Gravity waves generated by thunderstorms

Naren Athreyas, Kashyapa Bramha

2019

Naren Athreyas, K. B. (2019). Gravity waves generated by thunderstorms. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/85056>

<https://doi.org/10.32657/10220/49183>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

GRAVITY WAVES GENERATED BY THUNDERSTORMS

**KASHYAPA BRAMHA NAREN ATHREYAS
SCHOOL OF ELECTRICAL & ELECTRONIC
ENGINEERING**

2019

Gravity Waves generated by Thunderstorms

Kashyapa Bramha Naren Athreyas

School of Electrical & Electronic Engineering

**A thesis submitted to the Nanyang Technological University in
partial fulfilment of the requirement for the degree of Doctor of
Philosophy**

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

01/04/2019

Date

Kashyapa Bramha Naren Athreyas

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

1 April 2019
.....

Date

Erry Gunawan
.....

Erry Gunawan

Authorship Attribution Statement

This thesis contains material from three paper(s) published in the following peer-reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.

Chapter 3, 4 and 5 of this thesis comprise of the models and algorithms developed as published in the following journals/conferences:

- K Naren Athreyas, E Gunawan, and B. K. Tay, "Inertia-gravity waves generated by thunderstorms: South-East Asia tropical region study", *Advances in Space Research*, Available online on 21 February 2019.
- K. Naren Athreyas, E. Gunawan, and B. K. Tay, "Source tracing of thunderstorm generated inertia-gravity waves observed during the RADAGAST campaign in Niamey, Niger," *J. Atmos. Solar-Terrestrial Phys.*, vol. 172, no. January, pp. 1–9, 2018.
- K Narenathreyas, E Gunawan, and B. K. Tay, "Estimation of latent heat profiles of deep convective clouds using CloudSat", *IEEE Radar Conf.* 2018.

The contributions of the co-authors are as follows:

- Assoc. Prof. Erry Gunawan provided the necessary guidance and direction for the projects.
- I prepared the manuscripts which was reviewed and edited by both Assoc. Prof. Erry Gunawan and Dr. Tay Bee Kiat.
- I performed all the simulations, experiments and data analysis under the helpful guidance of both Assoc. Prof. Erry Gunawan and Dr. Tay Bee Kiat.

01/04/2019

Date

Kashyapa Bramha Naren Athreyas

Acknowledgements

I would like to thank Prof Erry Gunawan who has constantly guided me and provided me with critical comments which helped me to solve the problems during my PhD course. I would also like to thank him for giving me the opportunity to pursue this doctorate program and he was kind enough to tolerate me these past four years. I would like to thank Nanyang Technological University for giving me the opportunity to pursue my PhD with outstanding computing and study facilities.

I would like to thank the thesis advisory committee members: Dr. Tay Bee Kiat who has provided me with the direction and made my transition from an engineering background to developing knowledge in meteorology, relatively easy. He has constantly helped me with by judging my work and also getting involved during my paper submissions. I would also like to thank my other thesis advisory committee member Prof Guan Yong Liang who has guided me in my work.

I would like to thank the various atmospheric data providers – University of Wyoming, atmospheric sounding data centre, CloudSat data centre, MODIS data centre, NOAA data centre. I would like to thank the researchers for free software for calculations – Multiscatter and SSGRA software for calculating scattering properties by Robin Hogan, and RAMS software for cloud-resolving simulations by Colorado State University.

Table of contents

ACKNOWLEDGEMENTS	I
TABLE OF CONTENTS	II
SUMMARY	V
LIST OF FIGURES	VII
LIST OF TABLES	XI
ABBREVIATIONS	XII
NOTATIONS	XIII
1. INTRODUCTION	1
1.1. Background and motivation	2
1.2. Objectives	4
1.3. Scientific contribution	5
1.4. Thesis structure	5
2. BACKGROUND ON GRAVITY WAVES	7
2.1. Gravity waves	10
2.1.1. Linear theory	11
2.1.2. Wave scales and propagation	13
2.1.3. Sources	14
2.2. Thunderstorm as a source	16
2.2.1. Thunderstorm development	16
2.2.2. Waves from the thunderstorms	19

2.3. Research synthesis	22
3. INERTIA- GRAVITY WAVES GENERATED BY THUNDERSTORMS IN SOUTH-EAST ASIA TROPICAL REGION	24
3.1. Data and methods	24
3.1.1. Wavelet analysis	26
3.1.2. Thunderstorm source tracing	30
3.1.3. Cloud-resolving model simulations	35
3.2. Results	38
3.2.1. Velocity perturbations	38
3.2.2. Wave amplitude	40
3.2.3. Diagnosing the source	42
3.2.4. Validation	43
3.3. Conclusions	45
4. ESTIMATION OF THE VERTICAL STRUCTURE OF LATENT HEAT GENERATED IN THUNDERSTORMS USING CLOUDSAT	47
4.1. Introduction	48
4.2. Algorithm description	50
4.2.1. Overview	50
4.2.2. Bayesian Monte Carlo method	51
4.3. Data products and methods	54
4.3.1. CloudSat data	54
4.3.2. Cloud-resolving model	55
4.3.3. Scattering models	56
4.3.4. Database properties	58
4.4. Results and discussion	60
4.4.1. Quantitative analysis	60
4.4.2. Case study	64

4.5. Conclusions	65
5. GRAVITY WAVES GENERATED BY THUNDERSTORM: SATELLITE BASED STUDY	67
5.1. Introduction	68
5.2. Latent heat maps using MODIS	69
5.2.1. Data and concept	69
5.2.2. Bayesian neural network	71
5.3. Modelling gravity waves generated by thunderstorms: MODIS based study	75
5.3.1. South-East Asia region	75
5.3.2. West Africa tropical region	79
5.4. Conclusions	85
6. CONCLUSIONS AND FUTURE WORK	87
PUBLICATIONS	91
REFERENCES	92
APPENDIX A – RAMS-CRM VALIDATION	107
APPENDIX B – JUSTIFICATION OF DATABASE SAMPLES FOR BMC METHOD	112

Summary

Gravity waves are considered to be one of the integral parts of the atmosphere which are responsible for energy and momentum distribution among various layers and regions in the atmosphere. These can be generated from various sources such as earthquakes, volcanoes and thunderstorms which can create disturbances in the atmosphere. Thunderstorms are one of the phenomena which occur quite frequently in the tropical region. Thunderstorms are known to generate a wide range of waves including electromagnetic, sound and gravity waves. As it is known, thunderstorms are one of the strongest disturbances in the atmosphere involving huge amount of energy, therefore the gravity wave perturbations created by this can carry large amount of energy and momentum into the middle atmosphere which can affect the circulation and constitution of the layer where the wave dissipates.

This thesis is focused on the quantifying the gravity wave perturbations created by the latent heat inside a thunderstorm which is the driving force of the storm, as these relationships have not been quantified based on a large-scale statistical study. Multi-disciplinary approaches have been used in order to develop a gravity wave-latent heat model by studying the ten-year radiosonde data from Singapore. The gravity waves are detected in the stratosphere using the radiosonde profiles, and the source of these waves are traced back using ray tracing technique. The sources are classified into thunderstorms using global outgoing radiation maps and these thunderstorms sources are further analysed for the latent heat generation. Since the latent heat is not a measurable quantity, it is realised using a cloud-resolving model. The simple linear model developed is applicable to the South-East Asia region. The model is able to estimate waves phase velocities and wave amplitudes which are validated using the radiosonde in Singapore when the thunderstorm is in the close vicinity of the station. The model shows reasonable performance with mean estimation error between 14-23% and standard deviation of estimation between 11-19%.

Since the cloud-resolving models are computationally expensive, it becomes quite impractical to perform these quantifications on a large/global-scale. Therefore, satellite-based latent heat estimations are deemed to be necessary. The latent heat estimations are developed for CloudSat's radar and MODIS imaging spectroradiometer which can be used to develop gravity wave models on a global scale. The cloud resolving models are used to develop a database of thunderstorms valid for a region (eg. South-east Asia). The database of reflectivity profiles of thunderstorms simulated using a cloud-resolving model is used for comparing with CloudSat radar reflectivity profiles of thunderstorms using Bayesian Monte-Carlo approach to determine the vertical profiles of latent heat. The performance of the algorithm is evaluated using the simultaneous cloud-resolving model simulations near Singapore in 2016. The estimated latent heat profiles are used for training a Bayesian neural network with cloud inputs from MODIS on-board Aqua. This approach exploits the synergy between the CloudSat and Aqua satellites which are on the similar orbit with 60 s time lag. The gravity wave models are developed using satellite based latent heat estimations for South-East Asia and West Africa regions with validation performed using local radiosonde profiles. The gravity wave models developed using satellites also show reasonable performance which gives encouragement and hope for developing global relationships. The models developed here represent the interactions that affect circulation and constituency of the atmosphere which can be used for improving the weather models.

List of figures

Figure 1-1: Synoptic and mesoscale dynamic processes and interactions in the atmosphere [9].	2
Figure 2-1: Schematic of atmospheric layers according to the temperature, and also atmospheric constituents and processes [34].	8
Figure 2-2: Phases of thunderstorm cell evolution. Light grey shading represents clouds; medium grey is the pre-storm boundary layer of warm humid air that serves as fuel for the storm. Dark shadowing is colder outflowing air. Diagonal white lines represent precipitation. Arrows show air motion. <i>Cu con</i> = cumulus convection [1].	17
Figure 2-3: The vertical cross-section of a mesoscale convective system [1].	18
Figure 2-4: Model simulation results showing the wave patterns generated by an isolated thunderstorm in the mid-latitudes. The portion of the cloud penetrating 10 km is shown by the lower surface. Isentropic surfaces at 15 and 20 km with waves are shown in the top two layers [11].	21
Figure 3-1: Schematic diagram showing the stages of the study.	25
Figure 3-2: (top panel) Time series of the kinetic energy density averaged from 19 to 25 km for the ten-year duration. (bottom panel) Wavelet power spectrum of kinetic energy density averaged from 19 to 25 km for the ten-year duration. Cone of influence highlighted in thick black line and the shaded region.	27
Figure 3-3: 100 cases of ray-traced paths are shown which are identified as gravity waves generated from thunderstorms. The ray paths are shown in red, the black lines represent geographic map and the contours indicate OLR data averaged from 2007 to 2016.	35
Figure 3-4: Variation of gravity wave phase velocity with instantaneous latent heat release (K) in thunderstorms. (left) Horizontal phase velocity (m/s) and (right) vertical phase velocity (m/s). The error bars represent 95% standard deviation. The velocities shown are absolute values.	39
Figure 3-5: Variation of gravity wave phase velocity with time derivative of latent heat release (K/s) in thunderstorms. (left) horizontal phase velocity (m/s) and (right) vertical phase velocity (m/s). The error bars represent 95% standard deviation. The velocities shown are absolute values.	39

Figure 3-6: Variation of gravity wave amplitudes with the instantaneous latent heat of thunderstorms. (left) vertical wave action flux, $\log_{10}(F)$ (kg/s^2) and (right) horizontal velocity power spectrum, $\log_{10}(E)$ (m^2/s^2). The amplitudes shown are the logarithm of absolute values.....	40
Figure 3-7: Variation of gravity wave amplitudes with the time derivative of latent heat of thunderstorms. (left) vertical wave action flux, $\log_{10}(F)$ (kg/s^2) and (right) horizontal velocity power spectrum, $\log_{10}(E)$ (m^2/s^2). The amplitudes shown are the logarithm of absolute values.	41
Figure 3-8: Plot showing the histograms of the rate of change of latent heat (K/s) for different bins of absolute latent heat (K).	42
Figure 4-1: Block diagram describing the algorithm flow.	51
Figure 4-2: Mean and standard deviation of reflectivity profiles in the database. ...	59
Figure 4-3: Mean % error in the latent heat retrieval with RAMS validation.	61
Figure 4-4: Linear regression plot comparing the CloudSat estimated values and collocated RAMS simulated values of column-integrated latent heat. The plot shows the linear fit (black lines), 95% prediction intervals (black dashed lines) intercept, R^2 , bias and root mean square error (RMSE) values. The blue thick line shows the new fit line after neglecting the estimated values < 2 , and linear equation for new fit in blue text.....	62
Figure 4-5: Linear regression plot comparing the CloudSat estimated values and collocated WRF simulated values of column-integrated latent heat. The plot shows the linear fit (black lines), 95% prediction intervals (black dashed lines) intercept, R^2 , bias and root mean square error (RMSE) values.....	63
Figure 4-6: Case study profiles, (from top to bottom) a) – cloud scenario profiles obtained from CLDCLASS product which shows the different types of clouds. Here is the number 8 represents the thunderstorms which is seen as the red patch in the plot. b) The corresponding reflectivity (dBZ) profiles obtained from GEOPROF product, c) retrieved latent heat (K).....	65
Figure 5-1: Schematic diagram of the A-Train satellite depicting their right-ascending node [105].....	69
Figure 5-2: Schematic depicting the MODIS cloud image along with coinciding CPR vertical cloud profiles. The collocated data points for neural network training are highlighted by black dotted lines.....	71
Figure 5-3: Block of the neural network with hidden and output layers.	72

Figure 5-4: Regression plot showing the performance of training in SEA region. The axes represent the latent heat values (K). The blue line represents the linear fit of the data.	74
Figure 5-5: Regression plot showing the performance of training in WA region. The axes represent the latent heat values (K). The blue line represents the linear fit of the data.	74
Figure 5-6: MODIS latent heat image on the geographical maps with ray path (Thick black line) of the gravity wave generated by the thunderstorm shown in the red dotted circle.	76
Figure 5-7: Variation of gravity wave phase velocity with instantaneous latent heat release (K) in thunderstorms. (left) Horizontal phase velocity (m/s) and (right) vertical phase velocity (m/s). The velocities shown are absolute values.	77
Figure 5-8: Variation of gravity wave amplitude with the instantaneous latent heat of thunderstorms. The values are shown in terms of wave energy, $\log_{10}(E)$ (m^2/s^2). The amplitudes shown are the logarithm of absolute values.	77
Figure 5-9: Ray tracing results in WA region. a) Shows the geographic ray paths which are generated from thunderstorms. The background colours show the average % of deep convective activity in 2006. b) time-altitude ray paths showing propagation times of waves generated from thunderstorms.	82
Figure 5-10: Variation of gravity wave phase velocity with instantaneous latent heat release (K) in thunderstorms. (left) Horizontal phase velocity (m/s) and (right) vertical phase velocity (m/s). The velocities shown are absolute values.	83
Figure 5-11: Variation of gravity wave amplitude with the instantaneous latent heat of thunderstorms. The values are shown in terms of wave energy, $\log_{10}(E)$ (m^2/s^2). The amplitudes shown are the logarithm of absolute values.	83
Figure A-0-1: Mean reflectivity error profiles near Singapore region.	108
Figure A-0-2: (from left to right) Profiles of mean percentage error of a) mixing ratio, b) temperature and c) potential temperature.	109
Figure A-0-3: Linear fit plots of comparison between RAMS data and radiosonde data of temperature, potential temperature and mixing ratio.	110
Figure A-0-4: Mean reflectivity error profile in the tropical WA region.	111
Figure B-0-1: Performance of the algorithm vs. number of samples plot. The performance indicator in this study is considered to be the gradient value between the retrieved latent heat from the algorithm and the true latent heat from RAMS.	113

Figure B-0-2: Time to retrieval a single profile of latent heat vs. number of samples

plot..... 113

List of tables

Table 3-1: RAMS model configuration.	37
Table 3-2: Comparison of model retrieved values and radiosonde values for gravity wave generated by thunderstorms near Singapore radiosonde station. M – GW-RAMS Model, R – Radiosonde.....	44
Table 5-1: Comparison of GW-RAMS retrieved values, GW-Sat-SEA retrieved values and radiosonde values for gravity wave horizontal velocity perturbation and vertical wavelength generated by thunderstorms near Singapore radiosonde station. M – GW-RAMS Model, R – Radiosonde, S – GW-Sat-SEA model.....	78
Table 5-2: Comparison of GW-RAMS retrieved values, GW-Sat-SEA retrieved values and radiosonde values of gravity wave amplitude generated by thunderstorms near Singapore radiosonde station. M – GW-RAMS Model, R – Radiosonde, S – GW-Sat-SEA model.....	78
Table 5-3: Comparison of GW-RAMS retrieved values, GW-Sat-WA retrieved values and radiosonde values of gravity wave properties generated by thunderstorms in the WA region. R – Radiosonde, S – GW-Sat-WA model. ...	84

Abbreviations

ACW	Anti-Clock Wise
ARM	Atmospheric Radiation Measurement
BMC	Bayesian Monte Carlo
BT	Brightness Temperature
CER	Cloud Effective Radius
COT	Cloud Optical Thickness
CPR	Cloud Profiling Radar
CRM	Cloud-Resolving Model
CTT	Cloud Top Temperature
CW	Clock Wise
DCC	Deep Convective Clouds
ECMWF	European Centre for Medium-range Weather Forecasts
GARP	Global Atmospheric Research Program
GROGRAT	Global or Regional Gravity-wave Ray Tracer
ISCCP	International Satellite Cloud Climatology Project
IWC	Ice Water Content
JAXA	Japan Aerospace Exploration Agency
LEAF	
LH	Latent heat
LWC	Liquid Water Content
MCS	Mesoscale Convective System
MODIS	Moderate-resolution Imaging Spectro-radiometer
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
OLR	Outgoing Long-wave Radiation
RAMS	Regional Atmospheric Modeling System
RMSE	Root Mean Square Error
SEA	South-East Asia
SSE	Sum of Squares Error
SSGRA	Self-Similar Rayleigh-Gans Approximation
TOGA-COARE	Tropical Ocean Global Atmosphere program- Coupled Ocean Atmosphere Research Experiment
TRMM	Tropical Rainfall Measurement Mission
UTC	Universal Coordinated Time
WA	West Africa
WKB	Wentzel-Kramers-Brillouin
WRF	Weather Research and Forecasting

Notations

c_g	Wave group velocity (m/s)
c_{gz}	Wave group velocity in vertical direction (m/s)
c_s	Speed of sound (m/s)
f	Coriolis parameter (Hz)
g	Acceleration due to gravity (m/s ²)
k	Wavenumber in x direction (m ⁻¹)
k_H	Total wavenumber in horizontal plane (m ⁻¹)
l	Wavenumber in y direction (m ⁻¹)
m	Wavenumber in z direction (m ⁻¹)
p	Pressure (Pa)
$\check{p}$	Probability distribution of BMC retrieval
$\bar{p}$	Mean pressure (Pa)
$\tilde{p}$	Wave pressure amplitude (Pa)
r	Correlation between the error elements in BMC method
t	Time (s)
u	Velocity component in x direction (m/s)
$\bar{u}$	Mean velocity in x direction (m/s)
u'	Velocity perturbation in x direction (m/s)
$\tilde{u}$	Wave amplitude in x direction (m/s)
v	Velocity component in y direction (m/s)
$\bar{v}$	Mean velocity in y direction (m/s)
v'	Velocity perturbation in y direction (m/s)
$\tilde{v}$	Wave amplitude in x direction (m/s)
w	Velocity component in z direction (m/s)
$\bar{w}$	Mean velocity in z direction (m/s)
w'	Velocity perturbation in z direction (m/s)
$\tilde{w}$	Wave amplitude in x direction (m/s)
x	Displacement in x direction (m)
x_r	Retrieved vector of latent heat
x_s	Simulated matrix database of latent heat
y	Displacement in y direction (m)
y_o	Observation vectors of reflectivity and temperature
y_s	Simulated matrix database of reflectivity and temperature
z	Displacement in z direction (m)
A	Wave action density (kg/ms)
A_{ext}	Hydrometeor extinction area (m ²)
B	Neural network database
C	Coefficient in the Stokes parameter equation
D	Stokes parameter
E	Energy density (kg/ms ²)

E_D	Neural network error function
E_k	Kinetic energy density (kg/ms ²)
E_p	Potential energy density (kg/ms ²)
E_W	Sum of squares of weights
F	Vertical flux of wave action (kg/s ²)
$G^{(b)}$	Neural network targets
H	Density scale height (m)
I	Stokes parameter
M	Total number of samples in the database for BMC
N	Buoyancy frequency (Hz)
N_c	Hydrometeor number concentration (#/m ³)
O	Error covariance matrix of observation
P	Stokes parameter
Pr	Prandtl number
Q	Stokes parameter
S	Error covariance matrix of simulated database
T	Temperature (K)
$\bar{T}$	Mean temperature (K)
T'	Perturbation in temperature (K)
T'_{+90}	Hilbert's transform of temperature perturbation (K)
U	Background velocity in x direction for ray tracing (m/s)
$\bar{U}$	Mean velocity in horizontal plane (m/s)
$\mathcal{U}_R$	Real part of Fourier transform of u' (m/s)
$\mathcal{U}_I$	Imaginary part of Fourier transform of u' (m/s)
V	Background velocity in y direction for ray tracing (m/s)
$\mathfrak{B}_R$	Real part of Fourier transform of v' (m/s)
$\mathfrak{B}_I$	Imaginary part of Fourier transform of v' (m/s)
$X^{(b)}$	Neural network inputs
Y	Neural network output
Z	Reflectivity (dBZ)
Z_e	Apparent reflectivity (dBZ)
α	$= 1/2H$, represents inverse of twice of scale height
β	Neural network regularisation parameter
β_e	Apparent backscatter
γ	Neural network regularisation parameter
δ	WKB approximation threshold variable
ϵ	Error between y_s and y_o
$\bar{\epsilon}$	Mean of ϵ
ε	Neural network weights
θ	Potential temperature (K)
$\bar{\theta}$	Mean potential temperature (K)
κ	Dielectric constant of water
λ	Wavelength (m)

λ_h	Horizontal wavelength (m)
λ_v	Vertical wavelength (m)
ρ	Density (kg/m ³)
$\bar{\rho}$	Mean density (kg/m ³)
$\tilde{\rho}$	Wave amplitude in density (kg/m ³)
σ	Standard deviation
σ_x	Uncertainty in the retrieval of x_r
ς_{ext}	Hydrometeor extinction coefficient (m ⁻¹)
τ	Damping timescale (s)
τ_D	Turbulent damping timescale (s)
τ_r	Radiative damping timescale (s)
ϕ	Latitude (rad)
ω	Eulerian frequency or ground-based frequency (Hz)
$\hat{\omega}$	Intrinsic frequency (Hz)
Φ	Wave propagation direction in horizontal plane (rad)
Ψ	Modified neural network error function
Ω	Earth's inertial frequency (Hz)

1. Introduction

Earth is a planet with complex systems connected and coupled with each other in order to maintain stability and equilibrium. Just like every other system, when a disturbance affects the system, the system tries to bring back the state of equilibrium; the earth's atmosphere also works in a similar way [1], [2]. Various vertical layers of earth's atmosphere are coupled with each other as the atmosphere is a continuous fluid which is horizontally stratified [1]. Even though the mixing of constituents do not occur at all altitudes levels, the distribution and propagation of energy and momentum happens continuously [3], [4]. One of the main phenomena which are responsible for energy and momentum distribution in the atmosphere are waves [5]. There are various types of waves which carry energy away from the source and distribute and dissipate energy as they propagate such as sound waves, electromagnetic waves and gravity waves [1]. The main focus of this study is gravity waves which operate at much lower frequencies than the sound and electromagnetic waves. Any disturbances in the atmosphere, ocean or land such as earthquakes, volcanoes, thunderstorms, jets and fronts etc., are capable of producing gravity waves [3], [5], [6]. This study focuses on gravity waves which are generated from thunderstorms, which occur frequently in the tropical region [7]. This chapter briefly introduces to the process of thunderstorm-generated gravity waves and their significance, followed by the objectives and contributions made in the study.

1.1. Background and motivation

Gravity waves are an important and integral part of the atmosphere which helps to maintain energy and momentum balance and drive hydro cycles [8]. Various processes in the atmosphere and land that can generate gravity waves and other interactions are shown in Figure 1-1. As seen in Figure 1-1, the most common sources of gravity waves in the troposphere are topography, volcanoes, thunderstorms and spontaneous imbalances etc.

Figure 1-1: Synoptic and mesoscale dynamic processes and interactions in the atmosphere [9].

Thunderstorms are one of the most commonly occurring disturbances in the tropical region which usually possess large amount of energy [1]. They affect diurnal human life in many ways: heavy rains, hailstorms, winds, and lightning hazards etc. [1]. Most of the thunderstorms studies few decades ago were mainly limited to the aspects mentioned above [10]. Since last few decades, the waves generated from thunderstorms are actively studied which is important for weather modelling and prediction. The gravity waves generated from thunderstorms can drive large winds which can alter the local weather

immediately, and also affect events at a large distance from them [11]. Typically, thunderstorms are formed when the warm humid boundary layer air rises up usually around late afternoons due to buoyancy [1], [12], [13]. Yet, thunderstorms are always encountered in the tropical region. It is speculated that the downward propagating gravity waves which drive the winds to cause lift in boundary layer air are one of reasons for convective initiation [11]. However, the major impact of gravity waves generated from thunderstorms is in middle atmosphere where the upward propagating waves usually dissipate and break causing rapid winds which affect the constituents of the layer [3], [4].

Over the years, various numerical studies have simulated the structures and characteristics of gravity waves generated by thunderstorms in both idealised and non-idealised environments [14]–[21]. Various observational studies have also identified the characteristics of the waves generated by the thunderstorms [22]–[25]. It has been noted that the latent heat release due to the phase change of hydrometeors in the thunderstorms is the driving force for wave generation. Although some simulated and theoretical studies have parameterised that relationship between the thunderstorms and the waves generated by them [26]; these have been limited to mainly qualitative studies (with case studies) and there are no observational studies which have quantified the relationship. It is also challenging because the latent heat is not a quantity which is directly measurable. These problems have given the main motivation for this study which aims to develop a quantified model between the latent heat release in the thunderstorms and the properties of gravity waves generated by them using long-term data analysis.

The studies involving latent heat are mainly model-based studies which can resolve the convection as latent heat is not a measurable quantity [27]. This type of model will be limited to a small region (few 10s of *km* only) as the dynamic models which are capable of resolving convection are computationally expensive and running them on a global level is impractical [28], so long-term studies will be computationally expensive and may take months to complete. Therefore, satellite-based methods are required which can estimate the latent heat of thunderstorms on a large scale which can reduce the computation time

considerably. Algorithms have been developed for only two satellite radars which can measure precipitation: Tropical Rainfall Measuring Mission (TRMM) and Global Precipitation Measurement (GPM) [29], [30]. However, these radars cover quite a narrow footprint, therefore it is beneficial to develop methods for other satellites to increase coverage. Developing satellites methods for latent heat estimation over a large region not only provides a good source of data for this study but is beneficial for monitoring the atmosphere in general which can help to study the effects of latent heat on a regular basis. Usually, only satellite images cover large scale and MODIS is one of them which can be used for thunderstorm studies. Since MODIS on Aqua is closely associated with CPR on CloudSat, it is sensible to develop methods for estimating latent heat for CPR which can actively sense the profiles of thunderstorms and use this estimation for training collocated MODIS inputs. Developing a method for MODIS is advantageous as it is on-board two satellites which can provide data consistently at different times of the day.

1.2. Objectives

Based on the previous literature, there are four main objectives of this study.

- To develop a quantified and validated model between the latent heat release in the thunderstorms and the properties of gravity waves generated by them using cloud-resolving models for realising latent heat of thunderstorms.
- To develop validated satellite-based techniques that can help to monitor the latent heat of thunderstorms consistently on a global level.
- To implement the satellite-based methods to develop a latent heat-gravity wave model in different tropical regions with ground-based validation.
- Compare the performance of satellite-based gravity wave-latent heat models with cloud-resolving simulation-based gravity wave-latent heat models.

1.3. Scientific contribution

Based on the objectives of this study, there are three main research contributions and outcomes. The first research problem addressed here is the gravity wave-latent heat quantification. The relationship has been quantified using cross-disciplinary approaches based on observational data, cloud-resolving numerical simulations and data analysis. This relationship is modelled in South-East Asia region which is validated using radiosonde profiles in Singapore where the importance and necessity of accurate weather modelling are growing due to increasing automated systems.

The second research problem identified here is the difficulty of developing such models on a global level using cloud-resolving numerical simulations (which are computationally expensive). This problem is solved by developing satellites techniques which can provide consistent global coverage of thunderstorm's latent heat on a large scale. Therefore, an algorithm is developed first to retrieve latent heat profiles from active cloud probing radar on CloudSat which is the second contribution. CloudSat is a part of A-Train satellite constellation alongside Aqua's MODIS. The CloudSat radar has a narrow footprint which is ideal for 2D profiling of clouds, but it does not cover across track distribution of clouds. By the training of a neural network system which uses Aqua-MODIS and CloudSat-radar's collocated footprints, a system is developed which can provide latent heat maps of thunderstorms on a much larger scale. By training MODIS images at various regions, the gravity wave-latent heat relationships can be monitored on a global level on a regular basis. This is the third contribution in which the gravity wave-latent heat models are developed using MODIS for two different regions – South-East Asia (SEA) region and West Africa (WA) region.

1.4. Thesis structure

This thesis describes the model and methods developed for study the latent heat-gravity wave coupling in the moist convective atmosphere. Chapter 2 provides

a thorough introduction to the concept of gravity waves along with the theory involved in this study. It also explains the fundamentals of thunderstorm formation which is necessary to understand the impact it can have in the atmosphere; followed by the relevant literature survey and motivation of this study.

Chapter 3 describes the study that is carried out in South-East Asia region in which the latent heat-gravity wave model is developed using a multi-disciplinary approach of using wavelet transform to identify the gravity waves, ray tracing methods to determine the source locations and cloud-resolving model simulations to obtain the latent heat of thunderstorms. The model is validated using the radiosonde profiles obtained from Singapore weather station.

Chapter 4 describes the method developed to estimate the latent profiles of thunderstorms using CloudSat radar profiles. This method is the extension of existing algorithms for radar-based latent heat estimation [27], [29]. The method involves estimation based on Bayesian Monte Carlo method which uses a database of thunderstorm profiles developed using a cloud-resolving model. This method is validated using the simultaneous cloud-resolving model simulations at the retrieval instance.

Chapter 5 describes the method for extending the CloudSat radar-based latent heat estimation to the Aqua-MODIS using Bayesian neural network training. The chapter discusses the data and method in detail along with the implementation of the method in two different tropical regions – SEA and WA; to develop the gravity wave-latent heat model for those regions which are verified using ground-based results. This is followed by conclusions derived from the thesis with some recommendations for future works based on the limitations identified in this study.

2. Background on gravity waves

The earth's atmosphere constitutes many different physical phenomena which effectively protect the surface from harmful radiation from outer space. Since the atmosphere engulfs the surface, it is being affected by both incoming solar radiation and the surface properties [6]. The atmospheric behaviour such as circulation and cloud activity are dependent on solar activity and surface properties such as land-sea ratio, vegetation etc. [31]. The various processes occurring in the atmosphere is illustrated in Figure 2-1. The figure also shows the temperature variation with altitude, it is seen that the layer tops are defined by a change in gradient sign. All the processes and events such as clouds, thunderstorms, fog, and haze occur in the lower atmospheric layer called the troposphere which extends until the height of about 20 *km* in the tropical region [1]. There are not many strong forcing events occurring in the stratosphere and mesosphere except for noctilucent dust clouds which are basically suspended meteorite dust trapped in that layer [4], [32]. The layer above mesosphere is completely ionised, therefore it is known as ionosphere [33].

The atmosphere of the earth is highly coupled internally among the different layers and externally with solar radiation and magnetic field [6]. The disturbances caused in one layer can affect and distort the state of another layer considerably. The meteorological processes occurring in the lower atmosphere and impacts of the Sun and the geomagnetic field are all responsible for the

constant coupling process in the whole atmosphere [8]. The disturbances usually travel in waveform affecting the different parts of the atmosphere or even the surface [5]. The scientists involved in both meteorology and space weather have developed an interest in this phenomenon since weather modelling will be inaccurate without considering the coupling between various interactions. Hence, it becomes important to model these coupled dynamics accurately for weather forecasting and prediction.

Figure 2-1: Schematic of atmospheric layers according to the temperature, and also atmospheric constituents and processes [34].

The atmosphere is a fluid that is pervaded by waves of various scales in terms of both space and time. The waves are generated by disturbances in the atmosphere due to various physical processes which are called internal gravity waves. Internal gravity waves can propagate large distances (10s – 1000s of *km*), and transfer momentum and energy which makes its understanding very important for weather analysis [3], [5], [25].

The main sources of internal gravity waves identified so far include orography (mountains), convection and jet/front systems [35]. The orographic effects have been extensively studied in the mid-twentieth century. The non-orographic effects (convection and jets-fronts) are less intense locally, but the net effect can be significant in the middle atmospheric circulation [36]. The process of cloud formation and evolution creates perturbations in the atmosphere because of the heating/cooling cycles due to the phase change of the hydrometeors and the magnitude of these perturbations depends on the microphysical and macrophysical structure of the cloud [37]. Thunderstorms or Deep Convective Clouds (DCC) are known to generate a wide spectrum of gravity waves which influence the momentum budget of the lower-middle atmosphere [23].

Cloud creation and growth itself is highly coupled with aerosol and other surface processes. The aerosols that rise up from the surface interact with the clouds and can change the cloud structure completely. The change in cloud depends on background atmospheric state, aerosol chemical composition, size and density. The aerosols when interacting with the water vapour, may attract them and become the nuclei for the cloud droplets which release latent heat due to condensation. If the amount of aerosols is very high, the condensing droplet size will be smaller due to the limited amount of water vapour in the atmosphere. These smaller droplets are pushed upwards due to updrafts from heat release, which freeze at lower temperatures releasing more latent heat; hence creating further updrafts. This aerosol-cloud interaction could trigger large convection leading into a severe thunderstorm [38]–[41]. The microphysics involved in the growth of thunderstorms is quite complicated and, in addition, it requires the understanding of chemical interactions [31], which is beyond the scope of this thesis. This chapter mainly provides the fundamentals of gravity waves along with the brief history of the current understanding of them. The chapter also highlights the existing work related to thunderstorms as the source of gravity wave generation along with basic physical relationships and the gap in the research leading to the current work.

2.1. Gravity waves

The atmospheric layers are highly coupled with each other. The dynamic processes in the troposphere affect the other layers of the atmosphere by momentum transport through gravity waves [22]. In order to understand and model these interactions, it is important to understand and quantify the role of gravity waves and their generation. The various types of processes and waves existing in the atmosphere are shown in Figure 1-1. The stable stratification of the atmosphere allows wave propagation in vertical and wave amplitudes increase with propagation altitudes due to decreasing density [42]. The origins and dynamics of planetary and gravity waves are different. Planetary waves are generated due to the equator to pole gradient of potential vorticity and by contrasts in temperature between the land and ocean. They are very significant in the middle and high latitudes and can lead to dramatic deviations of the flow from its climatological mean state, such as a sudden stratospheric warming when the normal, winter-time polar vortex breaks down due to the propagation and breaking of a planetary wave packet from the troposphere [6]. Gravity waves are very ubiquitous and exist on a range of spatial scales from planetary to a few kilometres. They are generated by various sources such as flow over orography, earthquakes and volcanic disturbances, spontaneous imbalances and by strong convection for example thunderstorms as shown in Figure 1-1.

Gravity waves play an important role both in setting the mean climate of the stratosphere and mesosphere and in generating the predictable tropical oscillations of mean wind speed which can lead, in general, to the enhanced predictability of the climate system [8], [25]. Their nature of increasing wave amplitude during propagation; infers that they dominate and control the motion spectrum at various spatial and temporal scales. While these are propagating possessing large fluxes of energy and momentum, they tend to experience filtering and dissipation by the medium, i.e. similar to electromagnetic wave propagation. The flux divergence due to wave dissipation induces significant forcing in the large-scale circulation, and thermal and constituent structure of the middle atmosphere [42].

Recent studies have described the middle atmospheric gravity waves dynamics on many fronts. Various observational studies using in-situ, ground-based and space-based data have helped greatly to realise the gravity wave scales, amplitudes, fluxes and spectra [10], [35], [51], [52], [43]–[50]. Numerous other studies have also highlighted the vertical propagation and the geographical variability. Theoretical studies have studied the source characteristics and scales, spectral characteristics, evolution and energy transfers, instability dynamics, wave-wave and wave-mean flow interactions [21], [23], [53]–[58]. This section will go through the fundamentals variables and theories that are necessary to understand before discussing the state of art.

2.1.1. Linear theory

The fundamental fluid equations which are universally used for solving fluid flow are derived from conservation of momentum, mass and energy. The gravity waves are usually realised using the linear perturbation which basically assumes that the flow is divided into mean background flow and wave perturbation. This simplification is done in order to reduce the computational effort, and also these are more comprehensible [3]. The simple form of linear perturbation theory can be represented by:

$$u' = u - \bar{u} \quad (2.1)$$

$$v' = v - \bar{v} \quad (2.2)$$

$$T' = T - \bar{T} \quad (2.3)$$

where u , v and T are the velocities in x , y directions and the temperature. The primed symbols represent perturbations and overbar represents mean background state which is applicable for all equations in this thesis. The unforced forms of these equations for velocity which are linearized about a horizontally uniform hydrostatic basic state with background wind $(\bar{u}, \bar{v}, 0)$, pressure $\bar{p}$, and

density $\bar{\rho}$ varying only in vertical direction z are [25]. In this simplification, it is also assumed that the mean vertical velocity is negligible.

$$\frac{Du'}{Dt} + w' \frac{\partial \bar{u}}{\partial z} - f v' + \frac{\partial}{\partial x} \left(\frac{p'}{\bar{\rho}} \right) = 0 \quad (2.4)$$

$$\frac{Dv'}{Dt} + w' \frac{\partial \bar{v}}{\partial z} - f u' + \frac{\partial}{\partial y} \left(\frac{p'}{\bar{\rho}} \right) = 0 \quad (2.5)$$

$$\frac{Dw'}{Dt} + \frac{\partial}{\partial z} \left(\frac{p'}{\bar{\rho}} \right) - \frac{1}{H} \left(\frac{p'}{\bar{\rho}} \right) + g \left(\frac{\rho'}{\bar{\rho}} \right) = 0 \quad (2.6)$$

where D/Dt represents a linearized total derivative; (u, v, w) is the fluid velocity vector. The remaining symbols represent atmospheric fluid variables: p is pressure, ρ is density, $f = 2\Omega \sin \phi$ is the Coriolis parameter (where Ω is the Earth's angular frequency and ϕ is latitude) which represents the Coriolis force, H is the density scale height, and g is acceleration due to gravity. The shear terms in equations (2.4) and (2.5) are neglected by assuming background states to vary very slowly over a wave cycle in the vertical (WKB approximation). The equations (2.4) to (2.6) form the basis for ray tracing equations explained in section 3.1.2. In a moist convective atmosphere, when a parcel of air is lifted to a certain altitude, the gravity forces it to move back down, followed by an uplift due to buoyancy, leading to an oscillation. This frequency of this oscillation is the buoyancy frequency, N as given by equation (2.7).

$$N = \sqrt{\frac{g}{\theta} \frac{\partial \theta}{\partial z}} \quad (2.7)$$

where θ is the potential temperature. The linear theory is only valid for monochromatic gravity wave solutions, which are given in the form:

$$\left(u', v', w', \frac{\theta'}{\theta}, \frac{p'}{\rho}, \frac{\rho'}{\rho}\right) = (\tilde{u}, \tilde{v}, \tilde{w}, \tilde{\theta}, \tilde{p}, \tilde{\rho}) e^{\left[i(kx+ly+mz-\omega t) + \frac{z}{2H}\right]} \quad (2.8)$$

where (k, l, m) represent wavenumber vectors, ω is the Eulerian frequency and the tilde over variables represent the wave amplitude.

2.1.2. Wave scales and propagation

Gravity waves occur in a wide range of spatial and temporal scales and it is often difficult to associate a particular wave scale with any particular type of source. Generally, gravity waves classified into three groups based on intrinsic frequency, $\hat{\omega}$. The intrinsic frequency is related to the Eulerian frequency by:

$$k_H = \frac{\omega - \hat{\omega}}{\bar{U}} \quad (2.9)$$

where $k_H = \sqrt{k^2 + l^2}$, is the total horizontal wavenumber, $\bar{U} = \sqrt{\bar{u}^2 + \bar{v}^2}$ is the mean horizontal velocity. So, based on this the three wave scales groups are: 1) high-frequency: $N \approx \hat{\omega} \gg f$, 2) medium-frequency: $N > \hat{\omega} > f$, and 3) low-frequency waves: $\hat{\omega} \sim f$. For high-frequency waves, the influence of Coriolis force is almost negligible. These waves are observed in middle atmosphere at a range of altitudes based on the sources such as meteor airglow, and mechanical oscillator effect of convection etc., which are seen in airglow images, infrared and microwave image of stratosphere and radar [25]. For medium-frequency waves, group velocity c_g , vertical wavelength λ_v , and $\hat{\omega}$ increase as the wave propagates up-shear, i.e. altitude region where intrinsic phase velocity is increasing. This implies that the rate of vertical velocity growth will be larger than the rate of horizontal velocity growth and vice versa for down-shear. Low-frequency waves are significantly influenced by Coriolis force, which are also called inertia-gravity waves. The velocity perturbations in the direction perpendicular to the propagation are not negligible unlike the other two cases but are inversely

proportional to $\hat{\omega}$ as it approaches f . The ratio of vertical to horizontal group velocity is quite small for these waves compared to the other two cases, which allows them to propagate large horizontal distances from their sources before reaching the middle atmosphere [25], [42], [53], [59].

2.1.3. Sources

The gravity waves sources are of wide varieties which include disturbances in the ocean, atmosphere and surface. The most commonly identified sources include topography, convection and wind shear etc. In mid-latitudes, the geostrophic adjustment also generates waves in the jet-front systems [60], [61]. Few other sources of lower significance are body forcing accompanied by localised wave dissipation, wave-wave interaction, auroral heating and eclipse cooling (significant at higher altitudes) [42].

Topography: this is the most studied and understood source of gravity wave generation is the topography or more commonly known as the mountain waves. The waves are generated due to the fact that the topography acts as an obstacle to the basic atmospheric flow, creating perturbations and turbulence [62]. Over the last four decades, extensive research has been done on this subject matter through numerous theoretical, numerical and observational studies. Various 2D and 3D idealised numerical experiments have been conducted over the years using non-linear wave dynamics and hydraulic theory which helped in understanding the response to topography, yielding estimates of amplitudes, scales and fluxes [58], [63], [64]. The vertically propagating waves generated from large-scale topography have been associated to be responsible for triggering polar stratospheric clouds and ozone depletion in the polar region [25]. The momentum fluxes typically vary in the range of 0.1 to a few m^2s^{-2} in the upper troposphere and lower stratosphere based on the in-situ aircraft and radar observations of pressure distributions [63].

Convection: convection causes clouds to grow vertically which are known to generate gravity waves since half a century ago [11]. The vertical motions inside the clouds, driven by latent heat, or by the airflow over clouds in a sheared

environment can excite gravity waves [42]. Deep Convective Clouds (DCC) or thunderstorms are especially major sources of gravity waves in the tropical region. Waves generated by convection are not associated with any particular range of frequencies similar to topographic waves [25]. Low-frequency waves are observed in the middle atmosphere at large horizontal distances from the source which makes correlation studies very difficult [11], [25]. In the past literature, in the tropical region, the inertia-gravity waves are linked to convection as the source when origins are traced away from topography and regions of baroclinic instability [65], [66]. Various model-based studies have enhanced the understanding of convectively generated waves and linked to the thermal forcing associated with the latent heat release that interacts with the stable layers above the cloud [25], [26], [67]. Section 2.2 will elaborate the mechanism in more detail.

Shear: this source mechanism is by far the least characterised and understood among all the others. It is a challenge to distinguish the gravity waves generated from the shear layers, from the rapidly growing Kelvin-Helmholtz instability. It is extremely challenging to quantify and study this effect using the observations; nevertheless, there are few observational evidences of linear modes excited by the shear instabilities [25], [43], [68], [69].

Spontaneous adjustment: the basic concept behind this effect is that while the atmosphere is trying to stabilise itself after a disturbance, the gravity waves are emitted in order to distribute the momentum, energy and potential vorticity [25]. Most of the evidence comes from theoretical and numerical experiments of the adjustment processes. Fronts and baroclinic are the two main sources of the imbalance created in the atmosphere mainly in mid-latitude regions. Typical vertical scales range within few *kms* with horizontal scales can be 10-100 times larger [20], [35], [60], [61], [70].

All the sources discussed above significantly affect the upper and middle atmospheric circulation through gravity waves excitation. This thesis focuses on one particular type of source which is predominant in the tropical region – thunderstorms. The rest of the chapter describes the thunderstorm formation,

gravity waves excitation mechanism and the research gap found post literature survey.

2.2. Thunderstorm as a source

Thunderstorms or DCCs occur frequently in the tropical region in the Hadley cells [1]. They are one of most impactful phenomena affecting the general circulation on a diurnal basis. Violent winds are driven from them causing major hazards across the globe in the tropical region [1]. It is well documented that thunderstorms have a significant impact on the surface through rain and winds, but in the last few decades, studies have shown the impact they have in the atmosphere which affects the large-scale circulation causing major changes in the middle and upper atmosphere [14], [18], [19], [55], [67], [71]. It is important to understand and quantify every type of effect they have on the environment as in the recent years, various literature have highlighted the coupling of atmosphere, land and ocean. The changes in one layer or region of the earth are not limited to adversities in that area but affect the global climate slowly over the time [37], [38]. This subsection focuses on the basic theory of thunderstorm development and highlights some major impact it has in the atmosphere which is relevant to this study.

2.2.1. Thunderstorm development

Thunderstorms are convective clouds which grow vertically up to a large extent, often with tops near the tropopause and bases near the boundary layer, hence called deep convective clouds. The most basic form of a thunderstorm has a vertical stem of a diameter which is nearly equal to its vertical thickness or sometimes much greater. Typically, the lifetime of a thunderstorm has three stages: 1) towering cumulus; 2) mature, and 3) dissipation, as shown in Figure 2-2 [34].

The towering cumulus stage consists only of updrafts due to latent heat release which can grow very rapidly. At this stage, the cloud doesn't have any

downdrafts, anvils or precipitation. The warm humid boundary layer air is served as fuel to the storm [1], [72]. In this stage, the water vapour condenses and water droplets freeze releasing an enormous amount of latent heat due to phase change which continues during the mature stage [72].

Figure 2-2: Phases of thunderstorm cell evolution. Light grey shading represents clouds; medium grey is the pre-storm boundary layer of warm humid air that serves as fuel for the storm. Dark shadowing is colder outflowing air. Diagonal white lines represent precipitation. Arrows show air motion. *Cu con* = cumulus convection [1].

In the mature stage, the thunderstorm has an anvil and cloud top is near the tropopause. It consists of both updrafts and downdrafts along with large and heavy drops of precipitation. This is the stage during which the storm is most dangerous with lightning, turbulence and winds. The downdrafts drive cooler air towards the surface causing the surrounding air to be pushed outwards [1], [31], [72].

In the dissipating stage, the storm consists of downdrafts, precipitation and large glaciated anvil. The hazardous aspects of the storm diminish with no updrafts. Two factors affect the lifetime of the storm: 1) the amount of fuel available for the storm to keep going, and 2) the colder outflow from downdrafts restrict access to the surrounding warm boundary layer air. The afterwards stage shown in Figure 2-2d demonstrates the remains of a burnt-out storm which is a former anvil consisting of ice crystals which fall and evaporate before reaching the surface. The cold outflow, meanwhile can encounter a fresh warm boundary layer air, which it lifts along its leading edge to trigger a new cumulus convection [1]. This is basically called storm propagation.

There are several classes of thunderstorms based on cells and organisation. Single cell thunderstorms (which are the main focus in this study) can form in the middle of warm humid airmass usually in the late afternoon. These storms are triggered by rising thermals which are buoyant enough to break through the boundary layer inversion. The typical structure looks like the storm demonstrated in Figure 2-2. When a weather radar spots multiple cells by their cores and winds, while the human eyes can only see one cloud; such kind of system is called multi-cell thunderstorms. Multi-cell storms also occur in organised form in which the anvils from the thunderstorms are merged into a single large stratiform cloud, while the cores are separated often formed in a narrow line causing heavy precipitation. This type of cell is called a Mesoscale Convective System (MCS) shown in Figure 2-3. These systems can be triggered by mountains or synoptic systems such as gust fronts or cold fronts [1].

Figure 2-3: The vertical cross-section of a mesoscale convective system [1].

In summary, four environmental conditions are required to form a severe thunderstorm:

- High humidity in the boundary layer
- Non-local conditional instability, i.e. rising airmass is saturated
- Strong wind shear
- A trigger to lift the airmass above boundary layer inversion.

It is beyond the scope of this report to elaborate the thunderstorm formation theory or cell organisations in detail. The information given here should be sufficient for the reader to follow the studies presented in the later chapters. The following section will discuss the effects of the thunderstorm in the atmosphere relevant to this study.

2.2.2. Waves from the thunderstorms

The effects of thunderstorms on the surface are in the form of rain, floods and heavy winds which are more understood and monitored. As mentioned in the previous section, thunderstorms have a major impact in the atmosphere; driving winds which contribute to the circulation, and generate gravity waves which distribute energy and momentum in the upper atmosphere [1], [37].

One of the earlier observational studies which highlighted the gravity wave spectrum generated by thunderstorm was performed in 1974. It was suggested that a necessary condition for wave generation is the rate of kinetic energy generated from the convective updrafts to the stable layer should be large compared to the transfer rates in the surrounding atmosphere [45]. The strong heating in the boundary layer leading to shallow convection in the lower tropospheric stable layer acts as an obstacle to the flow; hence creating drag. These act like small hills which excite gravity waves [55]. This mainly occurs during the growing stage of convection.

Convection in the atmosphere is associated with the time-varying latent heat release (mainly due to condensation and freezing of water content). This thermal

forcing interacts with the overlying stable layers causing perturbations. There are three fundamental mechanisms suggested through which the convection is able to generate gravity waves: 1) thermal forcing, 2) the obstacle effect, and, 3) mechanical oscillator [25]. Most of the time, a single convective event can generate waves which are generated via a combination of all the three mechanisms. The waves generated by convective updrafts overshooting the equilibrium level and oscillating until reaching back the equilibrium state; have frequencies close to the Brunt-Väisälä frequency, periods of which can be as low as just a few minutes [11], [19], [73]. The horizontal wavelengths of this type of waves are proportional to the width of the convective updrafts. The convective clouds which span the entire troposphere such as thunderstorms contain updrafts strong enough to overshoot the tropopause inversion, which can increase the amplitude and frequency of the waves generated [11]. However, these waves due to convective overshoot are generated in the lower stratosphere; can be argued to be generated from the displacement of the stable air rather than the latent heat in the thunderstorm. This vertical displacement of air that oscillates to generate waves is called the mechanical oscillator (the third mechanism mentioned earlier). These are usually high-frequency waves with period comparable to the updraft oscillation [47]. There are low-frequency waves also generated by deep convective events mainly by thermal forcing [25], [36], [48], [73]. The waves propagating in the troposphere generated by thermal forcing usually have stabilising or destabilising effects in the environment, which can suppress or induce further convection; referred to as storm propagation. This type of coupled behaviour is frequently encountered in the tropics where the MCS drives a chain of convective systems [36]. Waves that propagate to stratosphere have their characteristics enhanced as the stratosphere is much more stable compared to the troposphere. The high-frequency waves generated by the convective overshoot at the tropopause propagate almost vertically with high group velocity, directly above the convective systems [11] as shown in Figure 2-4. The figure shows the simulated pattern of waves generated by a convective system in the mid-latitudes where the tropopause occurs at 12 *km* approximately. It is seen that the wavefronts are almost circular with the convective system as the centre. These waves usually have horizontal and vertical wavelengths of the same scale.

Convective clouds also generate waves with much lower frequencies that are not associated with the convective overshoot. These low-frequency waves also reach the stratosphere, are mainly generated by large MCS with horizontal wavelengths much larger than the individual cloud cell. As the clouds become organised, the spectrum of waves contains longer wavelengths and lower frequency waves. Since the atmosphere is a continuous fluid; the spectrum may contain high and low-frequency waves from both the MCS and the individual clouds.

Figure 2-4: Model simulation results showing the wave patterns generated by an isolated thunderstorm in the mid-latitudes. The portion of the cloud penetrating 10 km is shown by the lower surface. Isentropic surfaces at 15 and 20 km with waves are shown in the top two layers [11].

A radiosonde analysis of the wind vectors in the stratosphere can encounter a spectrum of waves with wavelengths from 10 to 1000 km, with many convective

systems contributing towards it. Some of the waves are influenced by the Coriolis force, which are called inertia-gravity waves as mentioned in section 2.1.2. The intrinsic period of these waves can be several hours or even days. The phase lines of these waves are oriented nearer to the horizontal leading to a large difference in the horizontal and vertical propagation speeds. These waves can travel large distances away from their source (even $> 1000 \text{ km}$) with high horizontal speeds while possessing very low vertical speeds; taking hours or days to reach stratospheric altitudes [11], [25], [36], [42].

This subsection provides an insight into the convectively generated gravity waves. Various studies have shown the significant impact thunderstorms can have in the troposphere and in the middle atmosphere which can change the constituency of layer significantly [19], [23], [73]. Therefore, it is important to quantify and model these interactions which have indirect effects that control the human life.

2.3. Research synthesis

Previous sections in this chapter discuss various types of processes that generate gravity waves and the significance of convectively generated gravity waves. This thesis's main focus is on the convectively generated gravity waves as deep convection or thunderstorms are frequent in the tropical region. This section discusses the existing work in the field of the convectively generated gravity waves briefly, followed by the state of art and identified gaps for research.

The convectively generated gravity waves have been studied for more than four decades, and numerous methods have been developed to understand the generation and propagation of these waves. Various numerical studies have provided the platform needed to analyse the physics in an isolated condition [14]–[16], [18]–[21], [55]. These studies have helped to understand and formulate the mechanisms of generation mentioned in 2.2.2. Some of the studies have delivered quantified relationships between heating depths and gravity wave phase velocities and wavelengths [16], [21], [26], but these are mainly based on

simulated cases in the sub-tropical or extra-tropical environment where various assumptions are involved. However, the real atmosphere is far from being isolated and the processes are complexly coupled with each other. Therefore, observational studies play an important role in not only understanding the behaviour but also to validate the model results. Gravity waves can be characterised using various instruments such as radar, satellite images, radiosonde, and *in-situ* rocket soundings. There are numerous observational studies which have characterised the gravity waves and associated them with convective events. Yet, these observational studies have not been able to quantify or model the relationship between the gravity waves and the thunderstorms. The numerical studies have highlighted the perturbation caused due to latent heat release in the thunderstorms, but the observational studies have not been able to quantify or validate this. This is mainly because of the limited instrumental capabilities and limitations.

The current study aims to extend the current understanding of gravity waves generated by thunderstorms in the tropical region. To quantify the relationship, multi-disciplinary methods will be utilised and applied which involves radiosonde analysis, cloud-resolving model simulations and satellite methods. The quantification here is based on extensive statistical analysis of various events, unlike the previous studies which involved few case-studies only. The study is conducted for two tropical regions with frequent thunderstorm activities and abundant data for study – South-East Asia (SEA) and West Africa (WA). The following chapters will explain and discuss the methods used and the results obtained for the two regions and conclude with a discussion on future works.

3. Inertia- gravity waves generated by thunderstorms in South-East Asia tropical region

The previous chapter discussed the research gap in the field of convectively generated gravity waves and need for a quantified model of the relationship between the thunderstorms and the gravity waves. In this chapter, a simple model describing the spectral properties of gravity waves excited from the thunderstorms is developed using long-term statistics. The aim is to deliver a simple model between the latent heat generated in the thunderstorms and the corresponding wave properties such as phase velocities and amplitude.

3.1. Data and methods

The study in this chapter is aimed to model the gravity waves in the South-East Asia region. Frequent occurrences of thunderstorms in SEA region always make the weather prediction challenging. The waves and circulation excited from these events alter local and global flows leading to unexpected weather patterns. Since the trend of smart cities are growing in SEA region, it is essential to have accurate understanding of the weather couplings. Ten years (2007-2016) of radiosonde

data at Singapore (103.8° , 1.38°) are studied for inertia-gravity waves in the lower stratosphere. The individual profiles are interpolated on to an equal spacing vertical grid and averaged at each altitude level. Radiosonde based gravity wave detection is very common; and is one of the well-established methods that have been developed for gravity wave detection purposes [74]–[76]. The radiosonde data are available every 12 *hours*, and therefore only the waves with periods longer than that interval can be detected accurately. The method of detection is based on the linear perturbation theory (explained in section 2.1.1). The perturbations are identified over a defined mean background state, which in this case is the mean of the ten years of data.

Figure 3-1: Schematic diagram showing the stages of the study.

The radiosonde profiles are available consistently only until 25 *km* altitude, therefore the detection of waves is limited to the lower stratosphere from 19-25

km. The perturbation profiles are averaged between 19 and 25 *km* in order to detect the waves in the lower stratosphere. The tropopause altitude for Singapore is approximately 18 *km*. It is more sensible to analyse the lower stratosphere for waves instead of upper troposphere because the stratosphere is much more stable, and the waves are readily seen in the signal with minimal filtering [11]. There are several stages in the study with cross-disciplinary approaches which are required to develop this model. Once the waves are identified in stratosphere, the sources of the waves are tracked using the ray tracing method. The traced ray paths are analysed further for convective sources along the path, and the cloud-resolving model simulations are performed at the identified convective locations. Figure 3-1 shows the stages of the study in a schematic diagram.

3.1.1. Wavelet analysis

Wavelet transform is used very commonly in the atmospheric studies in order to identify the spectrum of waves present in any signal using the time-frequency analysis [46], [48], [50], [77]. By decomposing a time series into time-frequency space, one is able to determine both the dominant modes of variability and how those modes vary in time. The conventional method of windowed Fourier transform suffers with inaccuracies of time-frequency localisation as it imposes a scale into the analysis. This happens due to the aliasing of high and low frequency components that are outside the window specifications [78]. In addition, several different windows must be analysed to determine the most appropriate choice. For analyses where a predetermined scaling may not be appropriate because of a wide range of dominant frequencies, a method of time-frequency localization that is scale independent, such as wavelet analysis, should be employed. Wavelets such as complex Morlet and Mexican Hat are used especially in gravity wave studies. Complex Morlet wavelet is a complex non-orthogonal wavelet which has been shown to work very well for the band of inertia-gravity waves [48], [50], [79]; and since this study is also focused on the inertia-gravity waves, complex Morlet wavelet is used for the wavelet transform.

Figure 3-2: (top panel) Time series of the kinetic energy density averaged from 19 to 25 km for the ten-year duration. (bottom panel) Wavelet power spectrum of kinetic energy density averaged from 19 to 25 km for the ten-year duration. Cone of influence highlighted in thick black line and the shaded region.

It is common to analyse the time-frequency spectrum of the horizontal kinetic energy density of the perturbation in order to identify the significant waves [48], [50], [77]. The wavelet analysis is performed on the averaged time series signals of u'^2 and v'^2 (between 19 and 25 km) and summed the wavelet transforms of them to obtain the time-frequency spectrum of the horizontal kinetic energy density. Previous studies have shown that this approach of summation is better than performing the wavelet transform directly on the energy density signal [46], [48]. The top panel in Figure 3-2 show the time series of the kinetic energy density signal and the bottom panel shows the corresponding wave spectrum of it. From the times series itself, it can be seen that there is a wide range of oscillations with periods ranging from 12 hours to few years which is also reflected in the wave

spectrum in terms of very low frequencies. From the wavelet analysis, the waves with significant power are identified which are considered to have a significant impact in the atmosphere but it is difficult to deduce that with the spectral power alone without any filtering. Therefore, from the spectrum shown in Figure 3-2, the waves with significant power are identified using the 95% significance test [79], based on the red noise spectrum [79]–[81]. This method is very common for wavelet analysis in atmospheric sciences. The one drawback of the wavelet transform is the loss of spectral accuracy near the ends of the signal which is defined using the cone of influence highlighted in Figure 3-2 by thick black line. The spectral power outside the cone is usually inaccurate and not considered during the analysis (highlighted by grey shaded region). Based on the red noise spectrum, it is likely that more high-frequency waves will be significant as the red noise spectral power at lower frequencies is very high. Here the identified frequencies are Eulerian or ground-based frequencies, ω .

Once the dominant Eulerian frequencies are identified, the dispersion relations are used to calculate the wave characteristic variables such as intrinsic frequency using the equation (2.9) and wave numbers. Firstly, the Coriolis parameter, f , is calculated for the Singaporean latitude as discussed in section 2.1.1. Based on earth's inertial frequency and the latitude of Singapore radiosonde, f is calculated to be 3.315×10^{-6} Hz. The intrinsic frequency can be calculated using the ratio between the kinetic and potential energy densities. This method has been demonstrated by [49] using the dispersion relations. The expression for this ratio is as shown in equation (3.1):

$$\frac{\overline{\overline{E_k}}}{\overline{\overline{E_p}}} = \frac{\left(1 + \left(\frac{f}{\omega}\right)^2\right)}{\left(1 - \left(\frac{f}{\omega}\right)^2\right)} \quad (3.1)$$

The double overbar over E_k and E_p represent averaging between the altitude range (19-25 km in this case). Using the dispersion relations shown in equations

(2.6) and (3.2), the horizontal and vertical wavelengths and wavenumbers were calculated [25], [46]:

$$m^2 = \frac{(k^2 + l^2)(N^2 + \hat{\omega}^2)}{(\hat{\omega}^2 - f^2)} - \frac{1}{4H^2} \quad (3.2)$$

It is important to calculate the horizontal components of k_H for the ray tracing calculations; which is obtained by calculating the direction of phase propagation using equation (3.3) [46], [48], [50], [52].

$$\Phi = \tan^{-1} \left(\frac{\overline{u'T'_{+90}}}{\overline{v'T'_{+90}}} \right) \quad (3.3)$$

where T'_{+90} is the Hilbert transform of the temperature perturbation signal. Now, all the variables required for tracing calculations are obtained using the dispersion relations. However, before proceeding further, it is important to verify that the energy propagation is in the upward direction at the stratospheric altitude (which implies that the waves are propagating from a tropospheric source). Stokes parameters can be used for such analysis which defines the polarisation state of the propagating wave [77], [82]. The polarisation ellipse gives the information regarding the energy propagation direction based on the ratio between the anti-clockwise and clockwise rotation of horizontal wind perturbations [82]. Stokes parameters are shown in equations (3.4) – (3.7):

$$I(m) = C \left[\overline{\mathfrak{U}_R^2(m)} + \overline{\mathfrak{U}_I^2(m)} + \overline{\mathfrak{V}_R^2(m)} + \overline{\mathfrak{V}_I^2(m)} \right] \quad (3.4)$$

$$D(m) = C \left[\overline{\mathfrak{U}_R^2(m)} + \overline{\mathfrak{U}_I^2(m)} - \overline{\mathfrak{V}_R^2(m)} - \overline{\mathfrak{V}_I^2(m)} \right] \quad (3.5)$$

$$P(m) = 2C \left[\overline{\mathfrak{U}_R^2(m)} \overline{\mathfrak{B}_R^2(m)} + \overline{\mathfrak{U}_I^2(m)} \overline{\mathfrak{B}_I^2(m)} \right] \quad (3.6)$$

$$Q(m) = 2C \left[\overline{\mathfrak{U}_R^2(m)} \overline{\mathfrak{B}_I^2(m)} + \overline{\mathfrak{U}_I^2(m)} \overline{\mathfrak{B}_R^2(m)} \right] \quad (3.7)$$

where C is a constant, and subscripts R and I denote the real and imaginary part of the Fourier transform, and $\mathfrak{U}$ and $\mathfrak{B}$ are the Fourier transforms of u' and v' . The overbar denotes the average between the altitude range (19-25 km in this case). The clockwise (CW) and anticlockwise (ACW) rotations of horizontal wind perturbations with height can be used to determine the propagation direction [82].

$$\begin{aligned} CW &= 0.5[I(m) - Q(m)] \\ ACW &= 0.5[I(m) + Q(m)] \end{aligned} \quad (3.8)$$

If the ratio of ACW to CW is greater than 1, it can be concluded that the direction of propagation is upwards and these can be calculated using equation (3.8) [48]. When the ACW rotation is higher than the CW rotation, it is found that the propagation is in the upward direction [59], [77], [82], [83]. The ratio varies from 0.3 to 3.8 over the altitude range of 16 to 25 km . At stratospheric altitudes, the value of the ratio is mainly greater than 2; indicating upward energy propagation.

3.1.2. Thunderstorm source tracing

The gravity waves variables identified at the stratospheric altitudes are now analysed for their sources using ray tracing method. This is performed using the Gravity wave Regional Or Global Ray Tracer model (GROGRAT) [84]. This model basically applies the electromagnetic wave ray tracing method to the gravity waves by modifying the dispersion relations suitable to the application [84], [85]. It solves a nonhydrostatic dispersion relation appropriate for small-amplitude gravity waves on a slowly varying basic flow $(\bar{u}, \bar{v}, \theta)$ in a rotating,

stratified and compressible atmosphere. By solving the differential equations representing the wavenumbers and positions of the wave, it is possible to trace the propagation path of the wave. GROGRAT model is not valid for high-frequency waves for which a value of $\hat{\omega}$ has been defined, called critical intrinsic frequency [84], and all the waves analysed here are lower than the critical intrinsic frequency of 2.541×10^{-3} Hz. Various global studies have used this method to retrieve the source location of the gravity waves [20], [48], [61], [77]. The background atmosphere is required to be varying slowly in order to facilitate consistent wave propagation. The six-hour interval European Center for Medium-range Weather Forecasts (ECMWF) ERA-Interim ensemble model with the grid resolution of 0.3° with 37 pressure levels is used for background atmosphere. This method requires the gradient of background variables to change smoothly in order for the Wentzel-Kramers-Brillouin (WKB) approximation to remain valid which is shown in equation (3.9). Therefore, as suggested in [84], a spherical harmonic smoothing has been applied in the horizontal and cubic spline smoothing in the vertical to the variables in order to obtain the mean background values.

$$\delta \approx \left| \frac{1}{c_{g_z} m^2} \frac{dm}{dt} \right| \quad (3.9)$$

where c_{g_z} is the vertical group velocity. The GROGRAT model requires the initial values of k, l, m, x, y , and z , (obtained from dispersion calculations) along with the initial values of the gradients of the background variables (obtained from ECMWF model) at 25 km altitude which is the ray launching altitude. The GROGRAT model is based on dispersive waves in which the frequency of the wave is a function of the wavenumber and the position. The method basically solves six differential equations in order to obtain the wave path in a reverse tracing method. The ray tracing equations are shown in equation (3.10) to (3.15) [84].

$$\frac{dx}{dt} = U + \frac{k(N^2 - \hat{\omega}^2)}{\hat{\omega}[k^2 + l^2 + m^2 + \alpha^2]} \quad (3.10)$$

$$\frac{dy}{dt} = V + \frac{l(N^2 - \hat{\omega}^2)}{\hat{\omega}[k^2 + l^2 + m^2 + \alpha^2]} \quad (3.11)$$

$$\frac{dz}{dt} = -\frac{m(\hat{\omega}^2 - f^2)}{\hat{\omega}[k^2 + l^2 + m^2 + \alpha^2]} \quad (3.12)$$

$$\frac{dk}{dt} = -k \frac{dU}{dx} - l \frac{dV}{dx} - \frac{\left[\left(\frac{dN}{dx} \right)^2 (k^2 + l^2) - \left(\frac{d\alpha}{dx} \right)^2 (\hat{\omega}^2 - f^2) \right]}{2\hat{\omega}[k^2 + l^2 + m^2 + \alpha^2]} \quad (3.13)$$

$$\begin{aligned} \frac{dl}{dt} = & -k \frac{dU}{dy} - l \frac{dV}{dy} - \frac{\left[\left(\frac{dN}{dy} \right)^2 (k^2 + l^2) - \left(\frac{d\alpha}{dy} \right)^2 (\hat{\omega}^2 - f^2) \right]}{2\hat{\omega}[k^2 + l^2 + m^2 + \alpha^2]} \\ & - \frac{\left[f \frac{df}{dy} (m^2 + \alpha^2) \right]}{\hat{\omega}[k^2 + l^2 + m^2 + \alpha^2]} \end{aligned} \quad (3.14)$$

$$\frac{dm}{dt} = -k \frac{dU}{dz} - l \frac{dV}{dz} - \frac{\left[\left(\frac{dN}{dz} \right)^2 (k^2 + l^2) - \left(\frac{d\alpha}{dz} \right)^2 (\hat{\omega}^2 - f^2) \right]}{2\hat{\omega}[k^2 + l^2 + m^2 + \alpha^2]} \quad (3.15)$$

where U, V are the background velocities in x, y directions, and $\alpha=1/2H$. The Eulerian frequency ω is assumed to be constant throughout the ray path, and this fact is used as a validity check of the integrations [84]. The wave amplitude is tracked by solving the ordinary differential form of the simplified wave action equation shown in equation (3.16) [84].

$$\frac{dF}{dt} = -\frac{2}{\tau}F - Fc_{g_z} \quad (3.16)$$

where F is the vertical flux of wave action defined by, $F = c_{g_z} A$. Here, A is the wave action density which is defined as the ratio of the wave energy density and the intrinsic frequency, $A = E / \hat{\omega}$, where E is a sum of E_k and E_p , and τ is the damping timescale given by equation (3.17) [84].

$$\tau^{-1} = \frac{\tau_r^{-1} \left(\frac{1 - \frac{f^2}{\hat{\omega}^2}}{1 - \frac{\hat{\omega}^2}{N^2}} \right) + \tau_D^{-1} \left(1 + \frac{f^2}{\hat{\omega}^2} + \frac{1 - \frac{f^2}{\hat{\omega}^2}}{\frac{N^2}{\hat{\omega}^2} - 1} + Pr^{-1} \left(\frac{1 - \frac{f^2}{\hat{\omega}^2}}{1 - \frac{\hat{\omega}^2}{N^2}} \right) \right)}{\left[1 + \frac{f^2}{\hat{\omega}^2} + \frac{N^2 + \hat{\omega}^2}{N^2 - \hat{\omega}^2} \left(1 - \frac{f^2}{\hat{\omega}^2} \right) \right]} \quad (3.17)$$

where τ_r is the radiative damping timescale, τ_D is the turbulent damping timescale which is given by $\tau_D^{-1} = \xi \left(k^2 + l^2 + m^2 + \alpha^2 \right)$, where ξ is the turbulent diffusion coefficient which varies linearly in vertical, and Pr is the Prandtl number. The complete explanation on obtaining these variables can be found in [84]. For waves with $m^2 \gg \alpha^2$ (as in this case), the variable coefficient of viscosity terms are negligible; therefore it is ignored in this study. The effects of horizontal deformation of the wave amplitude are also neglected which according to literature have higher order effects [84].

The initial altitude is given at 25 km and integrated values are calculated at each time step of 60 s. A fourth-order Runge-Kutta method was used for solving the differential equations (3.10) to (3.16). The WKB approximation shown in equation (3.9) is used to determine the end of the ray path. The integration is terminated when $\delta \geq 1$, which happens as $m \rightarrow 0$. Waves are reflected in such situations, and the reflected waves are not of interest in this case. It is taken in regard that some gravity waves with large δ , may propagate through the

atmosphere while violating the WKB approximation, but it is also difficult to analyse and identify the end of ray path in such cases because the ray-tracer method is not valid for cases in which the rate of change of m is too high. Therefore, the WKB approximation is used as the criterion to identify ray end.

Using the 10 years of data, 25000 ray tracing calculations were performed with 2500 rays from each year. Since this study is focused on finding the thunderstorm sources, only the rays which penetrated into the troposphere were analysed. The waves from possible tropopause overshooting were ignored as there is no established method to quantify the overshooting effect. Since the tropopause altitude varies with season and latitude, the average altitude of tropopause was calculated based on the latitude in the tropical region where the seasonal effect is negligible [86]. Out of the 25000 rays, only 4983 penetrated into the troposphere. These 4983 rays are analysed for convection along the ray paths with the help of Outgoing Longwave Radiation (OLR) data. The daily average interpolated data of OLR is available from the NOAA satellite images which can be used as the primary indicator of deep convective regions. Previous literature has suggested that the areas with deep convection have OLR values less than 220 W/m^2 , in general [77], [87], [88]. It is to be noted that this is only a primary indicator as there could be other processes also leading to low OLR such as strong aerosol forcing, but currently there are no other established methods to detect thunderstorms using passive image spectrometry or radiometers. The strong perturbations in the ray tracing fields are compared with the OLR maps to identify the possible areas of deep convection. Based on the OLR analysis, approximately 2000 cases of rays are identified which might have gravity wave perturbations due to the thunderstorms. Since the latent heat of the thunderstorms are realised using CRMs, considering the computation costs of each simulation; 100 instances of thunderstorms are chosen for deriving the model here. These 100 cases are selected randomly from each year with 10 instances representing each year which are shown in Figure 3-3.

Figure 3-3: 100 cases of ray-traced paths are shown which are identified as gravity waves generated from thunderstorms. The ray paths are shown in red, the black lines represent geographic map and the contours indicate OLR data averaged from 2007 to 2016.

Figure 3-3 shows the ray paths of 100 cases of thunderstorm-generated gravity waves. The ray paths (shown by red lines) mostly cross through the region shaded in green, where the occurrence of convective events is very high. The regions with low OLR are also investigated for possible mountains to avoid false alarms. The waves generated by balance adjustments are considered to be highly unlikely in these latitudes as the Coriolis force is not the dominant force in the tropical region [1], [35], [60]. Therefore, those ray paths having low OLR and not coinciding with other possible sources are concluded to convective sources. Once, these areas are identified, the cloud-resolving model (CRM) simulations are performed at those locations to determine the latent heat generated in the thunderstorms. The CRM description and setup is explained in the next section.

3.1.3. Cloud-resolving model simulations

For obtaining latent heat of thunderstorms, CRM simulations are performed at the identified locations. Cloud-resolving model simulations are a well-established method of analysing the cloud growth and other microphysical processes involved. In the last few decades, CRMs have evolved and are able to

simulate the microphysical processes reasonably well, even processes such as aerosol-cloud interactions are represented well [31], [89]. The Regional Atmospheric Modelling System (RAMS) is the CRM used in this study which is developed by Colorado State University. It employs a full nonhydrostatic compressible system representing the dynamics and thermodynamics of the atmosphere with conservation equations for moisture budgets [90]. These equations are supplemented with parameterisations for turbulent diffusion, solar and terrestrial radiation, moist processes including the formation and interaction of clouds and precipitating liquid and ice hydrometeors, kinematic effects of terrain, cumulus convection, and sensible and latent heat exchange between the atmosphere and the surface, consisting of multiple soil layers, vegetation, snow cover, canopy air and surface water [90]. RAMS is a regional model, designed for mesoscale or higher resolution scale grids. A two-way interactive grid nesting allows fine mesh grids to resolve small-scale atmospheric systems such as thunderstorms while simultaneously modelling the large-scale environment of the systems on a coarser grid [90]. Various studies have implemented RAMS for simulating the real events of moist convective events such as thunderstorms, hurricanes and cyclones [91]–[94]. These studies have shown that the RAMS can simulate the events realistically and quite reliable for simulating thunderstorms.

In this study, four two-way nested C-grids are used with a horizontal spacing of 25 km, 6.25 km, 1.5 km and 375 m covering an area of approximately, 62,500 km² in the outer grid, and 1,500 km² in the inner-most grid. The model consists of 41 vertical levels with variable spacing; the lower altitudes are assigned with higher resolution grid, and the model top extends until 20 km. A basic radiative condition was applied to the normal velocity components at the lateral boundaries of grids [95]. The surface processes were parameterised using the Land Ecosystem Atmosphere Feedback 3 (LEAF-3). Convection was explicitly resolved on the inner grids while parameterised on the outer grid using a modified Kuo cumulus scheme [96]. The mixing ratios and number concentration of the various hydrometeors were predicted through two-moment bulk microphysical scheme [97], with all the available water species activated (pristine ice, cloud water, snow, aggregates, graupel, hail and rain). Since Singapore

region does not have the available instrumental data of vertical aerosol profile, the default aerosol initiation and vertical profile inbuilt in RAMS was used; the equation for which is given in Table 3-1. The complete model configuration is given in Table 3-1.

Table 3-1: RAMS model configuration.

Model Aspect	Configuration
Grid	Arakawa C-grid Four-nested grids Horizontal grid: Grid 1: $\Delta x = \Delta y = 25 \text{ km}$; 25 nodes each Grid 2: $\Delta x = \Delta y = 6.5 \text{ km}$; 42 nodes each Grid 3: $\Delta x = \Delta y = 1.5 \text{ km}$; 82 nodes each Grid 4: $\Delta x = \Delta y = 375 \text{ m}$; 102 nodes each Vertical grid: Δz variable; 41 nodes Model top: 20 km
Initiation	ECMWF ERA-Interim 0.3° grid resolution data
Microphysics scheme	Two-moment bulk microphysics [97] Water species: vapour, cloud, rain, ice, snow, graupel, hail, aggregates and snow
Convection initiation	Kuo cumulus scheme for grid 1; Explicit convection on grid 2 and 3
Boundary conditions	Radiative lateral boundary [95]
Turbulence scheme	Deformation-based parameterisation [98], [99]
Aerosol initiation	Exponential profile with maximum concentration $C = C_{max} e^{\frac{z}{7000}}$, z is in m .
Radiation scheme	Harrington [100]
Surface scheme	Land Ecosystem Atmosphere Feedback 3
Time-step scheme	Forward time-differencing for thermodynamic Leapfrog differencing for velocity and pressure
Time step	10s in outer grid, increasing 5 times for each inner grid

The performance of RAMS is validated near Singapore region using radiosonde profiles for non-prognostic variables, and by comparing the reflectivity profiles with CloudSat reflectivity profiles (for prognostic variables) which is discussed in the appendix. The method to convert the RAMS prognostic variables to reflectivity is shown in section 4.3.3.

3.2. Results

The aim of this study is to develop the relationship between the latent heat generated in a thunderstorm and the gravity wave perturbations generated from it. As mentioned in the previous sections, the gravity wave perturbations are monitored along ray traced path, and thunderstorms are simulated at the identified locations of low OLR. When considering the waves generated by thermal forcing, previous studies have shown that the wave characteristics depend on the heating depth of the individual storm cores or cells [25], [42]. Therefore the integrated latent heat of storm cell collocated with the perturbations both in space and time are used for deriving relationships. RAMS basically provides latent heat in terms of instantaneous change in temperature or potential temperature (in units of Kelvin) at the location due to latent heat release. The relationships are developed based on the analysing correlations of wave properties with – a) absolute instantaneous latent heat release in the thunderstorm (K) and b) time derivative of latent heat release (K/s).

3.2.1. Velocity perturbations

The ray tracing calculations track k, l, m , and $\hat{\omega}$ at each integration step. These variables can be used to determine the horizontal and vertical phase velocities perturbations using the dispersion relations. These variables are compared because they are more intuitive for the readers.

Figure 3-4: Variation of gravity wave phase velocity with instantaneous latent heat release (K) in thunderstorms. (left) Horizontal phase velocity (m/s) and (right) vertical phase velocity (m/s). The error bars represent 95% standard deviation. The velocities shown are absolute values.

Figure 3-5: Variation of gravity wave phase velocity with time derivative of latent heat release (K/s) in thunderstorms. (left) horizontal phase velocity (m/s) and (right) vertical phase velocity (m/s). The error bars represent 95% standard deviation. The velocities shown are absolute values.

In Figure 3-4, the variation of gravity wave phase velocities is shown with instantaneous latent heat release (LH) in the thunderstorms. In order to make a generalised assessment, the data are binned into five latent heat (K) bins – 0-200, 200-400, 400-600, 600-800 and > 800 , which are the integrated LH values of individual thunderstorm cells nearest to the perturbation. In Figure 3-5, the variation of gravity wave phase velocities is shown with the rate of change of LH in the thunderstorms, which are also binned into five equally spaced values.

In both comparisons, it is seen that there is a general increasing trend of gravity wave phase velocity perturbation with absolute LH and rate of change of LH . In all four plots, the anomaly is seen when the averaged velocity values have a large deviation from the linear trend line, which is much more pronounced in the absolute LH plots. This peculiar behaviour can be traced back to the type of source mechanism discussed in section 2.2.2. The two possible reasons for higher velocities at low LH values may be: 1) the waves are generated from mechanical oscillator effect with relatively higher oscillator frequency, leading to higher phase velocities, and 2) the waves are generated from obstacle effect in which the LH plays no role in wave generation. Also, one of the reasons for such wide error bars is due to the different frequency waves; this dependency on frequency is discussed in the following section.

3.2.2. Wave amplitude

The ray tracing method calculates the vertical flux of wave action at each integration time-step as discussed in section 3.1.2. However, this quantity is not something that could be validated with the current study's instrumental capability. Therefore, F is converted into E which the wave energy density which can be validated with the radiosonde profiles; and both variables are discussed in this section.

Figure 3-6: Variation of gravity wave amplitudes with the instantaneous latent heat of thunderstorms. (left) vertical wave action flux, $\log_{10}(F)$ (kg/s^2) and (right) horizontal velocity power spectrum, $\log_{10}(E)$ (m^2/s^2). The amplitudes shown are the logarithm of absolute values.

Figure 3-7: Variation of gravity wave amplitudes with the time derivative of latent heat of thunderstorms. (left) vertical wave action flux, $\log_{10}(F)$ (kg/s²) and (right) horizontal velocity power spectrum, $\log_{10}(E)$ (m²/s²). The amplitudes shown are the logarithm of absolute values.

The differences in the behaviour of wave amplitudes and velocities may be due to the fact that the amplitude is also dependent on the group velocity which may be very different from phase velocity in many cases [26]. It is also seen that there is a larger standard deviation at the averaged points here compared to the velocity plots. It may be due to the fact that wave amplitudes are not only sensitive to the vertical depth of heating, but also to the horizontal span of heating [26]. Therefore, differences in the horizontal span of the thunderstorm might be creating the variance for similar values of vertical heating.

All the gravity wave variables show strong relationships with both absolute and rate of change of LH . As discussed in the previous section, the results seen here is a combination of all three wave generation mechanisms. However, it is also seen that wave properties are strongly proportional to the rate of change of LH which may be a combination of thermal forcing and mechanical oscillator simultaneously. The fact that the stronger updrafts leading to larger overshoot oscillation (hence stronger mechanical oscillator effect), is also driven by the latent heat makes these two effects highly coupled. The waves generated by the obstacle effect are mainly seen when the absolute LH is roughly between 200 and 400 K. These values are optimal for a convective cloud to be at a growing stage with the vertical development of few km ; able to create the same effect as a mountain [14], [16]. For the sake of reference, the model developed here will be referred to as GW-RAMS model throughout this thesis from here on.

3.2.3. Diagnosing the source

Previous literature has highlighted the different mechanisms through which a convective source could generate gravity waves. The three types of mechanisms are – thermal forcing, obstacle effect and mechanical oscillator [16], [25]. Based on the behaviour seen in Figure 3-4 to Figure 3-7, it is arguable that the trend seen here is a combination of all three types of mechanisms. The plots which show the gravity wave trends with instantaneous LH reveal that there is a strong relation of gravity wave properties with the latent heat, i.e. thermal forcing. Model-based simulations have shown that the vertical wavelength of large-scale waves induced by thermal forcing depends on the depth of the thermal forcing [101], which is generally seen in Figure 3-4 for any particular frequency.

The plots shown in Figure 3-5 and Figure 3-7 should enlighten more on the mechanical oscillator effect which depends on the updrafts and downdrafts frequency. The time derivative of LH gives the increase or decrease in the latent heat release, which actually defines the severity and frequency of the updraft-downdraft cycles.

Figure 3-8: Plot showing the histograms of the rate of change of latent heat (K/s) for different bins of absolute latent heat (K).

It is seen in Figure 3-8 that the majority of waves have slow rates of latent heat. The band of absolute latent heat (200–400 K) where the anomalies are seen in Figure 3-4 to Figure 3-7 have roughly 5% of the high rate of latent heat

occurrences which is lower than the low rate of latent heat. Therefore it can be concluded that the anomalies are dominated by the obstacle effect, but also contain instances from the mechanical oscillator effect. The trend seen here is also an increase in the gravity wave velocities with the rate of change of LH , especially for vertical phase speed. It is also seen that the higher absolute heat occurrences are dominated by low heating rates, which can be interpreted as the thermal forcing to be the most dominant source type in this study. In essence, it can be concluded that the latent heat in the thunderstorms does have a strong impact on the atmosphere by creating large perturbations in the stable regions.

3.2.4. Validation

The model developed in this study is validated using the local radiosonde measurements at Singapore station. The model basically calculates the phase velocity and amplitude of the wave generated from a thunderstorm. The individual vertical profiles of radiosonde can be used to determine the perturbation in the horizontal wind profiles and also the power spectrum of the signal [76], [82]. The power spectrum is analysed for significant waves based on the red-noise spectrum, and the most dominant frequency is used for the conclusion here. Using this frequency which is ω , and dispersion relations, $\hat{\omega}$ can be determined using the equations discussed in section 3.1.1. The radiosonde profile is analysed for waves when the thunderstorm is in the vicinity of the station, which can be identified by using the lightning maps provided by the Meteorological Service Singapore. These maps are used as an indicator or proxy for the thunderstorms, and only events with more than 4000 cloud-to-ground lightning hits per day are considered (cases of a strong thunderstorm).

The variables compared here are horizontal wavelength λ_h , vertical wavelength λ_v , and horizontal energy density E/ρ_0 . These variables can be obtained from the GW-RAMS model developed in the previous two sections with the help of little mathematical manipulations. This conversion formula can be used for obtaining wavelength $\lambda = (u' + \bar{u})/\omega$. When the thunderstorm event coincides with the radiosonde release times, the simultaneous RAMS simulations are performed near the station with the same model configuration as Table 3-1 and the latent

heat is obtained. The values compared here are the average values of velocities and amplitudes calculated from absolute and rate of change of LH equations. Ten events spanning the entire year are selected for performing RAMS simulation at the location and time and obtain the wave properties from the radiosonde profile. The perturbations are analysed above the cloud top altitude which is considered as the stable layer. It is important to note that the GW-RAMS model represents only the waves in the near vicinity of the thunderstorm and within several minutes of heat generation. As the radiosonde data is collected every 12 hours only, it is important to match the location and time of comparison as close as possible. For the radiosonde profiles, the perturbations are calculated based on mean state calculated using cubic spline smoothing in the vertical which is the same type of fit used for ray tracing calculations, to maintain consistency.

Table 3-2: Comparison of model retrieved values and radiosonde values for gravity wave generated by thunderstorms near Singapore radiosonde station. M – GW-RAMS Model, R – Radiosonde.

	$\lambda_h (km)$		$\lambda_v (km)$		$Log_{10} E/\rho_0 (m^2/s^2)$	
Date	M	R	M	R	M	R
22/3/2015	655.4	723.8	0.93	1.05	1.75	1.84
01/4/2015	514.7	486.5	4.56	4.12	0.43	1.42
21/4/2015	748.9	717.7	2.13	2.96	0.25	0.37
28/4/2015	359.9	404.1	2.43	5.23	2.45	2.31
7/11/2015	461.4	718.7	3.75	4.12	1.89	2.09
21/11/2015	681.6	991.2	5.21	4.91	4.56	3.89
10/12/2015	543.8	500.6	1.89	3.24	3.48	3.53
04/02/2016	396.1	544.8	3.45	6.83	5.97	4.356
03/04/2016	520.7	537.8	5.12	5.55	4.63	4.48
17/05/2016	756.9	846.4	5.89	5.24	2.11	2.02
Mean % error	14.27		23.08		18.61	
STD % error	11.37		17.66		19.74	

In Table 3-2, it is seen that the model calculated values have maximum and minimum mean error of 23.08% and 14.27%. The standard deviation of all three models is under 20%. The few instances in which the model and radiosonde derived values are quite different leading to higher % error, may be due to over fitted linear model which doesn't include the obstacle effect mechanism. It is

extremely difficult to classify and separate the mechanisms of generation with the limitations of the techniques and data used in this study. One of the future works for this study include having a supporting argument using the multi-scale modelling methods which can simulate sub-grid processes accurately within the framework of a large-scale general circulation model [102], [103].

3.3. Conclusions

Thunderstorms are a major source of gravity waves including all three types of mechanisms described; however, latent heat forcing seems to play a significant role in that. Previous model-based studies have highlighted the large perturbations created in the atmosphere due to thunderstorms. Three main mechanisms were discussed on how the thunderstorms can generate the gravity waves [15], [16], [19], [21], [55]. There are numerous observational studies which have showed the basic characteristics of the gravity waves generated by thunderstorms but have not developed any quantified relationship between the variables [10], [45], [46], [52], [74], [104]. In this study, a simple model has been developed between the latent heat of thunderstorms and the gravity waves generated by them based on a 10-year period of radiosonde collected data of gravity waves in the lower stratosphere. In order to cover the South-East Asia tropical region, a multidisciplinary approach was used to calculate and analyse the data.

A clear trend is seen that highlights that the gravity wave perturbations are directly proportional to the latent heat release in a thunderstorm. It was seen that not only the absolute amount of heat release but also the rate of change of latent heat release is also a deciding factor. Although the results seen here is a combination of all three mechanisms, it is clear that the latent heat is the driving force and thermal forcing is the most dominant effect. The waves generated by obstacle effect and the mechanical oscillator effect seem to have higher phase velocities helping them to higher altitudes much faster. The model developed here is validated using the individual radiosonde profiles coinciding with thunderstorm profiles at Singapore station. The model performance was mostly

good with mean % error of 14.27%, 23.08% and 18.61%, and standard deviation of % error is 11.37%, 17.66% and 19.74% for horizontal velocity perturbation, vertical velocity perturbation and wave amplitude respectively. This model provides a more intuitive result which can be used for weather forecasting and modelling which includes the gravity wave effects.

The findings concerning the gravity waves generated by the thunderstorms have numerous implications on climate and weather studies. The perturbations caused by thunderstorms can change regional and large-scale circulations and affect global climate [48], [50], [52], [77]. Gravity waves carrying large amounts of momentum and energy are known to have a significant impact in the middle atmosphere by changing ozone and other gaseous concentrations and also impacting ionosphere electron distributions [8]. This type of alterations in the atmosphere causes temperature variations leading to global heat and radiation exchanges which also affects hydro-cycles and global warming in the long run [6]. Models such as one presented in this study need to be studied and improved on a global scale in order to develop a weather forecasting approach which includes all types of processes that are not directly measurable. Satellite approaches are needed to be developed which can cover the regions which are not normally well equipped with ground-based instruments; and can also provide consistent high-quality weather data for research.

4. Estimation of the vertical structure of latent heat generated in thunderstorms using CloudSat

The previous chapter mentioned that the satellite-based methods are needed in order to increase the coverage of latent heat estimations. The convectively generated gravity waves mainly driven by the latent heat released during the phase change of hydrometeors in the thunderstorm [11]. Satellite instruments lack the ability to measure the latent heat directly which restricts such studies on a global scale. In this chapter, a method is discussed for satellite-based latent heat profile estimation which is built upon the previous research. The A-Train satellite constellation has revolutionised the earth observation research by flying six satellites with different instruments in formation in order to study the land, atmosphere and ocean [105]. These satellites with different instruments are aimed to have synergy with each other that complement the observations to obtain a complete scenario. Therefore, the benefit of using this satellite constellation is that the methods developed for one instrument can be extended to others based on some types of learning or other methods. The algorithm

developed here is aimed at the cloud profiling radar on-board CloudSat which can detect the thunderstorms reasonably well.

4.1. Introduction

The latent heat is known to be the energy source for convectively generated gravity waves which deposit a large amount of energy causing a major impact in the upper atmospheric circulation [6], [25]. The scientific community has been developing methods to estimate atmospheric heating for a few decades now. The total amount of surface precipitation can be used as a proxy for inferring the column integrated amount of latent heat released, but it does not represent the heat released from those hydrometeors which do not precipitate to the surface [106]. Numerous theoretical and modelling studies have been conducted which demonstrate the impact of coupling between cloud microphysical processes and latent heat [31], [37]. The better understanding of cloud microphysics has led to accurate cloud modelling with high-resolution micro-scale interactions [89]. These models are important for developing methods that can derive latent heat in the atmosphere from the measurements as latent heat cannot directly be measured by instruments.

Few methods have been developed to derive the latent heat from both ground-based and space-based measurements. To name a few, the studies with TOGA-COARE data retrieved diabatic heating fluxes using a combination of radiosondes and radar etc. Thompson in [107] used a network of ship observation data to relate the wave activity and diabatic heating during the GARP Atlantic Tropical Experiment. Many other ground-based and airborne campaign data were used to derive the heating profiles inside the clouds for a range of cloud types [108]–[110]. Apart from the regional studies, there is a need for consistent large-scale or global estimates of latent heat, which are only possible using satellites. The launch of Tropical Rainfall Measuring Mission (TRMM) provided essential data in the tropical region for developing methods for latent heat retrievals. Various algorithms have been developed over the years for TRMM satellite-based latent heat retrievals [29]. Over the years, TRMM

satellite has provided valuable data which has helped to understand the atmospheric heating better [29], [88], [108], [111]–[114]. Yet, there is a lack of complete picture because TRMM satellite was particularly designed to detect only the precipitation in the atmosphere [108], therefore there is a chance that TRMM retrievals under-estimate the latent heat in the thunderstorms [115]. In order to get the heating profiles of deep convective clouds accurately, the measurements of the complete cloud structure are necessary. Recent studies have highlighted that the smaller particles that freeze while rising up also generate a considerable amount of latent heat leading to severe thunderstorms [39], [116]–[119]. These smaller particles are not detected by the Ku-band radar in TRMM which operates at a larger wavelength. Therefore, methods are needed to develop for W-band cloud radars which can detect the structure of thunderstorms more accurately.

The launch of CloudSat contains the cloud profiling radar (CPR) operating at 94 GHz that has revolutionised the satellite remote sensing of clouds. It can detect a wide range of hydrometeors of different sizes from smaller supercooled cloud droplets to large raindrops and precipitation [120]–[122]. Using CPR to retrieve latent heat profiles can provide an advantage over TRMM as it can detect wider size-range of hydrometeors which undergo phase change [115], [121]. This is significant in thunderstorms as cloud top can extend until 18 km in the tropical region with a variety of hydrometeors in all size range such as rain, graupel, snow, hail, and ice etc. [123]. Nelson in [27] suggested an algorithm to retrieve the latent heat profiles of warm-rain clouds using the Monte Carlo method of estimation. This method proved that CPR profiles can be used to obtain the latent heat profiles with reasonable accuracy. Since the structure of a thunderstorm is different from warm-rain clouds, the effects of multiple scattering become significant especially for W-band radar [124], [125]. Since there have been no documented attempts to estimate latent heat of thunderstorms using CloudSat, this study attempts to perform the estimation with consideration of multiple scattering effects, which is the major novelty in this approach.

This chapter discusses an algorithm developed to estimate the latent heat profiles of thunderstorms in particular. Using the Bayesian Monte Carlo method,

CloudSat radar observations are compared with the simulated database of thunderstorm events to estimate latent heat vertical profiles of thunderstorm events in the region near Singapore. The estimations are validated with local radiosonde and simultaneous ECMWF based cloud-resolving model simulations.

4.2. Algorithm description

The basic idea of the algorithm is that the latent heat release is directly related to drop size and mass that is undergoing a phase change. Therefore, using model-based comparisons, the latent heat release could be indirectly related to the reflectivity. This basis has been proved to produce good results through TRMM studies [29], [108], [126]. This section gives a brief overview of the algorithm and the description of the Bayesian Monte Carlo estimation method which will help understand the requirement of models and approaches explained in section 4.3.

4.2.1. Overview

The CPR reflectivity observations and the corresponding temperature profiles from ECMWF are compared with the simulated database of reflectivity and temperature profiles using Bayesian Monte Carlo (BMC) method. The simulated database is developed using RAMS-CRM (Regional Atmospheric Modelling System – Cloud Resolving Model). The RAMS is initiated using the ECMWF ERA-Interim which provides the atmospheric profiles of general variables such as pressure, temperature and humidity etc. The RAMS is capable of simulating the cloud microphysical processes reasonably well (the details are provided in section 3). The simulation outputs various cloud variables such as hydrometeor concentration and mixing ratio, latent heat, surface and radiation variables etc.

Since the observational quantities are reflectivity and temperature profiles, the database that will be developed should also consist of reflectivity and temperature profiles. Therefore, the output from RAMS is converted into W-band radar reflectivity using the scattering models (more details are given in

section 4.3.3). Once the database is created, the individual observations are compared with the database using BMC method to obtain the estimates of latent heat. The algorithm flowchart is shown in Figure 4-1.

Figure 4-1: Block diagram describing the algorithm flow.

4.2.2. Bayesian Monte Carlo method

The algorithm to estimate the latent heating of thunderstorms is based on the Bayesian Monte Carlo (BMC) method. Numerous studies have applied this method for the retrieval of various atmospheric variables [27], [125], [127]–[129]. An advantage of using BMC is that it facilitates incorporation of variables of different types (even from different instruments) for the estimation. It also provides the uncertainty in each estimation which was developed by [130]. The method also does not need any initial guesses regarding the variables involved unlike the optimal estimation method [131]. The BMC method takes an *a priori* distribution and updates it to an *a posteriori* distribution by reducing the possible solution set to include only those states that provide a close match to the observations [27].

To explain the algorithm mathematically, it requires two vectors which are compared with the BMC method to get the latent heat estimation. The vectors contain the available cloud variables from the CloudSat products and the

corresponding cloud variables representing each member of the database developed using RAMS. The vector from database y_s and the vector from CloudSat measured event y_o comprise of same quantities:

$$\begin{array}{ccc}
 \text{Measurement database} & \text{Measurement} & \text{State database} \\
 y_{s,i} = \begin{bmatrix} Z_{1,i} \\ \vdots \\ Z_{q,i} \\ T_{1,i} \\ \vdots \\ T_{q,i} \end{bmatrix} & y_o = \begin{bmatrix} Z_1 \\ \vdots \\ Z_q \\ T_1 \\ \vdots \\ T_q \end{bmatrix} & x_{s,i} = \begin{bmatrix} LH_{1,i} \\ \vdots \\ LH_{q,i} \end{bmatrix}
 \end{array} \quad (4.1)$$

where Z and T refer to the reflectivity (dBZ) and the temperature (K). The subscript values ($1,...,q$) refer to altitude levels with q being the highest level, and i refers to the i^{th} member of the database. Here x_s represents the database of LH vectors corresponding to y_s .

The method is based on the theory that, the probability density function $pdf(x)$ of the variable x to be estimated, is proportional to the probability that x is the true set of parameters, x_{true} , given that y_s corresponding to x is equal to y_o , i.e. $pdf(x) \propto P(x=x_{true} | y_s = y_o)$. The probability of each database member, $\check{p}_i$, can be expressed as:

$$\check{p}_i = \frac{\exp \left[-\frac{1}{2} (y_o - y_{s,i})^T [O+S]^{-1} (y_o - y_{s,i}) \right]}{\sum_{j=1}^M \exp \left[-\frac{1}{2} (y_o - y_{s,j})^T [O+S]^{-1} (y_o - y_{s,j}) \right]} \quad (4.2)$$

where M is the number of database members, O and S are matrices representing the error covariance corresponding to the elements of y_o and y_s , and T refers to the transpose of the matrix. The observation matrix, O , is a diagonal matrix with the square of uncertainties provided in the CloudSat data products. The simulated error covariance matrix S , consists of the variances, σ^2 , of the error matrix, for all members of the database; in the diagonal components and correlation with the other elements multiplied with the standard deviations, σ of

both elements in the off-diagonal components [130] as shown in equation (4.3), where r_{ij} is the correlation between the i^{th} and j^{th} element. The method to calculate the variance and correlation is shown in equation (4.4) and (4.5).

$$S = \begin{bmatrix} \sigma_1^2 & r_{21}\sigma_2\sigma_1 & \dots & r_{q1}\sigma_q\sigma_1 \\ r_{12}\sigma_1\sigma_2 & \sigma_2^2 & \dots & r_{q2}\sigma_q\sigma_2 \\ \vdots & \vdots & \dots & \vdots \\ r_{1q}\sigma_1\sigma_q & r_{2q}\sigma_2\sigma_q & \dots & \sigma_q^2 \end{bmatrix} \quad (4.3)$$

$$\sigma^2 = \frac{1}{M} \sum_{i=1}^M (\epsilon_i - \bar{\epsilon})^2 \quad (4.4)$$

$$r_{12} = \frac{1}{M} \sum_{i=1}^M (\epsilon_{1,i} - \bar{\epsilon}_1)(\epsilon_{2,i} - \bar{\epsilon}_2) \quad (4.5)$$

where ϵ_i is the error in the i^{th} member of the database $\epsilon_i = y_{s,i} - y_o$, and overbar represents averaging over all database members. The error covariance matrix is one of the main aspects of the algorithm as this directly defines the accuracy of the estimation. The key to producing accurate estimations using BMC method is dependent on the accuracy of the database members' ability to represent the thunderstorm events close to the measured profiles by CloudSat and the variance of y_s which represents the error covariance matrix. This can be tricky as the CloudSat product variables are not 100% accurate as these are also estimations based on certain models and methods through which error can propagate. But since the uncertainties of these estimations are provided in the data products, these errors are considered in O . According to Bayes' theorem, as discussed earlier, the estimated vector (x_r) of latent heat is given by:

$$x_r = \sum_{i=1}^M x_{s,i} \check{p}_i \quad (4.6)$$

The uncertainty in each estimate is given with the consideration of the probability of each database member as shown in equation (4.7):

$$\sigma_x = \sqrt{\sum_{i=1}^M (x_{s,i} - x) \check{p}_i} \quad (4.7)$$

4.3. Data products and methods

The latent heat generated in the atmosphere is a quantity which is not directly measurable with the current state of instrument technology [27]. Therefore, dynamic models are used to derive the heat from the measurable quantities. This approach for immeasurable variable estimation is an established and common in atmospheric studies [129], [132], [133]. It is important that the model simulations are close to reality which ensures the accuracy of the algorithm. This section discusses the CloudSat data products, and the models used for generating the database of thunderstorm events.

4.3.1. CloudSat data

The CloudSat mission was launched in 2007 as a part of A-Train constellation by joint efforts of NASA-JAXA to collect data regarding clouds and aerosols. CloudSat contains a cloud profiling radar (CPR) operating at 94 GHz which is actively collecting cloud data on a global level while orbiting in Polar-Low Earth Orbit [134]. Over the years, various algorithms have been developed to retrieve different cloud variables such as liquid/ice water content, cloud phase, precipitation rates and liquid/ice water path using radar [121], [122], [142], [132], [135]–[141]. The existing data products corresponding to the two satellites provide profiles of liquid water content (LWC) and ice water content (IWC) which are derived from the reflectivity and backscatter measurements. The reflectivity is associated with the hydrometeors based on scattering information. Using the optimal estimation approach, the particle size distribution is retrieved, followed by calculating different moments of distribution which give water

content and effective radius etc. [143]. Since the latent heat is not directly related to the particle size distribution, optimal estimation method does not suit this application.

The main variables which are directly related to latent heat are the amount of hydrometeors present which can undergo the phase change. Since the hydrometeors are detected using the backscatter measurements, this study uses the reflectivity measured by CPR along with the temperature profile obtained by ECMWF ERA-Interim model as auxiliary information. The reflectivity profiles are obtained from the Level 2 GEOPROF data product and the collocated temperature profiles are obtained from ECMWF-AUX product, both the products are available in the CloudSat data processing centre website. The use of the temperature profile is helpful for retrieval as the phase change is partially dependent on the ambient temperature of the atmosphere. The CloudSat reflectivity profiles which are available in the Level 2 GEOPROF data product have a vertical resolution of 100 m.

4.3.2. Cloud-resolving model

The CRM used in this study is the same model described in section 3.1.3. RAMS-CRM solves full nonhydrostatic compressible equations which represent the dynamics and thermodynamics of the atmosphere with conservation equations for moisture budgets. A detailed description of the model and its configuration is given in section 3.1.3. The model is used in the exact same configuration which is initiated using ECMWF ERA-Interim model data with 0.3° grid resolution. The study is conducted nearby Singapore region covering approximately an area of $250,000 \text{ km}^2$. The initiation is supplemented with local radiosonde data whenever possible which are available at 12-hour intervals (<http://weather.uwyo.edu/upperair/sounding.html>). The initiation data is collected for events which are identified as thunderstorms through rain rate and lightning intensity data from Meteorological Services Singapore (MSS).

4.3.3. Scattering models

The database described in section 4.2.2 is populated separately for latent heat and the corresponding measurable quantities. It has to be populated with quantities which are measurable with CPR which is radar reflectivity, Z . The RAMS simulations can produce hydrometeor profiles (corresponding to latent heat profiles) in terms of mixing ratios (g/m^3) and diameter (m). These quantities can be converted into Z . A model to simulate the radar/lidar measurable quantities from the hydrometeor scattering information was developed by Hogan & Battaglia in [124] which shows good performance in estimating backscatter information for thunderstorm profiles. This fast model calculates the time-dependent multiple scattering returns from lidar or radar. It uses a hybrid approach with the small-angle multiple scattering returns of lidar, calculated by the photon variance-covariance method [144] and the wide-angle multiple scattering that occurs for both radar and lidar, calculated using time-dependent two-stream approximation [124]. This method developed needs extinction coefficient, extinction-to-backscatter ratio, asymmetry factor and scattering albedo information for hydrometeor profile of each type at each altitude level in order to get the corresponding apparent radar reflectivity. The extinction area and backscatter area can be calculated using the scattering theories such as Mie and Rayleigh theories [145]. There are methods developed which utilise versions of these theories for determining the scattering properties of cloud hydrometeors [30].

The method used in this study to calculate the scattering properties uses density based approximation because in the larger ice particles, the internal structure becomes important and the assumption of spherical shapes leads to the error in the estimation of scattering properties [145]. The density of large aggregates is generally low (for the size it projects), in such cases the Rayleigh-Gans Approximation has been found to be valid [145]. In this approximation, the electric field experienced at any point in the particle is approximated by the incident field, thereby neglecting interaction between dipoles. The backscatter cross-section may then be estimated numerically from a 1D description of the structure of the particle in the direction of the incident wave [145]. In [146], it is

shown that for realistic aggregates, this 1D function has a self-similar structure and hence its power spectrum can be represented by a power law. Therefore, it is referred to as Self-Similar Rayleigh-Gans Approximation (SSRGA). It is beyond the scope of this report for a detailed discussion regarding the method, the interested readers should refer to [145]. Therefore, to determine the individual particle scattering; Mie theory is used for cloud, rain, drizzle and hail and SSRGA is used for unrimed ice, graupel and snow particles. Both these methods are available as a ready-to-run program which is developed by Robin Hogan (<http://www.met.rdg.ac.uk/~swrhgnrj/>). Once the cross-sectional area is determined using the scattering model, it is possible to calculate the extinction coefficient ζ_{ext} using equation (4.7) [72]:

$$\zeta_{ext} = A_{ext} N_c \quad (4.7)$$

where N_c is the particle number concentration (m^{-3}) which is based on the size distribution function used in the RAMS model for the particle hydrometeor, and A_{ext} is the extinction cross-sectional area (m^2) which is obtained from SSRGA model output.

Primarily, the multi-scatter algorithm was developed for calculating lidar backscatters (which calculates the backscatter coefficient) and later extended it to radar [124], [125]. From the apparent backscatter coefficient, β_a , the apparent radar reflectivity, Z_e , can be calculated using equation (4.8) as described in [124]:

$$Z_e = \frac{4}{|\kappa|^2} \left(\frac{\lambda}{\pi} \right)^4 \beta_a \quad (4.8)$$

where $|\kappa|$ is a reference dielectric factor of water, and at millimetre wavelengths, it is temperature dependent [125] and λ is radar wavelength. From this, the calculated reflectivity values are populated in the database for each altitude level (simulated by RAMS) along with the temperature profiles obtained from ECMWF-AUX product.

4.3.4. Database properties

The database developed here contains two different sections: 1) the reflectivity and temperature profiles of simulated thunderstorms, and 2) the corresponding latent heat profiles of simulated thunderstorms. Ten different RAMS simulations were performed around Singapore region at different times at which thunderstorms were identified using lightning intensity maps produced by MSS which coincided with CloudSat orbital path in 2015. This is done in order to make sure that CPR can detect the thunderstorms seen in lightning maps. Since, this study concerns only with thunderstorms, only those profiles with identified thunderstorms are selected and the corresponding backscatter calculations were performed. The thunderstorm identification follows exactly the method used by the CloudSat CLDCLASS data product which classifies the cloud type [120], [121]. This ensures that the database will be populated with thunderstorm events which are recognised by the CloudSat scene classification algorithm. The database contains simulations to collect approximately 200,000 profiles of the concerned variables.

The database has to cover the most commonly occurring profiles of thunderstorms without having a large inherent variance in the matrix $\mathbf{S}$ (which will lead to higher estimation errors as explained in section 4.2.2). Since this is regional study, it is expected that the thunderstorm profiles will have similar reflectivity and temperature profiles for different events. Database with different variances is used for estimations to make the decision upon the acceptable range of database variance.

Figure 4-2: Mean and standard deviation of reflectivity profiles in the database.

The input components of the database consist of reflectivity and temperature profiles. As explained in previous sections, the reflectivity profiles are calculated from RAMS hydrometeor profiles using scattering models [124], [145]. The maximum of the mean reflectivity profile occurs approximately between 6 to 8 *km* as shown in Figure 4-2 (apart from the value near 0 *km* which represents the surface) which agrees with previous studies in South-East Asia region [123]. It is also seen that the latent heat also peaks in that altitude range indicating that reflectivity is indeed indirectly representative of latent heat. It is also to be noted that the lowest reflectivity value that can be detected by CPR is roughly -30 *dBZ* which refers to light cirrus clouds with small hydrometeors which are not dense [134]. Therefore the reflectivity values in the database at the top (above 17 *km*), which are very low, would not be highly matched with the CPR data. However, these altitudes can be ignored in this study as the latent heat is almost negligible as seen in Figure 4-2.

4.4. Results and discussion

4.4.1. Quantitative analysis

The estimated latent heat profiles are calculated for the CloudSat identified thunderstorms events during 2016 near Singapore region which are not part of the database, for testing purposes. The CloudSat CLDCLASS product is used to detect the thunderstorms with particular criteria to identify the pixels [121]. Approximately 150 profiles from four different events are used to test the algorithm in 2016. Once the estimations are obtained for CloudSat profiles, the validation of the profiles is done using simultaneous RAMS simulations initiated with ECMWF ERA-Interim atmospheric profiles at the particular location of the event. It is to be noted that, the RAMS based validation is the only option in this case because the nearest radiosonde station collects data every 12 *hours* at 00:00 and 12:00 UTC and CloudSat orbital pass time over this region is approximately at 06:00 and 18:00 UTC. Therefore the radiosonde cannot be used for validation, and there are no other instruments which are able to measure the vertical profile of the thunderstorm in the region. As to show the closeness of RAMS data with the true values, the RAMS simulations are validated using reflectivity profiles from CloudSat (details provided in the appendix). The performance of the algorithm is judged based on the error in vertical profile estimation and linear regression analysis between retrieval and validation values of the column-integrated heating.

Figure 4-3 shows the mean % error in the retrieved latent heat profiles. It is seen that the error is increasing with altitude. The reason may be due to the RAMS vertical resolution which does not allow higher resolution than 1 *km* at altitudes higher than 5 *km*. This may lead to an inaccurate representation of microphysics at higher altitudes. Also, the reflectivity differences between the database and CPR at altitudes higher than 16 *km* can be high as CPR can only detect the minimum reflectivity of -30 *dBZ*, while some of the simulated reflectivity profiles go up to -40 *dBZ*. It is also seen that largest error below 16 *km* occurs between 8 and 10 *km*, which is analogous to the peak standard deviation of the latent heat database. The database may need to be improved in order to obtain better results.

There are issues that need to be addressed in future: 1) the effect of different microphysics scheme used; 2) improving the database may solve some issues, but other estimation schemes may also be implemented to compare and optimise, and 3) to study the errors due to scattering models.

Figure 4-3: Mean % error in the latent heat retrieval with RAMS validation.

It is also sensible to compare the integrated values of heating in a column from estimation and true values to better judge the performance of the algorithm. Figure 4-4 shows the linear regression plots for column integrated latent heating. The plots show that the column integrated values have reasonably good linear fit with consistent negative bias; however due to some of the similar estimated values (highlighted in the red ellipse), the regression has a low value of 0.8022. The consistent negative bias is may be due to the limitation of the current scattering models which can represent the reflectivity of a W-band radar. The CPR itself gets attenuated heavily in the heavy precipitation and this attenuation may not always be represented accurately by the multi-scatter models. Therefore, the algorithm will underestimate the latent heat exchanges involving large particles such heavy rain or growth of graupel and hail.

The similar estimation values (highlighted in the red ellipse in Figure 4-4) may be due to over generalisation of the method when deriving the probability

distribution. This happens when the observation profile which is used may not find a similar profile in the database leading to a more general outcome. For example, when the BMC method cannot find any profile in the database similar to the observational profile, it creates a wide and general *pdf*, which on averaging will produce a very general estimation as per equation (4.6). If these types of events occur often, then this approach will produce estimations of similar magnitude which is seen in Figure 4-4, highlighted by red ellipse. This issue may be the limitation of the method described in this study. As mentioned before, an extended study of the effects of the microphysical schemes and scattering models is needed in order to improve the performance of the algorithm.

Figure 4-4: Linear regression plot comparing the CloudSat estimated values and collocated RAMS simulated values of column-integrated latent heat. The plot shows the linear fit (black lines), 95% prediction intervals (black dashed lines) intercept, R^2 , bias and root mean square error (RMSE) values. The blue thick line shows the new fit line after neglecting the estimated values < 2, and linear equation for new fit in blue text.

Although the performance is lacking, it is possible to improve it by neglecting the estimations lower than 2 K. As seen in Figure 4-4, the blue line represents the improved linear fit from neglecting the estimations below 2 K. This however limits the algorithm usage and range, it is sensible to do this to obtain a much better fit. Since the validation in this study is limited by models as latent heat is

not measurable quantity, it is sensible to compare the estimations with another CRM which is capable of similar functionalities and performance as RAMS. Weather Research and Forecasting (WRF) model is another well-established circulation model with CRM capabilities. Various studies have used WRF to simulate and analyse the cloud microphysics in the convective storms [28], [89], [147], [148]. The WRF-CRM was ran using the same configuration as RAMS-CRM described in section 3.1.3, except for microphysics scheme which is defined in [149]. This is also two-moment bulk microphysics scheme fully capable of simulating the micro-scale processes in a thunderstorm. The linear regression plot comparing the column integrated values latent heat of CPR based estimations and WRF-CRM is shown in Figure 4-5.

Figure 4-5: Linear regression plot comparing the CloudSat estimated values and collocated WRF simulated values of column-integrated latent heat. The plot shows the linear fit (black lines), 95% prediction intervals (black dashed lines) intercept, R^2 , bias and root mean square error (RMSE) values.

The linear regression (shown in Figure 4-5) comparison of the estimated values with the WRF model values show similar performance as with RAMS. The slope of the linear fit is slightly lower, and the distribution is wider than the RAMS comparison. These differences might mainly be due to the differences in the microphysics calculations as the models use different schemes. As expected, the

algorithm seems to over generalise the estimations below 2.5 K as lot of repeated values are seen. This comparison with WRF model shows a reasonable performance with no significant bias in the performance when compared with RAMS based validation. This helps to better comment on the robustness of the algorithm since an observational validation is not feasible. The next section demonstrates and discusses a case study of the retrieval done for a thunderstorm event.

4.4.2. Case study

A case study has been performed to demonstrate the results and retrieval limitations. The profile discussed in the case is a thunderstorm event on April 29, 2016, which was detected by CloudSat CPR. As seen in Figure 4-6, the retrieval algorithm has been applied to CLDCLASS identified thunderstorm event. It can be seen that the actual cloud extends beyond the identified thunderstorm (the anvil of the storm), yet the algorithm is restricted to perform the retrieval only in these regions of the storm tower. The reason for this mainly because the towering structure of the thunderstorm (shown as red path in cloud scenario panel of Figure 4-6) is believed to be the container for all the rapid heating, updrafts and downdrafts and the mass [21], [31], [93], [150], [151]. As this chapter is mainly concerned with estimating the thunderstorm strength, the estimation is performed only for the limited identified area. The latent heat profiles are exactly aligned with pixels where the reflectivity values are near the maximum.

The maximum reflectivity values for clouds are approximately 5 dBZ (the maximum of 20 dBZ colour-bar shown in Figure 4-6's reflectivity panel refers to the ground reflection). This aligns with the fact that bigger drops or higher drop concentration (higher reflectivity) undergoing phase change release a large amount of heat which is seen as circled in red in Figure 4-6. Also, having the temperature profile supports the retrieval which will help in a better probability distribution.

Figure 4-6: Case study profiles, (from top to bottom) a) – cloud scenario profiles obtained from CLDCLASS product which shows the different types of clouds. Here is the number 8 represents the thunderstorms which is seen as the red patch in the plot. b) The corresponding reflectivity (dBZ) profiles obtained from GEOPROF product, c) retrieved latent heat (K).

4.5. Conclusions

Thunderstorms are one of the largest disturbances in the atmosphere which are the source of gravity waves and changes in circulation. It can be considered as one of the phenomena in which a large amount of energy exchange is involved. The hydrometeors and ambient environment undergo interactions which generate heat, rain and waves affecting the human life. It is important to develop methods to observe and monitor the energies and heat released during a thunderstorm event which will be helpful in understanding the following the interactions and improving the weather models.

This chapter suggests a method for estimating latent heat release of a thunderstorm using CloudSat CPR measurements in the tropical region. The

method is based on the Bayesian Monte Carlo approach which calculates the probability of any particular event using a sample database. The sample database is generated using RAMS-CRM [90] in this study. The hydrometeor profiles from the model simulations are converted into radar backscatter values using the scattering models which also include the multiple scattering effects [124], [125], [145]. The CPR reflectivity profiles along with ECMWF temperature profiles are compared with the corresponding database profiles using the Bayesian Monte Carlo method to calculate the probability of each database member for a particular observation [130].

The algorithm performance is evaluated using approximately 150 CPR observations in 2016 near Singapore region which are identified as thunderstorm profiles using simultaneous RAMS and WRF simulations. The mean vertical error profile for latent heat has peak value less than 20% at higher altitudes where RAMS has poor vertical resolution and also the differences between reflectivity database and CPR reflectivity may differ at that altitude. The algorithm still needs improvements in many regards such as 1) vertical profile error reduction, 2) increase the quality of sample database to cover a wider range of scenarios with detailed error analysis of microphysical schemes and scattering models, and 3) to verify the performance in other tropical regions. Some of the future works identified for this study: 1) improve the database quality and quantity, 2) study the scattering model errors, and 3) application of other estimation methods such as variational analysis and machine learning for comparison. Overall, the study has proved that the latent heat can be estimated using CPR reflectivity profiles with reasonable accuracy for thunderstorms (for which CPR is subjected to multiple scattering effects). Gradually, the method will be extended to a global scale and other types of convective clouds.

5. Gravity waves generated by thunderstorm: Satellite based study

The latent heat is shown to have a definite relationship with the gravity wave velocity perturbations as discussed in chapter 3. It was mentioned that in order to have a global model, satellites methods need to be developed to estimate latent heat because RAMS-CRM simulations are computationally expensive and time-consuming. In chapter 4, a method was developed to estimate the vertical latent heat profiles from CPR reflectivity profiles. But this estimate is limited to a narrow distance across-track orbital path for studying the entire thunderstorm's latent heat. For studies similar to chapter 3, instruments need to have wide coverage region, enough to monitor the entire span of the thunderstorm. Satellite-based imaging is one such example which covers wide regional span. In this chapter, a method is described which uses the synergy between A-Train satellites in order to extend the method developed for CPR to Moderate Resolution Imaging Spectro-radiometer (MODIS) which is onboard Aqua [105]. MODIS is a spectro-radiometer imager which has been used for various land, atmosphere and ocean studies over the years. It has a wide span covering thousands of *km* along and across track with 1 *km* resolution which is highly

beneficial for atmospheric studies [152]. With the method developed for MODIS latent heat estimation, the similar gravity wave modelling study as GW-RAMS model study is performed in two tropical regions: SEA and WA, and the performance is tested with radiosonde results.

5.1. Introduction

A-Train satellite constellation was designed in such a way that not only the instruments involved can provide complementary data, but the satellites also have their orbital paths close to each other on almost the same ellipse [105]. Various studies over the years have developed methods utilising the synergy between the A-Train satellites to monitor the atmospheric variables [136], [138], [153], [154]. Since CloudSat is part of A-Train constellation; it is advantageous to extend the algorithm to estimate latent heat developed in chapter 4, for MODIS onboard Aqua exploiting the synergy between the satellites and instruments involved. MODIS has a very wide swath and coverage and it would be highly beneficial to have a method for estimating latent heat distribution using an instrument which can provide consistent data in the tropical region.

Over the years, MODIS has been providing valuable data on various land, atmosphere and ocean variables. Utilising the different channels available in MODIS, various cloud optical and microphysical variables are obtained at a resolution as high as 1 *km* in horizontal [152], [155], [156]. This chapter discusses a method to estimate the latent heat of thunderstorms for MODIS. Using the results from CPR's latent heat profiles using the method described in chapter 4, the column integrated collocated values of heating is used to train a neural network for a set of MODIS collocated cloud inputs.

Following the neural network training, the MODIS cloud images are used for estimating the latent heat for identified thunderstorm pixels. The estimated latent heat maps are used with the ray tracing results to obtain a gravity wave-latent heat model in a similar approach as chapter 3. In this chapter, the model is developed in two regions: 1) the South-East Asia (SEA) region, and 2) West

Africa (WA) tropical region. The model results in both regions are tested with collocated radiosonde based results.

5.2. Latent heat maps using MODIS

As mentioned previously, this study exploits the fact that Aqua and CloudSat fly on the similar orbital path as shown in Figure 5-1 with a time lag of 60 s; the algorithm assumes a stagnant atmosphere or very slowly changing atmosphere within this time lag. Since MODIS has channels which are used for retrieving various cloud top and optical properties, this information can be used to derive the thunderstorm latent heat. The latent heat of the cloud, the cloud top and optical properties may not be directly related to each other; therefore a learning algorithm is necessary in this case for the estimation. This section discusses the data and approach in detail along with the training results.

Figure 5-1: Schematic diagram of the A-Train satellite depicting their right-ascending node [105].

5.2.1. Data and concept

The A-Train satellites were launched to observe various aspects of the earth including land, ocean and atmosphere. The satellite orbits were designed in a way to complement each other to obtain the best possible observation of the events occurring [105], [157]. In this section, both the satellite instruments involved in the study are described along with the data products available; followed by the concept and approach for generation latent heat maps.

MODIS is an instrument developed by NASA which has a total of 36 radiation spectral bands with a wide swath of $2330 \times 2000 \text{ km}$ and temporal resolution of 5 *minutes* along the orbital track and is on-board Aqua satellite. The MODIS cloud product uses the radiation measurements at different frequencies to retrieve the cloud properties [152], [158]. The data product used in this study is the MODIS cloud product represented by MAC06S1 from which the cloud optical and microphysical variables are obtained [159], [160]. MODIS cloud products provide a wide range of cloud optical and macrophysical variables. In order to represent a thunderstorm, the variables need to represent the water content, the cloud vertical extent and microphysical properties if possible. The variables chosen for this study are:

- Cloud optical thickness (*COT*) – this represents the total water content in the cloud which can attenuate the optical frequencies.
- Cloud effective radius (*CER*) – this represents drop effective radius based on particle size distribution.
- Cloud top temperature (*CTT*) – this represents the vertical extent of cloud (an optically thick and high top cloud can have a lot of updrafts driven by latent heat).
- Brightness temperature at $11 \mu\text{m}$ (BT_{11}) – this is basically the water vapour channel used for retrieving cloud properties (mainly for redundancy purposes).

The horizontal resolution of these images varies from 1 to 5 *km*, which are all interpolated on to an equal grid resolution of 5 *km*. The CPR reflectivity profiles are used to calculate the latent heat profiles of thunderstorms using the method described in chapter 4. Since the CPR CLDCLASS product provides the cloud classification, the thunderstorm pixel from CPR and the corresponding geo-located MODIS pixels are obtained for training. The MODIS images have been collocated on to coinciding CPR footprints which can be accessed using the A-Train data depot website (<http://disc.gsfc.nasa.gov/atdd>) [105], [157], [161].

Figure 5-2: Schematic depicting the MODIS cloud image along with coinciding CPR vertical cloud profiles. The collocated data points for neural network training are highlighted by black dotted lines.

The CPR based latent heat profiles are integrated vertically in this study for training. This study is only limited to 2D map estimation of latent heat, therefore the vertical structure of heating is not considered. The CPR based estimation results are also corrected before the training. As seen in Figure 4-4, the estimation of column-integrated profiles has a consistent linear negative bias. This linearity of bias is used for correcting each estimated profile in this study. The next section explains the training algorithm used in the study along with the results of training.

5.2.2. Bayesian neural network

The real world problems are generally analysed with the help of statistics which are averaged over a large number of samples in order to make generalised predictions. In recent years, data learning algorithms have gained recognition for accurately modelling the nonlinear behaviour of complex systems. Various studies have been conducted in which the neural network has been used for atmospheric pattern recognition [162]–[164]. Since this study is also based on a large set of data from which the required patterns have to be recognised, the neural network is chosen which is easy to implement and verify the credibility of the concept which is highly non-linear.

Figure 5-3: Block of the neural network with hidden and output layers.

A two layer feed forward neural network as shown in Figure 5-3 with sigmoidal hidden neurons (the S-shaped curve in the first block) and linear output neurons are applied for training the variables the MODIS variables. This is a supervised network which can handle multi-dimensional mapping problems. This is commonly used for prediction, pattern recognition and nonlinear function fitting [162], [164]–[167].

The algorithm used to train the network to fit the input-target data is Bayesian Regularization back-propagation method. Back-propagation networks are able to learn the prediction and classification problems. These types of networks are considered to be a black box which fits and provides a solution to a normally incomprehensible problem [168]. The Bayesian approach, in contrast, depends on coherent inference based on defined relationships [169]. These two extremely opposite methods can be used to design a system which can have a controlled approach to the flexible system. The neural network is a type of regression technique with more non-linear adaptations. However, due to this non-linearity or flexibility, the network over-fits the data and loses the ability to generalise. Bayesian methods can be used to compensate for such over-flexibility [169], [170]. Since the problem at hand requires that the size of the network is sufficiently large; the functions created by the network will be more complex. In these complex networks, there is a possibility that the network also models the noise in the data as a part of nonlinear relation [164]. There are two ways to overcome the issue of over-fitting: 1) early stopping, and 2) regularisation. Early stopping the training can lead to large mean square errors or poor regression. Regularisation improves generalisation by constraining the size of network weights which is done using Bayesian methods [168]–[170].

The network in the study consists of MODIS cloud inputs, $X^{(b)}$ and latent heat targets, $G^{(b)}$, where the dataset $B=\{X^{(b)}, G^{(b)}\}$ is trained by adjusting the weights ε , in order to minimise an error function shown in equation (5.1):

$$E_B(\varepsilon) = \frac{1}{2} \sum_b \sum_i \left(G_i^{(b)} - Y_i(X^{(b)}; \varepsilon) \right)^2 \quad (5.1)$$

where b is the dataset member, and the output $Y(X; \varepsilon)$. The error function basically compares the closeness of the output with the target. The minimization is based on a repeated evaluation of the gradient of E_B using back-propagation along with the regularization, known as weight decay. The new error function is given by equation (5.2).

$$\Psi(\varepsilon) = \beta E_D + \gamma E_W \quad (5.2)$$

where Ψ is the modified error function, $E_W = \sum_i \varepsilon_i^2$, and γ and β are parameters which control the weight distribution and noise in the system. If $\gamma \gg \beta$, the training will lean more towards the optimisation of weights and if $\gamma \ll \beta$, then training will tend to reduce the error. Therefore Bayesian regularisation is designed to optimise these parameters [169].

As mentioned previously, the training is performed in two different tropical regions: 1) SEA region and 2) WA region. For training the network in SEA region, ten years (2007-2016) of collocated CPR and MODIS data are collected for CLDCLASS identified thunderstorm pixels in the tropical region between 100° E to 120° E longitude and -20° S to 20° N. Similarly, data for WA region is collected in 2006 in the region between -30° W to 30° E and -20° S to 20° N. The reason for the smaller data size for WA is explained in section 5.3. The performance of the training is shown in Figure 5-4 and Figure 5-5, which show the closeness of the predicted output with the target for the two regions.

Figure 5-4: Regression plot showing the performance of training in SEA region. The axes represent the latent heat values (K). The blue line represents the linear fit of the data.

Figure 5-5: Regression plot showing the performance of training in WA region. The axes represent the latent heat values (K). The blue line represents the linear fit of the data.

As seen in Figure 5-4 and Figure 5-5, the neural network is able to predict the latent heat values with good generalized linear fit with regression, R^2 and sum of squares error, SSE values on both regions. The number of hidden layers used for training is 100 which makes it computationally expensive, but since the training is a onetime event with good performance; this is acceptable. The next section describes the gravity waves models built using this trained neural network and MODIS images for the two regions.

5.3. Modelling gravity waves generated by thunderstorms: MODIS based study

This section discusses the gravity wave-latent heat model generated using satellite-based approach. Similar to the approach describes in chapter 3, the gravity waves are first identified in the stratosphere, and then traced back to their sources and check for the amount of latent heat at the identified locations. Here, instead of RAMS simulations at the convective source, MODIS images are used for estimating the latent heat. This approach is computationally and temporally more efficient than the RAMS simulations based approach.

Once the ray traced results are obtained, instead of analysing the ray paths with OLR maps from NOAA; the ray paths are analysed for perturbations near collocated MODIS images with $COT > 25$ [171]. The COT is just an indicator for clouds with very high hydrometeor density or large drop size. This along with the $BT_{11} < 200\text{ K}$ indicate the possible thunderstorm cloud [171]. The results and performance of the model developed using this method for SEA and WA are discussed in this section.

5.3.1. South-East Asia region

In this section, the gravity wave-latent heat model is developed using MODIS images based on the results from chapter 4 and section 5.2 for SEA region which will be referred to as GW-Sat-SEA model from now on. The data used here is the same data used in chapter 3 which is described in section 3.1. The same waves

are analysed for thunderstorms along the ray paths using MODIS COT and BT_{11} are mentioned previously. Since MODIS images have a very wide span, multiple storms and MCS can be seen in the images. Therefore, for the analysis to retain the same analytical consistency, the single thunderstorm core closest to the ray path (spatially and temporally) is considered for the model as shown in Figure 5-6 and the latent heat images are interpolated at 375 m horizontal resolution which is same as the RAMS grid resolution used for GW-RAMS model. This approach has a limitation that it cannot provide the rate of change of latent heat as the satellite is moving. Therefore the model developed here is only with the absolute latent heat information of the individual thunderstorm cells. Since the satellite images are not always coincident in time with the collocated ray paths; there is a threshold of time used to decide whether an image will be used. Based on the assumption in section 5.2, of stagnant atmosphere for approximately one minute; the same threshold is used here.

Figure 5-6: MODIS latent heat image on the geographical maps with ray path (Thick black line) of the gravity wave generated by the thunderstorm shown in the red dotted circle.

From 2007-2016, all the 2000 rays tracing calculations that led to convective origins are analysed in this study. Since the computational cost is much lower in

this case, more ray paths can be afforded for analysis. However, there is a trade-off because of the satellite orbital motion. Although the number of rays analysed is more; it is limited by the spatial and temporal collocation coverage. Therefore, only 240 rays satisfied the geo-spatial collocation and time threshold defined earlier. The GW-Sat-SEA model here shows the gravity wave phase velocities and amplitude relationships with the estimated absolute latent heat.

Figure 5-7: Variation of gravity wave phase velocity with instantaneous latent heat release (K) in thunderstorms. (left) Horizontal phase velocity (m/s) and (right) vertical phase velocity (m/s). The velocities shown are absolute values.

Figure 5-8: Variation of gravity wave amplitude with the instantaneous latent heat of thunderstorms. The values are shown in terms of wave energy, $\log_{10}(E)$ (m^2/s^2). The amplitudes shown are the logarithm of absolute values.

It is seen that the slopes of the linear relationships developed here are very similar to the ones developed in chapter 3. Yet, the slopes seem to have higher values in all three variables, which can be associated with the fact that the latent

heat estimated here is underestimated. This is consistent with the training results seen in Figure 5-4, where it is seen that the output values are underestimated to have a linear regression slope of 0.83. Since the latent heat is underestimated, the slope of the linear relationship developed here is higher. It is also noted that much higher latent heat values are observed here than before as satellites maps can provide the complete thunderstorm cell coverage which may be too big for RAMS grid sometimes (innermost grid which is able to resolve convection). The obstacle effect cases are not seen here mainly because most of the storms seen are in the mature stage; which is dominated by thermal forcing effect. The model results are validated with radiosonde profiles in a similar way as in section 3.2.4. The RAMS simulations are performed at the identified thunderstorms and the gravity wave variables are calculated based on the linear relationships.

Table 5-1: Comparison of GW-RAMS retrieved values, GW-Sat-SEA retrieved values and radiosonde values for gravity wave horizontal velocity perturbation and vertical wavelength generated by thunderstorms near Singapore radiosonde station. M – GW-RAMS Model, R – Radiosonde, S – GW-Sat-SEA model.

Date	$\lambda_h (km)$			$\lambda_v (km)$		
	M	S	R	M	S	R
22/3/2015	655.41	686.57	723.77	0.93	1.01	1.05
01/4/2015	514.67	524.73	486.53	4.56	5.23	4.12
21/4/2015	748.9	755.95	717.73	2.13	2.94	2.96
28/4/2015	359.87	374.95	404.1	2.43	3.42	5.23
7/11/2015	461.4	493.57	718.74	3.75	3.92	4.12
21/11/2015	681.55	708.69	991.16	5.21	5.86	4.91
10/12/2015	543.83	605.15	500.6	1.89	2.13	3.24
04/02/2016	396.06	423.2	544.83	3.45	3.76	6.83
03/04/2016	520.71	607.16	537.8	5.12	5.67	5.55
17/05/2016	756.94	825.29	846.4	5.89	6.54	5.24
Mean % error	14.27	14.39		23.08	21.46	
STD % error	11.37	9.98		17.66	15.14	

Table 5-2: Comparison of GW-RAMS retrieved values, GW-Sat-SEA retrieved values and radiosonde values of gravity wave amplitude generated by thunderstorms near Singapore radiosonde station. M – GW-RAMS Model, R – Radiosonde, S – GW-Sat-SEA model.

$\log_{10} E/\rho_0 (m^2/s^2)$

Date	M	S	R
22/3/2015	1.75	1.84	1.94
01/4/2015	0.43	1.42	0.61
21/4/2015	0.25	0.37	0.42
28/4/2015	2.45	2.31	2.85
7/11/2015	1.89	2.09	2.13
21/11/2015	4.56	3.89	5.02
10/12/2015	3.48	3.53	3.81
04/02/2016	5.97	4.356	6.35
03/04/2016	4.63	4.48	4.94
17/05/2016	2.11	2.02	2.43
Mean % error	18.61	19.73	
STD % error	20.74	17.11	

The validation results show that the model developed using satellite has surprisingly better performance. Both the mean and standard deviation of percentage error are almost equal or lower for all estimation except the mean percentage error of wave amplitude. The consistent underestimation of latent heat from satellites methods may have led to higher gravity wave variable based on the higher linear gradient of the models. One of the other reasons may be that in some instances, the RAMS innermost grid may not cover the thunderstorm core completely which may lead to incomplete modelling which is improved in the satellite estimations of latent heat.

5.3.2. West Africa tropical region

The West Africa region is geographically one of the most unique places on earth, which is surrounded by ocean, vegetation and a large desert which leads to unique meteorological events and behaviour. The climate is characterised clearly by two seasons – dry and wet. The dry season extends from October to April which experiences dust and sandstorms while the wet season extends from May to September with strong daily precipitations [52], [172]. Kafando in [172] shows the different frequencies of gravity waves present in the region using radiosonde data. Gravity waves structures were correlated with the thunderstorms in the

WA region using 10 year infrared sound stations [51]. Variation and characteristics of the gravity waves generated during the monsoon season were studied using several radiosonde stations covering WA which provided a much clear picture of the mesoscale waves generated by convection [52]. The waves identified by these studies have mean horizontal and vertical wavelengths of 2000 *km* and 2 *km*. These observational studies provided a general idea regarding the waves associated with the convective environment for this region. However, these studies do not emphasise the particular location of the sources associated with thunderstorms neither they quantify the relationship as mentioned in previous chapters. Therefore, the motivation for this study is to extend the knowledge gained by previous studies and develop a model defining the relationship between latent heat-gravity waves in the WA region which will be referred to as GW-Sat-WA model.

The data used in this study is based on a year-long campaign conducted in Niamey, Niger called RADAGAST by Atmospheric Radiation Measurement (ARM) group. Unlike the SEA region study where the regular radiosonde profiles are used for identifying the gravity waves in the stratosphere based on a ten-year study; here the study is conducted only for one year in 2006. The reason is that the regular radiosonde profiles available in the WA region are inconsistent at the stratospheric altitudes. Therefore, the RADAGAST campaign data are used which provides atmospheric profiles every six hours during 2006 [173]. Vertical profiles of temperature, wind speed and direction, are retrieved from MERGESONDE data product which assimilates the radiosonde, high-resolution microwave radiometer data with interpolated model data. The vertical profiles are produced at 315 altitude levels between 0-60 *km* with a temporal resolution of 60 *s* [173]. The study is limited to analyse up to the altitude of 23 *km* for investigating the gravity wave presence, above which the data is inconsistent. The mean profiles calculated for the entire year based on which the perturbation profiles are obtained as explained in section 3.1.1.

The waves are identified using the wavelet transform of the averaged kinetic energy density signal in the stratosphere based on the perturbations profiles of wind speed as described in section 3.1.1. After going through the similar process

as described in section 3.1.1, the wavenumbers and intrinsic frequencies of the waves are obtained for ray tracing calculations. Using GROGRAT along with ECMWF ERA-Interim model with 0.3° grid resolution to provide the background atmosphere, the ray tracing calculations are performed as described in section 3.1.2.

The ray tracer has been applied for 1000 identified waves. Based on the monthly mean tropopause altitude, the waves terminating below 17 *km* are analysed for WA region [86]. Around 31.18% terminated in the mid/high-troposphere, 32.35% traced back all the way to the surface and 36.47% terminated in the stratosphere. Since the study is conducted only in the tropical region (-25° S to 25° N), the waves generating in other latitudes are ignored. After ignoring the waves that are not originating within the tropical region, the number waves left for the analysis is approximately 300. Out of these 300 waves, approximately 100 waves are identified to originate from thunderstorms. The thunderstorm activity is obtained using the ISCCP (International Satellite Cloud Climatology Project) data which uses a cluster of NOAA geostationary and polar orbiting satellites to build global 3-hourly maps of cloud cover and activity, instead of OLR maps [174]. This dataset is better as it provides 3-hourly maps compared to the daily mean OLR maps used in the previous studies. The ISCCP data is available only until 2006, which is why it is not used in the SEA region studies.

It is seen from Figure 5-9b that the vertical propagation speed of waves in the WA region is very low on average. Most of the waves are taking longer than two days to reach stratospheric altitudes while most of the waves generated by thunderstorms in SEA region took less than a day to reach the stratosphere. This may be related to the general circulation of the mean flow which differs for different regions. The waves associated with thunderstorms mainly have a period from 20 to 40 *h*. The horizontal phase speeds at the source point vary from 2 to 10 *m/s* with the highest occurrence at 3 *m/s* and the vertical phase speeds vary from 0.02 to 0.1 *m/s* with a peak occurring at 0.04 *m/s*.

Figure 5-9: Ray tracing results in WA region. a) Shows the geographic ray paths which are generated from thunderstorms. The background colours show the average % of deep convective activity in 2006. b) time-altitude ray paths showing propagation times of waves generated from thunderstorms.

The gravity wave-latent heat model is developed for the WA region just like the SEA region model developed in section 5.3.1. As seen in Figure 5-9a, most of the waves generated by thunderstorms are mainly arising from the storms in the South Atlantic Ocean as shown by yellow pixels. The GW-Sat-WA model developed is here based on the seventy geo-located cases of ray-traced events and the availability of MODIS images from both Aqua and Terra satellites.

Figure 5-10: Variation of gravity wave phase velocity with instantaneous latent heat release (K) in thunderstorms. (left) Horizontal phase velocity (m/s) and (right) vertical phase velocity (m/s). The velocities shown are absolute values.

Figure 5-11: Variation of gravity wave amplitude with the instantaneous latent heat of thunderstorms. The values are shown in terms of wave energy, $\log_{10}(E)$ (m^2/s^2). The amplitudes shown are the logarithm of absolute values.

Figure 5-10 shows the variation of the gravity wave phase velocities with the absolute latent heat release. The relationships seen here are of the same order of the relationships developed for the SEA region shown in Figure 5-7. It is therefore expected to have the behaviour of similar order and the differences in the values of slopes and intercepts may be due to different mean flow and geographical influences such as air density and temperature because of the different geographical features. The same type of trend is seen for the wave amplitude relationship which is shown in Figure 5-11.

The GW-Sat-WA model is validated using the local radiosonde data at two different stations – Pointe-Noire and Ouessou. Unlike the stratospheric data, the radiosonde data until 19 km is quite consistent and can be used to identify the perturbations in the troposphere which are close to the station. Since the radiosonde profiles are available every twelve hours at 00:00 am and 12:00 pm daily, the thunderstorms are identified which are collocated with that. Aqua satellite's equator crossing time for that region overlaps roughly with the radiosonde collection time at 12 pm. Therefore, clouds with $COT > 25$ are searched in the vicinity of the two stations around that time.

Table 5-3: Comparison of GW-RAMS retrieved values, GW-Sat-WA retrieved values and radiosonde values of gravity wave properties generated by thunderstorms in the WA region. R – Radiosonde, S – GW-Sat-WA model.

	$\lambda_h(km)$		$\lambda_v(km)$		$Log_{10} E/\rho_0 (m^2/s^2)$	
Date	S	R	S	R	S	R
Pointe-Noire						
22/3/2007	587.79	502.31	1.91	3.41	3.77	3.21
01/4/2007	629.33	431.25	2.04	5.34	3.83	4.12
21/4/2007	1244.16	756.12	3.01	4.12	4.27	2.45
28/4/2007	444.31	612.34	4.32	5.65	4.86	3.02
Ouessou						
21/11/2007	527.66	586.45	4.63	4.12	5	3.56
10/12/2007	709.38	945.12	1.91	3.02	3.78	2.97
04/02/2007	289.69	412.54	2.63	3.56	4.11	3.02
03/04/2007	567.63	845.12	1.91	1.45	3.78	2.86
17/05/2007	551.69	645.12	2.48	1.78	4.03	2.45
Mean % error	29.02		31.26		31.14	
STD % error	16.62		15.26		19.73	

It is seen in Table 5-3 that the WA region has larger error compared to SEA region estimations. The RAMS has been simulated in WA region near Niamey, Niger near the RADAGAST campaign area. Since this campaign has a cloud radar [175], it is easy to detect the thunderstorms' locations at which RAMS simulations are performed. One of the noticeable differences between the two regions is the

considerable lack of ground-based data collection which also affects the accuracy of ECMWF data which is used for RAMS initiation. The RAMS simulations in this region may have inherent errors due to this sparse data coverage. The effects of this on RAMS accuracy are discussed in the appendix-A, where the RAMS profiles are converted into reflectivity profiles to compare with satellite and ground-based cloud radars.

5.4. Conclusions

This chapter describes a method to estimate the latent heat maps of thunderstorms using MODIS images. The integrated latent heat values from CPR reflectivity profiles are used to train the neural network for geo-located MODIS cloud inputs. The Bayesian neural network design is used for training which has an efficient method for minimising the cost function while not over-fitting the model. The training results show good performance based on the linear regression between the targets and outputs for both SEA and WA regions.

The MODIS latent heat maps are used for developing gravity wave-latent heat models in two tropical regions – SEA and WA. The aim is to develop a model similar to GW-RAMS model in SEA region using MODIS is to compare and validate the performance of the MODIS based model, and the aim to develop a model for WA region is to test the performance of this approach in a relatively unexplored region that has unique geophysical features. The model developed in the SEA region has a similar order of coefficients with differences values because of the satellite estimates' consistent underestimation of latent heat. However, due to this difference, the GW-Sat-SEA model has better performance than GW-RAMS model when compared with the radiosonde results with similar or lower mean and standard deviation of % errors. The model developed in the WA region also has similar trendlines with the same order of coefficients. However, the validation results show the larger mean and standard deviation of percentage errors compared to SEA region. One of the reasons for may be the performance of RAMS simulations in this region which affects the accuracy of CPR based latent heat estimations. The ground-based coverage in this region is

quite sparse which may lead to inaccurate ECMWF model profiles used for RAMS initiation. Since there are no ground-based instruments which can directly measure atmospheric latent heat, it is difficult with the current state-of-art to derive concrete conclusions. However, it is certain that the results may be improved by increasing the density of spatial coverage which will eliminate the interpolation errors which is evident in the SEA region results.

Developing the database for BMC-based latent heat estimation for CPR profiles and using neural network for training collocated MODIS and CPR pixels for latent heat estimation are one-time procedures. Using these, the time consumption for the estimation of column integrated latent heat for a single thunderstorm cell is reduced to less than a second. However, it is also to be noted that this method is limited to the instance of satellite orbital passage over the thunderstorm, unlike the RAMS simulations which can provide the complete process of growth and dissipation of a thunderstorm event.

This chapter demonstrates the possibilities and limitations of satellite-based latent heat monitoring and analysing the impact of it in the atmosphere. The neural network seems to provide a good bridge which can be used for modelling the variables with uncertain relationships. Further work is needed to improve the performance of gravity wave models developed in both SEA and WA region.

6. Conclusions and future work

Gravity waves are responsible for spontaneous energy and momentum distribution in the atmosphere. In the tropical region, one of the most common sources of gravity waves is thunderstorms which occur on a diurnal basis in most of the regions [11]. Thunderstorms are a major source of gravity waves including all three types of mechanisms described in section 2.2.2; however, latent heat forcing seems to play a significant role in that. Previous model-based studies have highlighted the large perturbations created in the atmosphere due to thunderstorms [15], [16], [19], [21], [55]. There are numerous observational studies which have showed the basic characteristics of the gravity waves generated by thunderstorms but have not developed any quantified relationship between the variables involved using any long-term data for developing generalised relationships [10], [45], [46], [52], [74], [104].

In this study, a generalised model has been developed between the latent heat of thunderstorms and the gravity waves generated by them based on a 10-year period of radiosonde data, analysed for gravity waves in the lower stratosphere. In order to cover the South-East Asia tropical region, a multidisciplinary approach was used to calculate and analyse the data. The GROGRAT ray tracing method is used to locate the source and RAMS-CRM is used for simulating the

thunderstorms at the identified source location. A clear trend shows the increase in gravity wave velocity perturbations and amplitudes with an increase in integrated latent heat of the thunderstorms. Although the waves seen in the data is a combination of thermal forcing, obstacle effect and the mechanical oscillator, it is possible to analyse them further using absolute latent heat and rate of change of latent heat values. Statistics in this study shows that thermal forcing is the dominant source as more than 50% of the waves seem to be associated with it. It is also seen that the phase velocities of waves from the obstacle effect and the mechanical oscillator effects are higher velocities compared to waves from large thermal forcing. The model has been validated using the Singapore radiosonde data which shows reasonably good match with the mean error of less than 24% and standard deviation of error less than 20%. This kind of study is computationally expensive and takes several months to obtain a decent set of data for analysis due to the running time of CRMs. Therefore, methods are required which can deliver latent heat on a large scale with minimal processing on a regular basis for continual studies similar to this.

In order to obtain latent heat estimation from MODIS, a method is first developed to estimation latent heat vertical profile using CPR reflectivity profile. A probabilistic approach using BMC method is used to retrieve the latent heat profiles of thunderstorms based on the database developed using RAMS-CRM. The peak retrieval error is approximately 15% in the higher altitudes which may be mainly due to the low vertical resolution of RAMS at higher altitudes. The comparison of column-integrated values of estimation and true values are done with RAMS and WRF models which show that the proposed method consistently underestimates the latent heat with a gradient of 0.78 and 0.75. The consistent underestimation may be related to the inaccurate representation of attenuation of the W-band radar reflectivity profiles during heavy precipitation. However, this is the current limitation of the algorithm given the existing data. The results from this study showed that the latent heat can be estimated using radar reflectivity profiles with reasonable accuracy.

The MODIS cloud inputs collocated with CPR based latent heat values are trained using the Bayesian neural network in order to extend the CPR-latent heat

estimation to MODIS. The training results show a very good match between the network output and the training target data. The MODIS-based latent heat estimates are used for comparing and analysing the performance by developing gravity wave-latent heat relationships using satellite-based latent heat maps. The satellite-based models are developed in SEA and WA region. The SEA region model showed better performance compared to WA region. The main reason may be due to the lower ground-based coverage in the WA region, which may lead to larger interpolation errors in the ECMWF data; hence leading to erroneous estimations of latent heat profiles using CPR which is dependent on ECMWF driven RAMS simulations.

One of the reasons to develop the satellite-based methods to estimate latent heat is to reduce the time and computation consumption in simulating the RAMS at every thunderstorm event. Once the database required for BMC method is constructed, and the neural network is trained using CPR-based latent heat estimations, the time taken to obtain the latent heat of individual thunderstorm cells for a given MODIS image is less than 1 *second*. It is to be noted that the BMC database and neural network training can be updated with newer profiles to keep it updated with every-changing weather details using RAMS. This doesn't need to be done very frequently, it is only necessary when there is a prior knowledge that the weather and climate have changed significantly since the development of the database. However, it is also to be noted that this method is limited to the instance of satellite orbital passage over the thunderstorm, unlike the RAMS simulations which can provide the complete process of growth and dissipation of a thunderstorm event. Overall, the results here have shown that the gravity wave quantifications using long-term statistics are possible using the multi-disciplinary approaches, which are very intuitive with relatively low errors for further use.

After completing this study, there are few key points that are identified which might improve and extend this study in future. 1) In this study, the direction of perturbation is not modelled which may be complexly dependent on other variables than latent heat. This might be very challenging to perform without using the dynamic models. Therefore, studying the coupling using a multi-scale

modelling is one of the approaches which can help realise these relationships. 2) It is also interesting to determine the impact of latent cooling which is also a spontaneous effect of phase change but in the opposite direction. 3) This study was carried out in a physical perspective and much of the chemistry involved is neglected because that is beyond the scope of the current availability of instrument and data for analysis.

The future extension of this study also involves improving the performance of satellite-based latent heat estimations. The thunderstorm database developed with RAMS needs to be improved by studying the effects of different microphysical schemes, scattering model conversion errors and different averaging methods. These atmospheric couplings are complex and the current state of equipment cannot measure many variables which always create the gap in the knowledge. In the future, these issues have to be addressed in order to build a better and complete atmospheric model.

Publications

Journals

K Naren Athreyas, E Gunawan, and B. K. Tay, "Inertia-gravity waves generated by thunderstorms: South-East Asia tropical region study", *Advances in Space Research*, Available online on 21 February 2019.

K Naren Athreyas, E Gunawan, and B. K. Tay, "Estimation of latent heat profiles of thunderstorms using CloudSat radar", *Meteorological Applications*, (under review).

K. Naren Athreyas, E. Gunawan, and B. K. Tay, "Source tracing of thunderstorm generated inertia-gravity waves observed during the RADAGAST campaign in Niamey, Niger," *J. Atmos. Solar-Terrestrial Phys.*, vol. 172, no. January, pp. 1–9, 2018.

Conferences

K Narenathreyas, E Gunawan, and B. K. Tay, "Estimation of latent heat profiles of deep convective clouds using CloudSat", *IEEE Radar Conf.* 2018.

References

- [1] R. Stull, *Meteorology for Scientists and Engineers*, 3rd ed. Canada: University of British Columbia, 2011.
- [2] W. W. Vaughan, *Standard Atmosphere*, Second Edi., vol. 5, no. 003. Elsevier, 2003.
- [3] J. Holton, *An Introduction to Dynamic Meteorology*, 4th ed. Elsevier Inc. 2004.
- [4] D. Andrews, J. Holton, and C. Leovy, *Middle Atmosphere Dynamics*, Academic Press, 1987.
- [5] B. R. Sutherland, *Internal Gravity Waves*. 2010.
- [6] E. Yiğit and A. S. Medvedev, "Internal wave coupling processes in Earth's atmosphere," *Adv. Sp. Res.*, vol. 55, no. 4, pp. 983–1003, Feb. 2015.
- [7] A. L. R. Retiree, *Classification of Clouds*, Second Edi., vol. 2. Elsevier, 2015.
- [8] E. Yiğit, P. Koucká Knížová, K. Georgieva, and W. Ward, "A review of vertical coupling in the Atmosphere–Ionosphere system: Effects of waves, sudden

stratospheric warmings, space weather, and of solar activity," *J. Atmos. Solar-Terrestrial Phys.*, vol. 141, pp. 1–12, Apr. 2016.

- [9] "ARISE - Atmospheric dynamics," *ARISE reports*. [Online]. Available: <http://arise-project.eu/atmospheric-dynamics.php>.
- [10] R. D. Hunsucker, "The sources of gravity waves," *Nature*, vol. 328, no. 6127, pp. 204–205, 1987.
- [11] T. P. Lane, *Gravity Waves: Convectively Generated Gravity Waves*, Second Edi., vol. 3. Elsevier, 2015.
- [12] L.T. Matveev, *Cloud Dynamics*. .
- [13] M. Twain and H. Finn, "Mesoscale Convective Systems," in *Cloud Dynamics*, vol. 104, 2014.
- [14] I.-S. Song, H.-Y. Chun, R. R. Garcia, and B. a. Boville, "Momentum Flux Spectrum of Convectively Forced Internal Gravity Waves and Its Application to Gravity Wave Drag Parameterization. Part II: Impacts in a GCM (WACCM)," *J. Atmos. Sci.*, vol. 64, no. 7, pp. 2286–2308, 2007.
- [15] E.-H. Choi and H.-Y. Chun, "Generation mechanisms of convectively induced internal gravity waves in a three-dimensional framework," *Asia-Pacific J. Atmos. Sci.*, vol. 50, no. 2, pp. 163–177, 2014.
- [16] I.-S. Song, H.-Y. Chun, and T. P. Lane, "Generation Mechanisms of Convectively Forced Internal Gravity Waves and Their Propagation to the Stratosphere," *J. Atmos. Sci.*, vol. 60, no. 16, pp. 1960–1980, 2003.
- [17] T. L. Clark, T. Hauf, and J. P. Kuettner, "Convectively forced internal gravity waves: Results from two-dimensional numerical experiments," *Q. J. R. Meteorol. Soc.*, vol. 112, no. 474, pp. 899–925, 1986.
- [18] B. Khouider and Y. Han, "Simulation of convectively coupled waves using WRF: A framework for assessing the effects of mesoscales on synoptic scales," *Theor. Comput. Fluid Dyn.*, vol. 27, no. 3–4, pp. 473–489, 2013.
- [19] J. Yano and T. P. Lane, "Convectively generated gravity waves simulated by NAM-SCA," *J. Geophys. Res. Atmos.*, vol. 119, 2014.
- [20] Y. Lin and F. Zhang, "Tracking gravity waves in baroclinic jet-front systems," *J. Adv. Model. Earth Syst.*, vol. 7, no. 1, pp. 67–91, 2015.
- [21] J. H. Beres, "Gravity Wave Generation by a Three-Dimensional Thermal

- Forcing," *J. Atmos. Sci.*, vol. 61, no. 14, pp. 1805–1815, 2004.
- [22] C. S. . P. K. S. Bretherton, "Gravity waves, compensating subsidence, and detrainment around cumulus clouds," *J. Atmos. Sci.*, vol. 46, pp. 740–759, 1989.
 - [23] T. P. Lane and M. W. Moncrieff, "Stratospheric Gravity Waves Generated by Multiscale Tropical Convection," *J. Atmos. Sci.*, vol. 65, no. 8, pp. 2598–2614, Aug. 2008.
 - [24] C. McLandress, M. J. Alexander, and D. L. Wu, "Microwave limb sounder observations of gravity waves in the stratosphere: a climatology and interpretation," *J. Geophys. Res.*, vol. 105, no. 11, 2000.
 - [25] D. C. D. C. D. C. Fritts and M. J. Alexander, "Gravity Wave Dynamics and Effects in the Middle Atmosphere," *Rev. Geophys.*, vol. 41, no. 1, pp. 1–64, 2003.
 - [26] J. H. Beres, M. J. Alexander, and J. R. Holton, "A Method of Specifying the Gravity Wave Spectrum above Convection Based on Latent Heating Properties and Background Wind," *J. Atmos. Sci.*, vol. 61, pp. 324–337, 2004.
 - [27] E. L. Nelson and T. S. L'Ecuyer, "Toward an Algorithm for Estimating Latent Heat Release in Warm Rain Systems," *J. Atmos. Ocean. Technol.*, vol. 33, pp. 1309–1330, 2016.
 - [28] W. Tao, "Cloud-System Resolving Modeling and Aerosols," *Encyclopedia of Atmospheric Sciences 2nd Edition*, vol. 4. Elsevier, pp. 222–231, 2015.
 - [29] W.-K. Tao *et al.*, "TRMM Latent Heating Retrieval: Applications and Comparisons with Field Campaigns and Large-Scale Analyses," *Meteorol. Monogr. - Multiscale Convect. Syst. Trop. A Tribut. to Dr. Michio Yanai*, no. Code 612, pp. 1–34, 2016.
 - [30] A. Battaglia, S. Tanelli, S. Kobayashi, D. Zrnic, R. J. Hogan, and C. Simmer, "Multiple-scattering in radar systems: A review," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 111, no. 6, pp. 917–947, 2010.
 - [31] W. Tao, J. Chen, Z. Li, C. Wang, and C. Zhang, "Impact of Aerosols on Convective Clouds and Precipitation," *Rev. Geophys.*, vol. 50, no. 11, 2012.
 - [32] G. P. Brasseur and S. Solomon, *Aeronomy of the Middle Atmosphere*. Springer, 2013.
 - [33] E. V Appleton, "Fine structure of ionosphere," *Nature*, vol. 224, 1969.
 - [34] M. K. Politovich, *Encyclopedia of Atmospheric Sciences*, Second Edi., vol. 1.

Elsevier, 2015.

- [35] R. Plougonven and F. Zhang, "Internal gravity waves from atmospheric jets and fronts," *Rev. Geophys.*, vol. 52, pp. 33–76, 2013.
- [36] T. P. Lane and F. Zhang, "Coupling between Gravity Waves and Tropical Convection at Mesoscales," *J. Atmos. Sci.*, vol. 68, no. 11, pp. 2582–2598, Nov. 2011.
- [37] S. S. Lee, W. K. Tao, and C. H. Jung, "Aerosol effects on instability, circulations, clouds, and precipitation," *Adv. Meteorol.*, vol. 2014, 2014.
- [38] U. Lohmann and J. Feichter, "Global indirect aerosol effects: a review," *Atmos. Chem. Phys. Discuss.*, vol. 4, no. 6, pp. 7561–7614, 2004.
- [39] Z. Li, F. Niu, J. Fan, Y. Liu, D. Rosenfeld, and Y. Ding, "Long-term impacts of aerosols on the vertical development of clouds and precipitation," *Nat. Geosci.*, vol. 4, no. 12, pp. 888–894, 2011.
- [40] B. A. Albrecht, "Aerosols , Cloud Microphysics , and Fractional Cloudiness," *Science (80-.)*, vol. 245, no. 4923, pp. 1227–1230, 1989.
- [41] D. Rosenfeld *et al.*, "Global observations of aerosol-cloud-precipitation-climate interactions," *Reviews of Geophysics*, vol. 52, no. 4. 2014.
- [42] D. C. Fritts, "Gravity Waves: Overview," *Encycl. Atmos. Sci. Second Ed.*, vol. 3, pp. 141–152, 2014.
- [43] M. Pramitha, M. Venkat Ratnam, A. Taori, B. V. Krishna Murthy, D. Pallamraju, and S. Vijaya Bhaskar Rao, "Evidence for tropospheric wind shear excitation of high-phase-speed gravity waves reaching the mesosphere using the ray-tracing technique," *Atmos. Chem. Phys.*, vol. 15, no. 5, pp. 2709–2721, 2015.
- [44] M. O. Ki and H. Y. Chun, "Characteristics and sources of inertia-gravity waves revealed in the KEOP-2007 radiosonde data," *Asia-Pacific J. Atmos. Sci.*, vol. 46, no. 3, pp. 261–277, 2010.
- [45] M. J. Curry and R. C. Murty, "Thunderstorm-Generated Gravity Waves," *Journal of the Atmospheric Sciences*, vol. 31, no. 5. pp. 1402–1408, 1974.
- [46] S. Evan and M. J. Alexander, "Intermediate-scale tropical inertia gravity waves observed during the TWP-ICE campaign," *J. Geophys. Res. Atmos.*, vol. 113, no. 14, 2008.

- [47] M. C. N. Hankinson, M. J. Reeder, and T. P. Lane, "Gravity waves generated by convection during TWP-ICE: 2. High-frequency gravity waves," *J. Geophys. Res. Atmos.*, vol. 119, no. 9, pp. 5257–5268, May 2014.
- [48] M. C. N. Hankinson, M. J. Reeder, and T. P. Lane, "Gravity waves generated by convection during TWP-ICE: I. Inertia-gravity waves," *J. Geophys. Res. Atmos.*, vol. 119, no. 9, pp. 5269–5282, May 2014.
- [49] M. A. Geller and J. Gong, "Gravity wave kinetic, potential, and vertical fluctuation energies as indicators of different frequency gravity waves," *J. Geophys. Res. Atmos.*, vol. 115, no. 11, pp. 1–8, 2010.
- [50] K. Naren Athreyas, E. Gunawan, and B. K. Tay, "Source tracing of thunderstorm generated inertia-gravity waves observed during the RADAGAST campaign in Niamey, Niger," *J. Atmos. Solar-Terrestrial Phys.*, vol. 172, no. January, pp. 1–9, 2018.
- [51] E. Blanc, T. Farges, A. Le Pichon, and P. Heinrich, "Ten year observations of gravity waves from thunderstorms in western Africa," *J. Geophys. Res. Atmos.*, vol. 119, no. 11, pp. 6409–6418, 2014.
- [52] P. Kafando, F. Chane-Ming, and M. Petitdidier, "Stratospheric variability of wave activity and parameters in equatorial coastal and tropical sites during the West African monsoon," *Clim. Dyn.*, vol. 47, no. 11, pp. 3433–3456, 2016.
- [53] Y. . Kim, S. D. Eckermann, and H.-Y. Chun, "An overview of the past, present and future of gravity-wave drag parametrization for numerical climate and weather prediction models," *Atmosphere-Ocean*, vol. 41, 2003.
- [54] F. Dias and C. Kharif, "Nonlinear gravity and capillary-gravity waves," *Annu. Rev. Fluid Mech.*, vol. 31, no. 1, pp. 301–346, 1999.
- [55] T. L. Clarke, T. Hauf, and J. P. Kuettner, "Convectively forced internal gravity waves: results from two-dimensional numerical experiments," *Q. J. R. Meteorol. Soc. J. R. Meteorol. Soc.*, vol. 112, 1986.
- [56] R. Plougonven, A. Hertzog, and M. J. Alexander, "Case studies of nonorographic gravity waves over the Southern Ocean emphasize the role of moisture," *J. Geophys. Res. Atmos.*, vol. 120, no. 4, pp. 1278–1299, 2015.
- [57] D. C. Fritts, "A review of gravity wave saturation processes, effects, and variability in the middle atmosphere," *Pure and Applied Geophysics PAGEOPH*, vol. 130, no. 2–3, pp. 343–371, 1989.
- [58] G. D. Nastrom and D. C. Fritts, "Sources of Mesoscale Variability of Gravity Waves. Part I: Topographic Excitation," *Journal of the Atmospheric Sciences*, vol.

49. pp. 101–110, 1992.

- [59] S. D. Eckermann and R. A. Vincent, “Falling sphere observations of anisotropic gravity wave motions in the upper stratosphere over Australia,” *Pure Appl. Geophys. PAGEOPH*, vol. 130, no. 2–3, pp. 509–532, 1989.
- [60] J. Wei and F. Zhang, “Tracking gravity waves in moist baroclinic jet-front systems,” *J. Adv. Model. Earth Syst.*, pp. 67–92, 2014.
- [61] J. Wei and F. Zhang, “Mesoscale Gravity Waves in Moist Baroclinic Jet–Front Systems,” *J. Atmos. Sci.*, vol. 71, no. 3, pp. 929–952, 2014.
- [62] A. Dörnbrack, M. Leutbecher, R. Kivi, and E. Kyrö, “Mountain-wave-induced record low stratospheric temperatures above northern Scandinavia,” in *Tellus, Series A: Dynamic Meteorology and Oceanography*, 1999, vol. 51, no. 5, pp. 951–963.
- [63] D. K. Lilly and P. J. Kennedy, “Observations of a Stationary Mountain Wave and its Associated Momentum Flux and Energy Dissipation,” *Journal of the Atmospheric Sciences*, vol. 30, no. September. pp. 1135–1152, 1973.
- [64] D. R. Durran and J. B. Klemp, “Another Look at Downslope Winds. Part II: Nonlinear Amplification beneath Wave-Overturning Layers,” *Journal of the Atmospheric Sciences*, vol. 44, no. 22. pp. 3402–3412, 1987.
- [65] D. J. Karoly, G. L. Roff, and M. J. Reeder, “Gravity wave activity associated with tropical convection detected in TOGA COARE sounding data,” *Geophys. Res. Lett.*, vol. 23, no. 3, pp. 261–264, 1996.
- [66] R. a. Vincent *et al.*, “Gravity wave generation by convection and momentum deposition in the mesosphere-lower thermosphere,” *J. Geophys. Res. Atmos.*, vol. 118, no. 12, pp. 6233–6245, Jun. 2013.
- [67] C. Stephan, M. J. Alexander, and J. H. Richter, “Characteristics of Gravity Waves from Convection and Implications for Their Parameterization in Global Circulation Models,” *J. Atmos. Sci.*, vol. 73, no. 7, pp. 2729–2742, 2016.
- [68] D. C. Fritts, “Shear Excitation of Atmospheric Gravity Waves,” *J. Atmos. Sci.*, vol. 39, no. 9, pp. 1936–1952, 1982.
- [69] O. Bühler and M. E. McIntyre, “On Shear-Generated Gravity Waves that Reach the Mesosphere. Part II: Wave Propagation,” *J. Atmos. Sci.*, vol. 56, pp. 3764–3773, 1999.
- [70] R. Plougonven and F. Zhang, *Gravity Waves Excited by Jets and Fronts*, Second

Edi., vol. 3. Elsevier, 2015.

- [71] S. Eswaraiah *et al.*, "A case study of convectively generated gravity waves coupling of the lower atmosphere and mesosphere-lower thermosphere (MLT) over the tropical region: An observational evidence," *J. Atmos. Solar-Terrestrial Phys.*, vol. 169, no. January, pp. 45–51, 2018.
- [72] J. H. Seinfeld and S. N. Pandis, *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. John Wiley & Sons, Inc., 2006.
- [73] T. P. Lane, M. J. Reeder, and T. L. Clark, "Numerical modeling of gravity wave generation by deep tropical convection," *J. Atmos. Sci.*, vol. 58, 2001.
- [74] J. Gong and M. A. Geller, "Vertical fluctuation energy in United States high vertical resolution radiosonde data as an indicator of convective gravity wave sources," *J. Geophys. Res. Atmos.*, vol. 115, no. 11, pp. 1–16, 2010.
- [75] W. F. Dabberdt and H. Turtianen, "Radiosondes," *Encyclopedia of Atmospheric Sciences 2nd Edition*, vol. 5. 2015.
- [76] G. Dutta, P. V. Kumar, and S. Mohammad, "Retrieving characteristics of inertia gravity wave parameters with least uncertainties using the hodograph method," *Atmos. Chem. Phys.*, vol. 17, no. 23, pp. 14811–14819, 2017.
- [77] M. Pramitha, M. Venkat Ratnam, P. P. Leena, B. V. Krishna Murthy, and S. Vijaya Bhaskar Rao, "Identification of inertia gravity wave sources observed in the troposphere and the lower stratosphere over a tropical station Gadanki," *Atmos. Res.*, vol. 176–177, pp. 202–211, Jul. 2016.
- [78] G. Kaiser and L. H. Hudgins, "A Friendly Guide to Wavelets," *Phys. Today*, 1995.
- [79] C. Torrence and G. P. Compo, "A Practical Guide to Wavelet Analysis," *Bull. Am. Meteorol. Soc.*, vol. 79, no. 1, pp. 61–78, 1998.
- [80] Z. Can, Z. Aslan, O. Oguz, a H. Siddiqi, and S. Arabia, "Wavelet transforms of meteorological parameters and gravity waves," *Ann. Geophys.*, vol. 23, pp. 659–663, 2005.
- [81] B. K. Woods and R. B. Smith, "Energy Flux and Wavelet Diagnostics of Secondary Mountain Waves," *J. Atmos. Sci.*, vol. 67, no. 11, pp. 3721–3738, 2010.
- [82] S. D. Eckermann, "Hodographic analysis of gravity waves: Relationships among Stokes parameters, rotary spectra and cross-spectral methods," *J.*

Geophys. Res., vol. 101, pp. 70–86, 1996.

- [83] E. Becker, *Dynamical control of the middle atmosphere*, vol. 168, no. 1–4. 2012.
- [84] C. J. Marks and S. D. Eckermann, “A Three-Dimensional Nonhydrostatic Ray-Tracing Model for Gravity Waves: Formulation and Preliminary Results for the Middle Atmosphere,” *Journal of the Atmospheric Sciences*, vol. 52, no. 11, pp. 1959–1984, 1995.
- [85] S. D. Eckermann and C. J. Marks, “GROGRAT: A new model of the global propagation and dissipation of atmospheric gravity waves,” *Adv. Sp. Res.*, vol. 20, no. 6, pp. 1253–1256, 1997.
- [86] D. J. Seidel and W. J. Randel, “Variability and trends in the global tropopause estimated from radiosonde data,” *J. Geophys. Res.*, vol. 111, no. April, pp. 1–17, 2006.
- [87] G. Hong, G. Heygster, J. Miao, and K. Kunzi, “Detection of tropical deep convective clouds from AMSU-B water vapor channels measurements,” *J. Geophys. Res. D Atmos.*, vol. 110, no. 5, pp. 1–15, 2005.
- [88] Y. Zhang *et al.*, “On the diurnal cycle of deep convection, high-level cloud, and upper troposphere water vapor in the Multiscale Modeling Framework,” *J. Geophys. Res. Atmos.*, vol. 113, no. 16, pp. 1–19, 2008.
- [89] W. Tao, “Cloud Resolving Modeling,” *J. Meteorol. Soc. Japan*, vol. 85, pp. 305–330, 2007.
- [90] J. Cotton and R. a. Pielke, “The RAMS modelling system: Technical Description,” *RAMS Tech. Man.*, 2000.
- [91] O. Altaratz, T. Reisin, and Z. Levin, “Simulation of the electrification of winter thunderclouds using the three-dimensional Regional Atmospheric Modeling System (RAMS) model: Single cloud simulations,” *J. Geophys. Res. D Atmos.*, vol. 110, no. 20, pp. 1–12, 2005.
- [92] S. C. van den Heever, G. G. Carrió, W. R. Cotton, P. J. DeMott, and A. J. Prenni, “Impacts of Nucleating Aerosol on Florida Storms. Part I: Mesoscale Simulations,” *J. Atmos. Sci.*, vol. 63, no. 7, pp. 1752–1775, 2006.
- [93] S. C. van den Heever and W. R. Cotton, “Urban aerosol impacts on downwind convective storms,” *J. Appl. Meteorol. Climatol.*, vol. 46, no. 6, pp. 828–850, 2007.
- [94] S. R. Herbener, S. C. van den Heever, G. G. Carrió, S. M. Saleeby, and W. R. Cotton, “Aerosol Indirect Effects on Idealized Tropical Cyclone Dynamics,” *J.*

Atmos. Sci., vol. 71, no. 6, pp. 2040–2055, 2014.

- [95] J. B. Klemp and R. B. Wilhelmson, "The simulation of three-dimensional convective storm dynamics," *J. Atmos. Sci.*, no. 35, 1978.
- [96] H. L. Kuo, "Further studies on the parameterization of the influence of cumulus convection on large-scale flow," *J. Atmos. Sci.*, no. 31, pp. 1232–1240, 1974.
- [97] M. P. Meyers, P. J. DeMott, and W. R. Cotton, "New RAMS cloud microphysics parameterization. Part II: The two-moment scheme," *Atmos. Res.*, no. 45, pp. 3–39, 1997.
- [98] D. K. Lilly, "On the numerical simulation of buoyant convection," *Tellus*, vol. XIV, no. 2, pp. 148–172, 1962.
- [99] G. E. Hill, "Factors controlling the size and spacing of cumulus clouds as revealed by numerical experiments," *J. Atmos. Sci.*, no. 31, p. 646, 1974.
- [100] J. Y. Harrington, "The effects of radiative and microphysical processes on simulated warm and transition season Arctic stratus," Colorado State University, 1997.
- [101] M. L. Salby and R. R. Garcia, "Transient response to localized episodic heating in the tropics. Part I: excitation and short-time near field behavior," *J. Atmos. Sci.*, vol. 44, no. 2, pp. 458–498, 1987.
- [102] J.-H. Jung and A. Arakawa, "Modeling moist-convective atmosphere with a Quasi-3-D Multiscale Modeling Framework (Q3DMMF)," *J. Adv. Model. Earth Syst.*, vol. 6, pp. 513–526, 2014.
- [103] W. K. Tao and M. W. Moncrieff, "Multiscale cloud system modeling," *Rev. Geophys.*, no. 2008, pp. 1–41, 2009.
- [104] C. J. Davis and C. G. Johnson, "Lightning-induced intensification of the ionospheric sporadic E layer," *Nature*, vol. 435, no. 7043, pp. 799–801, 2005.
- [105] A. Savtchenko, R. Kummerow, P. Smith, A. Gopalan, S. Kempler, and G. Leptoukh, "A-Train Data Depot: Bringing Atmospheric Measurements Together," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2788–2795, Oct. 2008.
- [106] D. L. Hartmann, H. H. Hendon, and R. A. Houze, "Some implications of the mesoscale circulations in tropical cloud clusters for large-scale dynamics and climate," *J. Atmos. Sci.*, vol. 41, 1984.

- [107] R. M. Thompson, S. W. Payne, E. E. Recker, and R. J. Reed, "Structure and properties of synoptic-scale wave disturbances in the Intertropical Convergence Zone of the eastern Atlantic," *J. Atmos. Sci.*, no. 36, pp. 53–72, 1979.
- [108] C. Schumacher, M. H. Zhang, and P. E. Ciesielski, "Heating structures of TRMM field campaign," *J. Atmos. Sci.*, no. 64, pp. 2593–2610, 2007.
- [109] C. Schumacher, P. E. Ciesielski, and M. H. Zhang, "Tropical cloud heating profiles," *Mon. Weather Rev.*, no. 136, pp. 4289–4300, 2008.
- [110] S. R. Guimond and J. M. Reisner, "A latent heat retrieval and its effects on the intensity and structure change of Hurricane Guillermo (1997). Part II: Numerical simulations," *J. Atmos. Sci.*, no. 69, pp. 3128–3146, 2012.
- [111] M. B. Rosa, A. P. Filho, and P. Satyamurty, "Hydrometeor and Latent Heat Profiles of Tropical Cyclones Conson , Ivan AND Catarina Using PR/TRMM Data," *Rev. Bras. Meteorol.*, vol. 25, no. 2, pp. 156–174, 2010.
- [112] S. Shige, Y. N. Takayabu, W.-K. Tao, and D. E. Johnson, "Spectral Retrieval of Latent Heating Profiles from TRMM PR Data. Part I: Development of a Model-Based Algorithm," *J. Appl. Meteorol.*, vol. 43, no. November 1992, pp. 1095–1113, 2004.
- [113] J. Indu and D. Nagesh Kumar, "Evaluation of Precipitation Retrievals From Orbital Data Products of TRMM Over a Subtropical Basin in India," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6429–6442, Dec. 2015.
- [114] X. Wang and M. Zhang, "The coupling of mixed Rossby-gravity waves with diabatic heating during the TRMM-KWAJEX field campaign," *Geophys. Res. Lett.*, no. July, pp. 166–169, 1999.
- [115] A. M. Tompkins and A. A. Adebiyi, "Using *CloudSat* Cloud Retrievals to Differentiate Satellite-Derived Rainfall Products over West Africa," *J. Hydrometeorol.*, 2012.
- [116] S. S. Lee, "Effect of Aerosol on Circulations and Precipitation in Deep Convective Clouds," *J. Atmos. Sci.*, vol. 69, no. 6, pp. 1957–1974, 2012.
- [117] Z. Li *et al.*, "East Asian Studies of Tropospheric Aerosols and their Impact on Regional Climate (EAST-AIRC): An overview," *J. Geophys. Res. Atmos.*, vol. 116, no. 4, pp. 1–15, 2011.
- [118] a P. Khain, D. Rosenfeld, and a. Pokrovsky, "Aerosol impact on the dynamics and microphysics of deep convective clouds," *Q. J. R. Meteorol. Soc.*, vol. 131, p. 2639–2663 ST–Aerosol impact on the dynamics and, 2005.

- [119] I. Koren, L. A. Remer, O. Altaratz, J. V Martins, and A. Davidi, "Aerosol-induced changes of convective cloud anvils produce strong climate warming," *Atmos. Chem. Phys.*, vol. 10, 2010.
- [120] K. Sassen and Z. Wang, "Classifying clouds around the globe with the CloudSat radar: 1-year of results," *Geophys. Res. Lett.*, vol. 35, no. 4, pp. 1–5, 2008.
- [121] R. Marchand, G. G. Mace, T. Ackerman, and G. Stephens, "Hydrometeor detection using Cloudsat - An earth-orbiting 94-GHz cloud radar," *J. Atmos. Ocean. Technol.*, vol. 25, no. 4, pp. 519–533, 2008.
- [122] M. Stengel *et al.*, "The Clouds Climate Change Initiative: Assessment of state-of-the-art cloud property retrieval schemes applied to AVHRR heritage measurements," *Remote Sens. Environ.*, vol. 162, pp. 363–379, 2015.
- [123] J. B. Dodson, D. A. Randall, and K. Suzuki, "Comparison of observed and simulated tropical cumuliform clouds by CloudSat and NICAM," *J. Geophys. Res. Atmos.*, vol. 118, no. 4, pp. 1852–1867, 2013.
- [124] R. J. Hogan and A. Battaglia, "Fast Lidar and Radar Multiple-Scattering Models. Part II: Wide-Angle Scattering Using the Time-Dependent Two-Stream Approximation," *J. Atmos. Sci.*, vol. 65, no. 12, pp. 3636–3651, Dec. 2008.
- [125] R. J. Hogan, "Fast approximate calculation of multiply scattered lidar returns," *Appl. Opt.*, vol. 45, no. 23, p. 5984, Aug. 2006.
- [126] T. S. L'Ecuyer and G. L. Stephens, "An Estimation-Based Precipitation Retrieval Algorithm for Attenuating Radars," *J. Appl. Meteorol.*, vol. 41, no. 3, pp. 272–285, 2002.
- [127] W. S. Olson *et al.*, "Atmospheric Latent Heating Distributions in the Tropics Derived from Satellite Passive Microwave Radiometer Measurements," *J. Appl. Meteorol.*, vol. 38, no. 6, pp. 633–664, 1999.
- [128] M. Grecu, W. S. Olson, C. L. Shie, T. S. L'Ecuyer, and W. K. Tao, "Combining satellite microwave radiometer and radar observations to estimate atmospheric heating profiles," *J. Clim.*, vol. 22, no. 23, pp. 6356–6376, 2009.
- [129] A. Battaglia and S. Tanelli, "DOMUS: DOppler MUltiple-scattering simulator," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1 PART 2, pp. 442–450, 2011.
- [130] T. S. Lécuyer and G. L. Stephens, "An uncertainty model for Bayesian Monte Carlo retrieval algorithms: Application to the TRMM observing system,"

Quart. J. Roy. Meteor. Soc., vol. 128, no. 583, pp. 1713–1737, 2002.

- [131] O. Sourdeval, L. C.-Labonnote, A. J. Baran, and G. Brogniez, “A methodology for simultaneous retrieval of ice and liquid water cloud properties. Part I: Information content and case study,” *Q. J. R. Meteorol. Soc.*, vol. 141, no. 688, pp. 870–882, Apr. 2015.
- [132] J. M. Haynes *et al.*, “Rainfall retrieval over the ocean with spaceborne W-band radar,” *J. Geophys. Res. Atmos.*, vol. 114, no. 8, pp. 1–18, 2009.
- [133] F. S. Marzano, L. Roberti, S. Di Michele, A. Mugnai, and A. Tassa, “Modeling of apparent radar reflectivity due to convective clouds at attenuating wavelengths,” *Radio Sci.*, vol. 38, no. 1, 2003.
- [134] S. Tanelli *et al.*, “CloudSat’s cloud profiling radar after two years in orbit: Performance, calibration, and processing,” *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3560–3573, 2008.
- [135] C. Mitrescu, T. L’Ecuyer, J. Haynes, S. Miller, and J. Turk, “CloudSat precipitation profiling algorithm-model description,” *J. Appl. Meteorol. Climatol.*, vol. 49, no. 5, pp. 991–1003, 2010.
- [136] J. Delanoë and R. J. Hogan, “Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds,” *J. Geophys. Res.*, vol. 115, p. D00H29, Jul. 2010.
- [137] N. L. Pounder, R. J. Hogan, T. Várnai, A. Battaglia, and R. F. Cahalan, “A Variational Method to Retrieve the Extinction Profile in Liquid Clouds Using Multiple-Field-of-View Lidar,” *J. Appl. Meteorol. Climatol.*, vol. 51, no. 2, pp. 350–365, Feb. 2012.
- [138] M. Ceccaldi, J. Delanoë, R. J. Hogan, N. L. Pounder, a. Protat, and J. Pelon, “From CloudSat-CALIPSO to EarthCare: Evolution of the DARDAR cloud classification and its comparison to airborne radar-lidar observations,” *J. Geophys. Res. Atmos.*, vol. 118, no. 14, pp. 7962–7981, Jul. 2013.
- [139] D. Josset, S. Tanelli, Y. Hu, J. Pelon, and P. Zhai, “Analysis of Water Vapor Correction for CloudSat W-Band Radar,” *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 3812–3825, Jul. 2013.
- [140] N. Mathew, C. Suresh Raju, R. Renju, and T. Antony, “Distribution of Tropical Deep Convective Clouds from Megha-Tropiques SAPHIR Data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6409–6414, 2016.
- [141] Z. Yao, Z. Han, Z. Zhao, L. Lin, and X. Fan, “Synergetic use of POLDER and MODIS for multilayered cloud identification,” *Remote Sens. Environ.*, vol. 114,

no. 9, pp. 1910–1923, Sep. 2010.

- [142] Y. Liu, J. R. Key, S. A. Ackerman, G. G. Mace, and Q. Zhang, “Arctic cloud macrophysical characteristics from CloudSat and CALIPSO,” *Remote Sens. Environ.*, vol. 124, pp. 159–173, 2012.
- [143] R. T. Austin, A. J. Heymsfield, and G. L. Stephens, “Retrieval of ice cloud microphysical parameters using the CloudSat millimeter-wave radar and temperature,” *J. Geophys. Res. Atmos.*, vol. 114, no. 8, pp. 1–19, 2009.
- [144] R. J. Hogan, “Fast Lidar and Radar Multiple-Scattering Models. Part I: Small-Angle Scattering Using the Photon Variance-Covariance Method,” *J. Atmos. Sci.*, vol. 65, no. 12, pp. 3621–3635, Dec. 2008.
- [145] R. J. Hogan, R. Honeyager, J. Tyynelä, and S. Kneifel, “Calculating the millimetre-wave scattering phase function of snowflakes using the self-similar Rayleigh–Gans Approximation,” *Q. J. R. Meteorol. Soc.*, vol. 143, no. 703, pp. 834–844, 2017.
- [146] R. J. Hogan and C. D. Westbrook, “Equation for the Microwave Backscatter Cross Section of Aggregate Snowflakes Using the Self-Similar Rayleigh–Gans Approximation,” *J. Atmos. Sci.*, 2014.
- [147] G. Thompson *et al.*, “Explicitly-Coupled Cloud Physics and Radiation Parameterizations and Subsequent Evaluation in WRF High-Resolution Convective Forecasts,” *Atmos. Res.*, vol. 168, pp. 92–104, Sep. 2015.
- [148] A. P. Khain *et al.*, “Representation of microphysical processes in cloud-resolving models: Spectral (bin) microphysics versus bulk parameterization,” *Rev. Geophys.*, vol. 53, 2015.
- [149] H. Morrison, J. A. Curry, and V. I. Khvorostyanov, “A New Double-Moment Microphysics Parameterization for Application in Cloud and Climate Models. Part I: Description,” *J. Atmos. Sci.*, 2005.
- [150] V. Jewtoukoff, R. Plougonven, and A. Hertzog, “Gravity waves generated by deep tropical convection: Estimates from balloon observations and mesoscale simulations,” *J. Geophys. Res. Atmos.*, vol. 118, no. 17, pp. 9690–9707, 2013.
- [151] I. Koren, G. Dagan, and O. Altaratz, “From aerosol-limited to invigoration of warm convective clouds,” *Science*, vol. 344, no. 6188, pp. 1143–6, 2014.
- [152] R. E. Holz *et al.*, “Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and height evaluation using CALIOP,” *J. Geophys. Res.*, vol. 113, p. D00A19, Dec. 2008.

- [153] H. W. Barker, M. P. Jerg, T. Wehr, S. Kato, D. P. Donovan, and R. J. Hogan, "A 3D cloud-construction algorithm for the EarthCARE satellite mission," *Q. J. R. Meteorol. Soc.*, vol. 137, no. 657, pp. 1042–1058, Apr. 2011.
- [154] A. H. Young, J. J. Bates, and J. A. Curry, "Complementary use of passive and active remote sensing for detection of penetrating convection from CloudSat, CALIPSO, and Aqua MODIS," *J. Geophys. Res. Atmos.*, vol. 117, no. 13, pp. 1–13, 2012.
- [155] A. Heidinger, Y. Li, B. Baum, R. Holz, S. Platnick, and P. Yang, "Retrieval of Cirrus Cloud Optical Depth under Day and Night Conditions from MODIS Collection 6 Cloud Property Data," *Remote Sens.*, vol. 7, no. 6, pp. 7257–7271, Jun. 2015.
- [156] Z. Yao, Z. Han, Z. Zhao, L. Lin, and X. Fan, "Synergetic use of POLDER and MODIS for multilayered cloud identification," *Remote Sens. Environ.*, vol. 114, no. 9, pp. 1910–1923, 2010.
- [157] T. S. L'Ecuyer and J. H. Jiang, "Touring the atmosphere aboard the A-Train," *Physics Today*, no. July, pp. 36–41, 2010.
- [158] T. Yuan and Z. Li, "General macro- and microphysical properties of deep convective clouds as observed by MODIS," *J. Clim.*, vol. 23, no. 13, pp. 3457–3473, 2010.
- [159] M. King, S. Tsay, S. Platnick, M. Wang, and K.-N. Liou, "Cloud retrieval algorithms for MODIS: Optical thickness, effective particle radius, and thermodynamic phase," 1997.
- [160] S. Platnick *et al.*, "MODIS Cloud Optical Properties: User Guide for the Collection 6 Level-2 MOD06/MYD06 Product and Associated Level-3 Datasets," 2014.
- [161] J. Joiner *et al.*, "Detection of multi-layer and vertically-extended clouds using A-train sensors," *Atmos. Meas. Tech.*, vol. 3, pp. 233–247, 2010.
- [162] M. Picchiani, M. Chini, L. Merucci, S. Corradini, A. Piscini, and D. F. F., "Automatic Monitoring of Ash and Meteorological Clouds by Neural Networks," *IEEE*, pp. 4805–4808, 2015.
- [163] A. Sharma and M. K. Goyal, "Bayesian network model for monthly rainfall forecast," *2015 IEEE Int. Conf. Res. Comput. Intell. Commun. Networks*, pp. 241–246, Nov. 2015.
- [164] T. M. Gray and R. Bernartz, "Automatic volcanic ash detection from MODIS observations using a back-propagation neural network," *Atmos. Meas. Tech.*,

vol. 8, no. 12, pp. 5089–5097, Dec. 2015.

- [165] N. J. Rodríguez-fernández *et al.*, “SoilMoisture Retrieval Using Neural Networks: Application to SMOS,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 5991–6007, 2015.
- [166] R. Bhattacharya and S. S. Dey, “A multilayer fuzzy neural network approach for cloud classification,” *Indian J. Radio Sp. Phys.*, vol. 32, pp. 104–107, 2003.
- [167] S. Ghosh, R. Chatterjee, and P. Shanker, “Estimation of ash, moisture content and detection of coal lithofacies from well logs using regression and artificial neural network modelling,” *Fuel*, vol. 177, pp. 279–287, Aug. 2016.
- [168] R. M. Neal, *Bayesian Learning for Neural Networks*, vol. 118. 1996.
- [169] D. J. C. MacKay, “A Practical Bayesian Framework for Backpropagation Networks,” *Neural Comput.*, 1992.
- [170] D. J. C. MacKay, “Bayesian Interpolation,” *Neural Comput.*, 1992.
- [171] Q. Mu, A. Wu, X. Xiong, D. R. Doelling, and A. Angal, “Optimization of a Deep Convective Cloud Technique in Evaluating the Long-Term Radiometric Stability of MODIS Reflective Solar Bands,” *Remote Sens.*, no. May 2002, pp. 1–19, 2017.
- [172] P. Kafando, F. Chane-Ming, and M. Petitdidier, “Climatology of gravity wave activity during the West African Monsoon,” *Ann. Geophys.*, vol. 26, no. 12, pp. 4081–4089, 2008.
- [173] D. Troyan, “Merged sounding value-added product,” 2012.
- [174] J. R. Norris *et al.*, “Evidence for climate change in the satellite cloud record,” *Nature*, vol. 536, no. 7614, pp. 72–75, 2016.
- [175] K. B. Widener and K. Johnson, “W-band ARM Cloud Radar (WACR) Handbook,” 2006.
- [176] W. Tao, “Cloud Modeling,” *Encyclopedia of Atmospheric Sciences 2nd Edition*, vol. 2. Elsevier, pp. 112–120, 2015.

Appendix A – RAMS-CRM Validation

In this chapter, the RAMS-CRM performance and validation are discussed. Although it is only sensible to validate RAMS' performance in this study as it is being used for the first time the SEA and WA regions for thunderstorm studies. So it is better to know the expected errors in the simulation.

The main variable of focus in this study is the latent heat. However, latent heat is a quantity which is not measurable by the current state of instruments. Therefore, the validation of RAMS has to be done based on the quantities which can be measured directly and also which can be related to the latent heat. It is a known fact that the latent heat is generated from the hydrometeors undergoing phase change [31], [176]. It would be sensible to validate the hydrometeor profiles from RAMS with the measured ones. However, SEA and WA regions both do not have any available instruments which can measure the individual hydrometeor concentration or mixing ratio in the atmosphere. Since the prognostic variables are needed for validation, the radiosonde cannot be used for validation in this case. Therefore, the reflectivity profiles are validated with the satellite-based radar reflectivity profiles. The reflectivity profiles from RAMS-CRM are calculated using the approach described in chapter 4, which obtains reflectivity for 94 GHz radar. The CloudSat reflectivity profiles of identified

thunderstorms near Singapore region are collected on four different instances in 2016, other than the ones used for the database in chapter 4. Since the physical hydrometeor variables such as number concentration and size cannot be validated with the available equipment in the region, the reflectivity profile based validation should be sufficient to provide a reasonable representation of RAMS-CRM's performance.

Figure A-0-1: Mean reflectivity error profiles near Singapore region.

As seen in Figure A-0-1, the mean error in the reflectivity profile estimation varies between -5 to 20 *dBZ* at different altitudes. The reason for such large error in the higher altitudes is may be due to the lower resolution of RAMS-CRM grid. Above 5 *km* altitude, the vertical resolution of RAMS is around 1 *km* which might be inducing hydrometeor accuracy issues, but RAMS cannot handle higher resolution at those altitudes, based on the ECMWF initiation data which also has a low resolution at those altitudes. It is also to be noted that this error here contains the errors from scattering models. Therefore the RAMS model errors itself might be lower than this, which is difficult to quantify currently.

The RAMS prognostic variables such as latent heat and cloud variables are not possible to compare with any truth data in Singapore as there is no available ground-based equipment for collection of these data. Therefore, the RAMS atmospheric profiles are compared to radiosonde obtained atmospheric profiles.

Profiles of mixing ratio, temperature and potential temperature are compared to analysis mean percentage error at each altitude and linear regression plots.

Figure A-0-2: (from left to right) Profiles of mean percentage error of a) mixing ratio, b) temperature and c) potential temperature.

The Figure A-0-2 shows that temperature and potential temperature have much more accurate values compared to mixing ratio. RAMS which is driven by ECMWF profiles are interpolated profiles which are assimilated using various ground and space-based sources. The radiosonde profiles collected at Singapore also has missing values and coarse resolution of mixing ratio collections. Therefore, considering these uncontrollable errors, the maximum error of 20% in the mixing ratio considered acceptable for our study.

The linear regression plots between RAMS values and radiosonde values are shown in Figure A-0-3. This aligns with the previous results demonstrating the accuracy of temperature and the potential temperature is much higher than the mixing ratio comparisons.

Figure A-0-3: Linear fit plots of comparison between RAMS data and radiosonde data of temperature, potential temperature and mixing ratio.

For analysing the RAMS's performance in the WA region, the RAMS simulations are performed which are collocated with CPR thunderstorm events spanning the entire tropical WA region (from 20°W to 30°E). The reflectivity profiles from RAMS are calculated in the same way as described in section 4.3.3. Four different instances of CLDCLASS thunderstorms are collected during 2007 and compared with collocated simultaneous RAMS simulations.

Figure A-0-4: Mean reflectivity error profile in the tropical WA region.

As mentioned in section 5.3.2, the ECMWF data may be inaccurate in the WA region due to the lack of ground-based coverage. Since RAMS is initiated using ECMWF data, the interpolation errors have an enhanced effect in the final outcome. Therefore, as seen in Figure A-0-4, the mean reflectivity error profile in WA region has much larger error values compared to the mean reflectivity error profiles shown for the SEA region in Figure A-0-1. This issue cannot be improved significantly just by using advanced software techniques, because there will always be this lack of knowledge of the true values. Therefore, efforts need to be made to increase the ground-based coverage of the atmospheric measurements for validation purposes.

Appendix B – Justification of database samples for BMC method

In chapter 4, a database representing the thunderstorm reflectivity profiles is used for performing the BMC method based latent heat retrieval. One of the questions that come to mind, when considering the database is the number of samples which is ideal for producing optimal performance. Since there are no established methods to determine the number of samples required for any particular application, it is determined here by trial and error. The performance of the algorithm is tested for one particular event on 29th April 2016 near Singapore region detected by CloudSat. The algorithm is applied to CPR profiles of this event using databases with various sample numbers with standard deviation of 20 *dBZ* (representing the maximum reflectivity deviation). This type of analysis provides the information needed to decide upon on the number of samples required in the database.

As seen in Figure B-1, the performance increases with the number of samples but is saturating after 200,000 samples. The number of samples does have a significant effect on the algorithm's performance, but only until certain samples. It is also to be noted that, increasing the samples will also increase the processing time which is also to be analysed.

Figure B-0-1: Performance of the algorithm vs. number of samples plot. The performance indicator in this study is considered to be the gradient value between the retrieved latent heat from the algorithm and the true latent heat from RAMS.

Figure B-0-2: Time to retrieval a single profile of latent heat vs. number of samples plot.

It is seen that the time taken to process a single profile of reflectivity to obtain the latent heat profile increases exponentially with the number of samples after 10,000 samples. The time taken is considered for a Matlab program running on a computer with 64 GB Random Access Memory, with eight cores running in parallel. Based on Figure B-0-1 and Figure B-0-2, the time taken for processing is much larger after 200,000 samples with no significant improvement in the performance. Therefore, the algorithm in chapter 4 uses a database of 200,000 samples for the latent heat estimation.