figure 3.1(b) shows an LWP field as calculated using q_1 as estimated from a reference 'satellite' LWP field (fig. 3.1(a)). Their good liking is also demonstrated by the difference between the two fields (fig. 3.1 (c)), showing the estimated LWP is always slightly higher than the actual LWP.

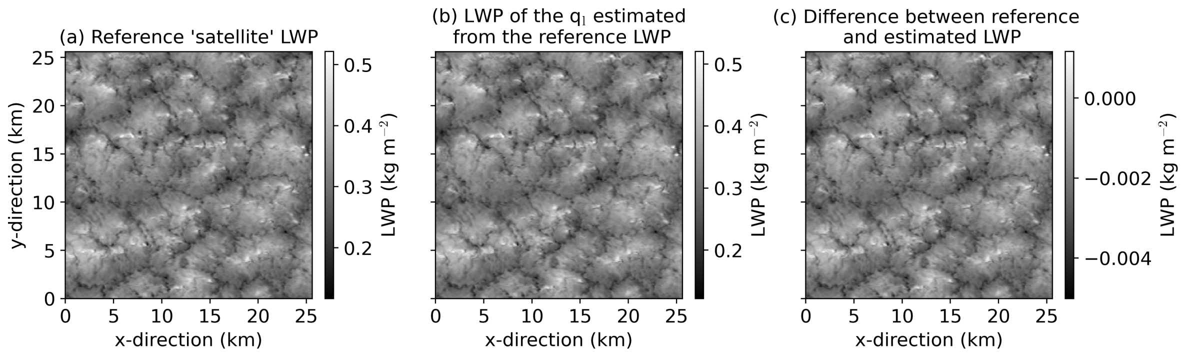


Figure 3.1: Snapshots of a reference LWP field, the LWP field calculated from the q_1 fields estimated using the reference LWP field, and the difference between the two fields. The reference LWP field is taken from a DALES run simulating the STBL but is treated as a satellite observation. Details of the used DALES run are given in sections 5.1 and 5.2.

3.3. Estimation of the thermodynamic fields

In turn, the q_1' field can be used to make estimations of the θ_1' and q_t' fields. The method proposed below does so by assuming a linear relationship between θ_1' and q_t' and estimates the slope in this relation using surface measurements. This is a novel method created especially for this research.

Previously q_1' was assumed to be constant at all heights in the cloud layer. Similarly, θ_1' and q_t' are assumed to be constant at all heights in the boundary layer. Research by de Roode and Los (2008) shows that, in the stratocumulus cloud layer, the fluctuations of the temperature and the total specific humidity satisfy an approximately linear relationship. The same investigation was done for the LES results used in this research. Figure 3.2 shows the resulting scatter plot of T' and q_t' in the middle of the cloud layer, also finding an approximate linear relation. The relation can be used to estimate the θ_1' and q_t' fields:

$$T' = c_{qT} q_\mathrm{t}', \tag{3.13}$$

where c_{qT} is a constant relating the two variables. As DALES uses the liquid potential temperature as a prognostic variable instead of the temperature, T' should be converted to θ_1'. Fluctuations of T can be expressed in θ_1' by applying Reynolds decomposition and averaging on equation 2.8:

$$T' \approx \theta' \approx \theta_1' + \frac{L_\mathrm{v}}{c_p} q_1'. \tag{3.14}$$

Here it is also assumed that $\Pi \approx 1$, a reasonable approximation for within the boundary layer (de Roode and Los, 2008, de Roode, 2021). To replace q_1' in this equation, the expression for the q_t (eq. 2.2) is rewritten in a similar way giving

$$q_1' = q_\mathrm{t}' - q_\mathrm{v}' = q_\mathrm{t}' - q_\mathrm{sat}'(T). \tag{3.15}$$

The fluctuations of the saturation specific humidity can be written in terms of the fluctuations of the temperature by using the Clausius-Clapeyron relation (Nicholls, 1984):

$$q'_{\mathrm{sat}} = \left(\frac{dq_{\mathrm{sat}}}{dT}\right)T' = \gamma T', \tag{3.16}$$

with $\gamma = dq_{\mathrm{sat}}/dT$. The expression for q'_1 (eq. 3.15) can then be rewritten using the Clausius-Clapeyron relation and the relationship between T' and q'_{t}, yielding

$$q'_1 = q'_{\mathrm{t}}(1 - \gamma c_{qT}) = \beta q'_{\mathrm{t}}, \tag{3.17}$$

where the β factor is introduced as $\beta = (1 - \gamma c_{qT})$. Note that this equation can be used to determine the q'_{t} field from q_1'. Finally, equations 3.13 and 3.17 are used to rewrite equation 3.14, finding a relation between q'_{t} and θ'_1

$$\theta'_1 \approx q'_{\mathrm{t}}(c_{qT} - \frac{L_{\mathrm{v}}}{c_p}\beta), \tag{3.18}$$

The method above now fully describes how to convert an LWP field as observed by satellites into thermodynamic fluctuation fields. In the STBL, the mean values for θ_1 and q_{t} are approximately constant (fig. 2.1) with height. Measurements of these variables near the surface can therefore be used to estimate their mean vertical profile. By adding the fluctuation fields to the mean vertical profiles, one can find the total thermodynamic fields at all heights.

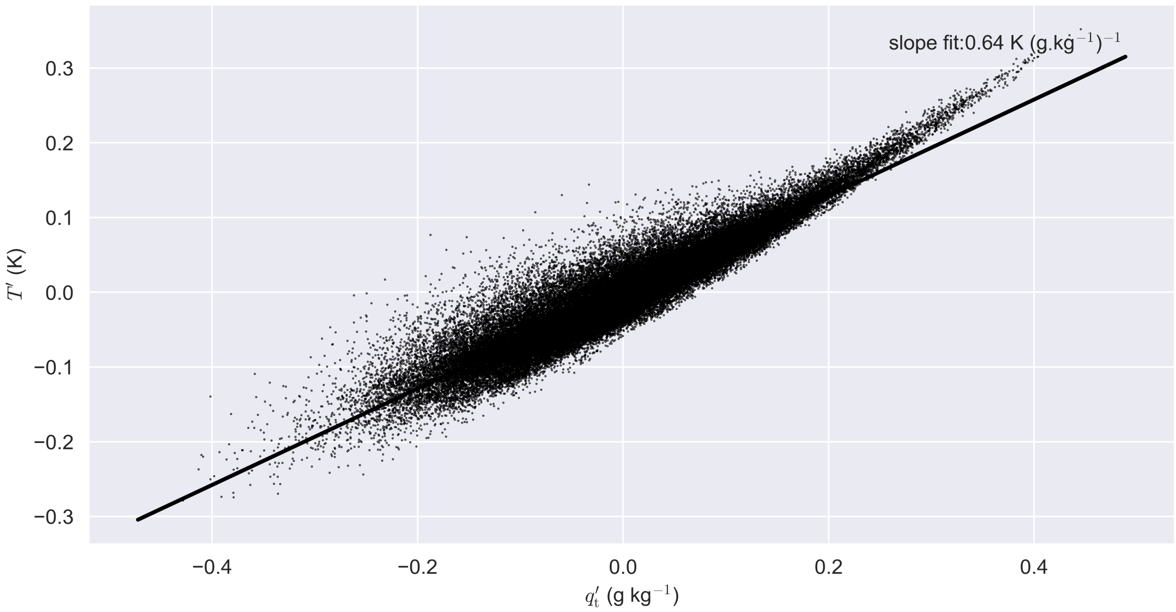


Figure 3.2: Scatter plot of T' and q'_{t} in the middle of the cloud layer. The plot has a fitted linear regression, shown by the black lines, the slopes of which are also given. Results for the fluctuations are obtained from a stratocumulus case LES run.

Realistically, the constant c_{qT} should be determined from observations instead of the slope of LES results. To this end equation 3.18 is rewritten to find

$$c_{qT} = \frac{\frac{\theta'_1}{q'_{\mathrm{t}}} + \frac{L_{\mathrm{v}}}{c_p}}{1 + \gamma\frac{L_{\mathrm{v}}}{c_p}}, \tag{3.19}$$

which leaves the unknown term $\theta'_1/q'_{\mathrm{t}}$. As the fluctuations are assumed to be constant at all heights, this term can be rewritten:

$$\frac{\theta'_1}{q'_{\mathrm{t}}} = \frac{\theta'_{1,\mathrm{sfc}}}{q'_{\mathrm{t,sfc}}} = \frac{\overline{w'\theta'_1}_{\mathrm{sfc}}}{\overline{w'q'_{\mathrm{t}}}_{\mathrm{sfc}}} = \frac{\mathrm{SHF_{sfc}}}{\mathrm{LHF_{sfc}}}\frac{L_{\mathrm{v}}}{c_p} = c_{\mathrm{sfc}}, \tag{3.20}$$

where the constant c_{sfc} is introduced for convenience and the subscript 'sfc' denotes the surface value of a variable. SHF and LHF give the sensible and latent heat fluxes, respectively. With this new term, equation 3.19 becomes

$$c_{qT} = \frac{c_{\mathrm{sfc}} + \frac{L_\mathrm{v}}{c_p}}{1 + \gamma \frac{L_\mathrm{v}}{c_p}}. \tag{3.21}$$

It is possible to measure both SHF and LHF the surface (Large and Pond, 1982, National Ecological Observatory Network (NEON), 2023), and thus it is possible to estimate the surface constant c_{sfc}, and from it c_{qT}, based on ground observations. Therefore, the fields of q_t and θ_l for replicating an observed LWP field can be estimated from observations only. The linear regression fit scatter plot of the LES case used in figure 3.2 gives a slope of $c_{qT} \approx$ 0.64 K (g kg^{-1})$^{-1}$. Using the above method on the SHF and LHF observed in the same LES case, one finds $c_{qT} \approx$ 0.85 K (g kg^{-1})$^{-1}$. Given the assumptions made, the value found from observations is a fair estimation of the actual value. The approach described above will be referred to as the 'linear estimation' in the remainder of this thesis.

To investigate the accuracy of the above method for determining θ_l' and q_t' from the LWP, two 'limiting' cases can be defined, as proposed by van der Dussen et al. (2014). In these, the fluctuation of one of the variables is assumed to be zero, thus making the fluctuation in the other variable fully responsible for fluctuations in LWP. When $\theta_\mathrm{l}' = 0$

$$q_\mathrm{l}' = \eta q_\mathrm{t}', \tag{3.22}$$

and when $q_\mathrm{t}' = 0$

$$q_\mathrm{l}' = -\eta \gamma \theta_\mathrm{l}', \tag{3.23}$$

with $(\eta = 1 + \gamma L_\mathrm{v}/c_p)^{-1}$. Introducing these limiting cases is useful, as the method described earlier in this section should, if accurate, outperform these limiting cases. The limiting approaches will from hereon be referred to as the 'θ_l' = 0 limit estimation' and the 'q_t' = 0 limit estimation'.

4

DALES and its governing equations

This chapter describes DALES, the LES model used in this study. It also notes the governing equations in DALES.

4.1. General description

The Dutch Atmospheric Large Eddy Simulation (DALES) model is a model written in the computer language Fortran based on filtered versions of the equations presented in chapter 2. DALES is an open-source model, maintained and used in research by the KNMI, the University of Wageningen, and of course by Delft University of Technology.

The prognostic variables in DALES are θ_l, q_t, u_i, and the subfilter-scale turbulence kinetic energy (SFS-TKE) e, the latter of which stems from the filtering and is applied for subgrid paramaterizations. As an input, DALES requires an initial state for these variables. This initial state can be determined using observations of the wind speed, pressure, temperature, and humidity (see sec. 5.2). The initial values are then propagated in time by numerically solving the governing equations. Integration in time is done using a third-order Runge-Kutta scheme (Heus et al., 2010). For a more detailed description of the mechanics and applications of the model than that given here, interested readers are referred to the paper by Heus et al. (2010).

4.2. Governing equations

The computation of the prognostic variables in DALES is divided into a resolved part and a sub-grid part. Physical processes that occur on scales larger than the DALES grid size are resolved, while smaller-scale processes are parameterized. The division between the two scales is referred to as filtering. Resolved quantities in the DALES governing equations are denoted by a tilde.

4.2.1. Momentum

The DALES governing equations for the wind speeds are found by applying the LES filter to equations 2.13 and 2.16, giving (Heus et al., 2010):

$$\frac{\partial \widetilde{u}_\mathrm{i}}{\partial x_\mathrm{i}} = 0, \tag{4.1}$$

$$\frac{\partial \widetilde{u}_\mathrm{i}}{\partial t} = -\frac{\partial \widetilde{u}_\mathrm{i}\widetilde{u}_\mathrm{j}}{\partial x_\mathrm{j}} - \frac{\partial \pi}{\partial x_\mathrm{i}} + \frac{g}{\theta_0}\widetilde{\theta_v}\delta_{\mathrm{i}3} + \mathcal{F}_\mathrm{i} - \frac{\partial \nu_\mathrm{ij}}{\partial x_\mathrm{j}}. \tag{4.2}$$

A modified pressure term is denoted by $\pi = \widetilde{p}/p_0 + 2/3\mathrm{e}$, and θ_0 and p_0 give θ_l and p at the surface. Furthermore, $\mathcal{F}_\mathrm{i}$ gives other forcings like large-scale forcings or the Coriolis acceleration, and ν_ij gives the deviatoric sub-grid momentum flux (Heus et al., 2010):

$$\nu_\mathrm{ij} = \widetilde{u_\mathrm{i}u_\mathrm{j}} - \widetilde{u}_\mathrm{i}\widetilde{u}_\mathrm{j} - \frac{2}{3}\mathrm{e}. \tag{4.3}$$

4.2.2. Thermodynamics

The LES filter is also applied to equations 2.18 and 2.19, giving the governing equations for the thermodynamic variables $\phi \in \{\theta_\mathrm{l}, q_\mathrm{t}\}$ (Heus et al., 2010):

$$\frac{\partial \widetilde{\phi}}{\partial t} = -\frac{\partial \widetilde{u}_\mathrm{i}\widetilde{\phi}}{\partial x_\mathrm{i}} - \frac{\partial \widetilde{u_\mathrm{i}''\phi''}}{\partial x_\mathrm{i}} + S_\phi, \tag{4.4}$$

where $\widetilde{u_\mathrm{i}''\phi''} = \widetilde{u_\mathrm{i}\phi} - \widetilde{u}_\mathrm{i}\widetilde{\phi}$ gives the SFS scalar flux. The term S_ϕ gives additional source terms for the variable, like radiation or microphysics.

4.2.3. Subfilter-scale model

To determine the subgrid fluxes in DALES, modeling through eddy diffusivity is used (Deardorff, 1980):

$$\nu_\mathrm{ij} = -K_\mathrm{m}\left(\frac{\partial \widetilde{u}_\mathrm{i}}{\partial x_\mathrm{j}} + \frac{\partial \widetilde{u}_\mathrm{j}}{\partial x_\mathrm{i}}\right), \tag{4.5}$$

$$\widetilde{u_\mathrm{i}''\phi''} = -K_\mathrm{h}\frac{\partial \widetilde{\phi}}{\partial x_\mathrm{i}}, \tag{4.6}$$

where K_m gives the eddy viscosity for momentum and K_h the eddy diffusivity for thermodynamic scalars ϕ (de Roode et al., 2017). As a function of the SFS-TKE, $\mathrm{e} = 1/2(\widetilde{u_\mathrm{i}u_\mathrm{i}} - \widetilde{u}_\mathrm{i}\widetilde{u}_\mathrm{i})$, these coefficients are determined respectively by (de Roode et al., 2017)

$$K_\mathrm{m} = c_\mathrm{m}\lambda \mathrm{e}^{\frac{1}{2}}, \tag{4.7}$$

$$K_\mathrm{h} = c_\mathrm{h}\lambda \mathrm{e}^{\frac{1}{2}}, \tag{4.8}$$

with c_m and c_h giving proportionality constants. The momentum proportionality constant is determined by

$$c_\mathrm{m} = \frac{c_\mathrm{f}}{2\pi}\left(\frac{3}{2}\alpha_\mathrm{K}\right)^{-\frac{3}{2}}, \tag{4.9}$$

where $\alpha_\mathrm{K} = 1.5$ is the Kolmogorov constant and $c_\mathrm{f} = 2.5$ is the filter width. Additionally, the two proportionality constants are related through the turbulent SFS Prandtl number Pr_T:

$$c_\mathrm{h} = \frac{c_\mathrm{h}}{\mathrm{Pr}_\mathrm{T}}. \tag{4.10}$$

In default DALES, $\mathrm{Pr}_\mathrm{T} = 1/3$. The characteristic length scale of SFS turbulent eddies λ is equal to the geometric mean l_Δ

$$\lambda = l_\Delta = (\Delta x \Delta y \Delta z)^{\frac{1}{3}}, \tag{4.11}$$

where Δx, Δy and Δz are the mesh sizes for the respective directions. Now, all that is needed to determine the SFS fluxes is the SFS-TKE. In DALES, the prognostic equation for the square root of e, after parameterization, is (Heus et al., 2010):

$$\frac{\partial \mathrm{e}^{\frac{1}{2}}}{\partial t} = -\widetilde{u}_\mathrm{i}\frac{\partial \mathrm{e}^{\frac{1}{2}}}{\partial x_\mathrm{i}} + \frac{1}{2\mathrm{e}^{\frac{1}{2}}}\left[K_\mathrm{m}\left(\frac{\partial \widetilde{u}_\mathrm{i}}{\partial x_\mathrm{j}} + \frac{\partial \widetilde{u}_\mathrm{j}}{\partial x_\mathrm{i}}\right)\frac{\partial \widetilde{u}_\mathrm{j}}{\partial x_\mathrm{i}} - K_\mathrm{h}\frac{g}{\theta_0}\frac{\partial \widetilde{\theta_v}}{\partial z}\right] + \frac{\partial}{\partial x_\mathrm{i}}\left(2K_\mathrm{m}\frac{\partial \mathrm{e}^{\frac{1}{2}}}{\partial x_\mathrm{i}}\right) - \frac{c_\epsilon \mathrm{e}}{2\lambda}, \tag{4.12}$$

where $c_\epsilon \approx 0.745$ gives the proportionality constant for viscous dissipation. So, by numerically solving equation 4.12, the subgrid flux terms for the momentum and thermodynamic prognostic equations can be found, allowing one to solve for these variables.

5 Methods and experimental set-up

This chapter showcases all the different ways in which experiments are done in DALES. First, it describes how a reference LES run was done to serve as 'observations' to initialize other experiments and eventually investigate their results. The reference run used in the present study is taken from van der Dussen (2015), who used observations from the Atlantic Stratocumulus Transition Experiment (ASTEX) to set up an LES model intercomparison case. The subsequent section gives all initial conditions for this case. Next, it is highlighted how regular LES experiments were set up to compare the new nudging and persistence methods to. This is followed by a description of how the persistence method was added to DALES and how persistence experiments were set up. In the next section, the same is done, but for the nudging method. Finally, the ways in which the experimental results are evaluated and compared are discussed.

5.1. Reference

A DALES run needs to be initialized with input such as the vertical profiles of θ_l, q_t, u, and v. To use the persistence and nudging methods as introduced later in this chapter, one also needs to supply the fields of θ_l and q_t. Finally, to evaluate the performance of a forecasting experiment, the results have to be compared to the actual situation for which the experiments are making predictions. In an ideal situation, all of this information is extracted from a combination of ground-based and satellite observations. However, such observations are not very detailed and generally lack information on the dynamics of the atmosphere, necessary for checking the turbulence in the experiments. Instead of using observations, this research therefore makes use of results from an LES run, from here on referred to as a 'reference' run. The advantages of using LES results instead of observations are that its dynamic fields can be used to check the turbulence of the experiments, and it gives the full 3D thermodynamic fields. This approach also has a flip side: extracting fields with the same amount of detail from observations is impossible. Hence, the performance of the used methods for actual solar forecasting will be lower than that shown in this research.

5.1.1. Experimental set-up

The reference run, which will provide the input to initialize the experiments, also needs inputs to be initialized. In this research, experiments are done using the LES model ASTEX intercomparison case created by van der Dussen (2015), which is based on observations of marine stratocumulus. Section 5.2 highlights the initial conditions for the ASTEX case. The ASTEX case STBL is evolving in time. This makes it an especially interesting case, as it allows one to investigate how well the different methods can predict changing a stratocumulus cloud structure.

DALES also allows the user to turn on and off or change certain calculation schemes such as those for radiation or microphysics, in the input file called 'namoptions'. The complete 'namoptions' file used for the ASTEX reference case can be found in appendix A. All runs are done using a simple radiation scheme that does not include interaction between cloud and solar radiation, greatly reducing computation time. Because solar radiation prediction is the topic of research, the LWP will be considered as a proxy for surface solar radiation.

Two different horizontal domain sizes are used in the experiments. To save computation time, a smaller domain was used for extensively testing the developed methods before applying these to a larger domain. The small domain covers an area of 3.2 by 3.2 km^2, with 64 grid points in each horizontal direction, resulting in a grid point size of 50 by 50 m^2. For the small domain, the reference run was simulated for 5 hours, which approximately took 52 hours of wall clock time using 8 processors. The larger domain consists of 256 grid points in each horizontal direction, each point covering 100 m in both directions, for a total horizontal area of 25.6 by 25.6 km^2. On the large domain, the total reference run time was 8 hours, taking roughly 180 hours of wall clock time using 32 processors. Figure 2.2 shows an example of the θ_l, LWP, and q_t fields as given by a reference run of an STBL on the larger domain. When describing the experimental setup for the remaining methods used in this thesis, only the setup for the larger domain will be discussed.

5.2. Initial conditions

A slab mean vertical profile for the thermodynamic variables is determined by matching measurements with the typical vertical profile and the observed cloud base of the atmosphere.

The input profiles used to set up the marine STBL reference run are based on the observations done during the second flight of the ASTEX first Lagrangian (van der Dussen, 2015). The input profiles of variables $\phi = \{\theta_l, q_t, u, v\}$ are given by

$$\phi(z) = \begin{cases} \phi_{ml} & z \leq z_i \\ \phi_{ml} + \Delta\phi\,(z - z_i)\,/\Delta z_i & z_i < z \leq z_i + \Delta z_i \\ \phi_{ml} + \Delta\phi + \Gamma_\phi\,(z - z_i - \Delta z_i) & z_i + \Delta z_i < z \leq 2\ \mathrm{km} \end{cases} \tag{5.1}$$

where the initial values ϕ_{ml}, $\Delta\phi$ and Γ_ϕ denote the mixed layer value, the inversion jump, and the free atmospheric gradient of the variable ϕ respectively. Table 5.1 presents an overview of these parameters for each variable, as used for the initial profiles of the reference case. Moreover, z_i = 662.5 m denotes the initial base height of the inversion layer, and Δz_i = 50 m denotes the initial thickness of the inversion layer. Figure 2.1 shows the initial ASTEX profiles for θ_l and q_t.

For the ASTEX profiles, the vertical grid is not equidistant and has 427 points in total. The first vertical point is defined at 7.5 m, and the step size for the lowest 500 m is a constant 15 m. Over the next 15 vertical points, the grid size changes gradually from 15 m to 5 m. The grid size remains at 5 m until the last 25 points, where the grid size gradually increases to about 60 m. The geostrophic wind speeds are set to be constant with height and equal u_g = −2 m s^{-1} and v_g = −10 m s^{-1}, and the surface pressure is constant in time and equal to 1029.0 hPa.

When a regular DALES run is started, pseudo-random perturbations are added to the fields created for ϕ by using the input slab mean vertical profiles, as described in detail in the introduction (ch. 1). The perturbations force the build-up of turbulence in the spin-up period. After approximately 2 hours, the turbulence has developed to a quasi-steady state. Results found during the spin-up period are usually discarded due to the nonphysical turbulence.

Variable (unit)	ϕ_{ml}	$\Delta\phi$	Γ_ϕ (km^{-1})
θ_l (K)	288.0	5.5	6.0
q_t (g kg^{-1})	10.2	-1.1	-2.8
u (m s^{-1})	-0.7	-1.3	0.0
v (m s^{-1})	-10.0	0.0	0.0

Table 5.1: Parameters for the calculation of the ASTEX case input profiles of the liquid potential temperature, total specific humidity, east-west wind speed, and south-north wind speed. Values in the table are copied from van der Dussen (2015).

5.3. Standard DALES runs

In this research, the 3D-nudging method as well as the persistence method are added to DALES. They aim to perform better than a standard DALES run. Therefore, standard DALES runs should be included in the skill comparison. The standard LES run is initialized using the results from the reference run, starting at the moment when these results are obtained. Consequently, no turbulence is yet present in the system, and the runs have to build it up. Turbulence in this period will be non-physical, causing unrealistic results to be found. As the runs have no turbulence initially, they are sometimes referred to as 'cold or 'cold start' runs.

Cold runs are initialized in two ways. The first uses the default DALES, pseudo-random initialization. Reference profiles at 7 hours were taken and converted to input profiles using the equations in section 5.2. These profiles are used as the input for this case. It is started in the regular DALES way using pseudo-random perturbations and thus needs to go through an initial spin-up period (fig. 1.1). Such runs will be referred to in this thesis as 'default spin-up' runs.

A second, more advanced way of initializing a cold start run is by setting the initial fields of θ_1 and q_t to be equal to the 3D reference run fields at 7 hours. Then, without the initial pseudo-random perturbations, the cold case starts with the correct thermodynamic fields. However, the lack of physical turbulence initially will likely still impact the evolution of its thermodynamic fields. Simulations with such initialization are referred to as '3D initial thermodynamics' runs. Both standard DALES cases have a total run time of 1 hour, simulating from hour 7 to hour 8 of the reference run.

5.4. Persistence

As discussed in the introduction, many current solar forecasts for short horizons are created using persistence methods (Srikrishnan et al., 2017). This modeling method stems from the frozen field hypothesis of Taylor (1938) which states that local changes are a product of just the wind stream velocity, provided it is much greater than the turbulent velocity. Solar forecasts based on this theory have proven more accurate than NWP solar forecasts on a horizon of 5 hours or shorter (Law et al., 2014, Wang et al., 2019). So, to properly assess the performance of a solar forecast with nudging, it should be compared to the performance of persistence solar forecasts. Prior to this research, DALES did not include a form of persistence modeling. Consequently, a persistence extension of DALES was developed in this study. This method serves as the benchmark solar forecast to compare the nudging results to. Below, an overview of the adapted equations in DALES is given.

If only advection is considered, the evolution of the thermodynamic properties $\phi \in \{\theta_1, q_t\}$ in time reads

$$\frac{\partial \phi}{\partial t} = -u_{\mathrm{hor}} \frac{\partial \phi}{\partial x_{\mathrm{hor}}}, \tag{5.2}$$

where the subscript 'hor' shows that only advection in the horizontal directions is considered ($u_{\mathrm{hor}} = (u, v)$). The vertical wind speed w is assumed to be zero. So, in this case, transport by horizontal wind is the only factor that can cause changes in the temperature or humidity. It is furthermore assumed that the horizontal wind speeds are constant throughout time ($\partial u_{\mathrm{hor}}/\partial t = 0$), and also that there are no fluctuations in the wind speed fields ($u'_{\mathrm{hor}} = 0$). The constant horizontal velocities are set equal to the mean value of the horizontal velocities in the cloud layer, as found in the reference results. Putting this mathematically:

$$u_{\mathrm{hor}}(\boldsymbol{x}, t) = c_{\mathrm{hor}}, \tag{5.3}$$

with $c_{\mathrm{hor}} = (\overline{u}_{\mathrm{cld}}, \overline{v}_{\mathrm{cld}})$, where the subscript 'cld' shows that the mean has been extracted over the cloud layer. An input switch was added to DALES which, if turned on, makes it use the calculations as described above. Moreover, in the persistence version of DALES, the user should supply the initial fields of θ_1 and q_t.

The persistence method runs on the large domain are initialized using the reference fields of θ_1 and q_t and the reference boundary layer horizontal velocities at the 7-hour mark. The total run time of the simulations is one hour, so they run until 8 hours reference time.

5.5. Nudging

As was explained before, turbulence build-up in DALES can cause the mean values of θ_1 and q_t as well as the horizontal fluctuations from the mean in an LES run to drift from observations. Hence, due to the lack of realistic turbulence, there will be an error in the results after spin-up. Among other researchers (ch. 1), Blossey et al. (2013) have implemented a one-dimensional nudging forcing term into various LES models, including DALES. This 1D-nudging pushes the thermodynamic quantities at a certain height in a column to the desired mean value. In default DALES, one-dimensional nudging is applied either to keep the mean state close to the desired mean state or to filter out gravity waves. The latter happens far above the boundary layer, outside the area of interest. Therefore, it does not interfere with the turbulence inside the boundary layer. Nudging to keep the state close to the desired state can be applied throughout the entire domain.

Mathematically, one-dimensional nudging $S_{\phi}^{\mathrm{n},1D}$ is given by:

$$S_{\phi}^{\mathrm{n},1D}(x,y,z,t) = -\frac{\overline{\phi}(z,t) - \overline{\phi_{\mathrm{n},1D}}(z)}{\tau_{1D}}, \tag{5.4}$$

where $\phi = \{\theta_{\mathrm{l}}, q_{\mathrm{t}}\}$ gives the considered flow variable, $\overline{\phi}$ gives the horizontally averaged value for ϕ in the simulation and $\overline{\phi_{\mathrm{n},1D}}$ gives the desired horizontal average value towards which nudging takes place. Finally, τ_{1D} gives the nudging time scale, which determines how strong the nudging effect is. The smaller the value for τ_{1D}, the larger the overall effect of nudging on the tendency of the thermodynamics. $S_{\phi}^{\mathrm{n},1D}$ is added to equation 4.4 as a source term. Default DALES also allows 1D-nudging of u, v, and w, but this is not used in this thesis.

One-dimensional nudging pushes the computed mean state to the desired mean state. This does not take the fluctuation from the mean that a horizontal point can have into account. Therefore, this research introduces a three-dimensional nudging during the spin-up period of DALES, which causes the fluctuations of the fields to be nudged to a desired 3D fluctuation $\phi'_{\mathrm{n},3D}(x,y,z)$. This is implemented in a similar way as the one-dimensional nudging, by adding an additional source term $S_{\phi}^{\mathrm{n},3D}$ to the tendency of the thermodynamic variables:

$$S_{\phi}^{\mathrm{n},3D}(x,y,z,t) = -\frac{\phi(x,y,z,t) - (\overline{\phi}(z,t) + \phi'_{\mathrm{n},3D}(x,y,z))}{\tau_{3D}}, \tag{5.5}$$

where $\phi'_{\mathrm{n},3D}$ is added to the LES horizontal slab average $\overline{\phi}$ before subtracting it from the actual value of ϕ at that position. An important difference between equations 5.4 and 5.5 is that the nudging term present in 5.4 has a value that is independent of horizontal position, whereas the nudging term in 5.5 does depend on the horizontal position.

A separate nudging time scale for the three-dimensional nudging is given by τ_{3D}. It is important to stress that the values for τ_{1D} and τ_{3D} are allowed to differ and can be chosen independently. As such, the effect of applying different nudging strengths for the 1D and 3D nudging in the same run can be investigated. However, the experiments in this research often use the same value for both nudging time scales. In this case, $\tau_{1D} = \tau_{3D} = \tau$, and the nudging terms can be combined to find:

$$S_{\phi}^{\mathrm{n},1D}(x,y,z,t) + S_{\phi}^{\mathrm{n},3D}(x,y,z,t) = -\frac{\phi(x,y,z,t) - (\overline{\phi_{\mathrm{n},1D}}(z) + \phi'_{\mathrm{n},3D}(x,y,z))}{\tau}, \tag{5.6}$$

where the addition of $\overline{\phi_{\mathrm{n},1D}}$ and $\phi'_{\mathrm{n},3D}$ gives the desired 3D field to be nudged to, in this case the reference field. When experiments in this research use the same nudging time scales for 1D and 3D-nudging, they are simply denoted as τ.

Some final characteristics of the nudging method should be mentioned. Nudging runs are started two hours before the specified times at which the desired reference fields are extracted, to allow the runs to go through the spin-up period with nudging. Nudging is only applied during the first 2 hours, however, it does not have to be activated immediately at the run start, and can also be activated at a later time during this period. Runs that receive nudging for the full 2 hours are referred to in this thesis as '2 hour nudge' runs and runs that receive nudging for a shorter time period as 'short nudge' runs. After two hours nudging is deactivated, allowing the simulation to evolve according to the governing equations only. Additionally, 3D-nudging is only applied until a height some tens of meters above the inversion layer. The region above does not significantly affect the STBL and is not considered in this thesis. Finally, it is possible to pass fields and profiles for multiple time points to the model. When this is done, the model nudges towards each given state until the specified time point to which they belong, after which nudging is done towards the next state. Such runs will be referred to as 'multiple time fields nudge' runs.

5.5.1. Experimental set-up

In this research, nudge runs are started at 5 hours of reference time. For their initialization, input profiles are calculated using the vertical profiles of the reference case at 5 hours and equation 5.1. During the 2 hour spin-up, until 7 hours of reference time, nudging is applied. Usually, this research applies nudging only to one time point, at 7 hours of reference time. In this case, reference slab mean vertical profiles at 7 hours are set as the desired mean states $\overline{\phi_{\mathrm{n},1D}}$. Desired fluctuation fields $\phi'_{\mathrm{n},3D}$ are found by subtracting the slab mean vertical profiles from the actual field at 7 hours: $\phi'_{\mathrm{n},3D} = \phi_{\mathrm{n}} - \overline{\phi_{\mathrm{n},1D}}$.

In some experiments in this research, multiple time fields nudging is done. A new desired state to nudge towards is then passed to the LES run every 10 minutes. The fields $\overline{\phi_{\mathrm{n},1D}}$ and $\phi'_{\mathrm{n},3D}$ are calculated using the same approach as above, but using the reference results at every 10-minute increment. The desired state is calculated by extrapolating the fields immediately before and after the current run time. To illustrate, consider a multiple time fields nudging run after 13 minutes of simulation. The 'desired' fields which it is being nudged towards at this point in time are 0.7 times the reference fields at 10 minutes, and 0.3 times the reference fields at 20 minutes. In this way, the thermodynamic fields are allowed to develop in a more 'natural' way, which also replicates the development of the reference field more accurately. Nudging runs are simulated for 3 hours. After two hours, they exit spin-up, and all nudging is deactivated. Results from the final hour are compared with the results of the other experimental runs. An overview of all the experimental runs is given in table 5.2 as well as in the schematic in figure 5.1.

Experimental run	**Start time (h)**	**Duration (h)**	**Description**
Reference (5.1)	0	8	Results from this run are used instead of observations.
Default spin-up (5.3)	7	1	Regular LES run, which is initialized using the reference vertical profiles (1D) at 7 hours.
3D initial thermodynamics (5.3)	7	1	LES run which is initialized using the thermodynamic reference fields (3D) at 7 hours.
Persistence (5.4)	7	1	Applies persistence method to the thermodynamic reference fields (3D) at 7 hours.
Nudging (5.5)	5	3	Initialized using reference vertical profiles (1D) at 5 hours. 1D and 3D-nudging applied towards desired thermodynamic fields (3D) during spin-up or nudging period (5h-7h). Nudging methods include 2 hour nudging, short nudging, and multiple time fields nudging.

Table 5.2: Overview of all the types of experiments done in this research. Behind the name of the run, in brackets, the section in which the run is described is given. Start times are given with respect to the reference run time. The overview is given for experiments on the large domain.

5.6. Evaluation methods

To compare the forecasts created by the experimental LES runs to the reference run and each other, several methods are employed, which are described here. Using the field for q_l given by DALES and equation 2.12, the LWP can be calculated. As the albedo of a cloud is for a large part indicated by the LWP, the LWP fields will be studied to determine the accuracy of the forecasts made by the experiments. A few statistical parameters are introduced to make the comparison of LWP fields easier. The first is

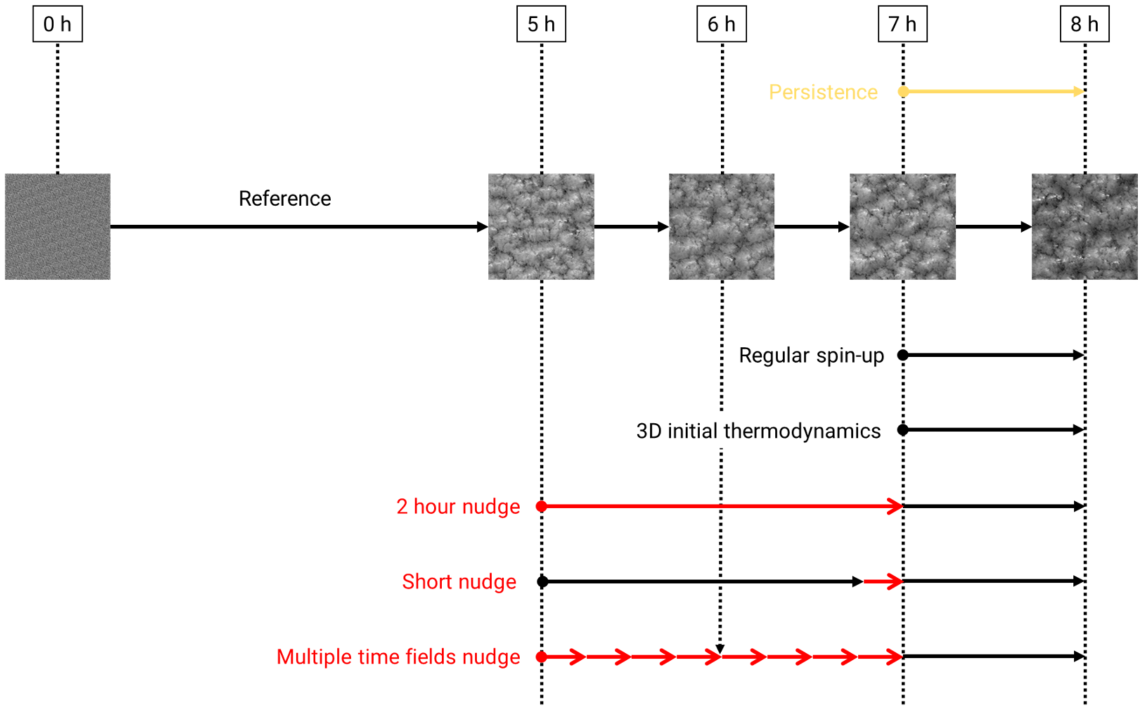


Figure 5.1: Schematic overview of the different experiments used in this research. LWP fields used in the schematic are found in the reference LES run. Indicated times are for the large domain.

simply the slab mean LWP:

$$\overline{\mathrm{LWP}}(t) = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \mathrm{LWP}(i, j, t), \tag{5.7}$$

in which N_x and N_y denote the total amount of grid points in both horizontal directions and i and j are the indices of a specific horizontal grid point. The slab mean LWP for the reference run and the experiments can be compared over time to see to what extent the forecasts follow the mean evolution of the cloud deck in the reference run.

However, the slab mean LWP does not provide much information about the correct horizontal field for the LWP. For solar forecasting, the cloud cover over a certain location must be predicted properly. To find the accuracy of the LWP forecast at each horizontal grid point, the root mean square error (RMSE) is used:

$$\mathrm{RMSE}_{\mathrm{LWP}}(t) = \sqrt{\frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} (\mathrm{LWP}_f(i, j, t) - \mathrm{LWP}_r(i, j, t))^2}, \tag{5.8}$$

where the subscript 'f' denotes the LWP obtained from forecasts, and the subscript 'r' indicates the LWP from the reference run. As can be seen in the equation, a forecast that perfectly reproduces the reference run LWP field will have an RMSE of zero.

The standard deviation of the reference run σ_r indicates the variation of the values of the reference LWP field from its mean LWP. It can be compared to the LWP RMSE by realizing that if a forecast gives the exact slab mean of the reference run at all locations, without any perturbations ($\mathrm{LWP}_f(i, j, t)$ = $\overline{\mathrm{LWP}_r}(t)$), the RMSE becomes equal to σ_r:

$$\mathrm{RMSE}_{\mathrm{LWP}}(t) = \sqrt{\frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} (\overline{\mathrm{LWP}_r}(t) - \mathrm{LWP}_r(i, j, t))^2} = \sigma_r(t). \tag{5.9}$$

A good forecast scores RMSE values below σ_{r}, by correctly predicting some characteristics of the field. As the RMSE shows if the predicted LWP in a certain horizontal grid point corresponds with the LWP as found in the reference run, it is a very strict metric, but a useful one nevertheless. Furthermore, the RMSE can also be plotted against time, showing a forecast's performance over a certain horizon.

Using the definition of the RMSE, another statistical property of the experiments can be introduced. The forecast skill (FS) can be used to compare the accuracy of one forecast to the accuracy of another. In this case, it investigates the accuracy of the new nudging method compared to the persistence method currently used in solar forecasting. So, for this research, the FS is determined by using the ratio of the RMSE of the mentioned methods (Wang et al., 2019):

$$\mathrm{FS}(t) = 1 - \frac{\mathrm{RMSE}_{\mathrm{LWP,e}}(t)}{\mathrm{RMSE}_{\mathrm{LWP,p}}(t)}, \tag{5.10}$$

where the subscript 'p' is used for persistence runs and subscript 'e' is used for experimental runs like regular LES or nudge runs. A positive value indicates that the experimental forecast performs better, whereas a negative value indicates a better performance by the persistence method. A time series can be made of this metric too, showing which method works better for different forecast horizons.

Besides the accuracy of the forecast made by the experiments, it is important to consider the effect of the nudging calculation on the turbulence build-up of the runs. For example, the intrusive nudging calculation can cause the turbulence generated in the spin-up period to be inaccurate. A wrong representation of the fluxes at the cloud top and base can show why experimental LWP evolves differently than reference LWP. On the other hand, investigation of the fluxes can show inaccuracies in an at-first-glance accurate LWP forecast. In other words, by investigating the turbulence of experimental runs, one can investigate if the LWP evolves correctly for the correct reasons.

To investigate the effect of the different experiments on the turbulence, vertical profiles of the vertical velocity variance $\overline{w'^2}$, and the buoyancy and total specific humidity turbulent flux $\overline{w'\theta_v'}$ and $\overline{w'q_{\mathrm{t}}'}$ are shown for different moments in time. Additionally, $\overline{w'\theta_{\mathrm{l}}'}$ and $\overline{w'q_{\mathrm{t}}'}$ at cloud base and top are plotted over time, as well as the difference in the net longwave radiation flux $\Delta F_{\mathrm{LW,net}}$ over the cloud. Surveying these fluxes at the base and top of the cloud allows for a so-called LWP budget analysis (van der Dussen et al., 2014). In all of these plots, results from the persistence method are not shown, as the frozen field approximation makes $w' = 0$ at all times (sec. 5.4).

6 Results

The first set of results compares 2 hour nudging experiments using different nudging time scales with each other and the persistence and regular LES methods, to investigate the effect of τ on the LWP forecast. Next, results will be presented for a short nudge run. Then, outcomes are illustrated for multiple time fields nudging. Finally, results from nudging experiments using thermodynamics fields estimated with the approaches in chapter 3 are given. Table 5.2 and figure 5.1 provide an overview of the different types of experiments used, and the names used to refer to them.

6.1. Nudging time scales

Equations 5.4 and 5.5 give the nudging calculations. The smaller the nudging time scale τ, the larger the effect of the nudging term on the tendency of the thermodynamic variables. Multiple 2 hour nudge experiments were performed using an array of different τ values to pinpoint the most effective τ for accurate forecasting, but also to investigate the effects of nudging on the turbulence build-up of the LES run.

Extensive testing was done on the small domain, using τ_{1D} and τ_{3D} values of 10, 20, 25, 30, 50, 60, 100, 300, 1000, 3600, 10^4 and 10^8 s. The use of different individual values for τ_{1D} and τ_{3D} was also explored, but the observed effects were little and not relevant to this research. Therefore, results shown here make use of the same value for τ_{1D} and τ_{3D}, henceforth referred to as τ. The experiments yielding the most insightful results were repeated on the large domain, and these results are given below.

6.1.1. LWP fields

To give the reader a better appreciation of the LWP, figure 6.1 shows the LWP fields of different simulations for various points in time. In the top left, the reference LWP field after 7 hours of simulation is given. The subsequent fields on the top row show the evolution of the reference LWP field over time, which the other experiments aim to replicate. Below the reference run fields, each row gives the LWP field and its evolution in time for a different experimental method.

The initial field of the persistence method is equal to the reference field. Because the persistence method freezes the thermodynamic fields, it is expected that the future LWP field will be identical to the initial field, but translated. Whereas this is largely the case, close inspection reveals a slight diffusion of the later fields. This has been attributed to numerical errors stemming from the advection calculation scheme. Numerical noise caused by the advection scheme is shown more clearly in appendix B.

Unsurprisingly, the initial field for the default spin-up method does not resemble the reference field. After 1 hour of simulation, the field still does not resemble the reference field in any way, clearly showing the negative effects of the lack of proper initialization of the cloud field and the spin-up period. The 3D initial thermodynamics method shows an initial field identical to that of the reference run, as expected. Additionally, its LWP evolution appears quite accurate when compared to the reference LWP.

The 7 hour field of the strong 2 hour nudge run (τ = 10 s) also looks identical to the reference field and maintains a good liking of the reference LWP over time. On the contrary, the experiment with weaker

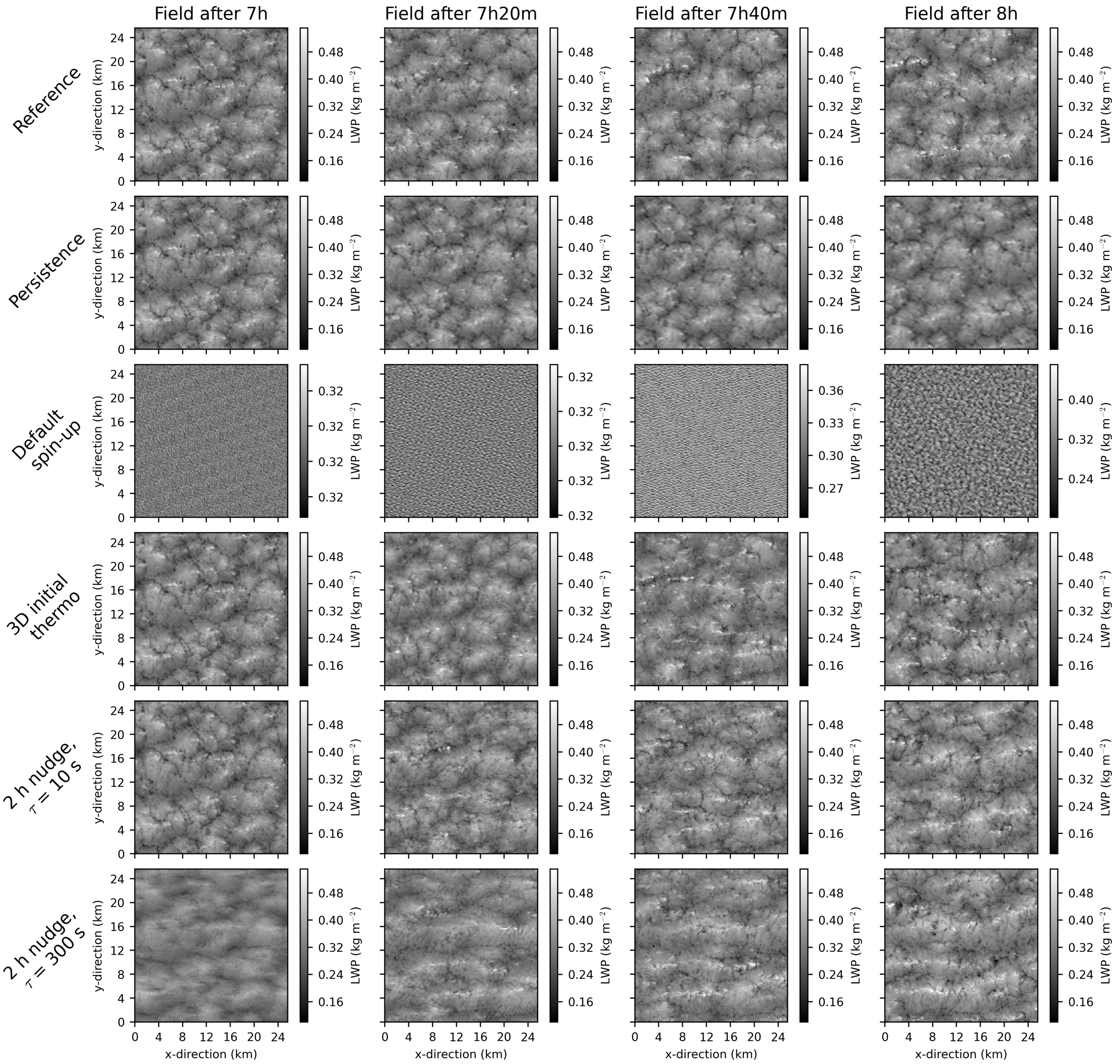


Figure 6.1: LWP fields at 20-minute increments for various LES runs. The title above each column gives the time which has passed since the start of the reference run. Each row is titled with the used experimental run, the methods for which are described in chapter 5. The reference field at 7 hours (large domain) is used as input for the other runs. All fields except the default spin-up use the same color bar, which ranges from 0.1-0.55 kg m^{-2}. Dark areas indicate a low LWP, and lighter areas a high LWP. The perturbations in the default spin-up LWP are so small that they are not represented in the colorbar of the first two fields.

nudging (τ = 300 s), as shown on the bottom row, displays an initial LWP field quite different from that of the reference run. However, the resemblance of the weakly nudged LWP field does seem to improve over time, but this is hard to say. A figure like 6.1 is aesthetically pleasing, but bar some basic observations, does not reveal much about the experiments. For example, one cannot conclude which of the experimental fields at 8h matches the reference field best by studying this figure. Hence, other techniques are used to evaluate the accuracy of the LWP forecasts (sec. 5.6).

6.1.2. Slab mean LWP

The first of these is the slab mean LWP (eq. 5.7), as shown in figure 6.2. The black line in the figure

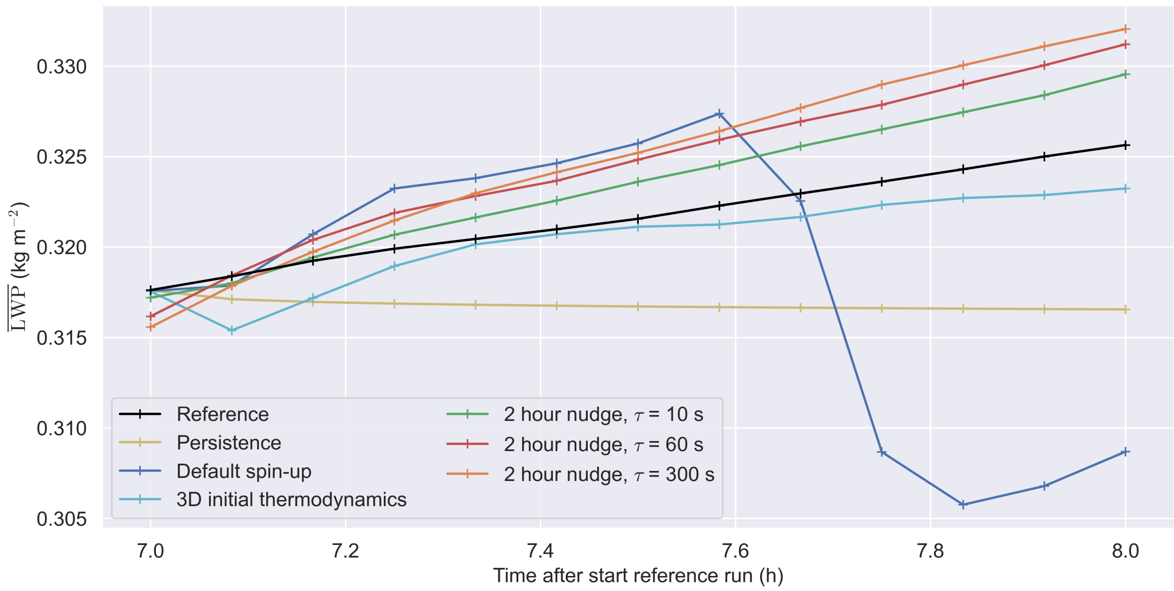


Figure 6.2: The slab mean LWP against time, time as measured from the start of the reference run. Analysis of the LWP forecast response to varying nudging time scales. Results are obtained on the large domain.

shows $\overline{\mathrm{LWP}}$ for the reference run. It grows slightly over the observed time, as evolving stratocumulus is simulated. At the 7-hour mark, all runs show good agreement with the reference $\overline{\mathrm{LWP}}$, except nudged runs with $\tau = 60$ s and $\tau = 300$ s, which have a $\overline{\mathrm{LWP}}$ slightly below the reference value. It seems that these runs are not nudged strongly enough to have the same initial slab mean LWP as the reference run. As the persistence method assumes a frozen field, it has a nearly constant $\overline{\mathrm{LWP}}$. Its small slope can be attributed to the numerical noise mentioned earlier. Because of this nearly constant $\overline{\mathrm{LWP}}$ forecast, its error with the reference $\overline{\mathrm{LWP}}$ grows over time. This underscores the limitation of the persistence method, as it cannot account for an evolving cloud.

The default spin-up $\overline{\mathrm{LWP}}$ has a good agreement initially, as it is initialized with the reference mean state and starts with only small perturbations in the field. However, after 15 minutes, it becomes evident that the spin-up period is causing the slab mean LWP to evolve differently from the reference run, as its growth is too large. After 35 minutes, a sudden drop in its $\overline{\mathrm{LWP}}$ is shown. Such a jump is typically observed during LES spin-up and is attributed to a swift rise in TKE at this point in time (also visible in figure 1.1).

After a dip in the first 5 minutes, the $\overline{\mathrm{LWP}}$ of the 3D initial thermodynamics method recovers and follows the reference state surprisingly well. Indeed, it provides the best forecast in the last 45 minutes of the observed hour. During the first 15 minutes, the strongly nudged run with $\tau = 10$ s gives the best forecast for the reference $\overline{\mathrm{LWP}}$. However, the slab mean LWP of this run has a slightly steeper slope than that of the reference run, causing the two lines to diverge. The other nudged runs have even steeper slopes of $\overline{\mathrm{LWP}}$ initially, causing them to diverge quicker. After about 30 minutes, the slopes of all nudge runs are similar. Overall, the 3D initial thermodynamics run gives the best $\overline{\mathrm{LWP}}$ forecast for horizons over 15 minutes. The nudging runs perform in the order of their nudging time scales, with the lowest nudging time scale giving the best performance. The persistence and cold case $\overline{\mathrm{LWP}}$ forecasts have the poorest performance.

6.1.3. LWP RMSE and FS

To further evaluate the performance of the forecast, the RMSE (eq. 5.8) of the experiments is taken, and plotted against time in figure 6.3(a). It is also used to determine the FS (eq. 5.10) of the regular LES and nudging methods with respect to the persistence method, as shown in figure 6.3(b).

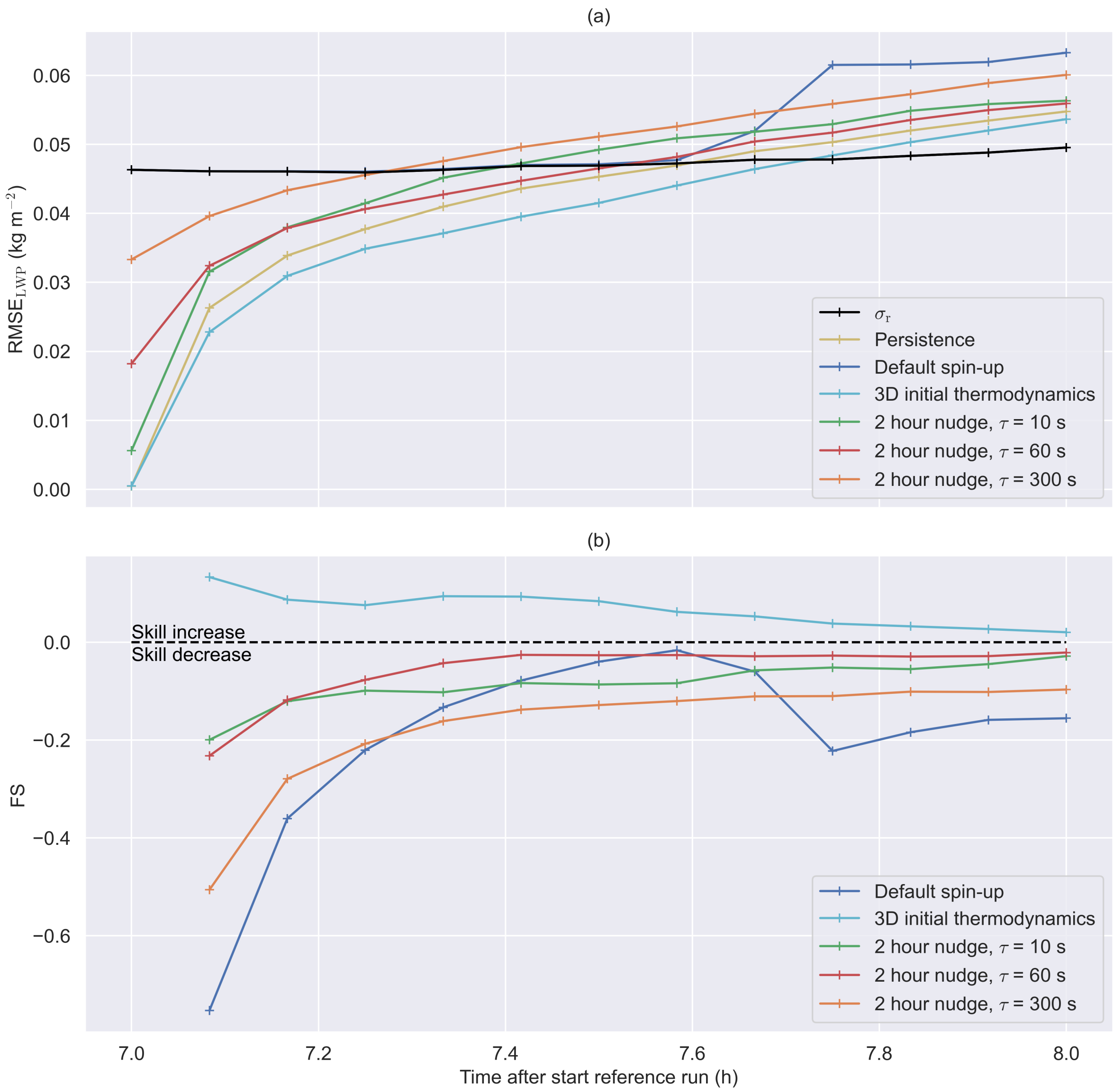


Figure 6.3: Effect of nudging time scale on LWP field forecast. Results are obtained on the large domain. In (a), the RMSE of the experimental runs is shown as a function of time. The black line shows the standard deviation of the reference run. (b) shows the FS of the RMSE of the experiments with respect to the persistence method as a function of time. The first FS data points (t = 7h) are omitted as the RMSE of the persistence method is zero initially, giving a negatively infinite FS.

Figure 6.3(a) gives a clear division of the performance of the different methods. Every experiment, except the default spin-up, has its lowest RMSE at the start of the comparison, below σ_r. The persistence and 3D initial thermodynamics method start with the exact thermodynamic reference fields and thus start with an RMSE of 0. For the nudging run with $\tau = 10$ s, the initial RMSE is also close to zero, suggesting that the strong nudging has very accurately replicated the reference field at the end of the nudging period. Conversely, the nudging run with $\tau = 300$ s, which experiences weaker nudging, exhibits a considerably higher RMSE compared to the strongly nudged run. In the nudge run with $\tau = 60$ s, the RMSE at 7 hours is approximately in between the RMSE of the other nudge runs. This shows that, as expected, stronger applied nudging leads to a closer resemblance between the field immediately after the nudging period and the desired field. However, the initial RMSE of all nudge runs lies below σ_r. This indicates that even nudging with a relatively high time scale of $\tau = 300$ s still manages to obtain some of the features of the desired field.

Default spin-up RMSEs are equal to the standard deviation of the reference run for the first 35 minutes. The reason for this is that the default spin-up starts with thermodynamic fields with the correct mean state, and with very small perturbations, which can be approximated as zero. Consequently, the RMSE becomes equal to σ_r (eq. 5.9). After 35 minutes, it undergoes the same sudden change as was observed for the slab mean LWP in figure 6.2, and the RMSE of the default spin-up increases significantly.

After the initial minimum value in RMSE, the other experimental runs all exhibit similar behavior. The RMSE increases sharply for the first ten minutes, after which it evens out to a somewhat constant increase. However, the magnitude of this first sharp increase and the approximately constant slope after differ between the runs. The 3D initial thermodynamics forecast consistently has the lowest RMSE, followed by the persistence method. Possibly, the 3D initial thermodynamics method performs better than the persistence method because it receives the appropriate wind profiles at the start, which makes its advection similar to that of the persistence method. However, as its fields are not frozen, it could be that some appropriate turbulence is quickly generated due to the correct initial thermodynamic fields, which improves its evolution slightly compared to the persistence method.

The LWP field forecasts of the nudging runs exhibit higher RMSEs compared to the persistence method. Interestingly, the run with $\tau = 10$ s goes through a much steeper initial increase than the runs with $\tau = 60$ and $\tau = 300$ s. This makes the RMSE of the run with $\tau = 60$ s lowest of the nudge runs after ten minutes of observations and onward. The initial increase in RMSE for $\tau = 300$ s is even less steep than for $\tau = 60$ s, but because of its higher RMSE at the start, its performance remains the worst of the three nudge runs. The differences in the magnitude of the initial RMSE increase for each of the nudging runs suggest that weaker nudging has produced more accurate turbulent fields, leading to a slower RMSE change. However, the more or less similar slopes in RMSE from 20 minutes onward indicate that the turbulence has recovered to roughly the same mean state for all nudge runs. The following subsections explore whether this conclusion holds by considering the turbulence in the experiments.

Figure 6.3(b) gives the FS of the regular LES and nudge runs with respect to the persistence method. It confirms many of the findings in the RMSE plot in 6.3(a), but offers a more comprehensive comparison of the experiments with the benchmark persistence method. The 3D initial thermodynamics run has levels above zero over the entire period, and thus its forecast outperforms that of the persistence method. Values for the other runs are below zero over the entire observed time. Interestingly, the FS of the nudged runs gradually inch upwards, indicating that the performance of the nudge forecasts relative to the persistence forecast slowly improves over time. FS scores also indicate that the nudge run with the best forecast is that with $\tau = 60$ s, closely followed by the run with $\tau = 10$ s. The weakly nudged run ($\tau = 300$ s) forecast has poor performance and even performs worse than the default spin-up forecast for a large period of time.

6.1.4. Turbulent vertical profiles

Figure 6.4 shows the vertical profiles of $\overline{w'^2}$, $\overline{w'\theta_v'}$ and $\overline{w'q_\mathrm{t}'}$ for the experiments at various moments in time during the observed hour. Vertical profiles found for the reference case are given in black. Both regular LES methods start with no initial turbulence. Therefore, their $\overline{w'^2}$ is zero initially. What is interesting to see is that $\overline{w'^2}$ for the 3D initial thermodynamics run grows towards the reference $\overline{w'^2}$ more quickly over time than the default spin-up $\overline{w'^2}$. This is due to its initialization with the reference thermodynamic fields, which seems to force the turbulence in this simulation toward reference values. After approximately 40 minutes, the vertical profiles of $\overline{w'^2}$, $\overline{w'\theta_v'}$ and $\overline{w'q_{\mathrm{t}'}}$ for the 3D initial thermodynamics run are similar to the reference profiles. Additionally, in the first 40 minutes, its profiles exhibit similar growth at all vertical levels. This is not the case for the default spin-up run, where $\overline{w'^2}$, when it finally starts growing after 40 minutes, initially grows more at the higher altitudes than the lower. This shows

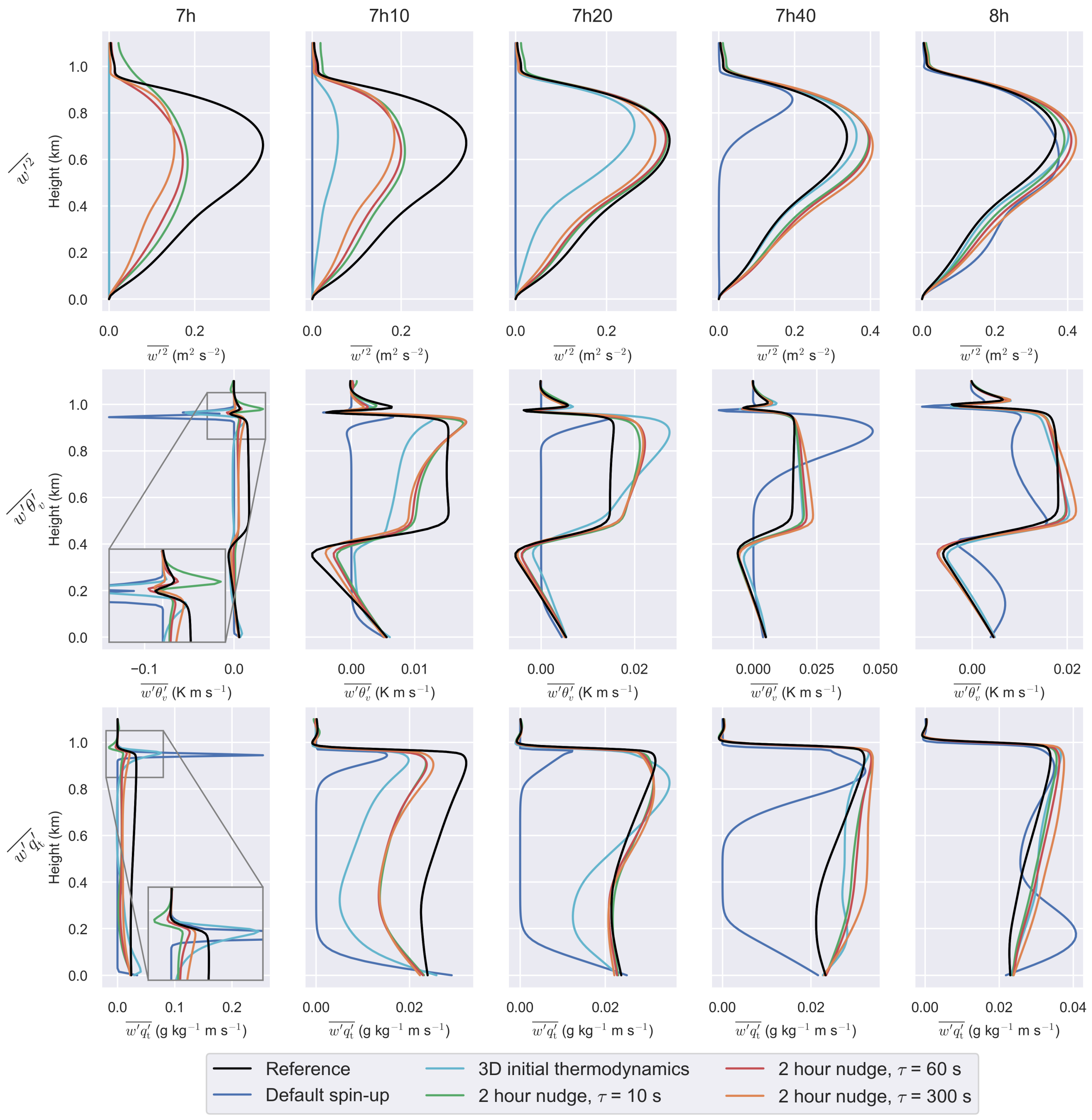


Figure 6.4: From top to bottom, the vertical profiles of the vertical velocity variance $\overline{w'^2}$, the buoyancy turbulent flux $\overline{w'\theta_v'}$ and the total specific humidity turbulent flux $\overline{w'q_t'}$. For the nudging time scale experiments. The persistence method is not shown in the plots, as it has zero vertical velocity.

the danger of forecasting using default spin-up LES runs. As $\overline{w'^2}$ starts its growth from the heights near the inversion layer, this can lead to an over-representation of entrainment. Excessive entrainment predicted in the simulation can cause the stratocumulus cloud to dissipate before the turbulence reaches a quasi-steady state, leading to a very erroneous solar forecast. Vertical profiles of $\overline{w'\theta_v'}$ and $\overline{w'q_{\mathrm{t}'}}$ already show this skewed growth grow from earlier time, with strongly growing fluxes at the top of the cloud layer. Finally, after 8 hours, the growth has shifted more to the lower altitudes, but the turbulent fluxes still are quite different than those observed in the reference run.

Curiously, the $\overline{w'^2}$ profiles for the nudged runs are very similar at all time points. Initially, their magnitude is less than half the reference $\overline{w'^2}$, but the profiles grow quickly until they are very similar to the reference profile after only 20 minutes. This is faster than the cold cases, showing the advantage of the nudging period. The only significant difference between the nudge runs is that the run with $\tau = 10$ s overestimates $\overline{w'^2}$ above the inversion layer. What is remarkable about the similar profiles is that it appears that the magnitude of the used nudging time scales has no large effect on $\overline{w'^2}$, except that nudging in general gives an underestimated vertical velocity variance.

Below the cloud layer, all nudge runs have a $\overline{w'\theta_v'}$ that is close to the reference $\overline{w'\theta_v'}$. In the cloud layer, however, the nudge runs have a similar profile compared to each other, but their predicted buoyancy flux is much smaller than the reference flux. This indicates that the nudging procedure reduces turbulent mixing in the cloud layer. This could be problematic because the absence of sufficient turbulent mixing for a sustained period of time could cause the nudge runs to predict a wrong evolution of the stratocumulus. Fortunately, the vertical profiles at later times grow closer to the reference case and become very similar to it from 40 minutes onward.

Much of the same can be observed for $\overline{w'q_{\mathrm{t}}'}$. The 2 hour nudge run $\overline{w'q_{\mathrm{t}}'}$ profiles at the 7-hour mark are lower than the reference profile over the entire BL. This shows that the nudge runs underestimate moisture transport in the BL, with possible negative effects on LWP evolution. The profiles begin agreeing with the reference profile over time until they are quite similar after 20 minutes.

A final note is that the strongly nudged run ($\tau = 10$ s) shows a large positive peak in $\overline{w'\theta_v'}$ and a large negative peak in $\overline{w'q_{\mathrm{t}}'}$ just above the inversion layer at the 7-hour mark. These peaks can be attributed to numerical artifacts and do not have a major influence on the simulation.

6.1.5. LWP budget analysis

Using an LWP budget analysis, it can be investigated if the accuracy of forecasts for the LWP shown earlier is caused by an accurate prediction of the evolution of the cloud layer, or by chance. Such an analysis for the τ experiments is given in figure 6.5. In figure 6.5(b), it is shown that the jump in net radiative flux over the cloud layer is the same for all runs and constant over time. This is due to the use of a cheap radiation scheme in the LES runs. A constant negative $\Delta F_{\mathrm{LW,net}}$ indicates constant radiative cooling, thus leading to the constant growth in slab mean LWP as observed for the reference case in figure 6.2. All experimental runs (except the persistence run) have the same $\Delta F_{\mathrm{LW,net}}$ and thus the same contribution of radiative cooling to $\overline{\mathrm{LWP}}$ growth. When a run deviates from the reference slab mean LWP growth, it is thus caused by the wrongly predicted $\overline{w'\theta_{\mathrm{l}}'}$ and $\overline{w'q_{\mathrm{t}}'}$ at cloud top or base. However, it is also possible that a run predicts wrong $\overline{w'\theta_{\mathrm{l}}'}$ and $\overline{w'q_{\mathrm{t}}'}$ at both the cloud top and the cloud base, but that the effects at these two heights compensate for each other. In such a case, a run might have a good $\overline{\mathrm{LWP}}$ evolution, but for wrongly predicted turbulence.

The LWP budget analysis once more accentuates the disadvantages of using a default spin-up LES run. In this run, all turbulent fluxes are zero for most of the first 40 minutes. It can therefore be concluded that for the first 40 minutes, the LWP develops wholly differently in the default spin-up run than in the reference case which it is supposed to replicate.

The same cannot be said for the reference cold run. By providing the reference field at the start of the run, it can be expected that the spin-up period gets a shove in the right direction. Initial values of the fluxes at the cloud base do show a small improvement over the default spin-up run, with values slightly closer to the reference values. However, the initial values of the turbulent fluxes at cloud top have a magnitude that is much too large compared to the reference case. Recall that in figure 6.2, the 3D initial

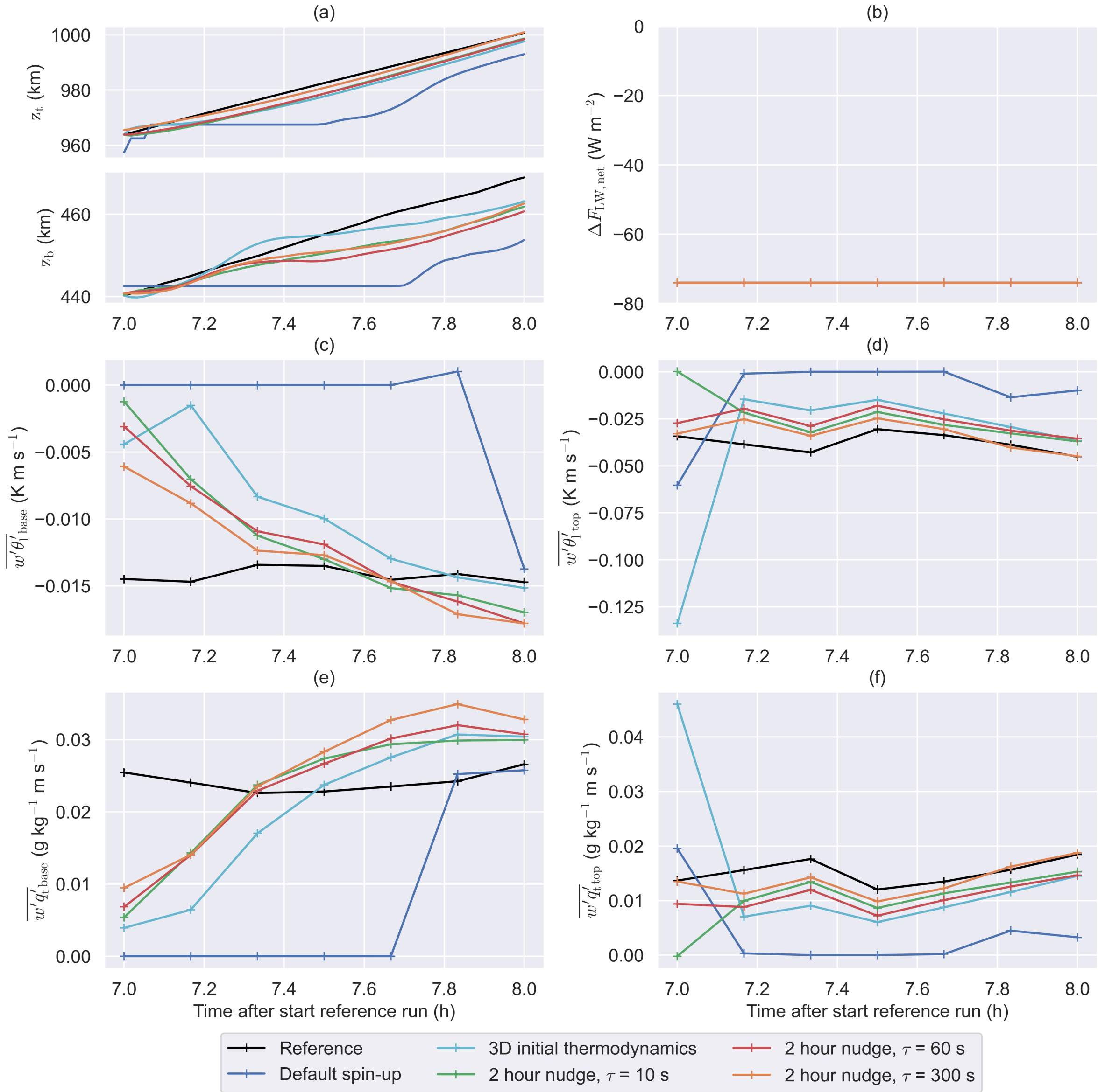


Figure 6.5: Plots showing the change in time for the cloud top and base height, the jump in net longwave radiation over the cloud layer, and the liquid potential temperature and total specific humidity turbulent fluxes at cloud base and cloud top. Used for LWP budget analysis of the nudging time scale experiments.

thermodynamics $\overline{\mathrm{LWP}}$ displays a dip in the first ten minutes. Initial turbulent fluxes at cloud top suggest a warming and drying effect that is much larger than in the reference case. This leads to a greater loss of LWP. Moreover, heat loss and moisture supply at the cloud base are underestimated with respect to the reference run, resulting in a smaller increase in LWP at cloud base. This explains the initial decrease in slab mean LWP for the 3D initial thermodynamics run. After ten minutes, the turbulent fluxes at cloud top change to values lower in magnitude than those of the reference case, which explains why the 3D initial thermodynamics $\overline{\mathrm{LWP}}$ recovers towards the reference $\overline{\mathrm{LWP}}$. Its turbulent fluxes at the cloud base also start resembling the reference turbulent fluxes more over time.

In the nudge runs, initial turbulent fluxes at the cloud base are similar to those in the 3D initial thermodynamics run, also suggesting a lower LWP growth than observed in the reference case. However, this underestimation is not accompanied by an overestimation of the decrease in LWP at the cloud top, like it is in the 3D initial thermodynamics run. Instead, the initial turbulent fluxes of the nudge runs suggest an LWP loss at cloud top that is more similar to that in the reference case, if slightly lower. The evolution of the fluxes over time is quite similar for all three nudge runs, but their magnitude is slightly different. Nudging fluxes continually underestimate LWP loss at the cloud top. In the first 30 minutes, the nudge runs underestimate LWP growth at cloud base, but after that, they tend to overestimate LWP growth. This likely causes their $\overline{\mathrm{LWP}}$ growth in this period to be higher than the reference $\overline{\mathrm{LWP}}$ growth.

Of the nudge runs, the run with the weakest nudging (τ = 300 s) predicts values for all turbulent fluxes except $\overline{w'q'_t}_{\mathrm{base}}$ that are closest to the reference values. Especially at the cloud top, its predictions are better than in the other nudge runs. Interestingly, this is likely what causes its $\overline{\mathrm{LWP}}$ in figure 6.2 to grow the most of all the nudge runs in the first 30 minutes. During this period the other nudge runs compensate for an underestimated LWP growth at cloud base with an underestimated LWP decrease at cloud top. This compensation occurs less for the τ = 300 s run as it has similarly underestimated LWP growth at the cloud bottom but its fluxes at the cloud top are very accurate, causing the LWP to grow more than in the other runs and the reference case. It can be concluded that strong nudging gives a poorer representation of the turbulent fluxes. However, the plots also show that in terms of predicted fluxes, the nudging methods are preferable over the 3D initial thermodynamics method, as they consistently predict fluxes that are closer to the reference fluxes than the 3D initial thermodynamics method.

6.2. Nudging start times

Up to this point, nudging is always applied during the full 2 initial hours of an LES run, when it is going through its spin-up period. However, the nudging calculation is expected to affect the build-up of turbulence negatively. Therefore, experiments have been done where nudging is not applied from the start of the spin-up period but from later time points during spin-up. As a result, the turbulence has an undisturbed, 'normal' development before nudging is started, with the aim of ending up with more realistic values for turbulence. It is important to strike a balance here between letting the turbulence develop freely for as long as possible and making sure the right initial field is enforced by nudging. Extensive testing was carried out on the small domain to explore various combinations of nudge start times and nudging time scales. From these tests, it was concluded that nudging with τ = 10 s only during the final 5 minutes of spin-up provided the best results, and as such, this is the short nudging experiment which was repeated on the large domain and which will be discussed in the remainder of this section.

6.2.1. Slab mean LWP

Figure 6.6 shows the $\overline{\mathrm{LWP}}$ for the short nudge run. Curiously, the short nudge run $\overline{\mathrm{LWP}}$ dips in the first 10 minutes, after having roughly the right value at 7 hours. It recovers after, gradually climbing back towards the reference $\overline{\mathrm{LWP}}$. The initial dip of the short nudge $\overline{\mathrm{LWP}}$ forecast leaves it outperformed by the 2 hour nudge forecast for the first half hour of observation. However, its steady climb back towards the reference results makes the former the more accurate of the two nudge forecasts in the second half hour. To boot, its forecast is even closer to the reference $\overline{\mathrm{LWP}}$ in the final 15 minutes than the 3D initial thermodynamics forecast. So, for the case used, the short nudge $\overline{\mathrm{LWP}}$ forecast is the best of all methods for a horizon of 45 minutes or more. Consideration of the turbulence of the simulations later

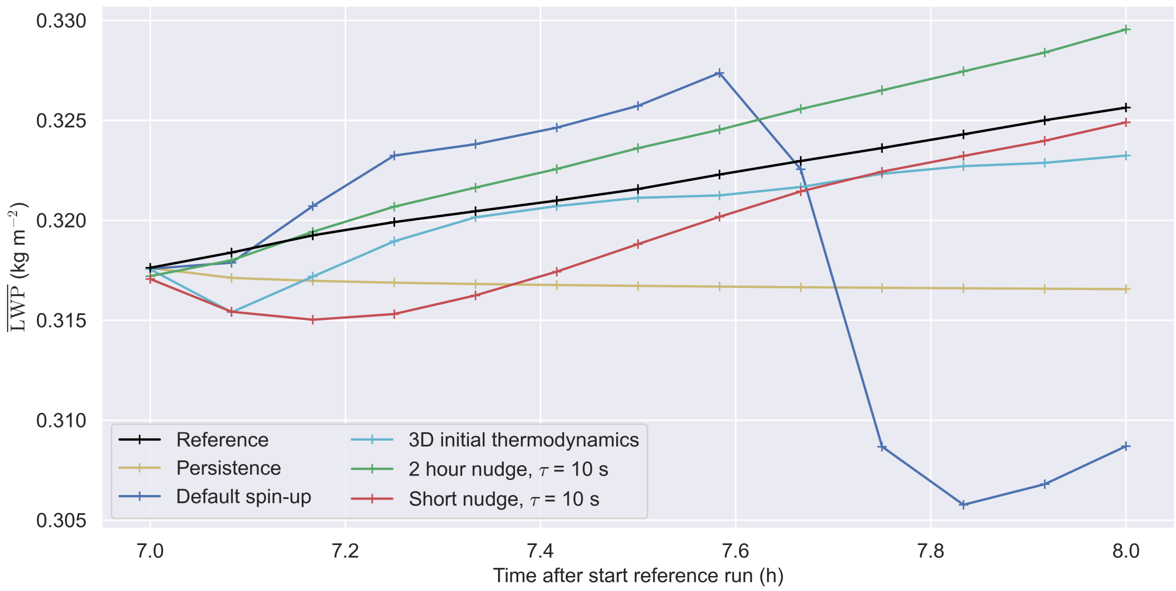


Figure 6.6: The slab mean LWP over time, researching the effect of a later nudging start time. Results are obtained on the large domain.

in this section will show if this accurate forecast is a result of more accurate turbulence, or chance.

6.2.2. LWP RMSE and FS

The RMSE and FS of the short nudge run in figure 6.7 do not show the same initial dip in performance as observed for the slab mean LWP. Figure 6.7 (a) illustrates that both nudge runs have the same RMSE at the start of the observation period. This means that the desired field has been introduced to a similar degree in both nudge runs, indicating that just 5 minutes of nudging with $\tau = 10$ s is sufficient to establish the desired field. The initial steep increase in RMSE of both nudge runs is roughly the same during the first 5 minutes. However, the RMSE growth after this time is much less steep for the short nudge RMSE, making the RMSE of their forecasts diverge quickly over time.

The short nudge RMSE stays below σ_r for 40 minutes, similar to the reference field cold method. Indeed, the slope of its increase is so small that despite its higher start, the short nudge RMSE is lower than the persistence method RMSE after 20 minutes and lower than the reference field cold method RMSE after 45 minutes. The plot of the FS in figure 6.7(b) illustrates this more clearly. Slight positive FS values after 20 minutes show the short nudge forecast outperforms the persistence method on this horizon. Also, the short nudge run FS is slightly higher than the reference field cold run FS for the final 15 minutes. Thus, as for the mean LWP, the short nudge LWP field forecast is the best forecast on horizons of 45-60 minutes, at least for the LES case used. These findings indicate that nudging for only 5 minutes produces significantly better results than nudging for a longer period of time. Only 5 minutes of nudging (with $\tau = 10$ s) is enough to reproduce the reference field to the same degree as a run that receives nudging of the same strength for the full two hours of spin-up. Moreover, a better evolution of the LWP field is achieved.

6.2.3. Turbulent vertical profiles

The improved evolution of the LWP demonstrated in the previous section hints at a more realistic turbulence for the short nudge run in comparison with the 2 hour nudge run. An examination of the turbulent fluxes and the vertical velocity variance in figure 6.8 confirms this hypothesis. The vertical velocity variance for the short nudge run at the 7-hour mark shows remarkable improvement with respect to the 2 hour nudge run. As remarked in the previous section, the 2 hour nudge initial $\overline{w'^2}$ is much smaller than the reference $\overline{w'^2}$. In contrast, the short nudge run initial $\overline{w'^2}$ is only slightly smaller than the reference profile at low heights, and almost matches it in the upper 0.3 km. So, the vertical transport is captured much more accurately from the start in the short nudge run. This explains why the evolution of the LWP

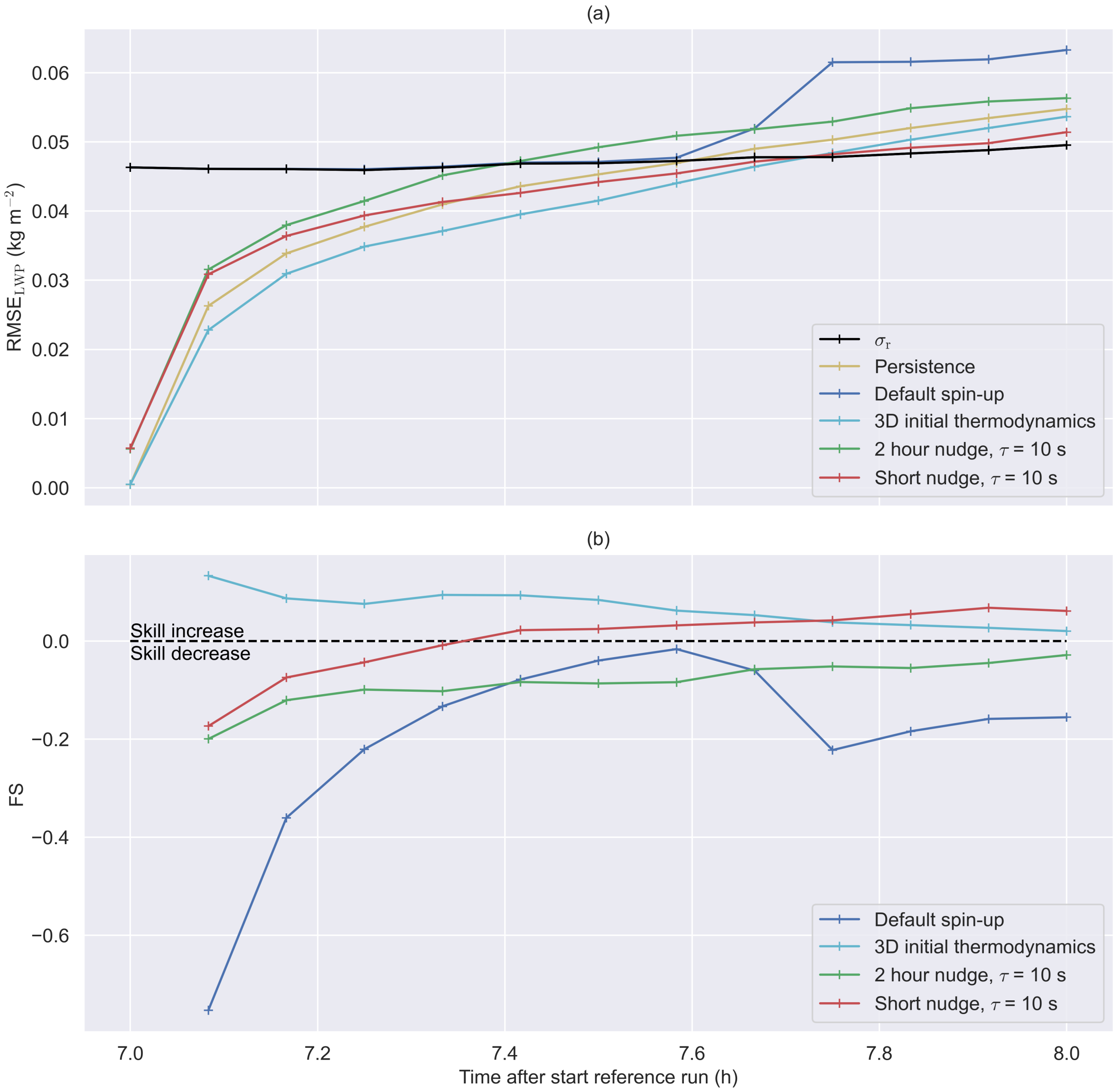


Figure 6.7: Effect of nudging start time on the LWP field forecast. Results are obtained on the large domain. In (a), the RMSE of the experimental runs is shown as a function of time. The black line shows the standard deviation of the reference run. (b) shows the FS of the RMSE of the experiments with respect to the persistence method as a function of time. The first FS data points (t = 7h) are omitted as the RMSE of the persistence method is zero initially, giving a negatively infinite FS.

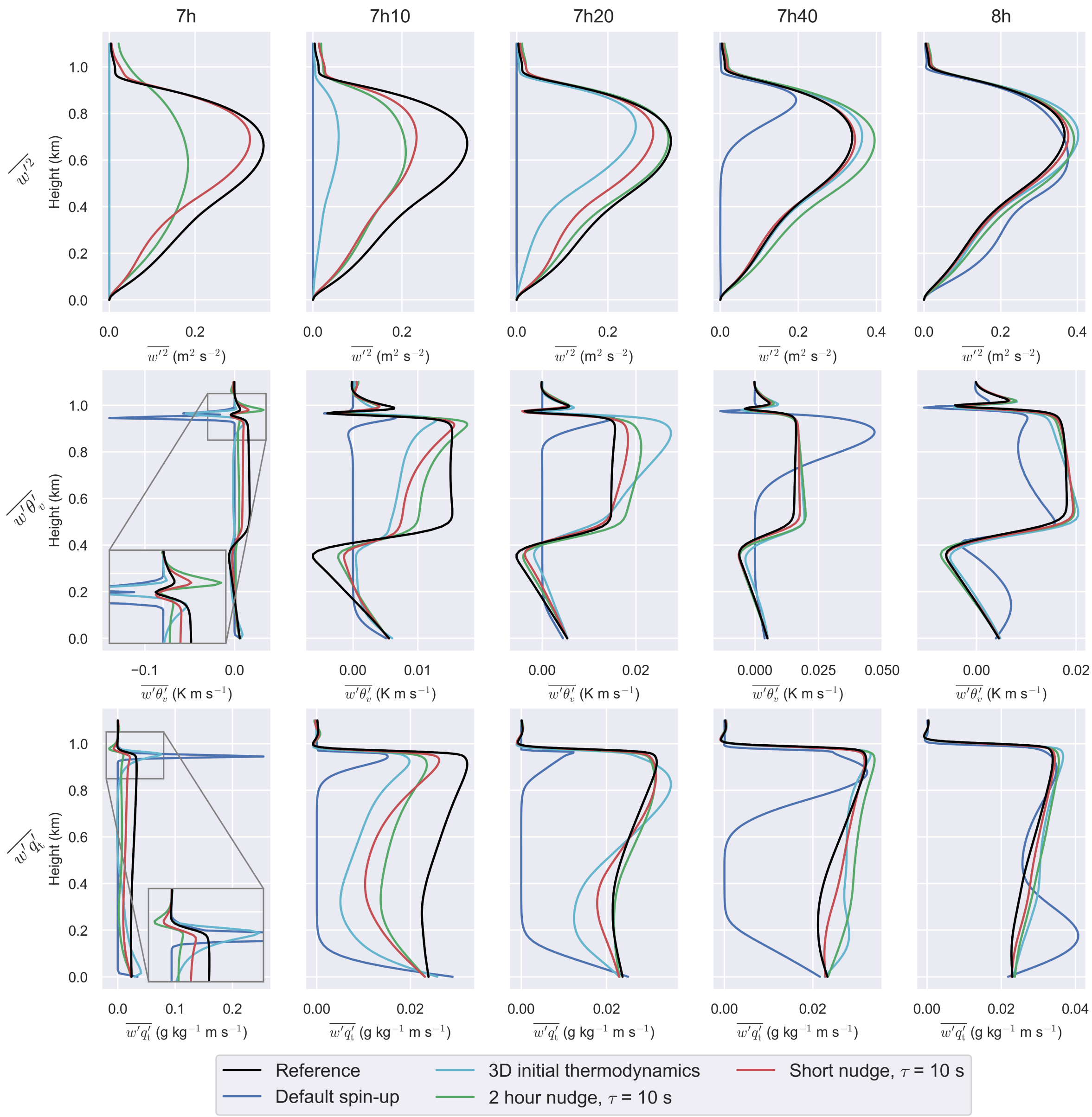


Figure 6.8: From top to bottom, the vertical profiles of the vertical velocity variance $\overline{w'^2}$, the buoyancy turbulent flux $\overline{w'\theta_v'}$ and the total specific humidity turbulent flux $\overline{w'q_t'}$. For the short nudging time experiments. The persistence method is not shown in the plots, as it has zero vertical velocity.

field for this run is better. After ten minutes, the magnitude of $\overline{w'^2}$ in the short nudge run has decreased somewhat, and the profiles for both the short and 2 hour nudging cases become similar. This decrease might be caused by the after-effects of nudging during the last 5 minutes. At 7 hours and 20 minutes, the short nudge run $\overline{w'^2}$ is still a little below the reference profile, whereas the 2 hour nudge run is closer. However, after 40 minutes and onward, the short nudge profile has a better agreement with the reference profile than the 2 hour nudge run.

Within the BL, both turbulent fluxes $\overline{w'\theta_v'}$ and $\overline{w'q_t'}$ of the short nudge run are in better agreement with reference fluxes than the 2 hour nudge run fluxes. The 2 hour nudge run shows an overestimation of $\overline{w'q_t'}$ in the BL after 7 hours and 40 minutes and after 8 hours. In the short nudge run, this overestimation is still present but is much smaller. As such, the surface moisture supply to the cloud is predicted more accurately, explaining why the evolution of the LWP is more accurate at these later times. Investigation of the turbulent vertical profiles shows improved resemblance with reference profiles for the short nudge run over the 2 hour nudge run. This indicates that nudging for a shorter time is less intrusive to the spin-up period and allows for more accurate (initial) turbulent fields.

6.2.4. LWP budget analysis

The analysis of the LWP budget shown in figure 6.9 can further illustrate why the short nudge run has such an improved LWP forecast over the 2 hour nudge run. Initial values for $\overline{w'\theta_l'}$ and $\overline{w'q_t'}$ at cloud top and bottom are all closer to the reference values in the short nudge run than in the 2 hour nudge run. This can explain why the initial increase in RMSE is lower in the short nudge run. At the cloud top, this increased agreement with reference fluxes is slight for both $\overline{w'\theta_l'}$ and $\overline{w'q_t'}$, yet it is sustained for 30 minutes. Afterwards, the fluxes for the short and 2 hour nudge runs are more or less equal. Values for $\overline{w'\theta_l'}$ at the base of the cloud are similar over time for both runs, except in the final ten minutes, where the short nudge run gives a more accurate flux. At cloud base, the 2 hour nudge run gives $\overline{w'q_t'}$ values more resembling the reference fluxes than the short nudge run between 5-20 minutes. Before and after this period, the short nudge run provides a more accurate flux.

The initial dip in $\overline{\mathrm{LWP}}$ for the short nudge run can be explained by observing that in the first 15 minutes, its forecasted LWP growth at the cloud base lies lower than the reference growth and lower than the growth forecasted by the 2 hour nudge run.

Overall, it can be concluded that the transport of heat and moisture in and out of the cloud layer as predicted by the short nudge run is more resembling of the transport in the reference run than the prediction of the 2 hour nudge run. This demonstrates that the LWP forecast accuracy improves between the two nudge runs because the short nudge run has induced more accurate turbulence, and not by chance. So, letting LES generate turbulence undisturbed and leaving nudging to a later time is a promising way to use the nudging method, which works better than all other methods investigated so far.

6.3. Multiple time fields nudging

In the experiments above, the simulations are only nudged toward the thermodynamic fields at the end of the spin-up period. As such, the evolution of the thermodynamic fields during the nudging period is restricted. For the multiple time fields nudging method, the reference thermodynamic fields were extracted for every ten minutes between the 5-hour and 7-hour mark. Nudging is applied to the subsequent fields using $\tau = 600$ s, the exact time between extracted fields. Multiple time fields nudging aims to induce turbulence that is more like that in the reference case, by going through a more similar development in thermodynamic states.

The results for the multiple time fields nudging are compared to a 2 hour nudge run with $\tau = 60$ s and the short nudge run with only 5 minutes of nudging with $\tau = 10$ s, as these are the nudge runs with the best performance in the previous investigations (sec. 6.1 and 6.2). As usual, the results will also be compared to the regular LES and persistence methods.

6.3.1. Slab mean LWP

Figure 6.10 shows $\overline{\mathrm{LWP}}$ for multiple time fields nudging. For these runs, nudge run results of $\overline{\mathrm{LWP}}$ are also shown during the nudging period, as it shows the differences between the approaches well.

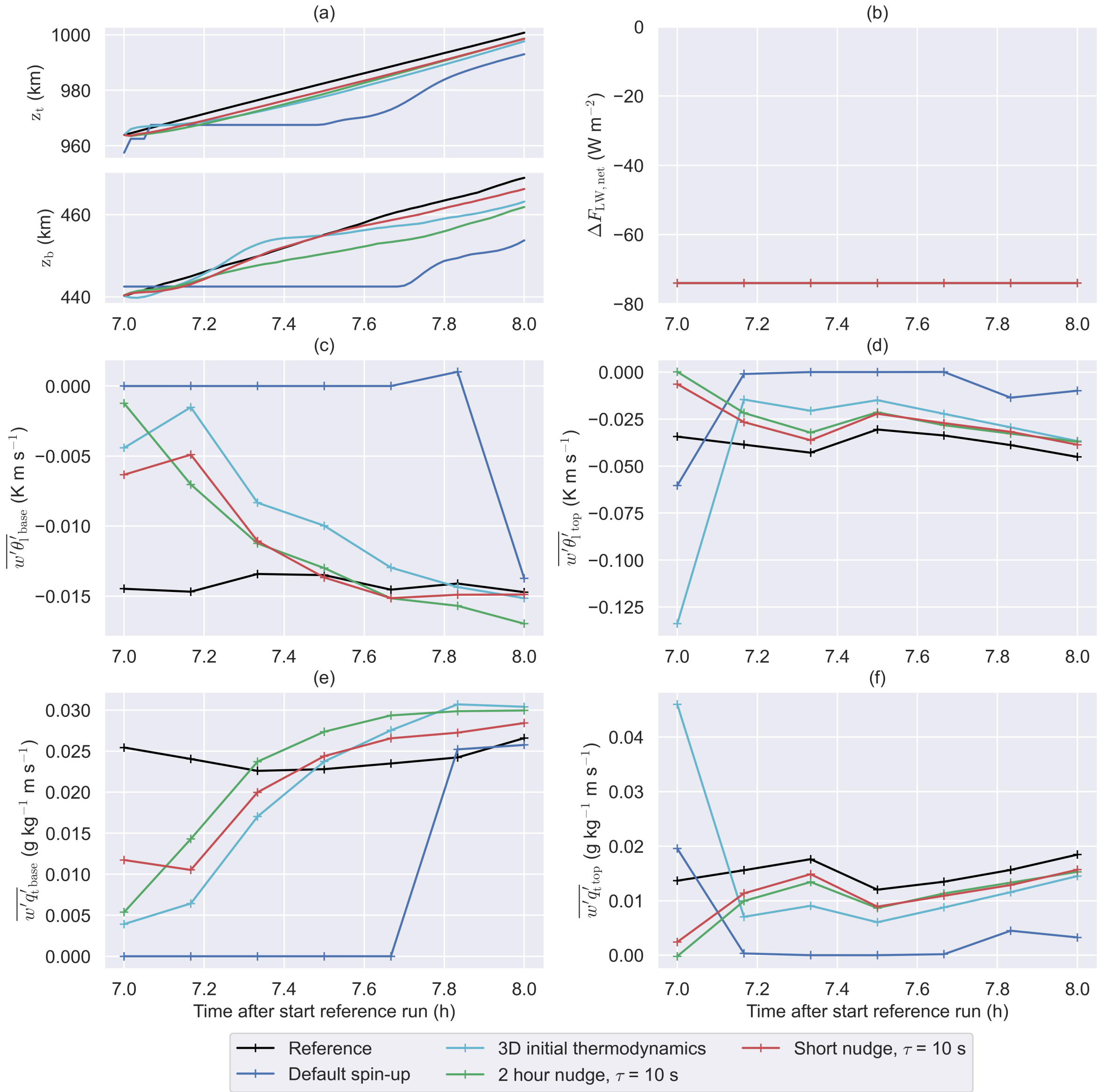


Figure 6.9: Plots showing the change in time for the cloud top and base height, the jump in net longwave radiation over the cloud layer, and the liquid potential temperature and total specific humidity turbulent fluxes at cloud base and cloud top. Used for LWP budget analysis of the short nudging experiments.

The 2 hour nudge run, to start with, shows approximately constant $\overline{\mathrm{LWP}}$ values throughout the nudging period. As it constantly receives fairly strong ($\tau = 60$ s) nudging towards the thermodynamic fields at 7 hours from the start, this is to be expected. The short nudge run, on the other hand, shows the typical behavior of an LES run in its spin-up period, with the characteristic fall in $\overline{\mathrm{LWP}}$ after 35 minutes. It also shows an immense increase in slab mean LWP in the final 5 minutes of the nudging period, when its nudging is activated and it grows towards the reference value. Finally, the multiple time fields nudge $\overline{\mathrm{LWP}}$ follows the reference $\overline{\mathrm{LWP}}$ very well, which is precisely the aim of nudging towards multiple fields in time. The reason its $\overline{\mathrm{LWP}}$ lies a little lower than reference values is likely because nudging is fairly weak, and the exact reference thermodynamic states are thus not achieved.

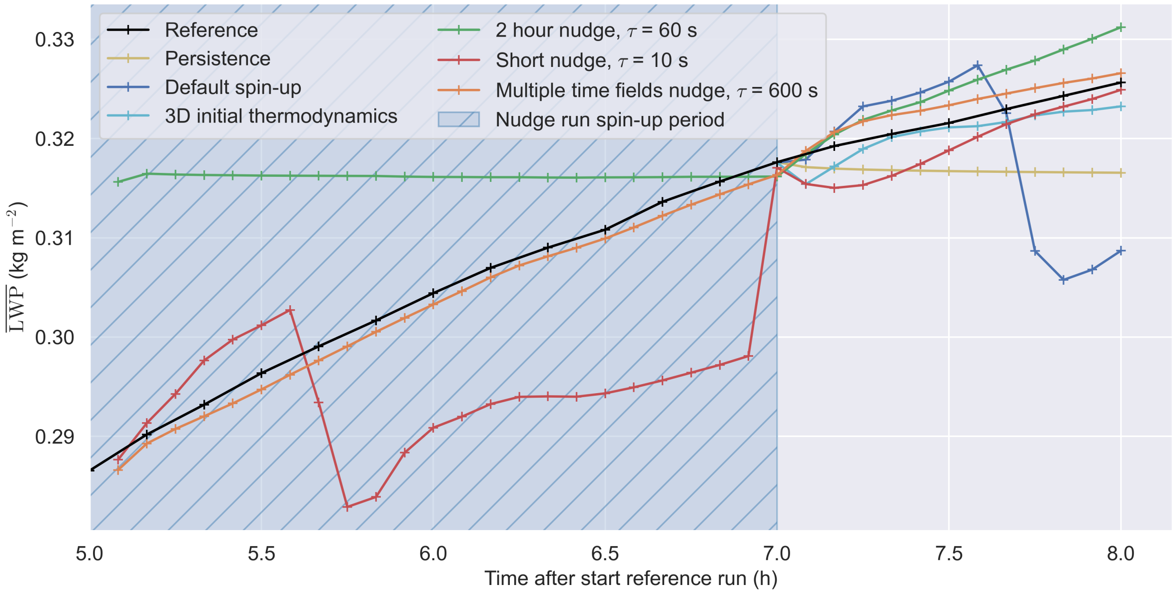


Figure 6.10: The slab mean LWP over time, researching the effect of multiple time fields nudging. Results are obtained on the large domain. The blue, arced area indicates the nudging period in the first two hours.

When nudging is deactivated at the 7-hour mark, the multiple time fields nudging $\overline{\mathrm{LWP}}$ has a value slightly lower than the reference value, similar to the value of the 2 hour nudge run. Indeed, the two runs follow the same pattern for the first 15 minutes, quickly overshooting the reference $\overline{\mathrm{LWP}}$ with a growth in $\overline{\mathrm{LWP}}$ that is too steep. However, where the $\tau = 60$ s run continues this steep growth after 15 minutes, the multiple time fields nudge run does not. Instead, its growth falls off, and its $\overline{\mathrm{LWP}}$ values grow more towards reference values. In the latter 45 minutes, it provides a better $\overline{\mathrm{LWP}}$ forecast than the reference cold run, but not quite as good as the short nudge run.

6.3.2. LWP RMSE and FS

Figure 6.11(a) gives the RMSE for the multiple time fields nudge run and the other experiments. The initial RMSE for multiple time fields nudging is quite high, as was observed before for nudging with higher τ, but it still lies well below σ_{r}. Actually, comparison with figure 6.3 shows that its initial RMSE lies below that of the 2 hour nudging run using $\tau = 300$ s. So, whereas the nudging time scale in multiple time fields nudging is two times as large, nudging towards subsequent fields has replicated the reference fields more accurately than the 2 hour $\tau = 300$ s run, already showing some of the benefits of multiple time fields nudging.

The sharp increase in RMSE during the first 5-10 minutes exhibited by all the other runs is not observed for multiple time fields nudging. Moreover, in the period after, the growth in its RMSE is smaller than that of the runs. As a result, the multiple time fields nudging RMSE lies below the RMSE of the other nudging runs after 15 minutes, below the persistence method RMSE after 20 minutes, and below the 3D initial thermodynamics RMSE after 35 minutes. Its RMSE only grows larger than σ_{r} after 55 minutes.

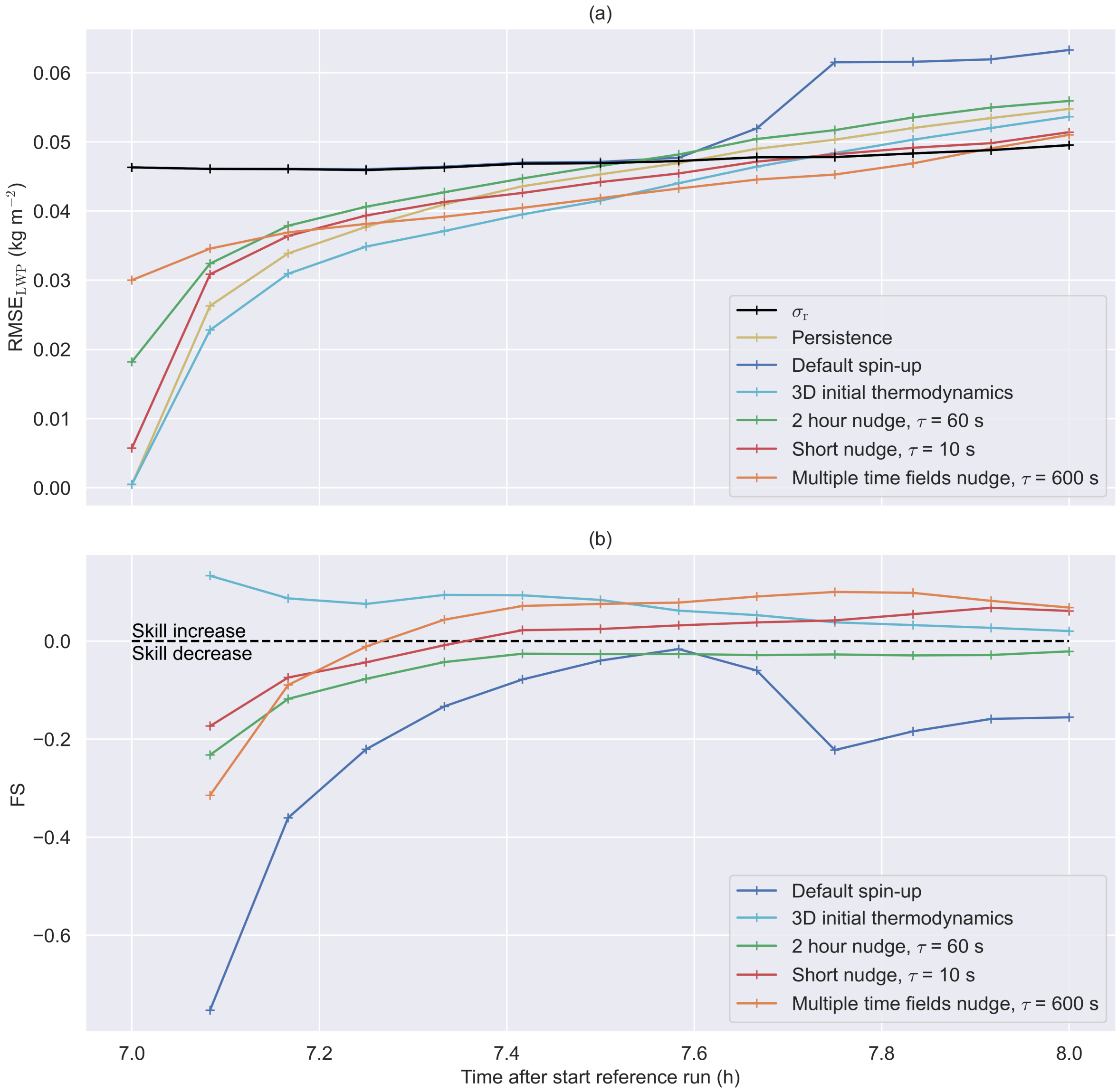


Figure 6.11: Effect of nudging start time on the LWP forecast. Results are obtained on the large domain. In (a), the RMSE of the experimental runs is shown as a function of time. The black line shows the standard deviation of the reference run. (b) shows the FS of the RMSE of the experiments with respect to the persistence method as a function of time. The first FS data points (t = 7h) are omitted as the RMSE of the persistence method is zero initially, giving a negatively infinite FS.

This makes it the best forecast for the LWP field for horizons of 35 minutes and beyond, and hints at accurately induced turbulent fields. Figure 6.11(b) underlines the above observations.

6.3.3. Turbulent vertical profiles

The vertical profiles for $\overline{w'^2}$, $\overline{w'\theta}$ and $\overline{w'q_t'}$ in figure 6.12 show an improvement of the induced turbulence in the multiple time field nudging run. In particular, the vertical turbulent fluxes $\overline{w'\theta_v'}$ and $\overline{w'q_t'}$ of the

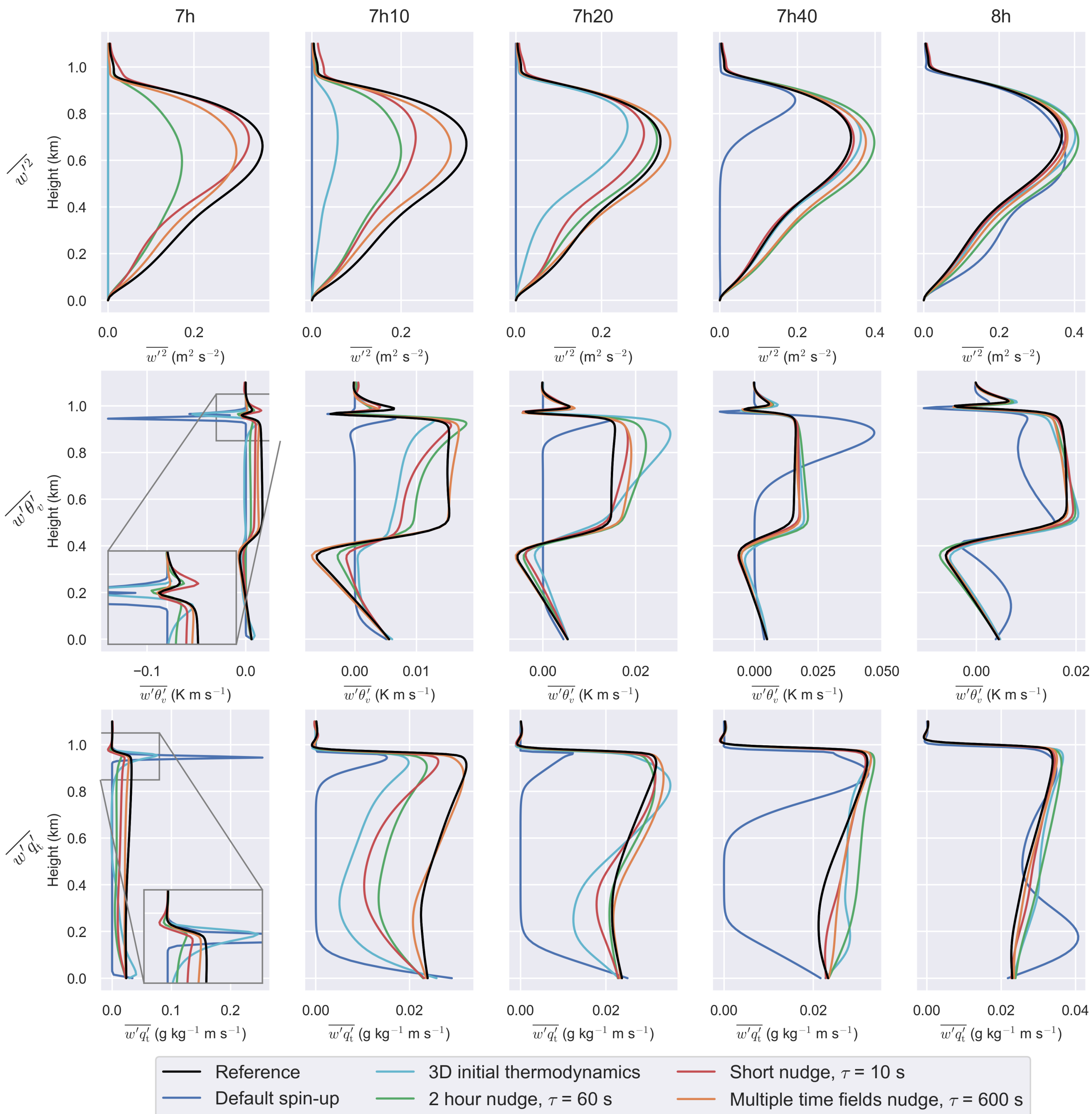


Figure 6.12: From top to bottom, the vertical profiles of the vertical velocity variance $\overline{w'^2}$, the buoyancy turbulent flux $\overline{w'\theta_v'}$ and the total specific humidity turbulent flux $\overline{w'q_t'}$. For the multiple time fields nudging experiment. The persistence method is not shown in the plots, as it has zero vertical velocity.

multiple nudge run show a remarkable improvement over the fluxes of the other nudge runs. At the 7-hour mark, multiple time fields fluxes resemble the reference fluxes much better. After 10 minutes, its fluxes are even almost the same as the reference fluxes. This is in stark contrast to the other runs, which have fluxes much smaller than reference values at this point. For later times, the turbulent fluxes are similar among all nudge runs. The accuracy of the turbulent fluxes in the first ten minutes can explain why the initial growth in RMSE is much less pronounced for multiple time fields nudging than for the other experiments, and its continued accuracy explains its low RMSE growth at later times.

6.3.4. LWP budget analysis

In the LWP budget analysis in figure 6.13, the improvement of the predicted fluxes at cloud base is remarkable. The other experiments exhibit an initial $\overline{w'\theta'_1}_{\mathrm{base}}$ and $\overline{w'q'_t}_{\mathrm{base}}$ that is 2-3 times smaller in

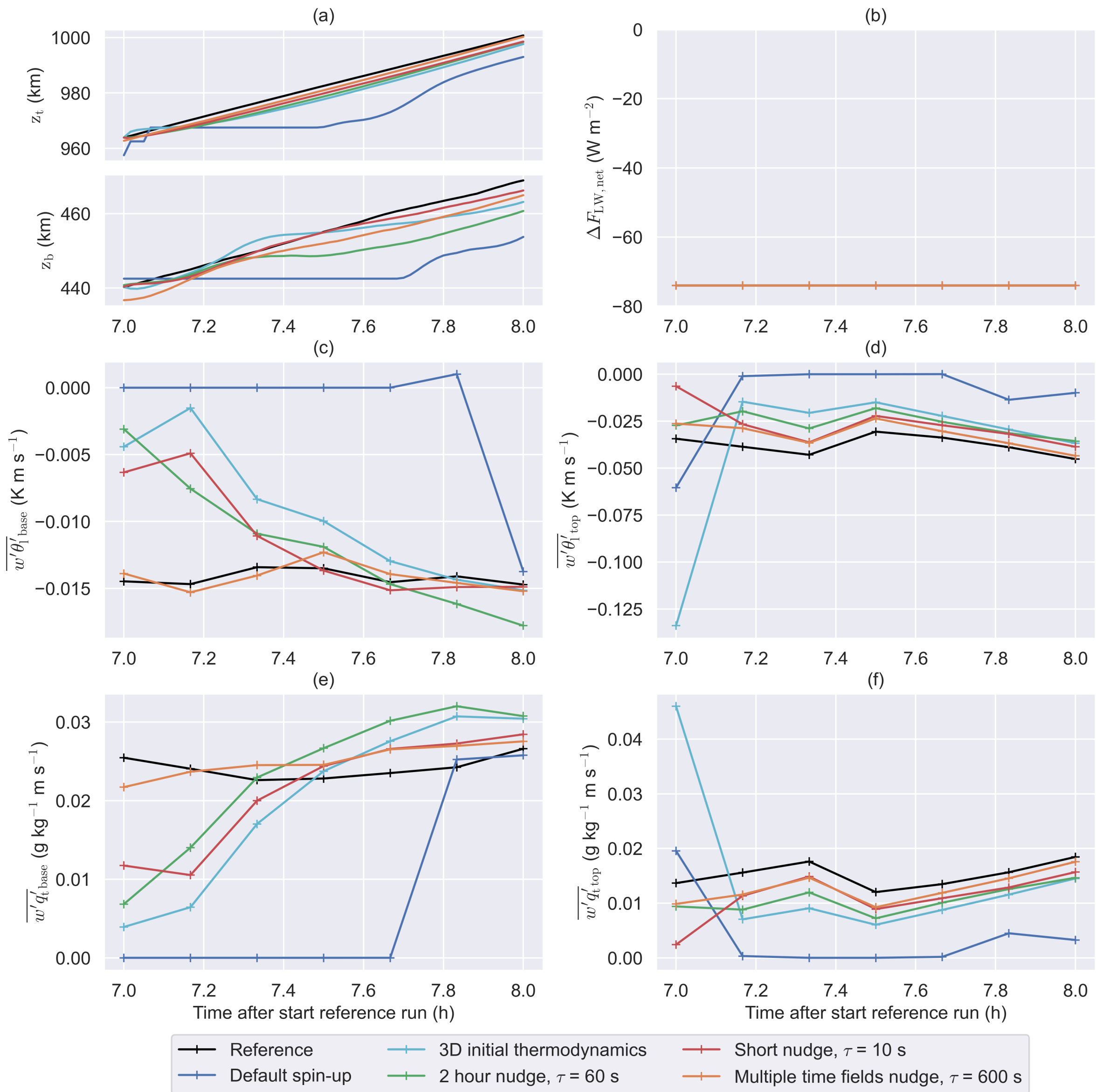


Figure 6.13: Plots showing the change in time for the cloud top and base height, the jump in net longwave radiation over the cloud layer, and the liquid potential temperature and total specific humidity turbulent fluxes at cloud base and cloud top. Used for LWP budget analysis of the multiple time fields nudging experiments.

magnitude than the values found in the reference case. In contrast, multiple time fields nudging gives values much closer to the reference value at 7 hours, with only a slight underestimation. It retains this close resemblance over time. Moreover, it gives the most accurate prediction of $\overline{w'\theta'_1}_{\mathrm{top}}$ and $\overline{w'q'_t}_{\mathrm{base}}$ of all runs over the entire observed period. Much better initial fluxes show why the multiple time fields nudging has a much less steep initial rise in RMSE. Furthermore, its growth in RMSE stays smallest of all runs over the rest of the period, because its fluxes are most accurate here as well. This proves that multiple time fields nudging provides the most accurate forecast as a result of having induced the most accurate turbulence of all considered runs.

6.4. Thermodynamic fields from satellite observations

In chapter 3, three approaches were described to estimate the thermodynamic fields of θ_l and q_t from a satellite LWP field. These were termed the linear estimation, the $\theta_l' = 0$ limit estimation, and the $q_t' = 0$ limit estimation. To test their validity, 2 hour nudging experiments were done using thermodynamic fields found by each of the approaches. Instead of an actual satellite LWP field, the approach was applied to the LWP field found in a reference LES run. Besides comparing the forecast performance of these approaches with each other, they are also compared with a nudge run that uses the exact 3D thermodynamic fields of the reference run. Because of a lack of time, these experiments were carried out on the small domain. On the small domain, reference fields are extracted after 2 hours reference time instead of 7, and thus nudge runs begin at 0 hours reference time instead of 5. The LWP field at the 2-hour mark was thus used in the estimations of the thermodynamic fields. Figure 6.14 shows the $\overline{\mathrm{LWP}}$ and the RMSE of the LWP field for these nudge runs.

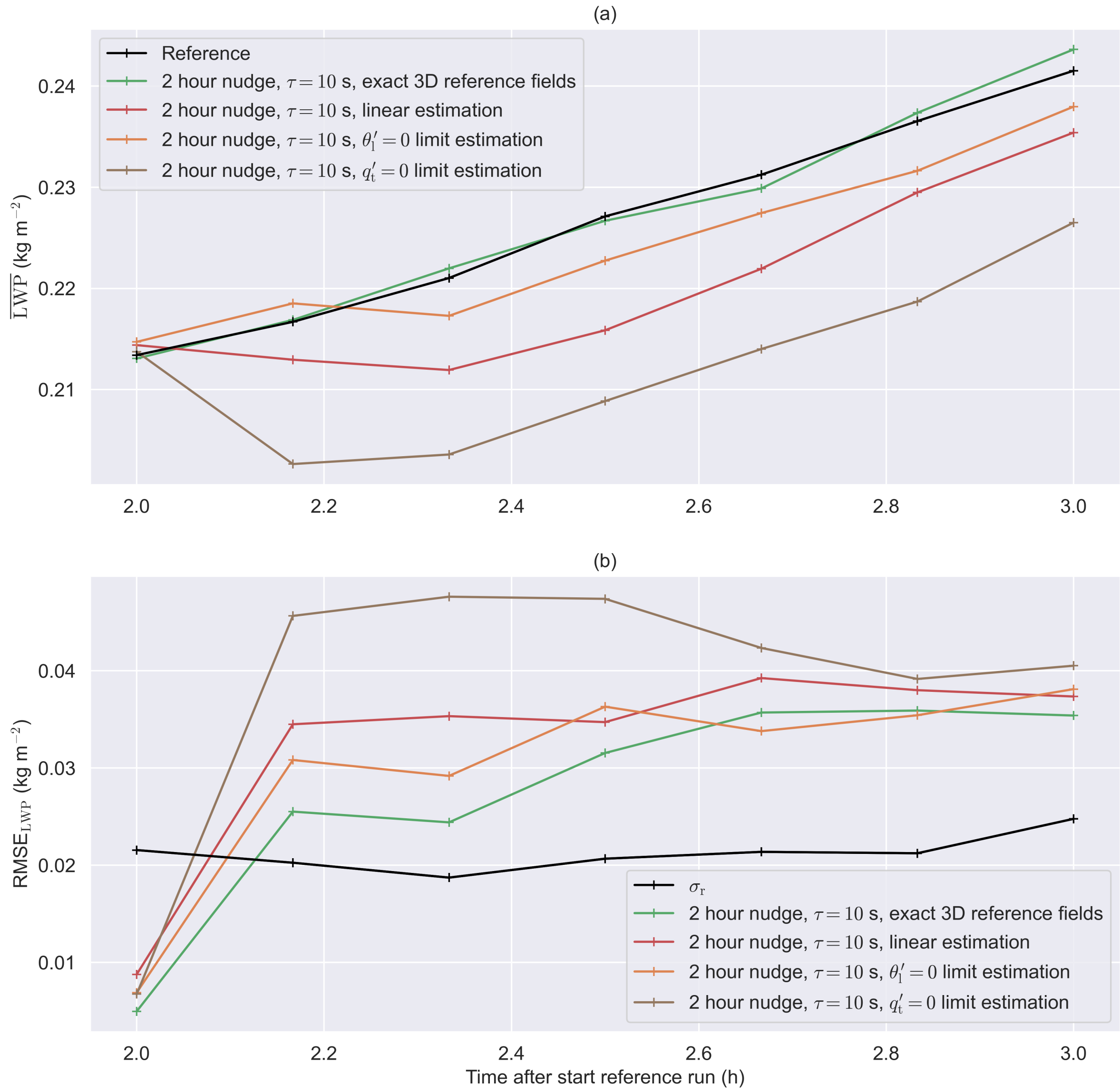


Figure 6.14: Slab mean LWP and the RMSE in the LWP field over time. For nudge experiments using thermodynamic fields as estimated from LWP fields. Experiments were done on the small domain.

The estimations were all designed to approximate thermodynamic fields that replicate the LWP field at the 2-hour mark. Therefore, the initial $\overline{\mathrm{LWP}}$ and RMSE are most important, as these indicate to what degree the found thermodynamic fields replicate the reference LWP. For $\overline{\mathrm{LWP}}$, unsurprisingly, the exact

3D fields run gives the best prediction. A close second best prediction is that given by the $q_t' = 0$ limit estimation run. Then, the linear estimation run $\overline{\mathrm{LWP}}$ follows, and finally, the value $\overline{\mathrm{LWP}}$ least like the reference value is given by the $\theta_l' = 0$ limit estimation run.

Perhaps even more important than the initial slab mean LWP is the initial RMSE. Again, the exact 3D fields run gives the most accurate LWP, as its RMSE is the lowest. The $\theta_l' = 0$ and $q_t' = 0$ limit estimations have the same initial RMSE, which is slightly higher than that of the exact 3D fields run. Finally, the linear estimation has the highest initial RMSE. It should be noted that all nudge runs have a relatively low RMSE initially.

From these results, it can be concluded that all three estimations produce thermodynamic fields that replicate the reference LWP well. However, both limiting estimations replicate the reference LWP more accurately than the linear estimation. This is unexpected, as the limiting estimations were meant to indicate the lower accuracy limits of the linear estimation. It seems therefore that there is an inaccuracy in the linear estimation. Two possible explanations are offered. The first is that the approximated linear relationship between θ_l' and q_t' does not hold sufficiently through the entire cloud layer. The second is that the constant c_{qT} used in the relation is not determined accurately through the use of c_{sfc}. Likely, a combination of both causes the inaccuracy of the fields in the linear estimation.

Whereas all of the estimation nudge runs replicate the reference LWP well initially, their predicted evolution of the LWP field is very different. After the exact 3D fields nudge run, the $q_t' = 0$ limit estimation nudge run predicts the reference LWP evolution most accurately. The next best prediction of the reference LWP evolution is made by the linear estimation nudge run, followed by the $\theta_l' = 0$ limit estimation nudge run. The evolution of the LWP forecast can indicate how intrusive the chosen estimation is on the build-up of turbulence during the spin-up period. Interestingly, fields assuming $\theta_l' = 0$ give a more accurate LWP evolution than the fields estimated using the linear estimation, which allows fluctuations in both θ_l and q_t. This once more indicates that the linear estimation is not optimal in its current form.

7
Discussion

In this chapter, the results of the previous chapter are discussed. The discussion is separated on the basis of the main and sub research questions formulated for this thesis.

7.1. Replicating the desired three-dimensional field

Nudging time scales under 300 s give a nudging strong enough to replicate at least some of the desired 3D fields. However, only nudging time scales of 10 and 60 s give an initial LWP field accurate enough to base a further forecast on. These conditions might be somewhat case-specific. It can be argued that for an LES case with a smaller or greater homogeneity in thermodynamic fields, the upper limit of τ where nudging obtains the desired fields so that its initial RMSE lies below σ_{r} might be different. Therefore, it is recommended that 3D-nudging be applied to more LES cases, stratocumulus or not, to test the above statement.

Nudging for a period of 5 minutes with $\tau = 10$ s is enough to replicate the desired LWP field to a similar degree as nudging with the same τ during the entire 2 hour spin-up period. It is expected that similar results can be found for runs with higher values of τ, but with larger time periods of nudging. Also, whereas the used parameters ($\tau = 10$ s, nudging for 5 minutes) might be the best combination for the stratocumulus case used in this research, other combinations might give better forecasts in different (stratocumulus) case studies. This is something to be wary of in future research using short nudging.

7.2. Effect on turbulence build-up and field evolution

As expected, implementing 3D-nudging clearly disrupts the generation of turbulence in the spin-up period. Forcing the thermodynamic fields to a single state during the entire spin-up period does not generate the turbulence characteristic to this thermodynamic state. The stronger the applied nudging, the more the turbulence after the nudging period differs from the reference state. Poorer initial turbulence is found to be connected to a less accurate evolution of the thermodynamic fields.

Undoubtedly, the methods devised to improve turbulence generation in the nudging period, short nudging and multiple time fields nudging, are successful. Especially multiple time fields nudging, which provides nudging to the subsequent reference thermodynamic states for every 10 minutes in the nudging period, gives greatly improved turbulent profiles. It is recommended to research the effect of increasing the time step between the fields on the accuracy of the induced turbulence in multiple time fields nudging.

7.3. Estimation of thermodynamic fields from observations

Three approaches of estimating thermodynamic fields to accurately reproduce a given LWP field were suggested in chapter 3: the linear estimation, the $\theta_{\mathrm{l}}' = 0$ limit estimation, and the $q_{\mathrm{t}}' = 0$ limit estimation. Initial values of the RMSE in figure 6.14 suggest that all three estimations are quite capable of approximating thermodynamic fields which accurately replicate the reference LWP field. However, the initial RMSE of the linear estimation nudge run is slightly higher than that of the limit estimation nudge runs. This indicates that the thermodynamic fields produced by the linear estimation are less capable of reproducing a reference LWP than those produced by the limit estimation methods. As the limiting estimations were designed to show the lowest limit in accuracy of the linear estimation, it is apparent that there are some inaccuracies in the linear estimation.

This claim is supported by the subsequent evolution in time of the LWP field for each of the estimation nudge runs. Of the estimation runs, the most accurate evolution is given by the $\theta_l' = 0$ limit estimation run, followed by the linear estimation run, followed by the $q_t' = 0$ limit estimation. This shows that the thermodynamic fields as given by the linear estimation are more disruptive to the spin-up of accurate turbulence than the fields given by the $\theta_l' = 0$ limit estimation. Reasons for the linear estimation to be inaccurate are that the approximate linear relationship between θ_l' and q_t' does not hold everywhere in the cloud layer or that the constant c_{qT} in the relation is not calculated accurately. It is recommended that in future research, it is investigated how to make the linear estimation produce more accurate thermodynamic fields. This can be done by checking the validity of the various assumptions used in the estimation and replacing them if possible. It could also be worthwhile to repeat the experiment on the larger domain, to bring the results more into perspective.

7.4. Solar forecasts

With the use of nudging in LES, more accurate solar forecasts can be created than when using the regular LES or persistence methods. However, 2 hour nudging is not enough, and either short nudging or multiple time fields nudging should be applied. In this research, short nudging uses $\tau = 10$ s and has only 5 minutes of nudging at the very end of the 2 hour nudging period, instead of during the entire nudging period. As mentioned before, short nudging in other case studies might require other τ or a nudging time period for the best performance of 3D-nudging.

Multiple time fields nudging performs even better than the short nudging, and provides the prevailing forecast for horizons of 35 minutes and larger. It should be denoted that the multiple time fields nudging has one drawback: the amount of observations needed to run the experiment. Whereas 2 hour and short nudging only need observations of the atmosphere at 5 hours (for initialization) and 7 hours (for the desired thermodynamic fields), the multiple nudging run uses observations of the fields for every 10 minutes during the period between 5 and 7 hours. However, actual satellite imagery like that used by Wang et al. (2019) has a temporal resolution of 15 minutes. Therefore, the effects of increasing the 10 minute time period to 15 minutes in multiple time fields nudging should be investigated in further research.

Currently, the initial RMSE of the multiple time fields nudging run is quite high, and its forecast becomes better than others mostly through a smaller slope in the error. It would be interesting to see the effects of using the same multiple time fields nudging method, but applying very strong nudging in the last period of time, like in the short nudging experiments. Possibly, this could combine the advantage of having a very accurate initial field, as observed in strongly nudged runs, with the superior turbulence generated by the multiple time fields nudging run, leading to even more accurate forecasts.

Another interesting result of the experiments is that the 3D initial thermodynamics run gives better forecasts than the persistence method consistently, and than the nudging methods in some cases. Analysis of its vertical turbulent profiles and LWP budget indicate that the 3D initial thermodynamics method has less accurate LWP development than the nudge runs, although already much better than that found in a default spin-up run. For further research, it would be interesting to investigate whether the 3D initial thermodynamics method performs well for any LES case study, or that situations can also occur where its turbulence is estimated wrongly resulting in a poor forecast.

8 Conclusions and recommendations

8.1. Conclusions

The aims of this thesis are to implement a novel three-dimensional nudging technique into LES in order to assimilate data on observed cloud fields into the model and to subsequently evaluate this technique by testing its produced solar forecast for a stratocumulus-topped boundary layer against current solar forecasting methods. Using these goals, the main research question was formulated as: *Does the implementation of three-dimensional nudging of the thermodynamic fields during the spin-up phase of an LES model give a better solar forecast for a stratocumulus-topped boundary layer than those generated by conventional methods?*

To answer this question, the sub-questions are addressed first:

- Is the three-dimensional nudging method capable of including the desired three-dimensional thermodynamic field in LES, and what nudging time scales are required to do so?

Initial values for the RMSE in LWP show that the three-dimensional nudging method is capable of replicating the desired thermodynamic fields to a high degree. The best agreement with the desired fields is found for the smallest nudging time scales, which have the strongest nudging. When nudging is applied during the entire 2-hour nudging period, nudging time scales as high as $\tau = 300$ s provide strong enough nudging to have an initial RMSE below σ_{r}. Values below σ_{r} indicate that at least some of the features of the desired fields are present in the simulation.

Moreover, applying very strong nudging ($\tau = 10$ s) in only the last 5 minutes of the nudging period gives an initial RMSE value that is just as low as when nudging of the same strength is applied during the entire 2 hour nudging period. Also, multiple time fields nudging, where the simulation is nudged to desired subsequent thermodynamic fields every 10 minutes, which has a nudging strength of only $\tau =$ 600 s, can replicate the desired field after the nudging period to a good degree. Its initial RMSE is lower than that observed for the 2 hour nudge run using $\tau = 300$ s. It can be concluded that the 3D-nudging method is quite capable of assimilating desired fields in LES, especially for low τ.

- Does the implementation of the three-dimensional nudging term have an effect on the build-up of turbulence during the spin-up period, and how does this affect the evolution of the thermodynamic fields in time?

Applying three-dimensional nudging during the spin-up period has a profound effect on the build-up of turbulence. In 2 hour nudging, nudging leads to a large underestimation in the magnitude of the turbulent fluxes with respect to reference fluxes upon nudging deactivation. This was observed for all investigated nudging strengths ($\tau = 10$, 60, 300 s), but the effect was largest for lower τ. As time elapses after nudging deactivation, the magnitude of the nudge fluxes grows, and after 20 minutes and onward the 2 hour nudge runs possess fluxes of a reasonable magnitude when compared to reference fluxes. The evolution of the thermodynamic fields is notably affected, as observed most clearly in plots of the RMSE over time. A large increase in RMSE is observed for all nudge runs in the first ten minutes, corresponding with the underestimated turbulent fluxes. This growth is largest for the lower

τ and decreases as τ increases. At later times, when fluxes grow to values more like the reference values, the growth of the RMSE decreases, and the evolution of the thermodynamic fields becomes more accurate.

To reduce the negative effect of nudging on the turbulence buildup, two variations on the nudging method were devised: short nudging and multiple time fields nudging. In short nudging, the LES spin-up period is left undisturbed for a longer period of time, before activating the nudging calculation. For short nudging of only 5 minutes with $\tau = 10$ s, initial profiles of $\overline{w'^2}$, $\overline{w'\theta_1'}$ and $\overline{w'q_t'}$ have much more resemblance with the initial reference profiles than the 2 hour nudging with $\tau = 10$ s. Moreover, the magnitude of initial fluxes at cloud base and top is underestimated less in the short nudge run. Over time, these indicators of turbulence and LWP evolution for both nudge runs grow closer together. The improved turbulence present in the short nudging run echoes through to the evolution of the LWP field, visualized by a smaller growth in RMSE over the entire observed time period when compared to the 2 hour nudge run.

In multiple time fields nudging, the thermodynamic development of the reference run is replicated more accurately by subsequently nudging (with $\tau = 600$ s) to the reference thermodynamic fields every ten minutes in the nudging period. As a result, the underestimations in magnitude previously observed for nudge runs in the initial profiles of $\overline{w'^2}$, $\overline{w'\theta_1'}$ and $\overline{w'q_t'}$, and the initial fluxes at cloud base and height are greatly reduced. For multiple time fields nudging, initial fluxes very closely resemble reference values. This improvement is greater than that observed for the short nudge run, especially in the fluxes at the cloud base. The close likeness to reference values is maintained over the entire 1 hour period of observation. Again, the effects of the improved turbulence are shown in the evolution of the thermodynamic fields. Multiple time fields nudging RMSE growth is remarkably low, especially for the first 30 minutes. Thus, whereas three-dimensional nudging has a large negative effect on turbulence build-up and the evolution of the thermodynamic fields, this effect can be mitigated significantly by adopting variations on the nudging method, in particular by using multiple time fields nudging.

- Can thermodynamic fields for use in LES methods accurately be determined from ground-based and satellite observations?

Chapter 3 suggests three approaches for estimating fields for q_t and θ_1 based on an observed LWP field and several ground-based observations: the linear estimation, the $\theta_1' = 0$ limit estimation and the $q_t' = 0$ limit estimation. Whereas all three approaches are fairly successful at reproducing the observed initial fields, the limiting estimation runs showed a better representation of the initial thermodynamic fields than the linear estimation run, contrary to expectations. This indicates that the linear estimation method makes some inaccurate assumptions.

Further evolution of the thermodynamic fields is best for the $\theta_1' = 0$ limit estimation run, followed by the linear estimation run, followed by the $q_t' = 0$ limit estimation run. Clearly, the thermodynamic fields estimated by the different methods have a varying effect on the build-up of turbulence. Thus, 3D thermodynamic fields can be determined from observations to accurately replicate an observed LWP field, but subsequent evolution of the LWP field shows that the estimated fields are different from actual fields and this affects the turbulence build-up.

- Does the implementation of three-dimensional nudging of the thermodynamic fields during the spin-up phase of an LES model give a better solar forecast for a stratocumulus-topped boundary layer than those generated by conventional methods?

The conventional methods referred to here are the persistence method commonly used in solar forecasting, and the default spin-up and 3D initial thermodynamics methods typically used in LES. As a proxy of surface solar radiation, this thesis uses the LWP, which is a good indicator of stratocumulus albedo. The reference LES case used as the state for which forecasts should be made simulates an evolving cloud field. As the persistence method freezes the cloud field, its slab mean LWP always remains constant, whilst the reference slab mean LWP grows. Therefore, all nudge runs, no matter what τ is used, give a better prediction of the mean state of the cloud than the persistence method. Nudge runs also always outperform the default spin-up method in predicting $\overline{\mathrm{LWP}}$. Forecasts of 2 hour nudge runs are most accurate for the lowest values of τ. Interestingly, the 3D initial thermodynamics method gives a very good prediction of the slab mean LWP, which outperforms all 2 hour nudging methods.

The improved nudging methods of short nudging and multiple time fields nudging only produce a better $\overline{\mathrm{LWP}}$ forecast than the 3D initial thermodynamics method on forecast horizons of 45 minutes and more. All in all, nudging methods provide better forecasts for the mean state of stratocumulus than the conventional persistence and default spin-up methods, and short and multiple time fields nudging produce better forecasts of the mean state than the 3D initial thermodynamics method on horizons of 45 minutes and larger.

Another way to rate the accuracy of the forecast is by using the RMSE of the LWP field. This measure is a lot more strict, but the right representation of the entire LWP field is important for solar forecasts, especially in instances of high cloud field heterogeneity. The RMSE is used to compute the forecast skill (FS) of the used methods in comparison to the persistence method, which is used as the benchmark forecast. Considering the FS, both the 3D initial thermodynamics and persistence forecasts outperform the 2 hour nudge forecasts over the entire observed period, with the 3D initial thermodynamics run providing the better forecast of the two. Of the 2 hour nudge runs, the run using $\tau = 60$ s performs the best, followed first by the $\tau = 10$ s and second by the $\tau = 300$ s runs. Because of the relatively poor evolution of the fields in 2 hour nudge runs, the adapted nudge runs were designed. The short nudging run ($\tau = 10$ s, 5 minutes of nudging) shows a better FS than the persistence method on horizons of 20 minutes and larger, and a better FS than the 3D initial thermodynamics method on horizons of 45 minutes and larger. An even more accurate performance is shown by the multiple time fields nudging run, which has an FS that outperforms the persistence method FS on horizons larger than 15 minutes, and the 3D initial thermodynamics method FS on horizons larger than 30 minutes. However, it should be noted that for the first ten minutes, the short nudging run gives a better forecast than the multiple time fields nudging run due to a more accurate initial LWP field.

Finally, it should be emphasized that the nudging, persistence, and 3D initial thermodynamics methods used in this thesis use exact thermodynamic reference fields. It is not possible to obtain thermodynamic fields in such detail from observations, as shown in this thesis. Therefore, when applying said methods to actual observations to create real solar forecasts, these forecasts will perform less well than remarked here. However, their performance relative to each other will stay similar.

To conclude, LES with 3D-nudging is capable of producing a more accurate solar forecast than conventional methods. For forecasts of the mean state of the cloud, all nudge runs outperform the persistence method and short and multiple time fields nudging runs outperform the 3D initial thermodynamics method on horizons of 45 minutes and larger. For forecasts of the cloud fields, 2 hour nudge runs cannot provide forecasts that are more accurate than the persistence and 3D initial thermodynamics methods, but short nudge and especially multiple time fields nudge runs can, on relatively short forecast horizons.

8.2. Recommendations

This section sums up all recommendations made in the discussion, as well as naming some more pointers for future research.

One of the most important recommendations for future research is that the experiments done in this thesis are repeated for other LES cases, simulating both stratocumulus and completely different atmospheric states. For the three-dimensional nudging method, this will gain insight into the effect of τ on replicating desired fields. Furthermore, this will show if the surprisingly good performance of the 3D initial thermodynamics method in this thesis is coincidental or not. It might also indicate if the τ or the time period used in short nudging (5 minutes) is sufficient for all LES cases, or if this depends on the atmospheric state that is being studied. In particular, it is suggested to use a case study with optically thin stratocumulus in further research. Here, perturbations of thermodynamic variables will have a more pronounced effect on the cloud cover, likely giving larger differences in the forecast accuracy of the different methods.

For the use of nudging in real situations, one obstacle remains, which has received limited attention in this work. This is the estimation of thermodynamic fields from observations. As demonstrated, approaches can be derived to estimate fields from observations that can accurately replicate an observed field. This being said, using these estimated fields in nudging runs showed adverse effects on the evolution of the LWP fields. It is recommended to conduct similar experiments on the large domain, to put the results into a better perspective. Also, further research is urged to improve the estimation approaches described in this thesis by working around the more invalid assumptions the suggested approaches make, like the assumption of constant fluctuations at all heights.

Multiple time fields nudging is the most promising of all (nudging) methods investigated in this thesis. It uses reference fields every 10 minutes. The temporal resolution of satellites is usually larger, on the order of 15 minutes (Wang et al., 2019). Further research should therefore focus on how increasing the time step in multiple time fields nudging affects the accuracy of its forecast.

Initial values of the RMSE for multiple time fields nudging show that its initial LWP field is less like the desired LWP field than the initial LWP field in most of the other methods. However, as its turbulence is very similar to that of the reference case, its RMSE growth is smaller than for other methods. It would be interesting to try to combine the low initial RMSE of strong nudging runs ($\tau = 10$ s) with the low RMSE growth of multiple time fields nudging. To do so, one could perhaps use multiple time fields nudging for most of the nudging period, and then apply very strong nudging for only the last few minutes (like in short nudging). Further research should show whether this method produces even better solar forecasts.

Mukherjee et al. (2016) investigate the predictability horizon of LES by considering the error growth of perturbations in two otherwise identical simulations. A similar investigation is possible for the methods used in this research, to compare their predictability horizons. Such research could also show to what extent the methods rely on the correct representation of the desired thermodynamic fields.

Experiments in this research make use of a simplified radiation calculation in DALES, allowing simulations to run much quicker. However, this calculation scheme does not take the interaction between solar radiation and the cloud into account. For future research, it would be interesting to include the full radiation calculation scheme in the experiments. This would not only make cloud evolution more realistic but would also allow one to observe the surface solar radiation and not have to use the LWP as a proxy for solar forecast accuracy. It should be noted that such runs would require a large computation time, and should therefore only be conducted if the necessary time and resources are available.

A

Namoptions files

To run, DALES needs an input file called 'namoptions'. In this appendix, the namoptions file that is used for the reference runs of the ASTEX case (as described in sections 5.1 and 5.2) is given. Key input parameters that were changed between the reference run and the experimental runs are highlighted. Also, as new functionalities were added to the DALES model in this research, the new input parameters associated with these functionalities are explained briefly. For a description of all the different input parameters, the reader is referred to the document written by Heus et al. (2015) which gives an overview of all namoptions. After the discussion of the most relevant input parameters, the full large domain ASTEX namoptions file is given.

A.1. Important inputs

In the block `&RUN` the number of the experiment is given in `iexpnr`. It is an important input as the names of all the other input files should end with it. Also, the time (in seconds) for which the simulation should run is given in `runtime`, which is different for each of the methods. Setting `irandom` is important as well. The number given here refers to the pseudo-random perturbations given to the thermodynamics at the start of the run. When the same number is used, the same perturbations are set. Therefore, a different number should be used between the reference run and the experiments, for a fair comparison. The same block also houses one of the input parameters created for this research. `lcoldstartfiles`. If this boolean switch is set to `.true.`, DALES assumes a cold start run is done which is initialized using the reference fields. The initial reference fields should be given in a file called `coldinifield.inp.iexpnr`, containing first the θ_l field and the q_t field second.

In the next block, `&DOMAIN`, one should give the inputs relating to the domain size, as described earlier. These can be used to switch between the large and the small domain. In `&PHYSICS`, another boolean input switch made for this research is given. When true, `ladvectonly` makes DALES use the persistence method. It turns off all tendency terms except advection and applies the method on the θ_l and q_t fields specified in `advfield.inp.iexpnr`. This file should be created in the same way as `coldinifield.inp.iexpnr`.

Nudging settings are given in the block `&NAMNUDGE`. The boolean switch `lnudge` indicates if DALES should use the nudging calculation or not. Parameters `knudgestart` and `knudgestop` indicate the vertical region where nudging should be applied. The nudging time scales for 1D and 3D nudging are given in `t1Dnudgefac` and `t3Dnudgefac` respectively. At what point in time the nudging calculation should start (in seconds) is given in `tnudgestart`, and when it should stop in `tnudgestop`. Finally, the amount of subsequent points in times to nudge to should be given in `ntnudge3D`. Two additional input files are required when using nudging. Firstly, `nudge.inp.iexpnr` should give the vertical profiles of the thermodynamic variables to which 1D nudging should be applied, for every time step. Secondly, `nudge3D.inp.iexpnr` should give the fluctuation fields to nudge the thermodynamic fields to. In this file, the array containing the points in time at which nudging fields are given should be printed first. After, the required fields of θ_l for all times and positions, and then q_t, should be printed.

Finally, an addition was made to the block `&NAMTIMESTAT`. When switch `llwp_pos` is set to true, DALES outputs the total LWP of a certain position in the field. The value is printed for every period of seconds as set in `dtav`. Four points in the field can be investigated in total, and their horizontal positions should be given in `ips1` and `jps1`, where the number at the end specifies the point in question. LWP values are printed against time in the output files `tmradps1.iexpnr`, where again, the number differs for each position used. This concludes the description of the input parameters most relevant to this research. If the reader is left with questions regarding other input parameters, they are once again referred to Heus et al. (2015).

A.2. ASTEX

```
&RUN
iexpnr     =  001
runtime    =  10800
dtmax      =  10.
ladaptive  =  .true.
irandom    =  43
randthl    =  0.001
randqt     =  2.5e-8
nsv        =  0
lcoldstartfiles = .false.
/

&DOMAIN
itot       =  256
jtot       =  256
kmax       =  427
xsize      =  25600.
ysize      =  25600.
xlat       =  34.
xlon       =  -25.
xday       =  164
xtime      =  17.
/

&PHYSICS
ps         =  102900.
thls       =  289.
lmoist     =  .true.
lcoriol    =  .true.
iradiation =  2
z0         =  2e-4
useMcICA   =  .false.
ladvectonly=  .false.
/

&NAMSURFACE
z0mav      =  2e-4
z0hav      =  2e-4
isurf      =  2
ps         =  102900.
albedoav   =  0.07
/

&NAMMICROPHYSICS
imicro     =  0
l_sb       =  .false.
```

```
l_rain      =  .true.
l_sedc      =  .true.
l_mur_cst   =  .false.
mur_cst     =  0
Nc_0        = 100e6
sig_g       = 1.2
/

/
&NAMBULKMICROSTAT
lmicrostat =  .false.
dtav        =  60
timeav      =  600
/

&DYNAMICS
lqlnr       =  .false.
cu          =  -1
cv          =  -7
llsadv      =  .false.

iadv_mom    =  2
iadv_tke    =  52
iadv_thl    =  52
iadv_qt     =  52
iadv_sv     =  52 52
/

&NAMSUBGRID
ldelta           =  .false.
cn               =  0.76
sgs_surface_fix =  .true.
/

&NAMNUDGE
lnudge         = .false.
knudgestart    = 1
knudgestop     = 130
t3Dnudgefac    = 10
t1Dnudgefac    = 10
tnudgestart    = 0
tnudgestop     = 7200
ntnudge3D      = 1
/

&NAMCHECKSIM
tcheck       = 5
/

&NAMTIMESTAT
ltimestat    = .true.
llwp_pos     = .true.
dtav         = 60
ips1         = 32
jps1         = 32
ips2         = 12
```

```
jps2        = 46
ips3        = 53
jps3        = 22
ips4        = 5
jps4        = 18
/

&NAMGENSTAT
lstat       = .true.
dtav        = 60
timeav      = 600
/

&NAMBUDGET
lbudget     = .false.
dtav        = 60.
timeav      = 600.
/

&NAMRADSTAT
dtav        = 60.
timeav      = 600.
lstat       = .true.
/

&NAMNETCDFSTATS
lnetcdf = .true.
/

&NAMFIELDDUMP
 DTAV        =   300
 LFIELDDUMP  = .true.
 LDIRACC     = .false.
 LBINARY     = .false.
 KLOW        =   1
 KHIGH       =   140
/
```

B

Persistence method numerical noise

During tests of the newly implemented persistence method (sec. 5.4) in DALES, an artifact in the results was encountered. The persistence method assumes that the fields for θ_l and q_t are frozen in time, and then transports these using only the horizontal winds. As the simulation employs periodic boundary conditions, one would expect the fields to be equal to the initial field at all times, except translated. However, upon testing the method, a sort of diffusion of the fields was observed. Figure B.1 shows the diffusion, where the field on the left is the initial field, and the field on the right is found after one hour of simulation. The field after one hour of persistence simulation has been translated to roughly match the

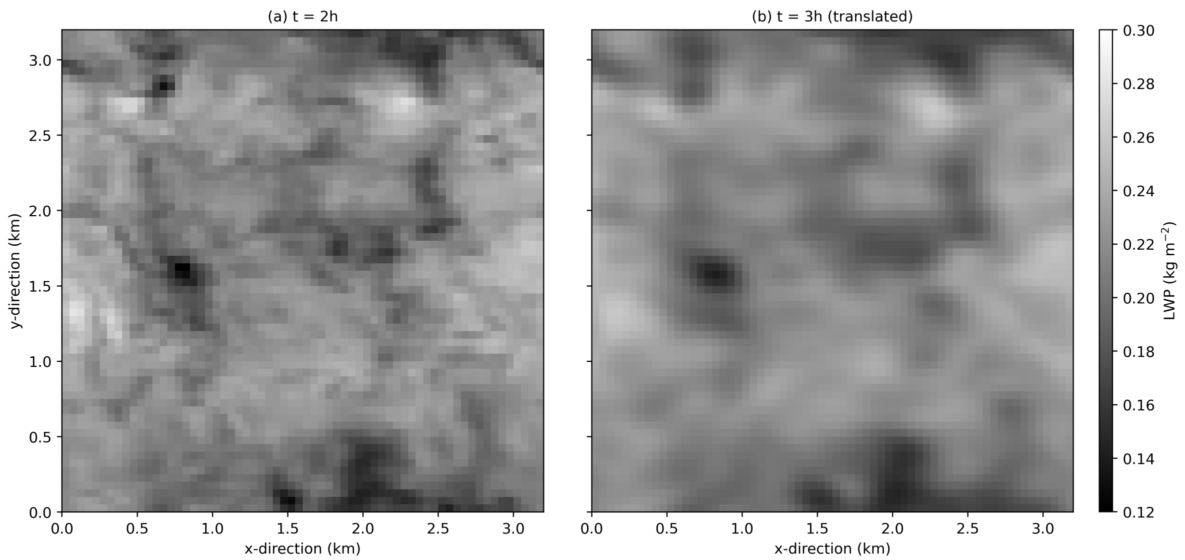


Figure B.1: LWP fields taken from the results of a persistence method on the small domain. (a) shows the initial field, two hours after the start of the reference run, and (b) shows the LWP field after applying the persistence method for one hour. In (b), the field has been translated to match the position of the field in (a), to highlight the numerical noise. The color range in (a) and (b) is the same, therefore only one colorbar is shown.

position of the initial field, so the comparison can be made more easily. The right field is significantly vaguer than the left. All processes other than advection have been turned off, so this diffusion must be caused by the advection process.

To further investigate the problem, a simple test was made for a dry CBL LES case. The passed initial field of θ_l to be advected is homogeneous and equal to 288 K, except for a few points, in which the letters TEST have been spelled by increasing θ_l by 2K (fig. B.2(a)). 90 second persistence method runs were done to investigate the diffusion. Figures B.2 and B.3 show the results of such runs. Again, the diffusion process can be noted from the evolving fields.

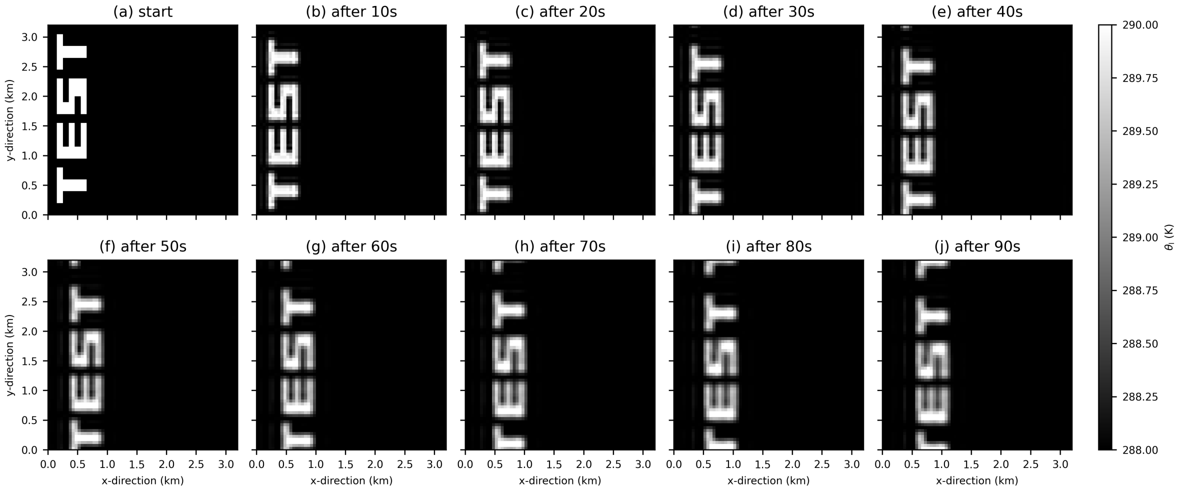


Figure B.2: Field of θ_1 every 10 seconds for a persistence method run of 90 seconds, on a small domain. In (a), the input field is given. Advection is calculated using the 5th-order central difference method.

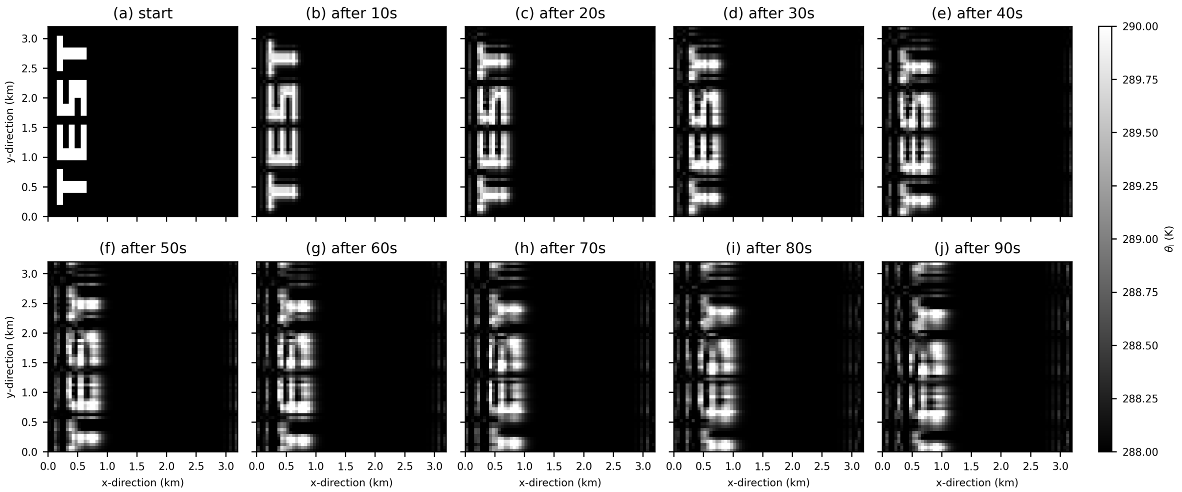


Figure B.3: Field of θ_1 for a persistence method run of 90 seconds, on a small domain. In (a), the input field is given. Advection is calculated using the 2nd-order central difference method.

Advection in DALES can be calculated using various calculation schemes. Most often, the central difference method is used. The central difference method is a way to estimate the slope in a variable ϕ at a point x by taking the difference between the value of ϕ at points just before and after x, and dividing it by the distance between the two points:

$$\frac{\partial \phi}{\partial x}(x) = \frac{\phi(x + \Delta x) - \phi(x - \Delta x)}{2\Delta x} \tag{B.1}$$

where the distance between two points is given by Δx. In similar ways, the higher-order derivative of a variable can be determined. The central difference method introduces a numerical error into the simulation, as it assumes the presence of a constant slope between the different points. This numerical error is what causes the diffusion of the field in the persistence method, which should actually be frozen. DALES allows one to change the order of the central difference method used for advection, further illustrating the effect. Figure B.2 shows the results of a run using the 5th-order central difference method, whereas figure B.3 shows the results of a run using the 2nd-order central difference method. The 2nd-order method generates a lot more numerical noise than the 5th-order method, showing the relevance of the advection calculation scheme. All experimental runs in this research use the 5th-order central difference scheme for horizontal advection.

Bibliography

Albrecht, B. A., Fairall, C. W., Thomson, D. W., White, A. B., Snider, J. B., & Schubert, W. H. (1990). Surface-based remote sensing of the observed and the adiabatic liquid water content of stratocumulus clouds. *Geophysical Research Letters*, *17*(1), 89–92. https://doi.org/10.1029/GL017i001p00089

Atlas, R. L., Bretherton, C. S., Blossey, P. N., Gettelman, A., Bardeen, C., Lin, P., & Ming, Y. (2020). How well do large-eddy simulations and global climate models represent observed boundary layer structures and low clouds over the summertime southern ocean? *Journal of Advances in Modeling Earth Systems*, *12*(11), e2020MS002205. https://doi.org/10.1029/2020MS002205

Barry, R., & Chorley, R. (2003). *Atmosphere, Weather and Climate* (8th). Routlegde.

Blossey, P. N., Bretherton, C. S., Zhang, M., Cheng, A., Endo, S., Heus, T., Liu, Y., Lock, A. P., de Roode, S. R., & Xu, K.-M. (2013). Marine low cloud sensitivity to an idealized climate change: The cgils les intercomparison. *Journal of Advances in Modeling Earth Systems*, *5*(2), 234–258. https://doi.org/10.1002/jame.20025

Chou, S.-H. (2011). An example of vertical resolution impact on WRF-Var analysis. *Electron. J. Oper. Meteorol*, *12*, 1–20.

Chow, C. W., Urquhart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J., & Washom, B. (2011). Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. *Solar Energy*, *85*(11), 2881–2893. https://doi.org/10.1016/j.solener.2011.08.025

Comstock, K. K., Bretherton, C. S., & Yuter, S. E. (2005). Mesoscale variability and drizzle in southeast pacific stratocumulus. *Journal of the Atmospheric Sciences*, *62*(10), 3792–3807. https://doi.org/10.1175/JAS3567.1

Dalton, J. (1801). Essay IV. On the expansion of elastic fluids by heat. In *Memoirs of the literary and philosophical society of Manchester* (pp. 595–602).

Das, U. K., Tey, K. S., Seyedmahmoudian, M., Mekhilef, S., Idris, M. Y. I., van Deventer, W., Horan, B., & Stojcevski, A. (2018). Forecasting of photovoltaic power generation and model optimization: A review. *Renewable and Sustainable Energy Reviews*, *81*, 912–928. https://doi.org/10.1016/j.rser.2017.08.017

Deardorff, J. W. (1980). Stratocumulus-capped mixed layers derived from a three-dimensional model. *Boundary-layer meteorology*, *18*, 495–527.

de Roode, S. R. (2021). *Atmospheric Physics (syllabus)*. Delft University of Technology.

de Roode, S. R., Frederikse, T., Siebesma, A. P., Ackerman, A. S., Chylik, J., Field, P. R., Fricke, J., Gryschka, M., Hill, A., Honnert, R., Krueger, S. K., Lac, C., Lesage, A. T., & Tomassini, L. (2019). Turbulent transport in the gray zone: A large eddy model intercomparison study of the CONSTRAIN cold air outbreak case. *Journal of Advances in Modeling Earth Systems*, *11*(3), 597–623. https://doi.org/10.1029/2018MS001443

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

de Roode, S. R., & Los, A. (2008). The effect of temperature and humidity fluctuations on the liquid water path of non-precipitating closed-cell stratocumulus clouds. *Quarterly Journal of the Royal Meteorological Society*, *134*(631), 403–416. https://doi.org/10.1002/qj.222

de Roode, S. R., Sandu, I., van der Dussen, J. J., Ackerman, A. S., Blossey, P., Jarecka, D., Lock, A., Siebesma, A. P., & Stevens, B. (2016). Large-eddy simulations of EUCLIPSE–GASS lagrangian stratocumulus-to-cumulus transitions: Mean state, turbulence, and decoupling. *Journal of the Atmospheric Sciences*, *73*(6), 2485–2508. https://doi.org/10.1175/JAS-D-15-0215.1

Draxl, C., Allaerts, D., Quon, E., & Churchfield, M. (2021). Coupling mesoscale budget components to large-eddy simulations for wind-energy applications. *Boundary-Layer Meteorology*, *179*(1), 73–98. https://doi.org/10.1007/s10546-020-00584-z

Dufresne, J.-L., & Bony, S. (2008). An assessment of the primary sources of spread of global warming estimates from coupled atmosphere–ocean models. *Journal of Climate*, *21*. https://doi.org/10.1175/2008JCLI2239.1

Duynkerke, P. G., de Roode, S. R., van Zanten, M. C., Calvo, J., Cuxart, J., Cheinet, S., Chlond, A., Grenier, H., Jonker, P. J., Köhler, M., Lenderink, G., Lewellen, D., Lappen, C.-l., Lock, A. P., Moeng, C.-h., Müller, F., Olmeda, D., Piriou, J.-m., Sánchez, E., & Sednev, I. (2004). Observations and numerical simulations of the diurnal cycle of the EUROCS stratocumulus case. *Quarterly Journal of the Royal Meteorological Society*, *130*(604), 3269–3296. https://doi.org/10.1256/qj.03.139

Duynkerke, P. G., & Teixeira, J. (2001). Comparison of the ECMWF reanalysis with FIRE I observations: Diurnal variation of marine stratocumulus. *Journal of Climate*, *14*(7), 1466–1478. https://doi.org/10.1175/1520-0442(2001)014<1466:COTERW>2.0.CO;2

Harshvardhan & Randall, D. A. (1985). Comments on “The parameterization of radiation for numerical weather prediction and climate models”. *Monthly Weather Review*, *113*(10), 1832–1833. https://doi.org/10.1175/1520-0493(1985)113<1832:COPORF>2.0.CO;2

Hartmann, D. L., Ockert-Bell, M. E., & Michelsen, M. L. (1992). The effect of cloud type on earth’s energy balance: Global analysis. *Journal of Climate*, *5*(11), 1281–1304. https://doi.org/10.1175/1520-0442(1992)005<1281:TEOCTO>2.0.CO;2

Heus, T., van Heerwaarden, C., van der Dussen, J., & Ouwersloot, H. (2015). *Overview of all namoptions in DALES*.

Heus, T., van Heerwaarden, C. C., Jonker, H. J. J., Pier Siebesma, A., Axelsen, S., van den Dries, K., Geoffroy, O., Moene, A. F., Pino, D., de Roode, S. R., & Vilà-Guerau de Arellano, J. (2010). Formulation of the Dutch Atmospheric Large-Eddy Simulation (DALES) and overview of its applications. *Geoscientific Model Development*, *3*(2), 415–444. https://doi.org/10.5194/gmd-3-415-2010

IEA. (2022a). Global energy and climate model [License: CC BY 4.0]. https://www.iea.org/reports/global-energy-and-climate-model

IEA. (2022b). Renewable electricity [License: CC BY 4.0]. https://www.iea.org/reports/global-energy-and-climate-model

Large, W. G., & Pond, S. (1982). Sensible and latent heat flux measurements over the ocean. *Journal of Physical Oceanography*, *12*(5), 464–482. https://doi.org/10.1175/1520-0485(1982)012<0464:SALHFM>2.0.CO;2

Law, E. W., Prasad, A. A., Kay, M., & Taylor, R. A. (2014). Direct normal irradiance forecasting and its application to concentrated solar thermal output forecasting – a review. *Solar Energy*, *108*, 287–307. https://doi.org/10.1016/j.solener.2014.07.008

Ma, C.-C., Mechoso, C. R., Robertson, A. W., & Arakawa, A. (1996). Peruvian stratus clouds and the tropical pacific circulation: A coupled ocean-atmosphere GCM study. *Journal of Climate*, *9*(7), 1635–1645. https://doi.org/10.1175/1520-0442(1996)009<1635:PSCATT>2.0.CO;2

Marquez, R., Pedro, H. T., & Coimbra, C. F. (2013). Hybrid solar forecasting method uses satellite imaging and ground telemetry as inputs to ANNs. *Solar Energy*, *92*, 176–188.

Martin, G. M., Johnson, D. W., & Spice, A. (1994). The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds. *Journal of Atmospheric Sciences*, *51*(13), 1823–1842. https://doi.org/https://doi.org/10.1175/1520-0469(1994)051<1823:TMAPOE>2.0.CO;2

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J., Maycock, T., Waterfield, T., Yelekçi, O., Yu, R., & Zhou, B. ((2021). *Climate change 2021: The physical science basis. contribution of working group I to the sixth assesment report of the intergovernmental panel on climate change*. IPCC.

Mathiesen, P., Collier, C., & Kleissl, J. (2013). A high-resolution, cloud-assimilating numerical weather prediction model for solar irradiance forecasting. *Solar Energy*, *92*, 47–61. https://doi.org/10.1016/j.solener.2013.02.018

McKee, T. B., & Cox, S. K. (1974). Scattering of visible radiation by finite clouds. *Journal of Atmospheric Sciences*, *31*(7), 1885–1892. https://doi.org/10.1175/1520-0469(1974)031<1885:SOVRBF>2.0.CO;2

Mukherjee, S., Schalkwijk, J., & Jonker, H. J. J. (2016). Predictability of dry convective boundary layers: An LES study. *Journal of the Atmospheric Sciences*, *73*(7), 2715–2727. https://doi.org/10.1175/JAS-D-15-0206.1

Muralikrishna, I. V., & Manickam, V. (2017). Chapter fourteen - air pollution control technologies. In I. V. Muralikrishna & V. Manickam (Eds.), *Environmental management* (pp. 337–397). Butterworth-Heinemann. https://doi.org/10.1016/B978-0-12-811989-1.00014-2

NASA. (2013). NASA worldview snapshots [Accessed 27th of July 2023]. https://go.nasa.gov/3qZe6Za

National Ecological Observatory Network (NEON). (2023). Bundled data products - eddy covariance (dp4.00200.001). https://data.neonscience.org/data-products/DP4.00002.001

Nicholls, S. (1984). The dynamics of stratocumulus: Aircraft observations and comparisons with a mixed layer model. *Quarterly Journal of the Royal Meteorological Society*, *110*(466), 783–820. https://doi.org/10.1002/qj.49711046603

Patz, J., Campbell-Lendrum, D., Holloway, T., & Foley, J. A. (2005). Impact of regional climate change on human health. *Nature*, *438*, 310–317. https://doi.org/10.1038/nature04188

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., & Hoff, T. E. (2010). Validation of short and medium term operational solar radiation forecasts in the US. *Solar Energy*, *84*(12), 2161–2172. https://doi.org/10.1016/j.solener.2010.08.014

Pincus, R., McFarlane, S. A., & Klein, S. A. (1999). Albedo bias and the horizontal variability of clouds in subtropical marine boundary layers: Observations from ships and satellites. *Journal of Geophysical Research: Atmospheres*, *104*(D6), 6183–6191. https://doi.org/10.1029/1998JD200125

Platt, C. M. R. (1976). Infrared absorption and liquid water content in stratocumulus clouds. *Quarterly Journal of the Royal Meteorological Society*, *102*(433), 553–561. https://doi.org/10.1002/qj.49710243305

Pörtner, H.-O., Roberts, D., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., Rama, B., Belling, D., Dieck, W., Götze, S., Kersher, T., Mangele, P., Maus, B., Mühle, A., & Weyer, N. (2022). *Climate change 2022: Impacts, adaptation and vulnerability working group II contribution to the sixth assessment report of the intergovernmental panel on climate change*. https://doi.org/10.1017/9781009325844

Randall, D. A., Coakley, J. A., Fairall, C. W., Kropfli, R. A., & Lenschow, D. H. (1984). Outlook for research on subtropical marine stratiform clouds. *Bulletin of the American Meteorological Society*, *65*(12), 1290–1301. https://doi.org/10.1175/1520-0477(1984)065<1290:OFROSM>2.0.CO;2

Randall, D. (1984). Stratocumulus cloud deepening through entrainment. *Tellus A.*, *36A*(5), 446–457.

Ritchie, H., Roser, M., & Rosado, P. (2020). CO_2 and greenhouse gas emissions [https://ourworldindata.org/co2-and-greenhouse-gas-emissions]. *Our World in Data*.

Romps, D. M. (2017). Exact expression for the lifting condensation level. *Journal of the Atmospheric Sciences*, *74*(12), 3891–3900. https://doi.org/10.1175/JAS-D-17-0102.1

Schalkwijk, J., Jonker, H. J. J., Siebesma, A. P., & Bosveld, F. C. (2015). A year-long large-eddy simulation of the weather over cabauw: An overview. *Monthly Weather Review*, *143*(3), 828–844. https://doi.org/10.1175/MWR-D-14-00293.1

Siebesma, A. P., Jakob, C., Lenderink, G., Neggers, R. A. J., Teixeira, J., Van Meijgaard, E., Calvo, J., Chlond, A., Grenier, H., Jones, C., Köhler, M., Kitagawa, H., Marquet, P., Lock, A. P., Müller, F., Olmeda, D., & Severijns, C. (2004). Cloud representation in general-circulation models over the northern Pacific Ocean: A EUROCS intercomparison study. *Quarterly Journal of the Royal Meteorological Society*, *130*(604), 3245–3267. https://doi.org/10.1256/qj.03.146

Srikrishnan, V., Young, G. S., & Brownson, J. R. S. (2017). Skill and skill prediction of cloud-track advection-only forecasting under a cumulus-dominated regime. *Journal of Applied Meteorology and Climatology*, *56*(3), 651–665. https://doi.org/10.1175/JAMC-D-16-0224.1

Stull, R. B. (1988). *An introduction to boundary layer meteorology*. Kluwer Academic Publishers. https://doi.org/10.1007/978-94-009-3027-8

Taylor, G. I. (1938). The spectrum of turbulence. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, *164*(919), 476–490. https://doi.org/10.1098/rspa.1938.0032

van der Dussen, J. J. (2015). *The transition of stratocumulus to cumulus during ASTEX: An LES study* (Doctoral dissertation). Delft University of Technology.

van der Dussen, J. J., de Roode, S. R., & Siebesma, A. P. (2014). Factors controlling rapid stratocumulus cloud thinning. *Journal of the Atmospheric Sciences*, *71*(2), 655–664. https://doi.org/10.1175/JAS-D-13-0114.1

van Zanten, M. C., Stevens, B., Vali, G., & Lenschow, D. H. (2005). Observations of drizzle in nocturnal marine stratocumulus. *Journal of the Atmospheric Sciences*, *62*(1), 88–106. https://doi.org/10.1175/JAS-3355.1

Wallace, J., & Hobbs, P. V. (2006). *Atmospheric science; an introductory survey* (2nd). Elsevier.

Wang, P., van Westrhenen, R., Meirink, J. F., van der Veen, S., & Knap, W. (2019). Surface solar radiation forecasts by advecting cloud physical properties derived from meteosat second generation observations. *Solar Energy*, *177*, 47–58. https://doi.org/10.1016/j.solener.2018.10.073

Warren, S., Hahn, C., Chervin, R. M., & Jenne, R. (1986). Global distribution of total cloud cover and cloud type amounts over land (no. NCAR/TN-273+STR). University corporation for atmospheric research. https://doi.org/10.5065/D6GH9FXB

Wood, R. (2015). CLOUDS AND FOG | Stratus and Stratocumulus. In G. R. North, J. Pyle, & F. Zhang (Eds.), *Encyclopedia of atmospheric sciences (second edition)* (Second Edition, pp. 196–200). Academic Press. https://doi.org/10.1016/B978-0-12-382225-3.00396-0

Wood, R., Bretherton, C. S., & Hartmann, D. L. (2002). Diurnal cycle of liquid water path over the subtropical and tropical oceans. *Geophysical Research Letters*, *29*(23), 7-1-7–4. https://doi.org/10.1029/2002GL015371

Wood, R. (2012). Stratocumulus clouds. *Monthly Weather Review*, *140*(8), 2373–2423. https://doi.org/10.1175/MWR-D-11-00121.1

Wood, R., & Taylor, J. P. (2001). Liquid water path variability in unbroken marine stratocumulus cloud. *Quarterly Journal of the Royal Meteorological Society*, *127*(578), 2635–2662. https://doi.org/10.1002/qj.49712757807

Wyant, M. C., Bretherton, C. S., Chlond, A., Griffin, B. M., Kitagawa, H., Lappen, C.-L., Larson, V. E., Lock, A., Park, S., de Roode, S. R., Uchida, J., Zhao, M., & Ackerman, A. S. (2007). A single-column model intercomparison of a heavily drizzling stratocumulus-topped boundary layer. *Journal of Geophysical Research: Atmospheres*, *112*(D24). https://doi.org/10.1029/2007JD008536