

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]

NOTE TO USERS

This reproduction is the best copy available.

UMI[®]

DISSERTATION

**TOWARD ASSIMILATION OF CLOUD RADAR DATA
FOR IMPROVEMENTS IN MESOSCALE FORECASTS**

Submitted by

ANGELA BENEDETTI

Department of Atmospheric Science

In partial fulfillment of the requirements
For the Degree of Doctor of Philosophy
Colorado State University
Fort Collins, Colorado
Fall 2001

UMI Number: 3038619

UMI[®]

UMI Microform 3038619

**Copyright 2002 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.**

**ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346**

COPYRIGHT BY ANGELA BENEDETTI 2001
All Rights Reserved

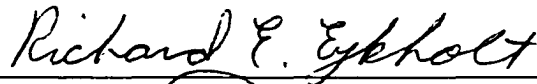
COLORADO STATE UNIVERSITY

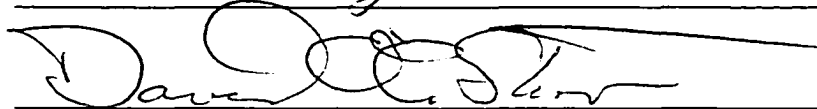
30th August 2001

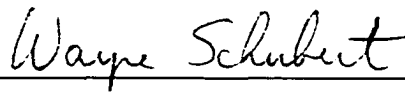
WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY ANGELA BENEDETTI ENTITLED "TOWARD ASSIMILATION OF CLOUD RADAR DATA FOR IMPROVEMENTS IN MESOSCALE FORECASTS" BE ACCEPTED AS FULFILLING PARTIAL REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

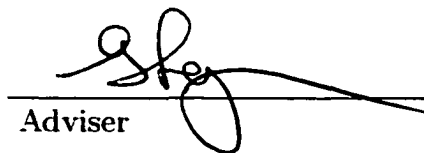
Committee on Graduate Work











Adviser



Department Head

ABSTRACT OF DISSERTATION

TOWARD ASSIMILATION OF CLOUD RADAR DATA FOR IMPROVEMENTS IN MESOSCALE FORECASTS

Unveiling the complex links between relevant physical processes at work in the Earth's atmosphere is a step toward understanding the weather and climate of Earth. Processes relating to the atmospheric branch of the hydrological cycle play a critical role in climate change through their impact on the atmospheric energy budget. Clouds, and in particular ice clouds, are an essential component of the atmospheric water cycle about which much uncertainty still exists. The paucity of observations and the inherent difficulties in modeling ice clouds both contribute to this uncertainty. Yet, the impact of this type of clouds on the atmospheric energy budgets through their influence on radiative and latent heat processes, especially in the upper-troposphere, is considerable. In recent years, much progress has been made in the description of clouds in numerical weather prediction models. Together with advancements in instrumentation and cloud observing capabilities, such as the ones afforded by short wavelength cloud radars, this has granted a better general understanding of ice cloud formation and evolution in relation to the ambient conditions. Improving data sets and models is a crucial part of a strategy to produce a more accurate representation of ice cloud processes and consequent improvements in weather forecast and climate predictions. In particular, the synergy of both these elements—models and observations—is the most promising way to address the open question about the role of ice clouds in climate change. Data assimilation offers an elegant mathematical framework to bring observations and model forecasts together. While it is no substitute for model improvement and development, it represents a powerful tool to enhance the potential of the observations through the use of the model, and to improve the model performance providing optimal model initialization through the use of measurements. In this larger picture, the focus of this research is to address some specific questions related to the combined use of model and cloud radar

observations in advancing modeling and prediction of cirrus clouds. The core of this work is the development of an assimilation system based on variational principles to perform both sensitivity studies and assimilation experiments with synthetic and real radar reflectivity measurements. To this end, an ice growth model and its adjoint are derived and used in one dimensional time-dependent variational assimilation studies. Sensitivity studies are performed and key parameters in cirrus cloud physics are identified. A more complex model—the Regional Atmospheric Modeling System— is also used in cirrus simulations and its skill assessed with the goal of using this model in four-dimensional variational experiments involving the use of radar data. Results are promising and show both the feasibility and the great potential of incorporating cloud radar observations into mesoscale models to improve cloud prediction.

Angela Benedetti
Atmospheric Science Department
Colorado State University
Fort Collins, CO 80523
Fall 2001

Alla mia famiglia

*To all the people who were in
Seattle, Quebec City, Prague and Genoa, and everywhere else
where expressing a different opinion is risking your life.*

ACKNOWLEDGMENTS

This research was financially sponsored under Department of Energy Research Grant DE-FG03-94ER61748, CloudSat NASA Contract NAS5-99237 and NASA Grants NCC5-288 and NAG5-6637.

Some of the people that supported this research in various ways are gratefully acknowledged:

Dr. D. Starr of NASA Goddard for introducing me to the world of cirrus modeling and assisting me throughout the whole length of my Ph.D.

Dr. T. Vukićević of CIRA for showing me the beauty of variational data assimilation, and supporting with enthusiasm my first attempts at it.

Dr. W. Cotton and his group at CSU, in particular Dr. R. Walko, for providing help and assistance with RAMS.

Dr. W. Schubert for transmitting to me his love for atmospheric dynamics as a teacher.

Dr. R. Eykholt for his invaluable help with revising this manuscript.

Prof. G. “Jay” Mace of University of Utah for his continuous assistance with ARM radar and in situ data.

Dr. E. Clothiaux of Penn State for first providing me with information on cloud radars.

Drs. C. Jakob and A. Beljaars of ECMWF for their help with the ECMWF model data.

Past and present components of the Stephens’ research group are also gratefully acknowledged. In particular, Sue Lini, Heather Jensen and Jillian L’Ecuyer for their assistance, Richard Engelen for his encouraging support and his friendship, Ian Wittmeyer for his help with hardware and software, Philip Gabriel for his love of research, Steve Miller for his generous help in many and diverse occasions, Tristan L’Ecuyer for numerous fruitful discussions and for helping me through the Ph.D. as a good officemate.

Last but not least, I thank Dr. G. Stephens for being an enthusiastic and supportive advisor who let me free to pursue my research interests and always stimulated my curiosity.

Special thanks to all my friends in Fort Collins that have constantly shown me their friendship and appreciation. Particularly Rich, for these four years of love and support. Un grazie di cuore a tutti i miei amici italiani per la loro costante presenza, anche da lontano, e per il loro sostegno. Un grazie affettuoso alla mia Raffaella, con la quale ho condiviso tanti momenti di scoramento e altrettanti di gioia in questa avventura americana.

Contents

ACKNOWLEDGEMENTS	v
LIST OF FIGURES	xii
LIST OF TABLES	xix
PREFACE	xx
1 Introduction	1
1.1 The global picture	1
1.1.1 State-of-the-art ice cloud modeling and observing systems . .	3
1.1.2 Bringing together observations and models	7
1.1.3 Issues in assimilation of cloud and precipitation	9
1.2 Overview of recent cloud and rainfall assimilation studies	13
1.3 On the potential of assimilation of cloud radar data	15
1.4 Outline of research	17
1.5 Thesis road map	19
2 ABC of data assimilation	21
2.1 Introduction	21
2.2 Data assimilation techniques	22
2.2.1 Nudging technique	23
2.2.2 Variational data assimilation	24
2.2.3 Link to estimation theory	28
2.3 Components of a VAR system	29
2.3.1 The dynamical model	29
2.3.2 The adjoint operator	31
2.3.3 The observational operator (mapping)	37
2.3.4 The background	38
2.4 The importance of error statistics	39
2.4.1 The modeling of background and observational errors	41
3 Cirrus cloud modeling	43
3.1 Current issues in cirrus modeling	43
3.2 Source of uncertainties in mesoscale models	49
3.3 Assessment of model skill	50
3.3.1 Description of the Cloud Resolving Model	51
3.3.2 Selected results from WG2 idealized test cases	55
3.3.3 Sensitivity studies	59
3.3.4 Concluding remarks on the sensitivity study	71

3.4	Real case simulations	72
3.4.1	The ARM-UAV Spring 1999 Experiment	72
3.4.2	Synoptic overview of April 30 cirrus case	72
3.4.3	Overview of analysis method	74
3.4.4	Synthetic radar reflectivities from model fields	75
3.4.5	Model implementation for the 3D channel-type LES runs	76
3.4.6	Results for the 3D-channel runs	77
3.5	RAMS error analysis ¹	79
3.5.1	Estimating model bias	79
3.5.2	Error Covariance Matrix	81
3.5.3	Error analysis results	82
3.6	Summary	89
4	Cloud radar data assimilation	93
4.1	Rationale for the use of a simplified model	93
4.1.1	Specific goals of this modeling effort	94
4.2	General overview of procedure	94
4.3	The microphysical module	96
4.3.1	Model equations	97
4.3.2	Model parameterizations	99
4.3.3	Numerical implementation	105
4.3.4	Observational operators	106
4.3.5	Some comparison with model and observed cirrus spectra	108
4.4	Sensitivity studies	127
4.4.1	Construction of the adjoint model	128
4.4.2	Sensitivity to model cloud-related variables	130
4.4.3	Sensitivity to environmental fields	131
4.4.4	Concluding remarks on the sensitivity study	135
4.5	Toward variational assimilation of radar data	136
4.5.1	On the choice of control variables	137
4.5.2	Discussion of radar observational and mapping error	138
4.5.3	Adding information	139
4.5.4	Including the background	139
4.5.5	Optimization procedure	140
4.6	Results of optimization using synthetic measurements	141
4.6.1	Cloud variables as control variables	143
4.6.2	Specific humidity as a control variable	153
4.6.3	Model parameters as control variables	162
4.7	Results of optimization using real measurements	163
4.7.1	Cloud variables as control variables	165
4.7.2	Specific humidity as a control variable	169
4.8	Summary and conclusions	169
5	Prelude to 4D-VAR assimilation using RAMS adjoint	174
5.1	Introduction	174
5.1.1	Variational data assimilation in regional models	175

¹This section is part of a paper entitled: “*Characterization of errors in cirrus simulations from a Cloud Resolving Model for application in Ice Water Content Retrievals*” by A. Benedetti and G.L. Stephens, accepted for publication in *Atmospheric Research*, 2001.

5.2	Requirements for VDA into limited-area models	176
5.3	RAMS adjoint construction	178
5.4	Experiment setup	179
5.5	System response to change in environmental variables	181
	5.5.1 Sensitivity results	182
5.6	System response to cloud synthetic data inclusion	187
	5.6.1 Experimental results for forced pristine ice cases	190
	5.6.2 Experimental results for forced snow cases	198
	5.6.3 Comparative discussion of pristine ice and snow results	204
5.7	Conclusion to the 4DVar prelude	212
6	Summary and concluding remarks	214
6.1	Summary of results	214
6.2	Relevance of this research to weather forecasting	219
6.3	Future work	220
	BIBLIOGRAPHY	223
	APPENDIX	
A	Radar retrievals of ice cloud properties	234
A.1	Introduction	234
A.2	Fundamentals of retrieval	234
A.3	Forward model equations	237
A.4	Application of ECM to the retrieval of cirrus cloud properties	238
	A.4.1 Use of RAMS model field as <i>a priori</i>	238
	A.4.2 IWC retrievals	239
B	More on ice growth model evaluation	245
B.1	Case study 3: Oklahoma, 26 September 1997	245
C	Properties of reflected sunlight derived from a Green's function method²	253
C.1	Abstract	253
C.2	Introduction	253
C.3	Theoretical Foundations	256
	C.3.1 Basic concepts	256
	C.3.2 Global Reflection and Transmission matrices	261
	C.3.3 Derivation of the Green's function matrix	263
C.4	Numerical Results	266
	C.4.1 Radiance calculations and comparison with the doubling method	267
	C.4.2 Green's function results: Contribution Function and Integrated Contribution Function	270
	C.4.3 Examples of reflection from vertically varying clouds	272
C.5	Further applications: the penetration optical depth	277
C.6	Summary and conclusions	280
C.7	Acknowledgments	282
C.8	Appendix I.	283

²To appear in *Journal of Quantitative Spectroscopy and Radiative Transfer*, 2001.

C.9	Appendix 2.	286
LIST OF ABBREVIATIONS	289

List of Figures

1.1	Six-year average of high cloud net radiative forcing over the Pacific Ocean. Figure is courtesy of Rich Moore.	2
1.2	The difference between OLR observed by ERBE minus that obtained by the ECMWF model for the DJF 1987/88 season. The upper panel is the difference between ERBE and the model with the diagnostic version of the cloud scheme and the lower two panels show differences between ERBE and the model with two different representations of fall speed of ice crystals. From Stephens et al. (1998).	4
1.3	Schematic of water vapor pumping, radiatively induced by cirrus clouds. From Sherwood (1999).	5
1.4	500mb temperature (upper panel), cloud occurrence quantified as millimeter-wave radar returns > -20 dBZ (central panel), and rain gauge measurements over the ARM SGP site. Data were obtained from the ARM Program sponsored by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Environmental Sciences Division. Figure courtesy of Ian Wittmeyer.	10
1.5	Synoptic data from worldwide weather stations on land and ships ingested in ECMWF model.	12
3.1	IWP (gm^{-2}) versus time for simulations of “cold” (upper panel) and “warm” (lower panel) with 16 cloud models (see text for explanation). Figure courtesy of Dr. D. Starr.	46
3.2	Warm cloud top height (upper panel), cloud base height (center panel), and cloud thickness (lower panel) given by various models. Figure courtesy of Dr.D.Starr.	47
3.3	Cold cloud top height (upper panel), cloud base height (center panel), and cloud thickness (lower panel) for the various models. Figure courtesy of Dr.D.Starr.	48
3.4	Flow diagram describing microphysical processes. From Cotton and Anthes (1989).	53
3.5	Idealized soundings for “warm” (solid line) and “cold” (dotted line) cirrus test cases. Left panel shows temperature in degree Celsius and right panel relative humidity in %. Data is courtesy of Dr. D. Starr.	56
3.6	Warm cirrus IWC snapshots at various simulation times. Time interval between plots is 1 hour, starting from the top left panel going counterclockwise.	57
3.7	Same as figure 3.6, but for the cold cirrus simulations.	58

3.8	Warm cirrus statistics. Upper left panel shows IWP (gm^{-2}) as a function of time (min); upper right panel shows vertically integrated particle number concentration ($\#\text{m}^{-2}$); left central panels shows mean zonal (solid line) and vertical (dotted line) velocities (m/s) and right panels shows root mean square zonal (solid line) and vertical (dotted line) velocities vertically averaged over cloud generation layer; bottom left panel shows environment potential temperature (K), and bottom right panel shows upward radiative flux (Wm^{-2}) at cloud top (solid line), and downward radiative flux at cloud base (dotted line).	60
3.9	Same as figure 3.8, but for "cold" cirrus.	61
3.10	IWP versus time for various model configurations (see text for explanations).	64
3.11	Cloud base at selected times for various model configurations (see text for explanations).	65
3.12	Fall speed velocities (ms^{-1}) vs IWC (gm^{-3}). x and y axis scale is logarithmic. Courtesy of P. Brown, U.K. Meteorological Office.	67
3.13	Fall speed velocities (ms^{-1}) vs size (μm) for the three RAMS ice categories pristine ice, snow, and aggregates.	68
3.14	Selection of fall speed velocities-IWC parameterizations chosen for this study (see text for explanations)	69
3.15	IWP versus time for various terminal velocities (see text for explanations).	70
3.16	Cloud base at selected times for various terminal velocities (see text for explanations).	71
3.17	GOES 10 visible images at 2200UTC on April, 30, 1999 (upper panel) and at 000UTC on May, 1, 1999 (lower panel) showing the advection of the cirrus layer over the experiment area. Images are courtesy of Steve Miller.	73
3.18	Co-located lidar and radar cirrus observation for April, 30, 1999.	74
3.19	Example of Altus II trajectory off the coast of Kauai, on April, 30 1999.	76
3.20	Observed (upper panel) versus modeled (lower panel) radar reflectivity for the 3D-channel runs.	78
3.21	Domain averaged radar reflectivity: comparison between observations and CRM results.	79
3.22	Observed (top) and model radar reflectivity (center) for one flight leg. Model bias is shown in lower panel.	84
3.23	As in figure 3.22, but for the Ice Water Content.	86
3.24	Domain average model and observed radar reflectivity.	87
3.25	As in figure 3.24, but for the domain average IWC.	87
3.26	Model IWC bias probability distribution function at selected vertical levels.	88
3.27	Profile of average model bias. Bias is expressed in IWC units (gm^{-3}).	89
3.28	Radar reflectivity Error Covariance Matrix.	90
3.29	IWC Error Covariance Matrix.	90
3.30	Profile of relative error on horizontally averaged IWC, expressed as percentage.	91

4.1	Radar reflectivity and total optical depth time series for a thick cirrus layer observed over the ARM SGP site, Oklahoma, on May 8, during the Spring 1998 Cloud IOP.	110
4.2	Relative humidity with respect to ice (upper panel), temperature (middle panel), and plane altitude (lower panel) recorded between 17:30 and 17:48 UTC on board of the UN Citation on May 8, 1998.	111
4.3	Ice particle mean size (upper panel), and number concentration (lower panel) measured between 17:30 and 17:48 UTC by the 2D-C probe on board of the UN Citation on May, 8 1998.	112
4.4	Temperature (upper panel) and relative humidity (lower panel) at 17:30 UTC on May, 8 1998. Sounding launched from the ARM-SGP central facility. Data were obtained from the ARM Program sponsored by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Environmental Sciences Division.	113
4.5	Comparison between observed (solid line) and model number concentrations and mean sizes (dashed line) averaged over a half hour interval.	114
4.6	Comparison between observed (solid line) and model reflectivities (solid line) averaged over a half hour interval.	114
4.7	Same as in figure 4.6, Decreased value of a parameter (see text for explanations).	115
4.8	Atmospheric state variable profiles between 6 and 7.5 km. Temperature is shown in the upper panel. Relative humidity (RH) in the lower panel. The various profiles represents respectively: original RH wrt water (solid line), original RH wrt ice (short dashed), increased RH wrt water (dot-dashed), and increased RH wrt ice (long dashed).	116
4.9	Comparison between observed (solid line) and model number concentrations and mean sizes (dashed line). Meyers et al. (1992) initialization for N_t and constant initialization for D_n ($100\mu m$). Increased value of RH wrt water (see text for explanations).	117
4.10	Same as figure 4.9, but for radar reflectivity.	118
4.11	Same as figure 4.9, but for initial value of $D_n = 50\mu m$	119
4.12	Same as figure 4.10, but for initial value of $D_n = 50\mu m$	119
4.13	Same as figure 4.9, but for ECMWF initialization for D_n	120
4.14	Same as figure 4.10, but for ECMWF initialization for D_n	120
4.15	Temperature (left) and relative humidity (right) at 23:20 UTC on April, 30 1999, Kauai, Hawaii.	121
4.16	Comparison between observed (solid line) and model reflectivities. Meyers et al. (1992) initialization for N_t and constant initialization for D_n ($100\mu m$).	122
4.17	Atmospheric state variable profiles between 6 and 12 km. Temperature is shown in the upper panel. Relative humidity (RH) is shown in the lower panel. The various profiles represent respectively: original RH wrt water (solid line), original RH wrt ice (short dashed), increased RH wrt water (dot-dashed), and increased RH wrt ice (long dashed).	123
4.18	Same as figure 4.16, but for the increased relative humidity profile of figure 4.17.	124
4.19	Same as figure 4.18, but for $D_n = 50\mu m$	124
4.20	Same as figure 4.18, but for ECMWF D_n initialization.	125

4.21	Same as figure 4.18, but for IWC D_n initialization.	125
4.22	Comparison of radar reflectivity fields for a cirrus case observed in Kauai, Hawaii, on April 30, during the Spring 1999 ARM-UAV campaign.	126
4.23	Sensitivity of radar reflectivity to 5% change in input parameters. In order, starting from the upper left corner: sensitivity of Z to initial characteristic diameter profile (left top panel); sensitivity to initial characteristic number concentration profile (top right panel); sensitivity to fall speed parameter (middle left panel), a ; sensitivity to fall speed parameter, b (middle right panel); sensitivity to mass parameter, α (bottom left panel); sensitivity to mass parameter, β (bottom right panel). See text for explanations.	132
4.24	Sensitivity of radar reflectivity to 0.5% change in atmospheric variables. In order, starting from the upper left corner: sensitivity of Z to initial temperature profile (left upper panel); sensitivity to initial specific humidity profile (right upper panel); sensitivity to pressure (left lower panel); sensitivity to vertical velocity (right lower panel).	134
4.25	Flowchart of the optimization procedure, beginning with the forward model computation from a first guess initial condition.	142
4.26	Characteristic diameter (top left panel), number concentration (top right panel) and synthetic reflectivity at final time (bottom left panel) as functions of iteration. The initial perturbation is -30%.	145
4.27	Same as figure 4.26, but for an amplitude perturbation of +30%.	146
4.28	Characteristic diameter (top left panel), number concentration (top right panel) and synthetic reflectivity at final time (bottom left panel) as functions of iteration. The initial profile is perturbed by 20% and misplaced by 1 km with respect to true profile.	147
4.29	Characteristic diameter (top left panel), number concentration (top right panel) and synthetic reflectivity at final time (bottom left panel) as functions of iteration. The initial perturbation is 20%. Observation plus mapping error is 1 dBZ ($\approx 23\%$).	149
4.30	Same as figure 4.29, but for an error of 5dBZ.	150
4.31	Same as figure 4.29, but with the inclusion of total optical depth information. The accuracy of the optical depth is same as the accuracy in radar reflectivity (23%).	152
4.32	Specific humidity (top panel) and synthetic reflectivity at final time (bottom panel) as functions of iteration. The initial perturbation is -20%.	155
4.33	Same as fig 4.32, but for an initial perturbation of +20%.	156
4.34	Specific humidity (top panel) and synthetic reflectivity at final time (bottom panel) as functions of iteration. Observational plus mapping error is 1 dBZ corresponding to $\approx 23\%$	157
4.35	Specific humidity (top panel) and synthetic reflectivity at final time (bottom panel) as functions of iteration. Background error is 50%, observational error is 1%.	160
4.36	Same as fig 4.35, but background error is 1%, and observational error is 50%.	161
4.37	Comparison of true synthetic reflectivity at final time (black) with reflectivity computed initial value of a (blue) and optimized value of a (red).	164

4.38	Characteristic diameter (top left panel), number concentration (top right panel) and synthetic reflectivity at final time (bottom left panel) as functions of iteration for assimilation of real measurements. Observation plus mapping error is 2 dBZ ($\approx 46\%$). Constant initialization for the characteristic diameter.	166
4.39	Same as figure 4.38, but with increased initial specific humidity. Constant initialization for the characteristic diameter.	167
4.40	Same as figure 4.38, but with increased initial specific humidity. ECMWF initialization for the characteristic diameter.	168
4.41	Specific humidity (top left panel), number concentration (top right panel) and synthetic reflectivity at final time (bottom left panel) as functions of iteration for assimilation of real measurements. Observation plus mapping error is 2 dBZ ($\approx 46\%$). Constant initialization for the characteristic diameter. Original specific humidity profile. . .	170
4.42	Same as figure 4.41, but for an increased initial specific humidity. . .	171
4.43	Same as figure 4.41, but for the SGP May 8 1998 cirrus case.	172
5.1	Domain averaged profile of IWC ($mg\ m^{-3}$). Relative contributions from pristine ice and snow are also shown.	180
5.2	Sensitivity of pristine ice mixing ratio to 1% change in ambient variables. In order, starting from the upper left corner: sensitivity of r_{ice} to specific humidity, q (top left panel); sensitivity to potential temperature, θ (top right panel); sensitivity to Exner perturbation function, Π' (middle left panel); sensitivity to vertical wind, w (middle right panel); sensitivity to zonal wind, u (bottom left panel); sensitivity to meridional wind, v (bottom right panel). See text for explanations. . .	183
5.3	Same as figure 5.2, but for radar reflectivity.	184
5.4	Same as figure 5.2, but for the snow mixing ratio.	187
5.5	Same as figure 5.3, but for radar reflectivity variations induced by perturbations in the snow mixing ratio.	188
5.6	West view three-dimensional snapshots of positive (orange) and negative (light grey) adjoint pristine ice solution computed with RAMAS for case (c.1) forcing. Isosurface value is $3 \cdot 10^7$ for positive and $-3 \cdot 10^7$ for negative fields. Units are $\frac{1}{kg/kg}$	192
5.7	Same as figure 5.6, but for $t = 30m$ (upper panel) and $t = 0$ (lower panel).	193
5.8	West view three-dimensional snapshots at $t = 30m$ (upper panel) and $t = 0$ (lower panel) of positive (dark yellow) and negative (green) specific humidity adjoint solution computed with RAMAS for case (c.1) forcing. Isosurface value is 10 for positive and -10 for negative fields. Units are $\frac{1}{kg/kg}$	195
5.9	West view three-dimensional snapshots at $t = 30m$ (upper panel) and $t = 0$ (lower panel) of positive (red) and negative (blue) potential temperature adjoint solution computed with RAMAS for case (c.1) forcing. Isosurface value is 0.1 for positive and -0.1 for negative fields. Units are $\frac{1}{K}$	196

5.10	West view three-dimensional snapshots at $t = 30m$ (upper panel) and $t = 0$ (lower panel) of positive (light violet) and negative (light blue) perturbation Exner function adjoint solution computed with RAMAS (c.1) forcing. Isosurface value is 1 for positive and -1 for negative fields. Units are $\frac{1}{J/(kgK)}$.	197
5.11	West view three-dimensional snapshots at $t = 30m$ (upper panel) and $t = 0$ (lower panel) of positive (light yellow) and negative (light green) zonal wind adjoint solution computed with RAMAS for case (c.1) forcing. Isosurface value is 0.1 for positive and -0.1 for negative fields. Units are $\frac{1}{m/s}$.	199
5.12	West view three-dimensional snapshots at $t = 30m$ (upper panel) and $t = 0$ (lower panel) of positive (light red) and negative (purple) vertical wind adjoint solution computed with RAMS adjoint for case (c.1) forcing. Isosurface value is 0.1 for positive and -0.1 for negative fields. Units are $\frac{1}{m/s}$.	200
5.13	Same as figure 5.8 for the specific humidity adjoint solution at time $t = 0$, but for a 50% mapping-observation error (case c.2).	201
5.14	Same as figure 5.9 for potential temperature adjoint solution at time $t = 0$, but for a 50% mapping+observation error (case c.2).	201
5.15	Same as figure 5.10 for perturbation Exner function adjoint solution at time $t = 0$, but for a 50% mapping+observation error (case c.2).	202
5.16	Same as figure 5.11 for zonal wind adjoint solution at time $t = 0$, but for a 50% mapping+observation error (case c.2).	202
5.17	Same as figure 5.12 for vertical wind adjoint solution at time $t = 0$, but for a 50% mapping+observation error (case c.2).	203
5.18	West view three-dimensional snapshots of positive (pink) and negative (very light yellow) snow adjoint solution computed with RAMAS for case (c.2) forcing. Isosurface value is $3 \cdot 10^7$ for positive and $-3 \cdot 10^7$ for negative fields. Units are $\frac{1}{kg/kg}$.	205
5.19	Same as figure 5.18, but for $t = 30m$ (upper panel) and $t = 0$ (lower panel).	206
5.20	Same as figure 5.8 for the specific humidity adjoint solution, but for snow mixing ratio as the forcing variable.	207
5.21	Same as figure 5.9 for the potential temperature adjoint solution, but for snow mixing ratio as the forcing variable.	208
5.22	Same as figure 5.10 for the perturbation Exner function adjoint solution, but for snow mixing ratio as the forcing variable.	209
5.23	Same as figure 5.11 for the zonal wind adjoint solution, but for snow mixing ratio as the forcing variable.	210
5.24	Same as figure 5.12 for the vertical wind adjoint solution, but for snow mixing ratio as the forcing variable.	211
A.1	Microphysical properties for April 30 cirrus, retrieved using CRM <i>a priori</i> .	241
A.2	As in figure A.1, but using constant <i>a priori</i> .	242
A.3	IWC from optimal estimation retrieval versus IWC derived from IWC-Z relationship. CRM fields were used as <i>a priori</i> .	243
A.4	As in figure A.3, but using the constant <i>a priori</i> profile.	244

B.1	Radar reflectivity and total optical depth time series for a cirrus cloud observed over the ARM SGP site, Oklahoma, on September, 26 during the Fall 1997 Cloud IOP.	246
B.2	Relative humidity with respect to ice (upper panel), temperature (middle panel), and plane altitude (lower panel) recorded between 17:30 and 17:48 UTC on board of the UN Citation on May 8, 1998.	247
B.3	Ice particle mean size (upper panel), and number concentration (lower panel) measured between 18:30 and 19:00 UTC by the 2D-probe on board of the UN Citation on September, 26 1997.	248
B.4	Atmospheric state variable profiles between 10 and 12 km. Temperature is shown in the upper panel. Relative humidity (RH) is shown in the lower panel. The various profiles represents respectively: original RH wrt water (solid line), original RH wrt ice (short dashed), increased RH wrt water (dot-dashed), and increased RH wrt ice (long dashed).	249
B.5	Comparison between observed (solid line) and model number concentrations and mean sizes (dashed line). Meyers et al. (1992) initialization for N_t and constant D_n ($100\mu m$) initialization for D_n . Increased value of RH wrt water (see text for explanations).	250
B.6	Same as figure B.5, but for radar reflectivities.	250
B.7	Same as figure B.5, but for initial value of $D_n = 50\mu m$	251
B.8	Same as figure B.6, but for initial value of $D_n = 50\mu m$	251
B.9	Same as figure B.5, but for ECMWF initialization for D_n	252
B.10	Same as figure B.6, but for ECMWF initialization for D_n	252
C.1	Vertical profiles of extinction coefficient for three hypothetical clouds (see text for details).	273
C.2	Profiles of Contribution Function and Integrated Contribution Function in the direction $\theta_1 = 8.35^\circ$ for the three cloud scenarios and two different single scattering albedoes, as labeled (upper panels, $\omega_o = 1$; lower panels, $\omega_o = 0.2$). See text for explanations. Optical parameters: $\tau = 1$, $g = 0.8$. Cosine of solar zenith angle, $\mu_\odot = 1.0$	275
C.3	As in figure C.2, except $\tau = 10$	276
C.4	As in figure C.2, except $\mu_\odot = 0.5$	277
C.5	As in figure C.2, except $\tau = 10$ and $\mu_\odot = 0.5$	278
C.6	Contour plot of penetration optical depth (τ_p) as a function of total optical depth and percentage of upwelling radiance at the upper boundary. Two single scattering albedo are shown: $\omega_o = 1$ (upper panel) and $\omega_o = 0.2$ (lower panel). The cosine of the solar zenith angle is 1.0.	279
C.7	As in figure C.6, except the cosine of solar zenith angle is 0.5.	280

List of Tables

3.1	Fall speed coefficients for RAMS ice categories (MKS units).	68
4.1	Sensitivity of cloud visible optical depth to a 5% change in input variables. Value of optical depth for unperturbed conditions is $\tau = 0.2338$	133
4.2	Sensitivity of cloud visible optical depth to a .5% change in atmospheric state variables. Value of optical depth for unperturbed conditions is $\tau = 0.01764$	136
4.3	Initial and final values of the cost function versus initial condition perturbation.	144
4.4	Initial and final values of the cost function versus magnitude of observational plus mapping errors.	148
4.5	Initial and final values of the cost function versus the magnitude of the perturbation to the specific humidity profile.	153
4.6	Initial and final values of the cost function as functions of error magnitude.	154
4.7	Value of a as a function of iteration. True value of $a = 11.7$ Initial perturbation -50%.	163
5.1	Maximum and minimum impact of 1% environmental parameter variation at cloud levels on pristine ice mixing ratio.	189
5.2	Same as table 5.5.1, but for snow mixing ratio.	189
C.1	Comparison of radiances computed with a doubling code (D) and with the global \mathcal{R} and \mathcal{T} technique presented in this study (RT). m is the index for the expansion terms of the azimuthal series. μ represents the cosine of the quadrature angles (observing directions). $I^+(\mathbf{H})$ and $I^-(0)$ represent respectively the upwelling and downwelling radiances summed up to the m -th term of the azimuth expansion series.	288

PREFACE

◇
*Sing to
me of the
man, Muse, the
man of twists and
turns driven time and
again off course, once he had
plundered the hallowed heights of
Troy. Many cities of men he saw
and learned their minds, many pains he
suffered, heartsick on the open sea, fighting
to save his life and bring his comrades
home.[...] Launch out on his
story, Muse, daughter of Zeus,
start from where you will
-sing for our time
too. Transla-
tion by R.
Fagles.*
◇

“*Odyssey*” has come to signify an adventurous and at times dangerous journey, usually marked by many changes of fortune. Another interpretation attributes to the term the meaning of intellectual or spiritual quest. It is in the spirit of both definitions that I lived this four years as a Ph.D. student at Colorado State University. In the work done during this time, I followed different paths, and explored different research alternatives, some successful, others less. Some of these research efforts are collected and presented in this dissertation in a hopefully coherent manner, others are listed as appendices since they did not quite fit with the rest, others have yet to be written, if they ever will be. The final goal of the Ph.D. journey is the same for everybody: improve oneself through knowledge. The journey itself is not the same for everybody. If you believe that it is the journey that matters, not only the final destination, then this should be the key to reading and judging this Ph.D. work.

Chapter 1

Introduction

1.1 The global picture

Understanding the climate of Earth and the way climate varies in time requires a quantitative knowledge of the way water is exchanged back and forth between its main reservoirs both within the atmosphere and at the surface. The influence of the hydrological cycle on the energy budget of Earth is central not only to understanding present climate but also to the prediction of climate change. Processes relating to the smallest of the reservoirs of water—the atmospheric branch of the hydrological cycle—play a critical role in climate change through their impact on the atmospheric energy budget. This is true both for water in vapor form, and water in condensed (liquid and ice) form.

A component of the atmospheric water cycle about which much uncertainty still exists is related to the ice phase. The paucity of observations and the inherent difficulties in modeling ice clouds both contribute to this uncertainty. Yet, the impact of this type of clouds on the atmospheric energy budgets through their influence on radiative and latent heat processes, especially in the upper-troposphere, is quite large (Liou, 1986). Figure 1.1 from Moore and Vonder Haar (2001) shows a six-year average for 1985–1990 of thin upper-tropospheric cloud net radiative forcing over the Pacific Ocean basin as computed using Earth Radiation Budget Experiment (ERBE) radiative fluxes at the top of the atmosphere (TOA) and the International Satellite Cloud Climatology Project (ISCCP) data. The Hartmann et al. (1992) regression technique is applied to isolate the high cloud contribution.

The forcing exerted by high clouds is quite large and positive between 10–25 degrees latitude in the Southern hemisphere and 5–15 in the Northern hemisphere

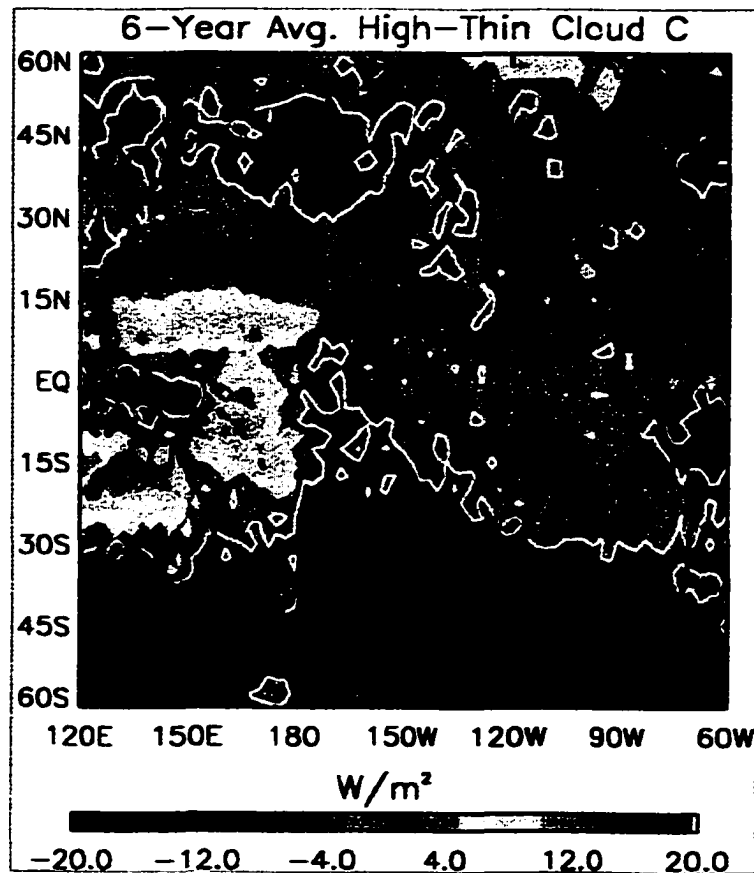


Figure 1.1: Six-year average of high cloud net radiative forcing over the Pacific Ocean. Figure is courtesy of Rich Moore.

over the Western Pacific area, both hemispheres, indicating that a heating of the atmosphere is occurring at low latitudes. This heating principally arises from cirrus clouds associated with convection (Stephens, 2001). Negative forcing is observed at high latitudes, especially over the oceans.

The vertical distribution of radiative heating between the surface and the atmosphere, which depends on location and amount of cloud condensate as shown in Stephens and Webster (1981), has a large impact on energy transport and atmospheric circulation (Randall et al., 1989).

Stephens et al. (1998) show how outgoing longwave radiation (OLR) at the top

of the atmosphere in the European Center for Medium-range Weather Forecast (ECMWF) model is influenced by the amount of ice and its vertical distribution. Figure 1.2 shows the difference between ERBE and ECMWF OLR global distribution averaged over the DJF 1987/1988 season for three model schemes with different treatment of ice phase. Substantial differences can be noticed from a model run with diagnostic (upper panel) and prognostic cloud scheme with two distinct fall speed parameterizations for the ice crystals (lower panels). Although the prognostic scheme provides better estimates of OLR, the degree of agreement is subject to tuning of the ice fallspeed parameterization (see also Jakob and Morcrette (1995)).

The tight link between cirrus clouds and upper tropospheric radiative and water budget is shown in the work of Sherwood (1999). He hypothesizes that cloud radiative forcing can be very important when dynamically converted into moisture perturbations (see figure 1.3). In particular, he analyzes the response of the atmospheric temperature to radiative forcing by cirrus clouds and relates this temperature change to vertical advection of moisture. According to his conceptual model, applicable in nonconvective areas but presupposing some deep convection to exist, cirrus act as water vapor “pumps”, and the strength of this pumping is a function of their optical depth (i.e., thinner clouds are more efficient pumps).

Another reason upper tropospheric clouds are important is through their influence on stratospheric water budget (Danielson, 1982), their connection to stratospheric wave activity (Boehm and Verlinde, 2000), and their role in stratosphere-troposphere exchange.

The above examples highlight some of the ways ice clouds influence the Earth’s climate system. In the next section we discuss most recent advances in ice cloud observing and modeling.

1.1.1 State-of-the-art ice cloud modeling and observing systems

In recent years much progress has been made in the description of clouds in atmospheric models. Most Numerical Weather Prediction (NWP) models and General Circulation Models (GCMs) now have cloud prognostic capabilities, as replacement

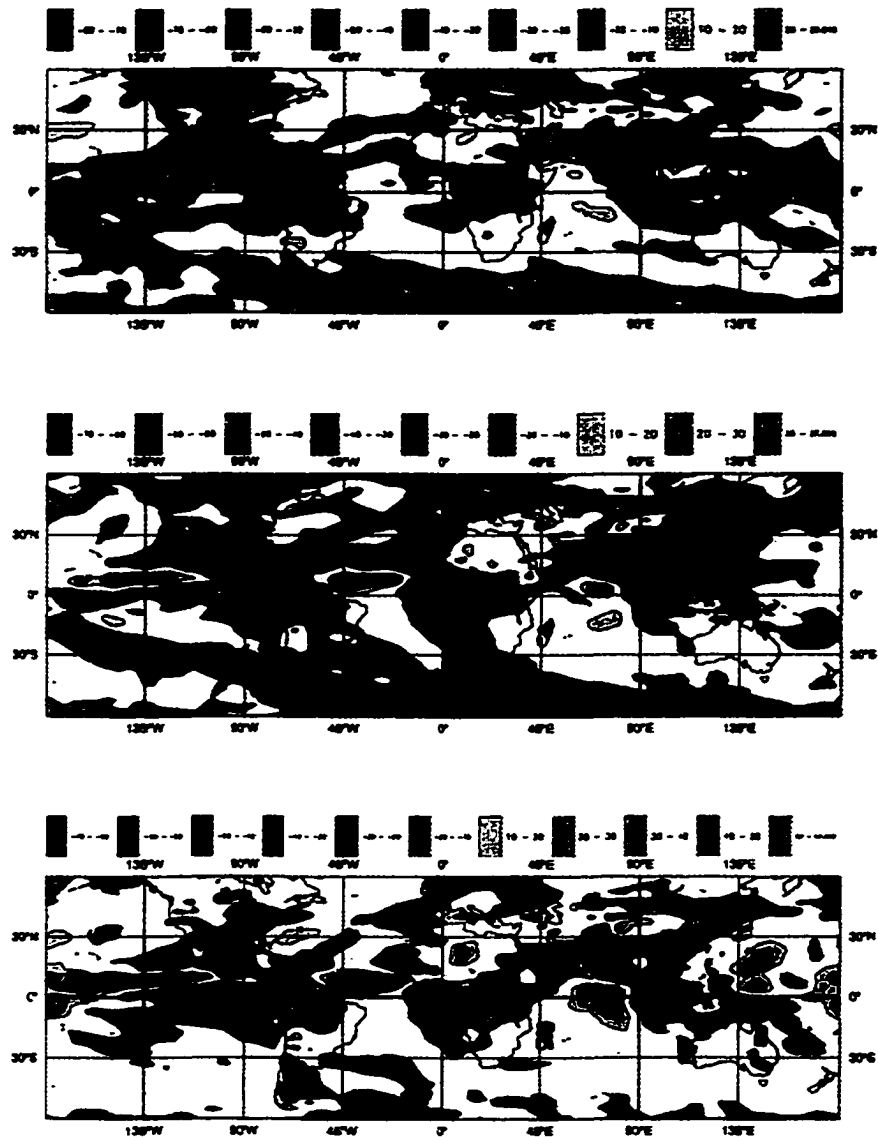


Figure 1.2: The difference between OLR observed by ERBE minus that obtained by the ECMWF model for the DJF 1987/88 season. The upper panel is the difference between ERBE and the model with the diagnostic version of the cloud scheme and the lower two panels show differences between ERBE and the model with two different representations of fall speed of ice crystals. From Stephens et al. (1998).

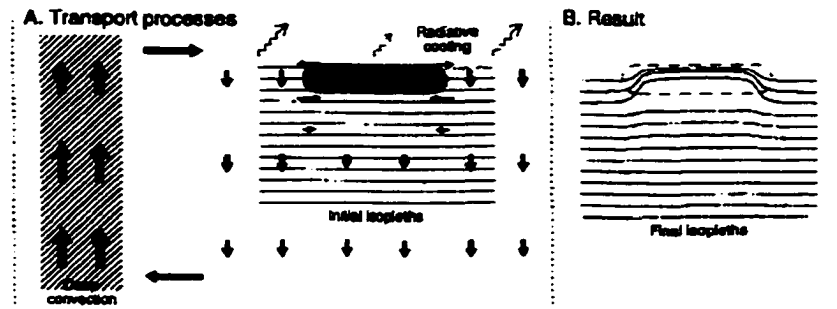


Figure 1.3: Schematic of water vapor pumping, radiatively induced by cirrus clouds. From Sherwood (1999).

to diagnostic schemes¹. The emergence of such schemes has been made possible by the availability of new data sets from field experiments, satellite observations and also from advancements in Cloud Resolving Models (CRM) and Large Eddy Simulation (LES) models. A number of these high-resolution cloud models now incorporate rather sophisticated microphysical and radiative treatments of cirrus (see for example, Khvorostyanov and Sassen (1998) and Meyers et al. (1997)). Some include explicit treatment of particle size distribution and bin microphysics (Lin, 1997) as well as detailed description of ice initiation processes (De Mott et al. (1994) and (1997)).

Amongst the existing datasets, a widely used one is represented by the above mentioned ISCCP database (Rossow et al., 1987). This database contains a cloud climatology based on cloud type, optical thickness and cloud top pressure, determined from visible and infrared satellite imagery. A similar philosophy is the basis of the Clouds and the Earth's Radiant Energy System (CERES) experiment (Wielicki

¹In diagnostic schemes, cloud description is based on diagnosis of fractional cloud coverage and condensate from large-scale conditions. The interaction of clouds with radiation is handled separately with respect to condensation and evaporation processes. At present, many NWP and GCMs include sophisticated prognostic cloud schemes which ensure a physical connection between large-scale forcing and cloud formation and maintenance processes (see for example Tiedke (1993) and Jakob (2001) for a description of the ECMWF prognostic cloud scheme; and Fowler et al. (1996) for the description of the CSU-GCM prognostic cloud scheme).

et al., 1996). The CERES instrument measures both solar-reflected and Earth-emitted radiation from the top of the atmosphere. Cloud properties are determined using simultaneous measurements by other instruments such as the Moderate Resolution Imaging Spectroradiometer (MODIS) on board the Earth Observing System (EOS) Terra and Aqua satellites. As for active sensing of clouds from space, the Lidar In-space Technology Experiment (LITE) provided samples of global lidar measurements of thin cirrus clouds (Winker et al., 1996). Future satellite missions such as CloudSat (Stephens et al., 2000b) and ESSP-3 (Winker and Wielicki, 1999) will fly active sensors (radar and lidar respectively) for the detailed mapping of global cloudiness.

In addition to these satellite programs, many ground and aircraft-based field experiments were deployed with the goal of understanding cirrus clouds and their interaction with radiation. We mention only one of them, the First ISCPP Regional Experiment (FIRE), that is now entering in its fourth phase with the upcoming Cirrus Regional Study of Tropical Anvils and cirrus Layers-Florida Area Cirrus Experiment (CRYSTAL-FACE) deployment.

Ground-based cloud measurements are taken routinely with passive and active sensors at the sites established by the Atmospheric Radiation Measurement² (ARM) Program. The ARM programs maintains three main sites, in Oklahoma (Southern Great Plains (SGP) site), in Alaska (North Slope of Alaska (NSA) site) and on Manus and Nauru islands in the Pacific Ocean (Tropical Western Pacific (TWP) site), and instruments in a number of ancillary locations.

Together with an increasing number of observing missions, advancements in

²This multi-laboratory and interagency program was created in 1989 with funding from the U.S. Department of Energy (DOE), as an effort to resolve scientific uncertainties about global climate change, with a specific focus on improving the performance of general circulation models used for climate research and prediction. In pursuit of its goal, the ARM Program operates field research sites, called Cloud and Radiation Testbeds (CARTs), in several climatically significant locations. Scientists collect and analyze data obtained over extended periods of time from large arrays of instruments to study interactions of radiation and clouds, and their effect on weather and climate.

Source: <http://www.arm.gov>

instrumentation and cloud observing capabilities have also been fundamental toward a better representation of ice cloud properties. The use of millimeter wavelength radars for cloud detection and monitoring has become very common over the past years. Retrieval techniques for determining Ice Water Content (IWC) from radar observations have been developed (Mace et al., 1998a; Matrosov et al., 1992 and 1994), and are still being evaluated and refined. There have been significant advances in airborne *in-situ* instrumentation including particle impactors/replicators (and the data processing thereof, see for example Heymsfield and Iaquina, 2000), holographic particle imaging (Cloud Particle Imager, CPI), and Counter-flow Virtual Impactors (CVI). This progress in *in situ* instrumentation is essential for the evaluation of radar-based retrieval methods. In turn, the radar-based methods provide highly suitable means to obtain the data sets required for development and validation of space-based techniques.

1.1.2 Bringing together observations and models

Better data sets and improved models are keys toward a more accurate representation of ice cloud processes and consequent improvements in weather forecast and climate predictions. In particular, the synergy of both these elements—models and observations—is the most promising way to address the open question about the role of ice clouds in climate change. In this perspective, two main ways of using data in conjunction with models can be identified: one involves the use of data for the evaluation of models, and the development and testing of cloud parameterizations; the other involves the assimilation of observed cloud fields into models. These two approaches are not opposed, rather they are complementary. Data assimilation offers an elegant mathematical framework to compare observations and model forecasts. While it is no substitute for model improvement and development, it represents a powerful tool to obtain optimal initial or/and boundary conditions that insure more realistic (i.e., closer to observations) model results. Once the model error induced by sub-optimal initialization and/or forcing is corrected, specific aspects of the model, either related to parameterizations or to numerical representation of physical pro-

cesses, can be more closely scrutinized and evaluated making use of observations. Ultimately, by combining these two approaches, the model performance as a whole is improved and a more reliable prediction is obtained.

As a testimony to the importance of the first approach, the entire GEWEX Cloud System Study (GCSS) science plan (Randall et al., 2000) is centered around the development of new and improved parameterizations for large-scale GCMs using results from models that explicitly resolve the cloud scales and treat the cloud formation from first principles (CRMs and LES). Before the latter can be reliably used for this purpose they have to be themselves tested and improved using observations. Evaluation of NWP model prognostic cloud schemes using observations has been the object of many recent studies (Mace et al., 1998b; Miller et al., 1999b) and has been instrumental in gaining better knowledge of cloud formation and evolution in relation to large-scale weather patterns. Comparison of model results with observations has been also a strong component of regional cirrus studies (Brown and Field (2000); Benedetti and Stephens (2001); Fouilloux and Iaquina (1998)).

On the other hand, an indication of the importance of data assimilation comes from its long-standing application in operational forecast models. The initial state of NWP models (analysis) is obtained by combining in an optimal way all possible sources of information on the state of the atmosphere (observations of various types, atmospheric equilibria, short-range forecasts). This procedure ensures a better quality of both the forecast and the reanalysis data products, often used in climate studies as global data sets (Trenberth and Caron, 2000). Data that are routinely assimilated at the operational level include conventional and satellite observations of pressure, temperature, wind and water vapor. Requirements for operational, as well as non-operational, assimilable data include good spatial and temporal coverage. Although most NWP models have difficulties in producing realistic cloud and precipitation fields, particularly at the beginning of the forecast period, operational data assimilation systems do not currently assimilate any quantity related to the condensed phase of water. Progress has recently been made toward the assimilation of precipitation fields both in large-scale models (Marécal and Mahfouf (2000); Hou et al. (2000)) and in regional forecast models (see for example Zupanski

and Mesinger (1995) and Zou and Kuo (1996)). Assimilation of cloud fields into forecast mesoscale models is ongoing research (Vukićević et al., 2001). This type of observations is more challenging to assimilate due to the non-linear nature of the parameterizations involved in their description, and their inherent high temporal and spatial variability. Problems related to precipitation and cloud field assimilation are discussed in section 1.1.3 with particular attention to operational applications.

Issues notwithstanding, there is a great potential in assimilation of rainfall and cloud measurements, and the exploration of this potential is the main motivation for this research. The specific focus of this work is on the still unexplored assimilation of cloud radar data into cloud-scale models, as a source of improvement in our current understanding of ice clouds and interaction with the other atmospheric physical elements. Motivation for the particular choice of radar data is offered in section 1.3.

1.1.3 Issues in assimilation of cloud and precipitation

Some of the main problems related to assimilation of precipitation and cloud fields are discussed in this section. Among these, it is the intermittent nature of these phenomena which makes them difficult to model, together with the scarcity of global observations, particularly of cloud fields. The latter issue is especially crucial for operational assimilation into NWP models.

Figure 1.4 shows a plot of 500mb temperature, occurrence of millimeter-wave cloud radar returns above -20dBZ , indicative of cloud presence, and rain gauge measurements at the ARM SGP³ site. From comparison of these time series for the different variables we note the following:

- the tropospheric temperature shows a limited variation over a month (except

³The U.S. Southern Great Plains Cloud and Radiation Testbed (CART) site was the first field measurement site established by DOE's ARM Program. The site consists of in situ and remote-sensing instrument clusters arrayed across approximately 55,000 square miles in north-central Oklahoma and south-central Kansas. Deployment of the first instrumentation to the SGP site occurred in the spring of 1992, just 24 months after the program was approved. Additional instrumentation and data processing capabilities have been incrementally added in the succeeding years. Source: <http://www.arm.gov>

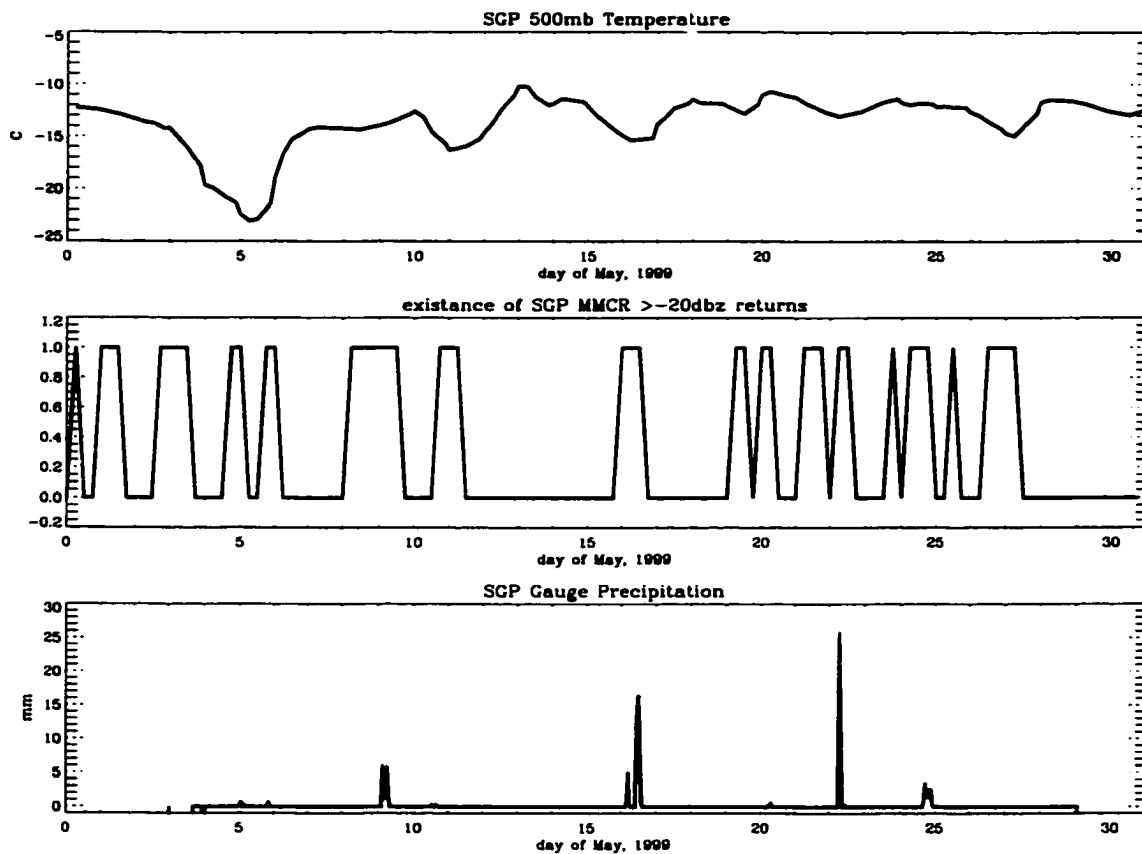


Figure 1.4: 500mb temperature (upper panel), cloud occurrence quantified as millimeter-wave radar returns $> -20\text{dBZ}$ (central panel), and rain gauge measurements over the ARM SGP site. Data were obtained from the ARM Program sponsored by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Environmental Sciences Division. Figure courtesy of Ian Wittmeyer.

for the beginning of the month);

- the presence of water condensed phase, i.e., cloud occurrence, is an intermittent process;
- precipitation events are sporadic, with large variations in amplitude.

Although somewhat obvious, these differences underscore the fact that assimi-

lation of fields which present a discontinuous behavior, such as cloud and precipitation, is inherently different from assimilation of continuous fields such as temperature. The physics of the two types of processes—continuous and discontinuous—is represented in different ways in numerical models. For discontinuous processes, parameterizations often consist of on–off switches activated by exceeding predefined threshold values of given variables (i.e., condensation is activated if the specific humidity is above a certain value). Creating data assimilation systems for this type of parameterizations is an art in of itself. Zupanski and Mesinger (1995) discuss problems connected to developing a variational assimilation system based on the adjoint for a regional forecast model—the National Centers for Environmental Prediction (NCEP) eta model—in particular related to the on–off switch for convective precipitation amount in the Betts–Miller cumulus parameterization (Betts and Miller, 1986). As a solution the authors suggest the introduction of a continuous function that substitutes for the discontinuous switch. A similar approach involving the elimination of discontinuities by modifying the full–physics model is proposed in Verlinde and Cotton (1993). Vukićević and Errico (1993) discuss requirements for obtaining a meaningful adjoint model when the dynamical model is highly nonlinear and presents discontinuities. The main requirement is that the sensitivity of functions that control regime transitions is small with respect to the model sensitivity for state vectors not near the critical states. They also make the point that the nonlinear model has to have good skill in representing the physical phenomenon if one hopes to obtain a well–behaved tangent linear and adjoint model. This is a non-trivial requirement, given the relative poor skill of numerical models in forecasting cloud occurrence and precipitation.

In terms of availability of direct observations, the difference between standard meteorological variables and clouds is also relevant. As an example, figure 1.5 shows surface synoptic data from weather stations worldwide (over land and ship) ingested in an assimilation cycle at ECMWF. Meteorological variables that are routinely measured include temperature, relative humidity, pressure, and wind speed and direction. There is no equivalent of figure 1.5 for cloud fields.

Even when global–scale cloud observations are available, as for example geosta-

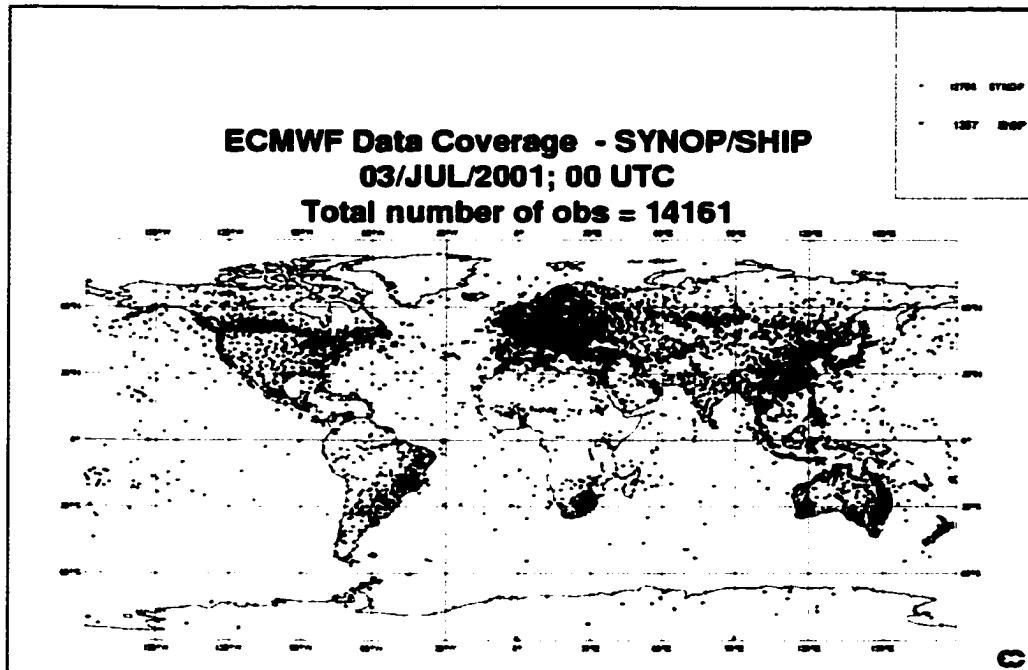


Figure 1.5: Synoptic data from worldwide weather stations on land and ships ingested in ECMWF model.

tionary satellite radiances, the inherent uncertainties in retrieving cloud information from passive sensors only (Miller, 2000) or in directly assimilating cloudy pixel radiances make them difficult to use in operational contexts. There is no general agreement whether it is best to assimilate cloud information retrieved from measurements or to import the measurements into the assimilation system directly. Both methods present difficulties. Assimilation of retrieved quantities has the limitation that the profiles contain not only observed information, but also prior information used to constrain the retrieval. On the other hand direct assimilation of measurements sensitive to cloud variables might pose problems related to nonlinearities in the observational operator (i.e., the Radiative Transfer model for radiances) and to the fact that model resolution might not be accurate enough to resolve cloud features present in the observations. This problem is particularly relevant for NWP models, and it is the reason why satellite radiances are assimilated operationally only in cloud-free conditions (Eyre, 1997). Cloud-sensitive observations seem more

apt to be assimilated into mesoscale models that resolve cloud-scale features.

Radar measurements of cloud fields are not taken routinely, and are not available for operational assimilation. The exception is represented by few locations where cloud data are collected and analyzed on a daily bases (e.g. the above-mentioned ARM sites). Individual institutions and groups maintain observational sites that take passive and active cloud measurements. An example is the Facility for Atmospheric Remote Sensing (FARS) cloud research station affiliated with the Department of Meteorology, University of Utah, and directed by Prof. K. Sassen. Instruments at FARS include lidars and polarimetric Doppler radar. The University of Reading also maintains an observational site at Chilbolton, UK, provided with a 94-GHz and 35-GHz radars for the study of cloud and precipitation systems (Hogan et al., 2000). McGill University, Canada, has operated a precipitation and cloud radar for about ten years, and data are almost continuously collected (Fabry and Zawadzki, 1995). To the best of our knowledge there is no coordinated effort to collect all radar data available from these different locations.

In this context, the advent of CloudSat will bring a very innovative and exciting contribution for the data assimilation community, with the first ever global cloud radar dataset.

1.2 Overview of recent cloud and rainfall assimilation studies

Although the potential of cloud observations for assimilation in mesoscale and large-scale models has been recognized for a number of years, the difficulties discussed in section 1.1.3 have limited the number of studies on cloud assimilation. Assimilation of rainfall has recently received much attention, due, in part, to the availability of observations from the Tropical Rainfall Measurement Mission (TRMM) satellite. Here, a few of these studies are presented for their relevance in the general discussion about assimilation of variables that are linked through the hydrological cycle.

In a recent study, Vukićević et al. (2001) investigate the use of radiances from the Geostationary Observational Environmental System (GOES) satellite to assimilate low-level cloud information into the Regional Atmospheric Modeling System

(RAMS) (Pielke et al., 1992). The authors introduce a Mesoscale Radiance Model inclusive of the mesoscale model and the Radiative Transfer observational operator necessary to arrive from model variables to model equivalent radiances. A verification of model fields shows that the cloud does not extend as far north and is significantly brighter in the southern portion of the domain than the observations show. Preliminary results from this study using the RAMS adjoint show a positive response of cloud solution to the initial condition adjustment. More details on this study are given in chapter 5.

Rainfall assimilation in a regional model is the focus of a work by Zou and Kuo (1996) who show the positive impact of assimilation of observations from the Severe Environmental Storm And Mesoscale Experiment (SESAME) on short-range forecast of an observed Mesoscale Convective System (MCS) with a nonhydrostatic mesoscale model (MM5). Model simulations of the same event without assimilation of rainfall data failed to correctly simulate location and intensity of observed precipitation. Assimilation of rainfall data led to a better forecast, and captured mesoscale features relative to the structure of MCSs, lower- and upper-level jets, the position of the dryline, the low-level moisture convergence, and the formation of a localized capping inversion.

Other studies have explored the impact of assimilation of information contained in precipitation and water vapor fields to improve global analysis with GCMs. Gerard and Saunders (1999) discuss the introduction of SSM/I-derived Total Column Water Vapor into the ECMWF model. As a result of the assimilation, the model-predicted global water-vapor is increased by 2%, particularly in areas where the model is notoriously too dry. There is also an impact on circulation with an increase of the strength of the Hadley cell. However, the use of TCWV information is subject to the model physics (i.e., the model cannot retain additional moisture in the Tropics and reacts through increased precipitation), and the authors imply that a more optimal use of observations might be achieved both through improvement in model shallow convection and vertical boundary-layer resolution, and through a better specification of background errors in humidity fields in the data assimilation system. Deblonde (1999) obtains similar results from the assimilation of SSM/I

retrievals into the Canadian Meteorological Centre (CMC) model.

Marécal and Mahfouf (2000) find that one-dimensional variational assimilation of TRMM rainfall rates into a single-column version of the ECMWF model improves the short-term forecast, but is not able to trigger precipitation where the model background field did not have precipitation to begin with.

The study of Hou et al. (2000) shows how assimilation of TRMM rainfall rates and Total Precipitable Water (TWP) derived from the TRMM Microwave Imager (TMI) into a single column version of the Goddard Earth Observing System (GEOS) GCM improves diagnosis of clouds and radiation in areas of active convection, and the latent heating distribution and large-scale dynamical fields in the Tropics, while reducing systematic errors in forecasts. In this study, it is implied that the combination of both types of observations, and/or the introduction of similar measurements, leads to the largest improvement in model prediction.

Two main conclusions can be drawn from these studies: the most obvious is that assimilation of atmospheric condensate fields, whether in precipitation or non-precipitating form can have beneficial impact on all aspects of model forecasts, even on those not directly related to the assimilated field itself. The second is that the more information the better. Assimilation of rainfall information and cloud information has great potential for positive impact on model forecasts, and improvements in understanding of moist physics processes.

1.3 On the potential of assimilation of cloud radar data

Application of weather radar data in conjunction with models is not a novelty. Doppler radar-derived wind fields and reflectivities, sensitive to precipitation-sized scatterers, have been assimilated successfully into cloud models, both for retrieval purposes, and for improvements in model initialization. The work of Sun and Crook (1997) shows the applicability of a four-dimensional variational assimilation of Doppler radar data into a cloud-scale model with warm rain parameterization for optimal retrievals of winds and rainwater mixing ratio. In a follow-up study (Sun and Crook, 1998), the authors show results from an observational case from

the Convection and Precipitation/Electrification Experiment (CaPE); the 3D wind, temperature and microphysical structure of the storm are obtained by minimizing the difference between radar-observed radial velocities and rainwater mixing ratios derived from radar reflectivities and their model equivalent. Verification with independent aircraft data shows good agreement in the general structure of the storm (vertical velocity, buoyancy and water vapor mixing ratio), demonstrating that inclusion of radar data has the potential to improve general model prediction of variables that are not directly observed. However, assimilation results for rainwater quantities are sensitive to assumptions in deriving the mixing ratios from reflectivities. In both these studies, the focus is more on optimal retrievals of winds and precipitation by combining Doppler radar observations and a cloud model, rather than purely improving initial conditions for better model forecast. Wung et al. (2000) show positive impact of radar data variational assimilation for simulation of a severe thunderstorm, but they find large algorithm sensitivity to model microphysical and turbulence parameterizations. They conclude that assimilation of radar observations demonstrates the need for improved parameterizations that would allow models to more accurately reproduce observed systems. A different approach for the use of Doppler radar reflectivity is suggested by Rogers et al. (2000). The radar data are used to tell the model where the deep moist convection is occurring, and consequently to activate the convective parameterization scheme. Although this type of data assimilation is non-optimal, it results in an improvement in model initialization and forecasts of precipitation and mesoscale environment. Even with the limitations indicated by the authors, related to specific model parameterizations, these studies emphasize that assimilation of radar fields is feasible, and that radar observables contain information regarding the general state of the atmosphere through coupling of microphysical and dynamical processes.

Moreover, recent technological progress has granted the possibility of using radar at different wavelenghts to obtain detailed cloud information for clouds normally undetected by standard weather radars. Cloud radar data are good candidates for assimilation purposes due to their high sensitivity to cloud bulk properties (particle mean size and number concentration, Ice/Liquid Water Content, particle fallspeed).

At the same time, the radar mapping is conceptually simple, and can be easily incorporated as the observational operator in an assimilation system. For most types of ice clouds, the radar equation is simplified due to the fact that the Rayleigh approximation can be invoked, attenuation is weak, and a semi-analytical expression for radar reflectivity can be derived. Most mesoscale models contain bulk microphysics parameters and prognostic variables from which the model equivalent radar reflectivity can be derived. There is a wide range of empirical relationships, based on both in-situ measurements and models that can be used to convert simply from model Ice Water Content (or similarly Liquid Water Content) to equivalent reflectivity (Liu and Illingworth (2000); Matrosov (1999); Liao and Sassen (1994)). For these reasons, assimilation into numerical forecast models appears as a viable novel application of cloud radar data.

1.4 Outline of research

This study is aimed at showing the feasibility and the benefits of assimilating cloud information derived from radar data into two cloud models— a one-dimensional time-dependent cloud ice growth model and a full-blown mesoscale model. Results are relevant for possible assimilation of radar data into large-scale operational models. The focus is on ice clouds, but the concepts and techniques presented are general and can be applied to water clouds⁴ as well. The steps taken to achieve these goals are summarized in what follows.

- A cloud resolving model, the CSU Regional Atmospheric Modeling System (RAMS) developed by scientists at Colorado State University and the *ASTER Division of Mission Research Corporation, was used in numerical simulations of cirrus clouds. Sensitivity of the model to changes in various parameters relevant to the formation and evolution of ice clouds was investigated using an idealized cirrus test case. Model results were compared with other cirrus models in the context of the GEWEX Cloud System Study (GCSS) Cirrus

⁴Application to mixed phase clouds, although conceptually possible, would need further refinements in the observational operator, and introduction of data other than cloud radar only.

Working Group (WG2) modeling and intercomparison activity (Starr et al., 2000). The focus of this research was to assess model skill at cirrus cloud prediction under “controlled” conditions.

- As a continuation of model skill assessment, the CRM was then used to simulate an observed cirrus case from the ARM–Unmanned Aerospace Vehicle (UAV) Spring '99 campaign (Stephens et al., 2000a). Airborne radar observations were used in the evaluation of the model performance. A statistical technique adapted from Hollingsworth and Lönnberg (1986) was introduced to identify model biases and perform an error analysis.
- The knowledge acquired from studying cirrus formation with RAMS was applied to develop an ice growth model based on Mitchell (1988). This model and its adjoint were used to investigate the feasibility of variational assimilation of radar reflectivities with a series of experiments. These experiments included the use of the adjoint to study model sensitivity to external inputs (atmospheric state variables), initial cloud variable conditions, and model parameters (fallspeed–diameter and mass–diameter parameterization coefficients). Synthetic radar reflectivities and real measurements from the above–mentioned ARM–UAV case as well as from ground–based observations from the ARM–SGP site were both used in the assimilation experiments.
- Based on results obtained with the one–dimensional model, further experiments leading to four–dimensional variational assimilation of radar data were performed using the adjoint of RAMS developed by Dr.T. Vukićević, coordinator of the Data Assimilation group at the Cooperative Institute for Research in the Atmosphere (CIRA).
- As part of the assimilation work, a semi–analytical radar “mapping” was introduced and tested. As a side result, a new retrieval of IWC using radiometric and radar observations was developed. A few results from this retrieval are presented in appendix A.

1.5 Thesis road map

The dissertation is structured in three parts. The first part comprises this introduction and Chapter 2 which provides basic data assimilation concepts. The second part includes chapter 3 which describes the cirrus modeling work accomplished with the Cloud Resolving Model. The third part, which is the most substantial and novel aspect of this research, is dedicated to the description of the one-dimensional cirrus model, and its use in variational assimilation experiments (Chapter 4), and the discussion of the adjoint experiments using RAMS (Chapter 5) which are the prelude to four-dimensional cloud data assimilation. Chapter 6 summarizes and concludes this work. A brief summary of every chapter is offered here.

In **Chapter 2**, some introductory concepts regarding data assimilation with specific focus on the variational approach are presented. All components of a variational assimilation system are introduced and described in detail, with particular care in the definition of the adjoint. An example using the potential vorticity equation is given to better illustrate theoretical concepts. The importance of error statistics is also discussed. The material of this chapter is largely of a review nature.

In **Chapter 3**, experiments performed using RAMS to assess model skill in cirrus prediction are presented. Section 3.1 describes the current problems in representation of cirrus clouds in cloud resolving models. Specific reference is made to the intercomparison activity of the GCSS Cirrus Working Group. In section 3.2 a general overview of source of uncertainties in cloud models is presented. Concrete examples are presented, using sensitivity studies performed with RAMS. In section 3.3.1, a description of the RAMS model is presented. A cirrus case study from the 1999 ARM-UAV observation campaign is presented in section 3.4 and used to evaluate the CRM performance. In section 3.5, an error analysis for the RAMS model is presented. The final section (3.6) summarizes the main conclusions regarding the cirrus modeling activity with RAMS. The material in this section is novel to the extent that the modeling study is applied to case study data analyzed specifically for this research.

Chapter 4 introduces the new ice cloud model that was developed specifically

for this study. Section 4.1 provides the rationale for the development and the use of a simple cloud model. Section 4.3.1 provides a detailed derivation of model equations. In section 4.3.5, observations and RAMS cirrus simulations are used as benchmark for the evaluation of the model. The derivation of the adjoint of the cirrus growth model is presented in section 4.4.1. Application of the adjoint for sensitivity studies is offered in section 4.4. Assimilation experiments using synthetic and real observations are discussed in the remainder of the chapter. The content of this chapter is original.

Chapter 5 describes sensitivity experiments performed with the RAMS adjoint. The adjoint forcing is derived from synthetic reflectivities and used to understand mesoscale model response to introduction of radar data. While the RAMS adjoint system was developed by other investigators, the application of the model adjoint and the inclusion of cloud radar data in a pre-4DVar study is original work.

Chapter 6 offers a summary of the entire dissertation work, and draws some conclusions concerning the relevance of these experiments to the assimilation of cloud radar data into NWP models.

Appendix A introduces the radar retrieval work that was generated from the preparation for the assimilation experiments.

Chapter 2

ABC of data assimilation

2.1 Introduction

Data assimilation has a famous ancestor in K. F. Gauss. In his *Theoria Motus Corporum Cælestium* (1809), he lays the basis for estimation theory in his attempt to evaluate the orbits of comets and planets from incomplete astronomical data. “*This problem*”, he writes, “*can only be properly undertaken when an approximate knowledge of the orbit has been already obtained, which is afterward to be corrected as to satisfy all of the observations in the most accurate manner possible*”. These words recognize both the importance of having a formal knowledge of the phenomenon (i.e. a “model”) and of having observations to correct and improve the model prediction. An analysis that includes model prediction and measurements is more accurate and powerful than either one alone. Data assimilation can hence be defined as an analysis technique in which the observed information is integrated into the model state with respect for the physical properties and time behavior of the system. As such, it is an essential component of any modeling activity. Each time model parameters are adjusted to better match observed fields, this represents, however rudimentary, a form of data assimilation. The difference between these empirical forms of data assimilation and the more sophisticated assimilation systems is only in the mathematical complexity, while the underlying goal remains to use observations and models together, thus obtaining a more realistic representation of the physical phenomenon, and with that, the sense that we are closer to understanding the physical world which we live in.

The specific object of atmospheric data assimilation (hereafter, DA) is to produce a regular, physically consistent representation of the state of the atmosphere from

a series of measurements often sparse and uneven in time and space. DA not only combines optimally observations with model estimates, but it confronts theory with reality, which can potentially lead to improvements in both model and observing systems.

In this section we introduce a general theory of data assimilation techniques in atmospheric science, with special emphasis on *variational* data assimilation.

2.2 Data assimilation techniques

In their overview of assimilation concepts, Ide et al. (1997) divide DA techniques into two broad categories: *sequential* assimilation and *variational* assimilation. In sequential assimilation, observations are integrated in the model at each time they are available, and used to “update” the model prediction. This procedure can thus be described as a sequence of analyses performed at observation times, and of integrations of the model between successive analyses. This feature makes sequential assimilation well adapted for NWP. The drawback of this type of DA is that observations only influence the estimated model state at later times, and the information is propagated from the *past* to the future. For real-time weather forecasting, this is not an issue, but it can be for *reanalysis* activities, in which it is desirable to use an algorithm capable of carrying information both forward and backward in time. Examples of sequential DA are, in order of complexity: interpolation of observations, dynamic initialization by repeated insertion of data (also known as Newtonian relaxation or “nudging”, following the terminology of MacPherson (1991)), Cressman Successive Corrections (Daley, 1991), Optimal Interpolation (Lorenc, 1981), and Kalman Filter (Cohn, 1997). A general overview of these methods is presented in Lorenc (1986).

In variational DA, all observations available over the assimilation period are incorporated at once in the adjustment process, so they all influence the final model state. Since the analysis is adjusted to the observations over the whole interval, the result is a smooth trajectory. For this reason, variational algorithms are also called *smoothing* algorithms. Examples of variational methods are offered by XD–

Var algorithms, where X represents the dimensionality of the var system, and can be equal to 1, 2, 3 or 4. Sometimes the designation 1+1D-Var, as opposed to 2D-Var, is also employed to remove the ambiguity related to the use of one spatial and one temporal coordinate (1+1D) versus the use of two spatial coordinates (2D). It is important to make this distinction since the inclusion of the time dimension brings about the true potential of variational methods, which is the propagation of information backward in time, allowing for the inclusion of *future* information in the assimilation. In this sense variational assimilation systems that do not include the time dimension, such as 3D-Var systems, are closer to sequential algorithms, although the details of actual assimilation procedure are different. Current NWP models implement operational 3D and 4D-Var assimilation systems (i.e. for the European Center for Medium-range Weather Forecasts model see Rabier et al. (2000), Mahfouf and Rabier (2000) and Klinker et al. (2000); for the National Centers for Environmental Prediction model see Parrish et al. (1997)).

The focus on the remainder of the chapter will be on two of the above-mentioned assimilation techniques: nudging and variational data assimilation (VDA). The reason for this choice comes from the fact that nudging is frequently used in mesoscale cloud modeling. An example is provided in chapter 3 with cirrus simulations obtained from the RAMS model nudged with a large-scale forecast provided by the European Center for Medium-range Weather Forecast. Variational assimilation is discussed in depth since this technique was used in the assimilation experiments presented in chapter 4 and 5. We feel it is important to provide a concise, yet complete description of the basic theory of variational assimilation to allow for a better understanding of the terminology used in the remainder of this work.

2.2.1 Nudging technique

The nudging data assimilation (NDA) technique (Anthes, 1974) consists in relaxing the model state to observations¹ by adding a non-physical diffusive term to the

¹In mesoscale models, “observations” can also be taken to be fields from a NWP large-scale analysis. Fields generally nudged are winds, temperature and pressure. This type of NDA was

model equations. The nudging terms are defined as the difference between the observation and the model solution multiplied by a nudging coefficient. In mathematical terms,

$$\frac{\partial x}{\partial t} = \mathcal{M}(x) + \frac{(x_{obs} - x)}{\tau_{ng}} \quad (2.1)$$

where the first term on the RHS represents the physical forcing terms, and the second term represents the nudging contribution. Solution of (2.1) depends both on the initial condition and on the nudging coefficient, τ_{ng} . The value of τ_{ng} , also called the *relaxation* coefficient, is generally chosen by numerical experimentation so as to keep the nudging term small in comparison with the dominant forcing terms in the governing equations, yet large enough to impact the simulation. As a general rule τ_{ng} should increase with increasing observational error and increasing spatial and temporal separation. The large sensitivity of the model solution to the choice of nudging coefficient, which can force the model solution away from its physics and violate physical constraints like mass and energy conservations, and the lack of an optimum way to define the nudging coefficient, as found by Stauffer and Seaman (1990), represent the major limitations of this technique. Another limitation is that one can only nudge model-predicted variables and not derived variables such as radar reflectivity, radiances, etc. Some investigators have suggested to estimate nudging parameters using variational data assimilation (Zou et al., 1992; Stauffer and Bao, 1993). Results from these studies show that optimal NDA is feasible, and performs well as far as quality of assimilated fields. General advantages of the NDA are the relative simplicity of implementation and the low computational cost.

2.2.2 Variational data assimilation

The variational technique applied to data assimilation was first introduced in optimal control theory as a means to “control” the output parameters of a complex model by acting on the input parameters (Lions, 1971). A central component in optimal control theory is represented by the *adjoint* of the tangent linear model as used in the RAMS simulations presented in chapter 3.

a tool to efficiently compute local gradients of a scalar function of the model input and output parameters. The idea of using this technique in meteorological problems was first suggested by Marchuk (1974). Since then several authors have used it in various applications (sensitivity studies as in Hall et al. (1982), quasi-geostrophic theory as in Lewis and Derber (1985), ocean modeling as in Bennett and McIntosh (1982), to mention only a few of the earlier studies). An excellent review of the variational methodology is presented in Le Dimet and Talagrand (1986).

At the core of variational data assimilation is the search for an optimal solution which minimizes the distance between model state and observations expressed in terms of a *cost* function. Optimization of the cost function is performed using the gradient information provided by integration of the adjoint model. The variational problem can be cast in the following way. The nonlinear dynamical model predicts the time evolution of the state vector x given an initial condition, x_0 :

$$\dot{x} = \mathcal{M}(x, x_0, \dot{x}). \quad (2.2)$$

where \dot{x} is the time derivative of x . This can also be expressed in the following way:

$$\dot{x} - \mathcal{M}(x, x_0, \dot{x}) = G(x, \dot{x}, x_0) = 0. \quad (2.3)$$

Let's assume that we have a discrete series of observations y_{obs} at given times, t_{obs} over the assimilation interval $[0, T]$. Over the same interval of time, the model solution is obtained by integration of equation (2.2). The cost function can be defined as follows:

$$\mathcal{J} = \frac{1}{2} \int_0^T [(P(x) - y_{obs})^T \mathbf{W}^{-1} (P(x) - y_{obs})] \delta_D(t - t_{obs}) dt \quad (2.4)$$

where $P(x)$ is a mapping from the state variable space into the observational space. The introduction of the function $\delta_D(t - t_{obs})$, Dirac's delta, allows us to represent, with a continuous function, observations that are discretely distributed in time. Errors in the mapping (*representativeness* errors), as well as observational errors, are included in the error covariance matrix \mathbf{W} . The inverse of \mathbf{W} , which appears in (2.4), represents a weighting of the distance between model prediction and observed state which takes into account uncertainties in observations and mapping.

The optimal model state is obtained by minimization of the cost function \mathcal{J} with respect to a set of control parameters (i.e. the initial condition, the boundary conditions, etc.), subject to the constraint \mathcal{G} (the dynamical model). The inclusion of the dynamical model as a constraint in the minimization process ensures consistency of optimal control parameters with model physics since the solution is forced to be in the model space. If the model is assumed perfect, and no model errors are included, as in the standard variational formulation, it is said that the model is used as a *strong* constraint (Sasaki, 1970).

Since the number of observations can never equal the number of possible model states, especially when the model is a global high-resolution NWP model, the analysis problem is underdetermined (Lorenz, 1986). To address this issue, common practice is the introduction of prior knowledge, expressed mathematically with the inclusion of a *background* term x_b and its error covariance matrix, \mathbf{B} in the definition of the cost function:

$$\begin{aligned} \mathcal{J} = & \frac{1}{2} \int_0^T (P(x) - y_{obs})^T \mathbf{W}^{-1} (P(x) - y_{obs}) \delta_D(t - t_{obs}) dt \\ & + \frac{1}{2} (x - x_b)^T \mathbf{B}^{-1} (x - x_b). \end{aligned} \quad (2.5)$$

In NWP, the background is represented by the model forecast prior to the analysis, i.e. prior to the inclusion of new observations at successive times. Depending on the specific application, the background information can also be provided either by climatological records or by a global model reanalysis. More general expressions of the cost function may include a model error term ((Ménard and Daley, 1995), (Zupanski, 1997)) and other so-called penalty terms to control spurious numerical modes introduced in the model initialization by the assimilation process (i.e. gravity wave control).

The quadratic form commonly chosen for the cost function allows establishment of a link between estimation theory and optimization in the context of the nonlinear least-squares technique. In general, it is also possible to choose a cost function that is non-quadratic, by defining a different measure for the distance between model and observations.

Proceeding with the discussion of the variational approach, the Lagrangian of the system is introduced:

$$\mathcal{L} = \mathcal{J} + \int_0^T \lambda \mathcal{S} dt \quad (2.6)$$

where λ is a parameter function called the Lagrange multiplier. The constrained optimization consists in finding the minimum of the cost function subject to the constraint provided by the physical model. This is obtained by minimizing the Lagrangian, which entails, in turns, solving the Euler–Lagrange equations. To obtain this set of equations, also known as the *adjoint* system, we have to evaluate $\delta\mathcal{L}$ and set it to zero, according to the least–action principle. Note that these equations are specific to the dynamical model under consideration and the preassigned cost function. An example of derivation of Euler–Lagrange equations for a relatively simple system is provided in section 2.3.2.

In the search for the minimum of \mathcal{L} , the adjoint system is integrated backward in time, and its solution at time $t = 0$ represents the gradient of the Lagrangian with respect to the initial condition. If the initial condition that was chosen at the beginning of the simulation were perfect, so that, at time T , the cost function is at its minimum, meaning that the model prediction is as close to the observations as permitted by the error statistics, then the gradient at time $t = 0$ would be zero, and there would be no need to correct the model prediction. Since this is rarely the case, the adjoint solution, which represents the *gradient* of the cost function with respect to the control variables, is used to march toward the minimum of the cost function. Many optimization methods make use of the gradient information to optimize a function, for example steepest descent or conjugate gradient algorithms. An accurate description of the optimization procedure used in this study is offered in chapter 4. The next section provides mathematical proof of the connection between estimation theory and variational optimization through the introduction of the quadratic cost function. In section 2.3 the components of a variational DA system are described in detail following Bouttier and Courtier (1999).

2.2.3 Link to estimation theory

In atmospheric data assimilation problems, the cost function describes the difference between model variables and observations weighted by their statistical significance. Provided that this function is a quadratic form, and the mapping operator is linear or weakly non-linear, a derivation of the cost function can be outlined from general estimation theory principles. This is demonstrated in this section.

The analysis problem in estimation theory can be generalized using the concept of conditional probability. According to Bayes' theorem, the *joint* probability of a given model state x and an observational state y_{obs} occurring together is given by:

$$p(x \wedge y_{obs}) = p(x|y_{obs})p(y_{obs}) = p(y_{obs}|x)p(x) \quad (2.7)$$

where $p(x|y_{obs})$ represents the probability of having a particular model state given the observational state (what we are after); $p(x)$ represents the probability of having a given model state independent of the observations (the prior knowledge); $p(y_{obs})$ is the probability of having a given observation (equal to 1 if the measurements have been made and we know the value of y_{obs}), and $p(y_{obs}|x)$ represents the probability of having an observational value, given a model state. If the model state is projected into the observational space via the mapping introduced in equation (2.4), $p(y_{obs}|x)$ peaks at $y_{obs} = P(x)$, but it is not exactly equal to a Dirac distribution because of representativeness and observational errors.

Two main criteria can be applied to optimize $p(x|y_{obs})$:

1. *maximum likelihood*
2. *minimum variance*

If we assume that the statistics are Gaussian, i.e. all variables are normally distributed around their mean with predefined variance, the best estimates given by the two criteria coincide, due to the fact that the Gaussian probability distribution function (pdf) is symmetric. The quadratic form of the cost function is derived making this assumption. The proof follows. Writing the pdf as Gaussian we have:

$$p(x) = c_1 \exp \left\{ - \left[\frac{1}{2} (x - x_b)^T \mathbf{B}^{-1} (x - x_b) \right] \right\} \quad (2.8)$$

$$p(y_{obs}|x) = c_2 \exp \left\{ - \left[\frac{1}{2} (y_{obs} - P(x))^T \mathbf{W}^{-1} (y_{obs} - P(x)) \right] \right\} \quad (2.9)$$

where c_1 and c_2 are normalization constants.

Substituting in (2.7), we have:

$$p(x|y_{obs}) = c_1 c_2 \exp \left\{ - \left[\frac{1}{2} (x - x_b)^T \mathbf{B}^{-1} (x - x_b) + \frac{1}{2} (y_{obs} - P(x))^T \mathbf{W}^{-1} (y_{obs} - P(x)) \right] \right\} \quad (2.10)$$

Optimization of (2.10) implies minimization of the exponent, which is exactly the cost function that is minimized in variational assimilation. The solution of the variational data assimilation problem represents the maximum likelihood/minimum variance estimate of the optimal state, assuming all error statistics are Gaussian. In this framework, and for linear (or semi-linear) observational operators², 3D-Var DA and statistical interpolation with least-square estimation, are exactly equivalent (Courtier, 1997). The equivalence ceases to exist, when the cost function is non-quadratic, i.e. error distributions are non-Gaussian. Bayes' theorem can still be applied for generic pdfs, but mathematical solution to the maximum likelihood problem might involve resorting to approximations of the Gaussian statistics (Lorenc, 1984).

2.3 Components of a VAR system

2.3.1 The dynamical model

In principle any model that describes the physical system under consideration can be a good candidate for the DA problem. In practice, some dynamical models or some parameterizations are more suitable than others. In NWP, for example, it is

²If the observational operator is nonlinear, the least square formulation needs to be extended to nonlinear-least squares.

customary to solve the dynamical equation in spectral space where differential equations are reduced to algebraic equations via the introduction of *ad hoc* transforms³. Likewise, for certain classes of problems, it may be more convenient to solve the assimilation problem in spectral space, whereas for others the analysis might give better results if carried out in physical space. A similar question can arise when considering the variable with respect to which the optimization is performed (control variable). There are several variables that are routinely used in models to describe the vapor distribution in the atmosphere (relative humidity, specific humidity, dew-point temperature); although they all represent the same physical field, one can be more convenient than the other to assimilate. Sometimes it can be a function of a model variable that gives the best results.

Regardless the choice of model description, or of control variable, two things are relevant before embarking on any type of DA:

1. assessment of model skill, i.e. the characterization of model performance in representing the physical phenomenon:
2. evaluation of model sensitivities to change in initial or boundary conditions, and/or model parameters (which are all possible candidates for the role of control variables).

The first activity is crucial if we hope to obtain any type of improvement in the model prediction via inclusion of data. If the model fields are far from what the observational data indicate, it may be impossible to obtain a solution which satisfies, even within the given uncertainties, both the model and the observations. This is particularly true in “standard” VDA, where the model error is not considered. The second point underscores the importance of understanding the response of the dynamical model to various forcings, which is in turn important when choosing control variables. The evaluation of model sensitivities can be carried out efficiently by

³An example of that is the use of Fourier transforms to solve for the zonal structure, and of the spherical harmonics basis functions to solve for the meridional structure of the atmospheric flow on the sphere.

introducing the adjoint model. A general description of adjoint theory is presented in the upcoming section.

2.3.2 The adjoint operator

Given a *linear* operator \mathbf{A} on the Hilbert space H and an inner product on H , the adjoint operator, \mathbf{A}^\dagger , is defined as:

$$(\mathbf{A}x, y) = (x, \mathbf{A}^\dagger y), \quad \forall x, y \in H \quad (2.11)$$

The adjoint is a linear operator which always satisfies conditions of uniqueness and existence in a finite-dimensional Hilbert space with respect to the specific choice of inner product⁴. For linear discrete operators that can be represented in matrix form, the adjoint is simply the transpose. The definition (2.11) can be used to derive the adjoint and establish its validity if a perfect (within numerical uncertainties) equality is recovered when the inner product is computed with \mathbf{A} and \mathbf{A}^\dagger .

In DA, the term "adjoint" is usually shorthand for adjoint of the dynamical model (or of its *tangent linear*, if the dynamical model is nonlinear). The role of the adjoint in variational assimilation is crucial. The adjoint solution allows for an effective and fast computation of the gradient of the cost function. In theory, it would be possible to compute the gradient by perturbing each point in time and space, and then applying a finite difference or similar technique (*forward* sensitivity). However, if the model domain is large, and the assimilation interval long, this may be a computationally intensive task. The solution to this problem is to compute the gradient with a single adjoint integration (adjoint or *backward* sensitivity). Provided that the adjoint equations can be derived and solved, this results in an efficient tool. The theoretical foundation for this use of the adjoint is discussed in detail by Talagrand and Courtier (1987), and summarized briefly here.

Consider a scalar function of x in the generic Hilbert space, and let it be the cost function⁵, \mathcal{J} , introduced in section 2.2.2. The differential of \mathcal{J} with respect to x , $\delta\mathcal{J}$,

⁴Most commonly, the Euclidean inner product is chosen.

⁵The function \mathcal{J} does not necessarily have to be the cost function usually introduced in vari-

can be expressed symbolically using the inner product notation introduced earlier

$$\delta\mathcal{J} = (\nabla_x\mathcal{J}, \delta x), \quad (2.12)$$

where $\nabla_x\mathcal{J}$ represents the gradient of \mathcal{J} with respect to x . Suppose x is a differentiable function of the variable y , $x = \mathbf{F}(y)$. The differential of x is equal to

$$\delta x = \mathbf{F}' \delta y \quad (2.13)$$

where $\mathbf{F}' = \frac{\partial \mathbf{F}(y)}{\partial y}$ is the Jacobian (tangent linear) operator. By rewriting $\mathcal{J}(x)$ as $\mathcal{J}(\mathbf{F}(y))$ and substituting equation (2.13) into (2.12), the definition of adjoint in (2.11) can be used to obtain

$$\delta\mathcal{J} = (\nabla_x\mathcal{J}, \mathbf{F}'\delta y) = (\mathbf{F}'^\dagger \nabla_x\mathcal{J}, \delta y) \quad (2.14)$$

where \mathbf{F}'^\dagger represents the *adjoint* of the tangent linear operator \mathbf{F}' .

By definition of gradient we have that

$$\nabla_y\mathcal{J} = \mathbf{F}'^\dagger \nabla_x\mathcal{J}. \quad (2.15)$$

Equation 2.15 provides the basis for the use of adjoint operators in variational assimilation, and more in general in linear sensitivity studies. If, for example, we equate \mathbf{F} with the dynamical model \mathcal{M} of equation (2.2), and we equate y with the initial condition x_0 in the same equation, we have that the variation of the cost function induced by a variation of initial condition (control variable) is equal to the product of the adjoint of the tangent linear of the dynamical model times the change of the cost function due to a variation in model solution x . To find the gradient of the cost function with respect to the initial condition x_0 , a series of simple steps can be taken: perturb x_0 , compute the new model solution x , compute the gradient of \mathcal{J} with respect to x , evaluate the adjoint of the tangent linear of the model and multiply it by $\nabla_x\mathcal{J}$. An example taken from a meteorological application will help clarify these concepts.

ational DA, but can be any scalar function of model variables, representative of the particular model aspect under consideration.

Example of derivation of adjoint equations

We consider a simple differential equation describing the temporal evolution of the linearized potential vorticity, ζ , for a shallow water system⁶ on an f -plane⁷:

$$\frac{\partial \zeta(t)}{\partial t} + Q\alpha^2 te^{-\alpha t} = 0 \quad (2.16)$$

Equation (2.16) is linear with respect to the unknown $\zeta(t)$ and all quantities that appear in it are nondimensional. The term $Q\alpha^2 te^{-\alpha t}$ represents a mass source/sink. The potential vorticity increases locally if there is a mass sink ($Q < 0$), and decreases if there is a mass source ($Q > 0$). Local conservation during the geostrophic adjustment process occurs for $Q = 0$. The parameter α controls the time-dependent part of the forcing (small α corresponds to slow forcing, large α corresponds to rapid forcing). The solution of equation (2.16) depends on the initial conditions, and on the specification of the mass source/sink parameters.

Two type of problems can be addressed using variational DA for the system described by (2.16):

1. Given a set of observations, $\zeta_{obs}(t)$, at observation times, t_{obs} , over the interval $[0, T]$, determine the initial condition on initial vorticity such that

$$J = \frac{1}{2} \int_0^T (\zeta(t) - \zeta_{obs}(t))^T \mathbf{W}^{-1} (\zeta(t) - \zeta_{obs}(t)) \delta_D(t - t_{obs}) dt \quad (2.17)$$

is minimized.

2. Using the same observations, determine the optimum value of α .

⁶The shallow water system describes the dynamics of an incompressible (constant density) fluid with a free upper surface. The fluid is supposed to be in hydrostatic balance, i.e. the vertical pressure gradient is in balance with the gravity force, $\partial p / \partial z = -\rho g$. Even though, through the hydrostatic assumption, the vertical structure is omitted, the shallow water equations are used extensively in geophysical fluid dynamics to describe the horizontal aspects of the flow. Examples include horizontal wave propagation and barotropic instability problems (Schubert, 1997).

⁷The f -plane approximation consists of assuming the Coriolis parameter f constant with latitude.

The initial condition and the parameter α represent the *control variables* in the optimization problem. The Lagrange function is written following (2.6):

$$L = J + \int_0^T \lambda(t) \left[\frac{\partial \zeta(t)}{\partial t} + Q\alpha^2 t e^{-\alpha t} \right] dt. \quad (2.18)$$

The Euler–Lagrange equations are derived by taking the variation of L with respect to the model solution $\zeta(t)$, in the neighborhood of a minimum, and setting the variation to zero, as discussed in the general variational theory section 2.2.2.

$$\delta L = \delta J + \int_0^T \lambda(t) \frac{\partial \delta[\zeta(t)]}{\partial t} dt, \quad (2.19)$$

where δJ represents the variation of the cost function induced by a variation in model solution and is equal to $\int_0^T \mathbf{W}^{-1}(\zeta(t) - \zeta_{obs}(t)) \delta_D(t - t_{obs}) dt$. Note that since (2.16) is linear, only the time derivative term “survives” in the variational equation (2.19).

If we integrate the second term on the RHS of (2.19) by parts we have:

$$\int_0^T \lambda(t) \frac{\partial \delta[\zeta(t)]}{\partial t} dt = [\lambda(t) \delta \zeta(t)]_0^T - \int_0^T \frac{\partial[\lambda(t)]}{\partial t} \delta \zeta(t) dt. \quad (2.20)$$

At the minimum, $\delta L = 0$, and since the variation $\delta \zeta$ is arbitrary, the following equation has to be satisfied:

$$-\frac{\partial \lambda(t)}{\partial t} + \mathbf{W}^{-1}[\zeta(t) - \zeta_{obs}(t)] \delta_D(t - t_{obs}) = 0 \quad (2.21)$$

subject to boundary conditions

$$\lambda(0) = \lambda(T) = 0 \quad (2.22)$$

Equations (2.21) and (2.22) represent the *adjoint* system. The adjoint equation (2.21) is linear in $\lambda(t)$, consistent with the definition of adjoint operator, and it involves a *backward* integration in time (the partial derivative with respect to time is in fact multiplied by a negative sign). The term including the observations, $\mathbf{W}^{-1}(\zeta(t) - \zeta_{obs}(t)) \delta_D(t - t_{obs})$, represents the derivative of the cost function with respect to $\zeta(t)$, and acts as an external forcing in the adjoint equation. The adjoint solution is dependent on the strength of this forcing, i.e. on the weighted difference

between model prediction and measurements. Note that, had the starting model equation for the vorticity not been linear, linearization would have been necessary in order to derive the adjoint equation. This linearization is always implicit in the variational procedure since, to write δL , we apply a small perturbation to the model solution, and we retain only the first order (linear) term, neglecting higher order ones.

Next step is to show explicitly that the adjoint solution at time $t = 0$ represents the gradient of the cost function with respect to the initial condition as shown for the general case in section 2.3.2. Consider the variation of the Lagrangian with respect to the initial condition on $\zeta(t)$:

$$\delta L = \frac{\partial L}{\partial \zeta(0)} \delta \zeta(0). \quad (2.23)$$

Assuming the Lagrangian is not at its minimum and that the adjoint equation (2.21) is still satisfied, (2.19) reduces to

$$\delta L = [\lambda(T) \delta \zeta(T) - \lambda(0) \delta \zeta(0)]. \quad (2.24)$$

By requiring that the adjoint solution is zero at time $t = T$, i.e. $\lambda(T) = 0$, and by taking the partial derivative of (2.24), it is easily seen that the adjoint solution at initial time represents minus the gradient of the Lagrangian with respect to the initial condition, i.e.

$$\frac{\partial L}{\partial \zeta(0)} = -\lambda(0). \quad (2.25)$$

The gradient provided by the adjoint solution is used to find the optimal initial condition, i.e. the initial condition that ensures the forecast of potential vorticity will be as close as possible to the observed values. There are a few optimization methods which make use of the gradient information⁸ to optimize a quadratic (or semi-quadratic) function. The most commonly used are Steepest Descent⁹ (SD)

⁸For highly nonlinear multidimensional problems, gradient methods are ineffective, and alternative methods, such as *simulated annealing* (Kirkpatrick et al., 1983) and *genetic adaptive* (Whitley, 1994) algorithms should be used.

⁹At present, the Steepest Descent method is more important as a theoretical rather than practical reference by which to test other methods. However, steepest descent “steps” are often incorporated into other methods.

(Luenberger, 1984) and Conjugate Gradient (CG) (Gilbert and Nocedal, 1992) algorithms. The principle on which both these methods are based is that the gradient direction (or the conjugate direction) is the one that allows to move in an “optimal” way toward the minimum. A mathematical proof of this, using the Lagrangian least action principle, is presented in Vukićević (2000). Other optimization methods currently used are Newtonian methods which are based on the necessary condition of zero gradient at the optimum point. An overview of this type of algorithms can be found in Jr. and Schnabel (1983). Regardless the specific choice of the method, the search for the optimal point is generally carried out through an iterative procedure. At each iteration, the new value of the control variable is updated as follows

$$\zeta_{n+1}(0) = \zeta_n(0) - \sigma_n \lambda(0) \quad (2.26)$$

where n is the iteration number, and σ_n is a variable stepsize¹⁰. As mentioned earlier, this approach to estimating the gradient is effective, since only one adjoint integration is required to obtain the gradient. The extension to the use of adjoint in linear sensitivity studies, i.e. to investigate model response to variations in initial and boundary conditions and/or model parameters, is straightforward.

Utility of the adjoint solution is further explored in the remainder of this section in which we address the problem of optimal parameter estimation using the variational technique. Consider the variation of the Lagrangian induced by a variation of the parameter α :

$$\begin{aligned} \delta L = \delta J &+ \int_0^T \lambda(t) \frac{\partial}{\partial \zeta(t)} \left(\frac{\partial \zeta(t)}{\partial t} + Q\alpha^2 t e^{-\alpha t} \right) \delta \zeta(t) dt \\ &+ \int_0^T \lambda(t) Q t^2 \alpha e^{-\alpha t} (2 - \alpha t) \delta \alpha dt. \end{aligned} \quad (2.27)$$

Assuming again that the adjoint equation holds and its solution is known, the variation of L with respect to α is

$$\frac{\partial L}{\partial \alpha} = \int_0^T \lambda(t) Q t^2 \alpha e^{-\alpha t} (2 - \alpha t) dt. \quad (2.28)$$

¹⁰The stepsize can also be optimized to have faster convergence, using the variational procedure. This requires two model evaluations, and can be expensive computationally.

If the spatial component of the mass source/sink term is known, we can integrate (2.28) to obtain the gradient $\frac{\partial L}{\partial \alpha}$ and find iteratively the optimum value of the parameter α :

$$\alpha_{n+1} = \alpha_n - \sigma_n \frac{\partial L}{\partial \alpha} \quad (2.29)$$

This simple example¹¹ illustrates the power of the variational technique whose mathematical beauty and generality make it applicable to a broad range of physical problems. In chapter 4 an application of adjoint and variational concepts to the study of cirrus clouds with a microphysical model describing the cloud time and vertical evolution is presented.

2.3.3 The observational operator (mapping)

In section 2.2.2 we introduced the notion of the observational operator, $P(x)$. This operator can be thought of as a mapping from the model space into the observational space¹². The equation that defines this mapping is

$$y_{obs} = P(x) + \epsilon_y, \quad (2.30)$$

where ϵ_y represents the observation error, which is the sum of the instrumental error and the representativeness error. The latter is due, for example, to discretization and/or interpolation, or to errors in the observational operator itself. A detailed discussion of errors is given in section 2.4.

There are many types of observational operators. For meteorological variables such as specific humidity and temperature, which are commonly carried as prognostic or diagnostic variables in any model, if the observations are co-located with the model grid-points, $P(x)$ is equal to the identity matrix. If observations are

¹¹The example presented here is an adaptation of a similar problem discussed in the class notes of the course on Data Assimilation held by Dr. Tomislava Vukićević at Colorado State University, Department of Atmospheric Science in Spring 2000.

¹²In optimum estimation retrievals, this is also called the *forward model* (Rodgers, 1976). Here we adopt the DA terminology to avoid confusion with the *dynamical model* \mathcal{M} introduced in section 2.2.2.

available at locations that are noncoincident with model points, an interpolation, usually linear must be performed. The interpolation operator belongs to the category of observational operators. For variables that are not directly carried in the model, such as satellite radiances or radar reflectivities, the observational operator is represented by a radiative transfer model or by a radar return model. These are examples of non-linear operators. Under the assumption that these operators are weakly nonlinear, a least square analysis can be still performed by considering the Jacobian of the mapping (tangent linear approximation) (Courtier, 1997; Bouttier and Courtier, 1999). Variational techniques allow also for the application of (highly) nonlinear operators, although the existence of multiple minima for the cost function can prevent convergence to the true absolute minimum, and the optimal solution might not be representative. In an operational context, to avoid the problem of an expensive search for the minimum, the nonlinear observational operators are linearized (e.g., Janiskova et al., 1999), and an incremental technique is implemented for which only a first order correction to the background is computed using the variational approach, and the analysis is updated using this correction (Mahfouf and Rabier, 2000).

In section 4.3.4, the observational operator (radar reflectivity mapping) chosen for the assimilation project is described and discussed. Another example of an observational operator is provided by the radiative transfer model introduced in appendix C.

2.3.4 The background

The background¹³ can be provided either by results from a previous model forecast, as it is in NWP models, or from climatological records. If the background is prescribed from a previous model forecast, biases and random errors in the forecast model have to be characterized in order to define the background error covariance matrix that appears in the definition of the cost function (2.5). Background errors are usually correlated both in time and space. A common assumption is that

¹³In retrieval theory, the background is also known as *a priori*.

the background error is stationary in time, and some parameterized form of the background error covariance matrix can be described as a function of the spatial coordinate only (Daley, 1991). The task of computing the background error covariance matrix is challenging. For standard meteorological variables, such as temperature and specific humidity, forecast background errors have been subject to accurate analysis. The general approach in operational models is to estimate error covariances based on statistical comparisons with observations (Hollingsworth and Lönnberg and Lönnberg and Hollingsworth, 1986).

However, if the analysis is performed with respect to a variable for which there is little background information available, such as is the case for cloud or precipitation fields, and this information is not reliable, it is safer to assume that the error covariance is infinite, i.e. $\mathbf{B}^{-1} = \mathbf{0}$, and just perform the optimization with the information provided by the measurements. The problem with this approach is that convergence could be slow if achievable at all, especially if the discrepancy between model fields and measurements is large. In general, the inclusion of a background term ensures more stability and efficiency in the convergence of the optimization algorithm, and it might be necessary when the problem is seriously undetermined due to the lack of observations. When some confidence can be put in the background fields, the results of the optimization of (2.5), with the background term included in the definition of the cost function, might reflect this confidence to the extent that the optimized fields will not depart too largely from the background, and the relative importance of the measurements will be decreased. This aspect of the DA system is controlled by preassigning a background error covariance matrix, and an observation covariance matrix of given relative magnitude with respect to each other. Implications on the results of the analysis are further examined in the following section.

2.4 The importance of error statistics

To the concept of error, a negative meaning such as in “mistake” is often associated. Errors are indeed an admission that we don’t know the “true” value of a physical

variable, and our representation of it, either through a model or through a measurement, is necessarily approximate. However, errors can also be regarded as a measure of how good our knowledge of a certain phenomenon is, and as an instrument to improve this knowledge.

In any assimilation activity, the definition of error statistics is important for a real improvement of the analysis through inclusion of measurements. If we specify observational errors that are much larger than, for example, the background errors, the influence of the latter will be disproportionate, and no real improvement with respect to the background will be seen in the analysis. This is especially undesirable if the background is not capturing certain aspects of the physical phenomenon that are instead present in the observations¹⁴. Vice versa if the background is assigned a large error, the correction will only depend on the measurements, and if observations are uncalibrated or biased, this can lead to an unrealistic departure from the background, and to a poor analysis. An accurate error analysis is especially important when assimilating fields that are not directly observed, but are *retrieved* from the measurements. Retrieved variables may have additional biases introduced by specific assumptions used in the retrieval. A discussion of the implication of using retrieved quantities as observations is found in Marécal and Mahfouf (2000), where the authors explore the sensitivity of assimilation results to specification of errors in surface rainfall rates from the TRMM satellite Microwave Imager. In general, it is preferable to assimilate measurements directly, although this may add the further complication of having to specify a complex observational operator. The advantage is that potential biases are reduced because there is no intermediate forward model to characterize.

Another important type of error is the error in the forecast model itself. In Kalman filter theory, the model error is explicitly included, but the high computational cost of this technique limits its applicability to operational systems. In 4D-Var

¹⁴A classical example of this is the delay in the observational proof of the ozone depletion over Antarctica, known as the “ozone hole”. Observations which were showing this depletion were discarded because they were not in accordance with climatological values used as background information!

systems this error was historically neglected. However, the proven importance of accounting for it has prompted the development of state-of-the-art variational systems that include the evaluation of model errors. Inclusion of model errors in 4D-Var assimilation has been explored by Zupanski (1997); in her study, the model is used as a *weak* constraint and model error is divided into a systematic and a random component (first-order Markov variable). The latter is used as a control variable and estimated as part of the assimilation process; the former is assigned as a function of the initial model bias, and its value at a given time depends only on the value at a previous time. Results from this study, show that relaxing the assumption of the model as a strong constraint improves the performance of the assimilation system, and the quality of the assimilated fields, at a reasonable computational cost. The constant increase of computing power and speed might render feasible the inclusion of model errors even in operational 4D-Var setups.

2.4.1 The modeling of background and observational errors

In general, errors are assumed distributed according to a Gaussian function, with expected value equal to zero (unbiased observations and background) and assigned variance. If there is a bias¹⁵ which can be estimated and removed, the assumption of unbiased fields can still be satisfied. Dee and Da Silva (1998) show how forecast bias can be estimated as part of the data assimilation process, based on an unbiased subset of the observing system. In chapter 3, an example of how to estimate biases in cirrus simulations from using actual radar observations is shown.

Error statistics for multidimensional systems are expressed in terms of error covariance matrices. The diagonal of this matrix contains variances for each model variable; the square root of the variances represent the standard deviations. Off-diagonal elements are the error covariances and represent the cross-correlations between each pair of variables in the model. In general, observations are assumed uncorrelated, i.e. the error covariance matrix \mathbf{W} , introduced in (2.4), is assumed

¹⁵Biases in observations may arise from incorrect calibration, whereas biases in background fields may be due to inherent model features related to parameterizations.

diagonal. Although commonly used, this assumption might not be true for observations that are taken from the same platform (radiosonde, aircraft or satellite) or when several successive reports from the same station are used. Observation preprocessing can also generate artificial correlations between transformed variables (for example when converting specific to relative humidity a correlation with temperature is introduced) or when a retrieval procedure is applied to data. Errors in the observational operators can also introduce correlations: this is definitely true for interpolation operators, especially when the spatial distribution of observations is on a finer scale than model resolution. For nonlinear observational operators, errors and error correlations can be due to the specific approximations used. For example, in radiative transfer modeling, a common approximation is to treat the atmosphere as plane-parallel as opposed to three-dimensional; while this may be acceptable for radiative transfer in a clear sky, it might not be right in radiative transfer through clouds (due to 3D effects and inhomogeneities). The presence of correlations between observations is not necessarily a negative feature. However, since such correlations are hard to estimate, it is general procedure to neglect them.

Background biases and error covariances are also hard to quantify because they may depend on many different factors that are case-dependent (i.e. background biases may vary with latitude and time of the year). If the model forecast is used as background, then this is almost equivalent to estimating model error covariances, very much like is done in a Kalman filter approach. In general, a crude estimate is made by taking an arbitrary fraction of the climatological variance. A better estimate is obtained by comparison with an analysis of good quality, or using an observational method as described in Hollingsworth and Lönnberg (1986). Ideally, background errors should be flow-dependent since they are expected to vary with the weather situation. This can be obtained in a Kalman filter approach or in 4D-Var with the model treated as a weak constraint. A correct specification of background errors is important not only for the goodness of the analysis but also because correlations in background errors spread information in data sparse areas, and contain information about model balance properties. A detailed account of these properties is given in Bouttier and Courtier (1999).

Chapter 3

Cirrus cloud modeling

3.1 Current issues in cirrus modeling

Cirrus clouds represents a challenge for modelers. In the past years, great improvements have been achieved in ice cloud representation both in large-scale and cloud-scale models. Prognostic cloud schemes in NWP models have been proven successful in qualitatively representing cirrus cloud cover associated with large-scale forcings (Miller, 2000; Mace et al., 1998b). Quantitative evaluation of microphysical cloud parameters is a topic of ongoing research. Climate models with prognostic ice schemes appear to give more realistic simulations of the current Earth's climate (Fowler et al., 1996), which is the premise for any type of inference on climate change. A number of regional two-dimensional and three-dimensional models include rather complex microphysical schemes to describe all relevant interactions between ice species and the other hydrometeors (Cheng et al., 2001). Improved radiation schemes are also an important element of modeling activities related to cirrus, due to the recognized importance of atmospheric ice amount and distribution in the atmospheric radiative budget, and the link of the latter to global precipitation and evaporation cycles (Wu et al., 2000). Despite this effort in advancing the representation of cirrus processes in numerical models, large discrepancies still appear when contrasting cirrus simulations with observations. Similarly, comparisons of cirrus simulations obtained with different models show a large spread in results. This points to the need for more work in improving and accurately testing model parameterizations related to the description of the ice phase.

Possibly the most extensive combined cirrus modeling effort of the last years is the one carried under the auspices of the GEWEX Cloud System Study program

(GCSS, GEWEX is the Global Energy and Water Cycle Experiment). This program counts four working groups that focus on different cloud systems (Boundary Layer, Cirrus, Extra tropical Layer, Precipitating Deep Convective). Many different models are involved in the project, with the mission of validating and developing new parameterizations related to cloud processes to be included in climate models. To achieve this goal, emphasis is put on gaining a better understanding of physical processes at work within the different types of cloud systems. The inter-comparison of high resolution Cloud Resolving Models (CRMs) with Single Column Models (SCMs) run for the same cloud scenarios, and the evaluation of both using observations is an integral part of the GCSS strategy.

One of the GCSS working groups focuses on Cirrus Cloud Systems (WG2). The main activity of the group has involved model intercomparison for a series of idealized cirrus simulations with relatively simple initial conditions and large-scale forcing coordinated by Dr. D. Starr. Results from sixteen models, ranging from 3-D Large Eddy Simulation (LES) models to 2-D CRMs and 1-D SCMs models, including the CSU RAMS, are presented and discussed in Starr et al. (2000). Here, a brief summary of results relevant to the discussion of current issues in cirrus modeling is presented. Amongst the participating models, two broad categories were identified: models that were primarily developed as cirrus models versus models originally developed to treat convective systems. As far as microphysical treatment, the major distinction was between models with bin-microphysics versus models with bulk-microphysics schemes¹. In the results, some grouping can be seen for models

¹Cloud scale models can be divided into two categories according to the treatment of microphysical processes: **bin microphysics models** in which the evolution of hydrometeor size and number concentration is described by dividing the particle spectra in size *bins*; **bulk microphysics models** in which an analytical size distribution is introduced and used to parameterize microphysical processes providing an integrated description of the cloud *bulk* properties. The bulk microphysics formulation is less accurate since it relies on the assumption about the hydrometeor size distribution, but it is usually preferred in CRMs because it is computationally more efficient. A degree of sophistication in bulk microphysical schemes is introduced by using prognostic equations for both mixing ratio and number concentration (second moment bulk microphysics). For details on the

with common heritage, but a large spread exists between ice field values from the various models. Figure 3.1 from Starr et al. (2000) shows the time-behavior of the horizontally-averaged Ice Water Path (IWP), which represents the vertically-integrated IWC.

Results from models built primarily to be cirrus models or with a strong cirrus heritage are shown by the heavy dashed (bulk microphysics) or heavy solid (bin microphysics) lines. Thin solid lines correspond to results from models originally developed to treat deep convective cloud systems and thin dashed lines correspond to results from SCMs. Specific models are not identified. It appears evident that the model predicted IWP can differ from model to model up to one order of magnitude, especially in the “cold” cirrus case.

Another measure of model response is the vertical positioning of the cloud (i.e. cloud top and base). Figures 3.2 and 3.3 shows the distribution of cloud top and cloud base and corresponding cloud thickness for the warm and the cold cirrus cases respectively. Once again the spread is noticeable and can be traced back to different treatments in ice crystal fall speed parameterizations, although differences in other model parameterizations are likely to play a role.

Based on results from the intercomparison activity some key points are identified as necessary for advancement in current cirrus cloud modeling:

- The necessity for a better understanding of physical processes at work in formation and evolution of cirrus leading to more accurate and physically based parameterizations.
- The necessity for direct observations of ice cloud parameters, such as ice particle fall speed and IWC.
- The necessity for a procedure to compare cirrus observations and model results quantitatively with the goal of evaluating newly developed parameterizations.

Recent progress in cloud observing capabilities on one hand, and computing resources to run models at high spatial and temporal resolution on the other, will subject, see Cotton and Anthes (1989).

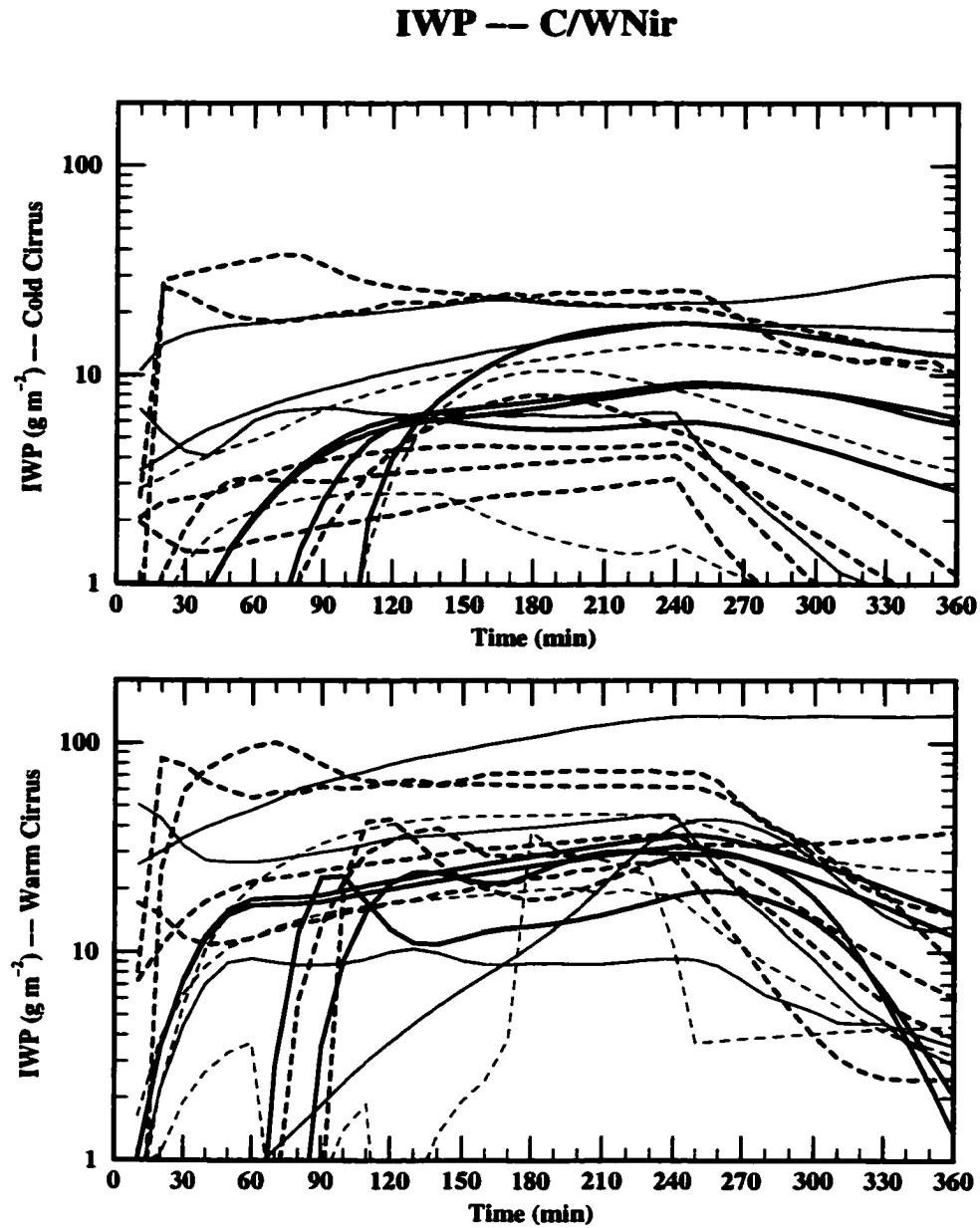


Figure 3.1: IWP (gm^{-2}) versus time for simulations of “cold” (upper panel) and “warm” (lower panel) with 16 cloud models (see text for explanation). Figure courtesy of Dr. D. Starr.

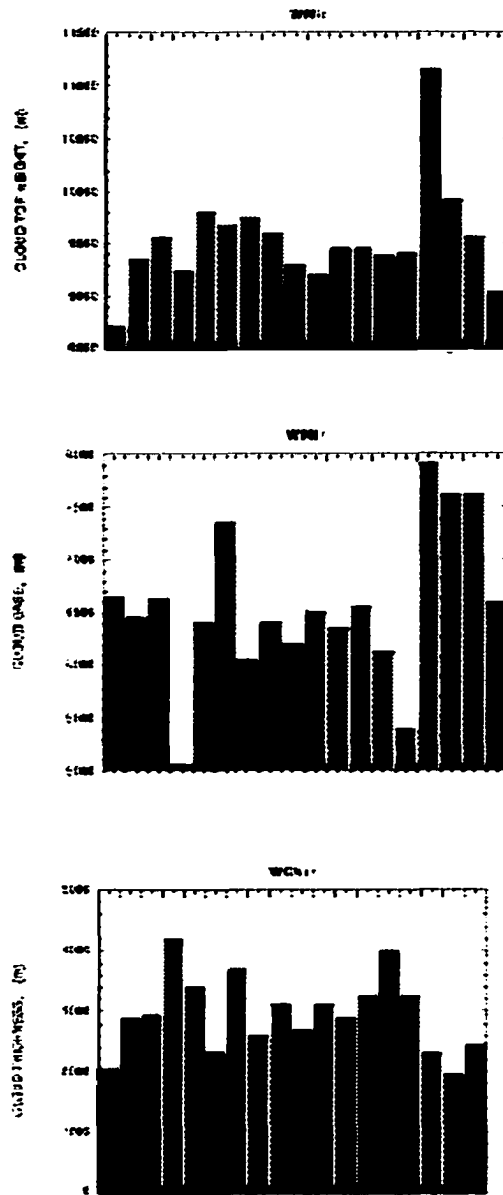


Figure 3.2: Warm cloud top height (upper panel), cloud base height (center panel), and cloud thickness (lower panel) given by various models. Figure courtesy of Dr.D.Starr.

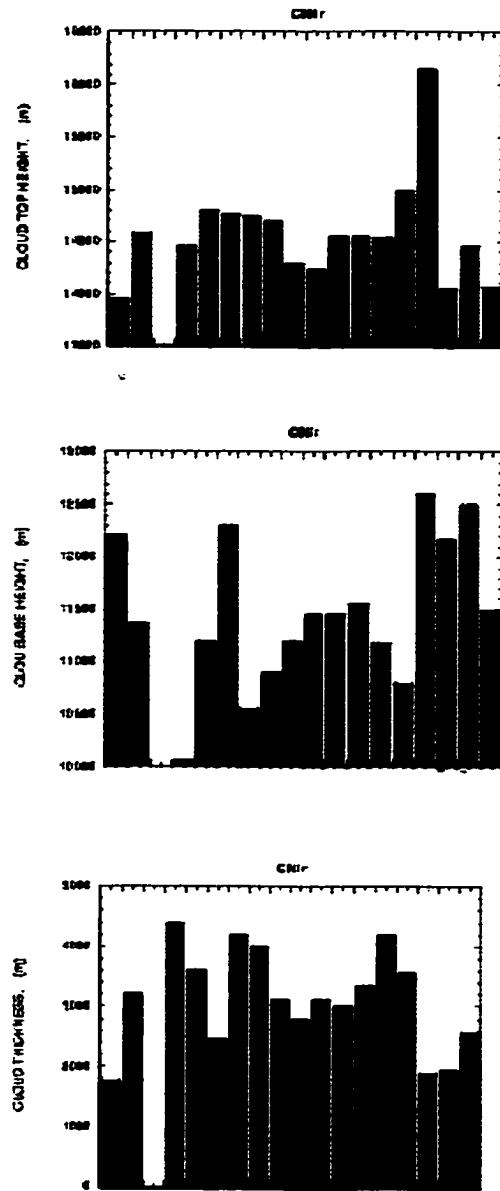


Figure 3.3: Cold cloud top height (upper panel), cloud base height (center panel), and cloud thickness (lower panel) for the various models. Figure courtesy of Dr.D.Starr.

allow these issues to be addressed in the near future.

3.2 Source of uncertainties in mesoscale models

Limited-area models are intrinsically different from global models for a number of reasons. While large-scale NWP models solve the atmospheric dynamics equations on the sphere, hence having a natural “closure” condition, mesoscale models require lateral boundary conditions that have to be specified externally. Based on this premise, three main sources of model errors in regional modeling can be identified:

- errors related to intrinsic model parameterizations (i.e. the use of a parameterization derived from laboratory or field experiments to describe real phenomenon at all latitudes and time of the year);
- errors related to description of physical processes (i.e. the use of approximations that do not apply to the specific system under examination, such as hydrostaticity, etc.);
- errors in externally specified initial or boundary conditions.

All three types of errors are reflected in both systematic and random errors in model prediction. The hardest to quantify are errors in model parameterizations and approximations to physical processes, since their quantification involve intensive comparisons with direct observations (which are not always available) and the dismembering of the model to pin down specific “faulty” parameterizations. Sensitivity studies can also help understand potential sources of model errors.

Errors in initial and boundary conditions, when specified for example from another model, are also generally unknown, and so is the impact that erroneous initial and boundary conditions might have on the mesoscale model forecast. More confidence can be placed in boundary conditions specified from observing systems, whose errors statistics are generally known. A general answer to the question of minimizing model errors deriving from “bad” initial and boundary conditions can be found in optimal assimilation of observations to correct them, and to ameliorate the model

performance. This aspect of the problem is explored in chapter 5. In this chapter, certain aspects related to quantification of model errors for RAMS are addressed.

3.3 Assessment of model skill

As discussed in chapter 2, an integral part of any assimilation activity is the understanding of the model behavior and the assessment of its skill in reproducing the physical phenomena. The focus of the research presented in this chapter is on acquiring familiarity with the mesoscale model (RAMS) to be used in the preliminary assimilation studies presented in chapter 5 by running it for a series of cirrus simulations. Previous studies performed with RAMS showed that the model has good skill in representing both thin nonbuoyant cirrus and thick convectively active cirrus (Cheng et al., 2001), and captures main processes operating in cirrus at cloud scales such as radiative/microphysics feedbacks (Wu et al., 2000) .

The approach taken here consists of first testing the model performance under controlled conditions (the WG2 idealized cirrus cases), then using those cases as a benchmark to study model response to changes in descriptions of specific processes (ice growth mechanisms) and parameterizations (ice crystal fall speed). The choice of focusing on these two aspects is based on results from the GCSS–WG2 intercomparison activity highlighted in the previous section. As a final step, an observed cirrus case is simulated to evaluate model skill by direct comparison with measurements. Results from 3D–channel and 3D model runs are presented. A statistical approach is also introduced and applied to results from the 3D simulation to quantify average model bias under the assumption of bias–free observations. An error covariance matrix associated with simulated fields is computed by adapting a procedure to estimate background error covariance used in NWP modeling, and used to identify model strengths and deficiencies. An application of derived error covariance matrices in retrievals of cloud properties from radar measurements is presented in appendix A. Before moving on to the analysis of results, a general overview of the CRM is offered in the following section.

3.3.1 Description of the Cloud Resolving Model

The RAMS model is built upon the full set of non-hydrostatic, compressible atmospheric dynamics and thermodynamics equations. Conservation properties of scalar quantities such as water vapor and liquid and ice hydrometeors are also included. The model includes a number of state-of-the-art parameterizations for relevant processes, such as turbulent diffusion, solar and infrared radiation, moist processes, cumulus convection, kinematic effects of terrain and sensible and latent heat exchange between the atmosphere and the surface. The model also has data assimilation capabilities such as nudging (operational) and 4D-Var (experimental). RAMS' ancestors are a non-hydrostatic cloud model (Tripoli and Cotton, 1982) and a mesoscale model originally developed to study sea breeze circulation (Mahrer and Pielke, 1977). The previous codes have been rewritten to accommodate a higher degree of flexibility, specifically the two-way interactive grid nesting capability of the current version. The model is mostly used as a limited area model, and many of its parameterizations have been designed to simulate mesoscale and high-resolution cloud-scale processes. However, the model can also operate in a global configuration by using two hemispheric grids. The two versions of the model used in this work are RAMS 3b and RAMS 4.2. The main difference between the two versions, which is relevant to cirrus simulations, is the implementation of a two-moment bulk microphysics scheme in RAMS 4.2, which also includes simultaneous computations of heat and vapor diffusion, along with hydrometeor and air temperature and water content. Other features include the incorporation of the computation of potential temperature into the implicit diffusion equations, and the use of pre-computed lookup tables using detailed bin calculations to improve accuracy of several processes including autoconversion (transition from cloud droplets or ice crystals to precipitation-sized particles due to collision and coalescence) and sedimentation. More details on RAMS' development history and present capabilities can be found in the RAMS 4.2 User's Guide (Walko and Tremback, 2000) and in recent paper by citetcottonetal01.

The main prognostic variables are wind components (u , v , w), the ice-liquid

water potential temperature (θ_{il}), the perturbation Exner function², ($\Pi - \Pi_0$), the total liquid water mass mixing ratio, and mixing ratios and number concentrations of different hydrometeor species. From these variables, temperature, potential temperature, vapor mixing ratio and cloud–water mixing ratio are diagnosed. The grid structure consists of a polar–stereographic grid in the horizontal and a terrain–following vertical coordinate ($\sigma_z = \ln(p/p_s)$). Vertical and horizontal resolution are user–defined.

The most current version (RAMS 4.4) includes a second moment bulk microphysics scheme (Walko et al., 1995; Harrington et al., 1995; Meyers et al., 1997), i.e. prognostic equations for both mixing ratio and number concentration for all hydrometeor species, included pristine ice, snow and aggregates. Hydrometeors in each category are assumed to be distributed according to a modified gamma function.

$$n_\gamma(D) = N_t \frac{1}{\Gamma(\nu)} \left(\frac{D}{D_n} \right)^{\nu-1} \frac{1}{D_n} e^{-(D/D_n)} \quad (3.1)$$

described by parameters N_t , total number concentration of ice crystals, D_n , characteristic diameter, and ν , width of the distribution. Assuming a given width of the distribution, and knowing the mixing ratio and the number concentration, the particle characteristic size can be retrieved inverting equation (3.1).

The prognostic equation for the mixing ratios of various hydrometeors, except cloud water and vapor, is:

$$\frac{\partial r}{\partial t} = ADV(r) + TURB(r) + SOURCE(r) - SINK(r) + SED(r), \quad (3.2)$$

where $ADV(r)$ and $TURB(r)$ represent the advective and turbulent transport of r . The $SOURCE(r)$ and $SINK(r)$ terms represent the possible generation or loss of the species r by microphysical processes, such as autoconversion, collisional break–up, secondary ice production due to riming, melting and shedding. A schematic

²The Exner function is defined as

$$\Pi = C_p \left(\frac{p}{p_0} \right)^{R/C_p},$$

where p is the atmospheric pressure and p_0 a reference value. R is the gas constant and C_p the heat capacity of air at constant pressure.

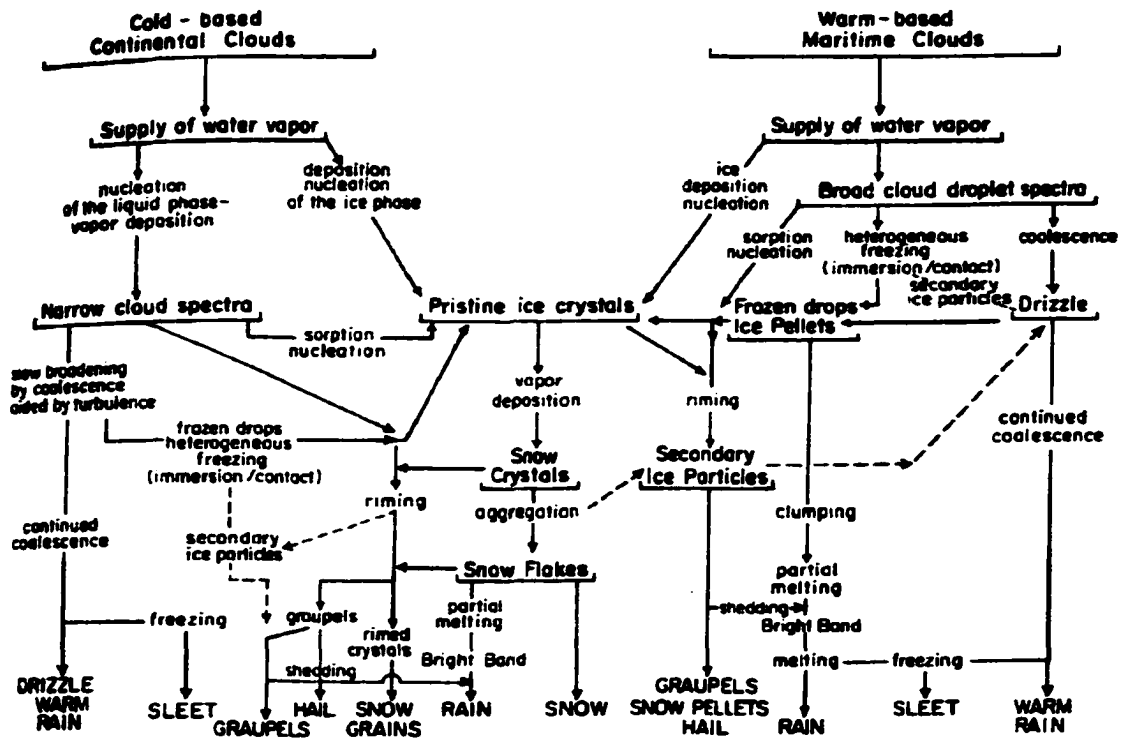


Figure 3.4: Flow diagram describing microphysical processes. From Cotton and Anthes (1989).

of possible interactions between ice and other type of hydrometeors is shown in figure 3.4. $SED(r)$ accounts for local loss or gain of species r due to gravitational sedimentation.

An equation similar to (3.2) is introduced in the second-moment bulk microphysics scheme for the number concentrations:

$$\frac{\partial n}{\partial t} = ADV(n) + TURB(n) + SOURCE(n) - SINK(n) + SED(n) \quad (3.3)$$

Details on parameterizations of specific processes can be found in Walko et al. (1995) and (Meyers et al., 1997). Hydrometeors categories include

- **pristine ice** defined as small unaggregated ice crystals whose main growth mechanism is vapor diffusion;
- **snow** defined as larger unaggregated ice crystals, whose main growth mecha-

nism is vapor diffusion and riming³;

- **aggregates**, whose main growth mechanism is self-explained by their definition;
- **graupel**;
- **hail**;
- **rain**;
- **cloud droplets** .

The threshold diameter for transition from pristine ice to snow is $D = 125\mu m$, but this threshold can be altered . Cloud droplets, which are assumed to be supercooled, and pristine ice are allowed to nucleate only from vapor, while all the other categories form from pre-existing hydrometeors and then grow by vapor deposition and collision-coalescence mechanisms.

Ice nucleation results from four primary processes:

- a combination of vapor deposition and condensation-freezing mechanism for which the number of activated nuclei is derived from the amount of supersaturation with respect to ice (Meyers et al., 1992);
- homogeneous freezing of droplets and haze particle (De Mott et al., 1994);
- contact nucleation for which the number of potential nuclei is derived from an observed temperature dependence;
- secondary ice particle production (Hallett and Mossop, 1974).

The radiation scheme used in this version of RAMS is a two-stream model with 3 bands in the shortwave and 5 in the longwave (Harrington, 1997). The required inputs for the radiation routine are: optical depth, single scattering albedo and asymmetry parameter of the cloud layer, derived from size-spectra of all hydrometeors as predicted in RAMS.

³Growth by *riming* involves the collection of frozen cloud droplets by the growing crystal. Due to its nature, this mechanism is not very important in cirrus cloud evolution.

3.3.2 Selected results from WG2 idealized test cases

The idealized test cases and relative sensitivity studies are performed with a 2D version of RAMS 3b. Profiles of temperature and relative humidity, shown in figure 3.5, were specified on the basis of the idealized “warm” and “cold” cirrus test cases defined by the WG2 intercomparison protocol⁴. To ensure prompt cloud formation, the initial humidity field is supersaturated with respect to ice over a layer about 1 km deep (120% in a 0.5 km layer). Cloud top occurs at about -47°C in the warm case, and -66°C in the cold case, respectively. Environment conditions correspond to Spring/Fall 45°N and Summer 40°N standards with tropopause occurring 1 km above the nominal cirrus layer.

The total length of the simulation was 6 hours. A large-scale forcing was implemented for four hours of the simulation, and turned off for the last two hours, and consisted of an adiabatic cooling ascent consistent with a rate of roughly 3 cm s^{-1} .

Examples of cirrus time evolution as captured by the 2D model are shown in figures 3.6 and 3.7 for the warm and cold cirrus cases respectively. Figures show time snapshots of IWC vertical cross-sections. Some similarities in the two figures are represented by the similar time evolution of the cloud from the initial formation to dissipation stage. The main difference is the smaller maximum IWC (1 order of magnitude) in the cold simulation, and the evolution of the cold cirrus into a homogeneous layer as opposed to the more dynamical evolution of the warm cirrus into separate cells.

More details on cloud and cloud environment structure are shown in figures 3.8 and 3.9. Upper panel plots show horizontally-averaged IWP and vertically integrated number concentration versus time. Despite the fact that the warm cirrus IWP is approximately an order of magnitude larger than the cold cirrus IWP, integrated ice crystal number concentrations are almost comparable in both cases (10^7 crystals per square meter at $t=240$ min into the simulation in the cold cirrus case, and 4×10^7 crystals per cubic meter at the same simulation time in the warm cirrus case), implying that ice crystals are larger in the warm case as opposed to the cold

⁴Details can be found on the web at http://eos913c.gsfc.nasa.gov/gcss_wg2

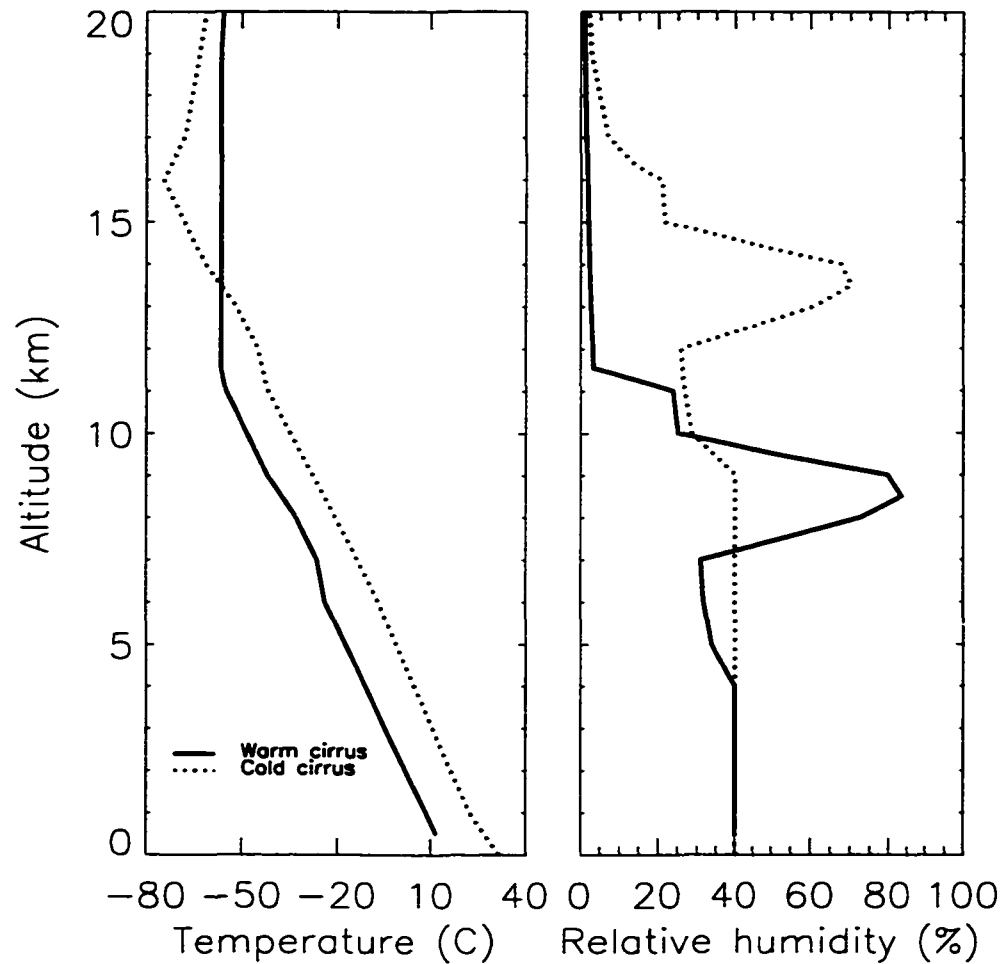


Figure 3.5: Idealized soundings for “warm” (solid line) and “cold” (dotted line) cirrus test cases. Left panel shows temperature in degree Celsius and right panel relative humidity in %. Data is courtesy of Dr. D. Starr.

case. This is to be expected from microphysical considerations since riming and aggregation processes are more efficient at warmer temperatures. Central panels show time series of environmental variables related to cloud dynamics, namely the average zonal (u) and vertical (w) velocities, and their root mean square. The latter is a measure of turbulent activity at cloud levels. The vertical average is performed over a thin layer around cloud top that represents the generation layer. Mean vertical motion at cloud levels is approximately zero in the cold cirrus case. However, a

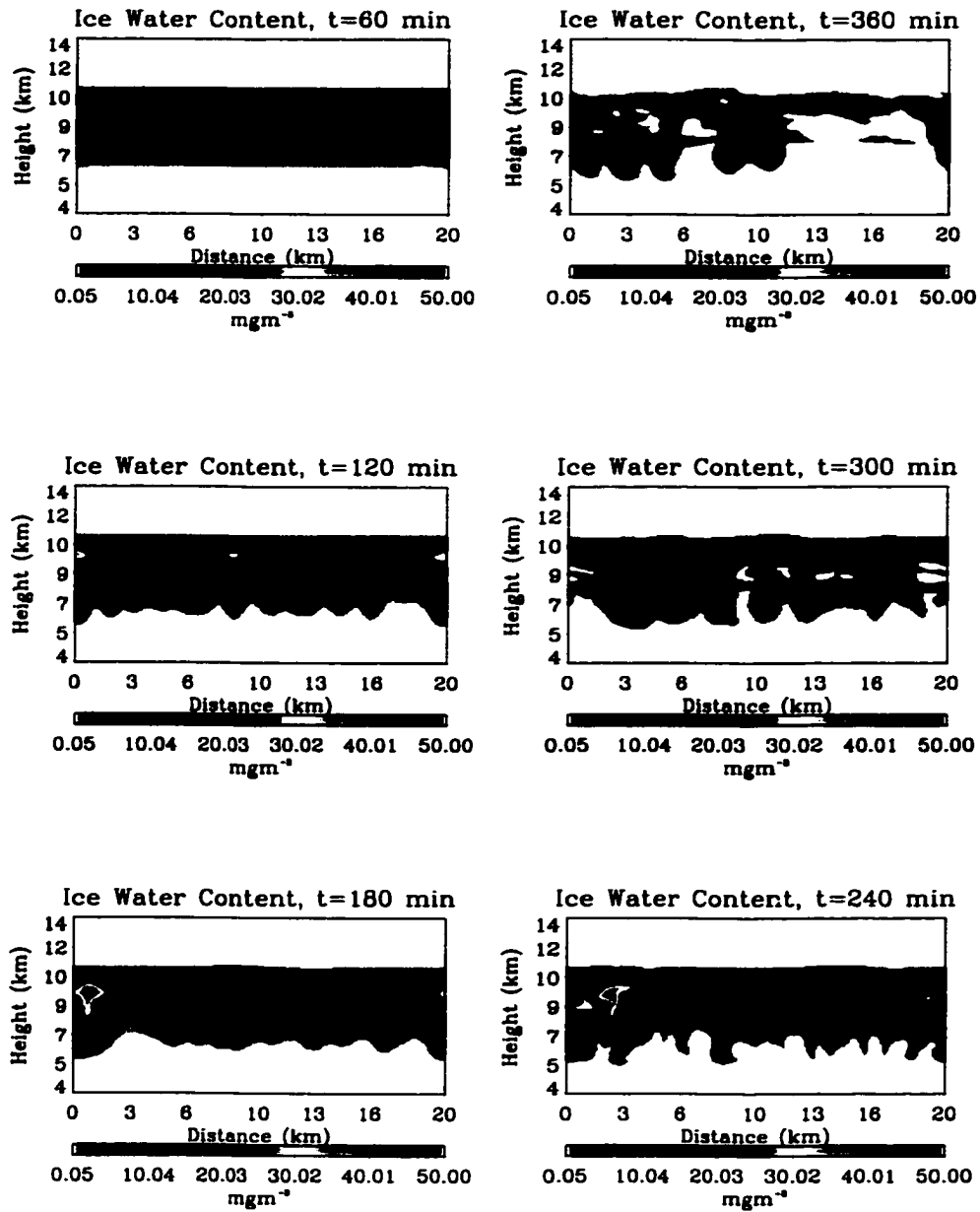


Figure 3.6: Warm cirrus IWC snapshots at various simulation times. Time interval between plots is 1 hour, starting from the top left panel going counterclockwise.

mean vertical velocity of -2 cm s^{-1} is observable in the warm cirrus case (figure 3.8, left middle panel) which seems to persist until the forcing is turned off at $t=240$ min into the simulation. This appears to be an unphysical model response to balance

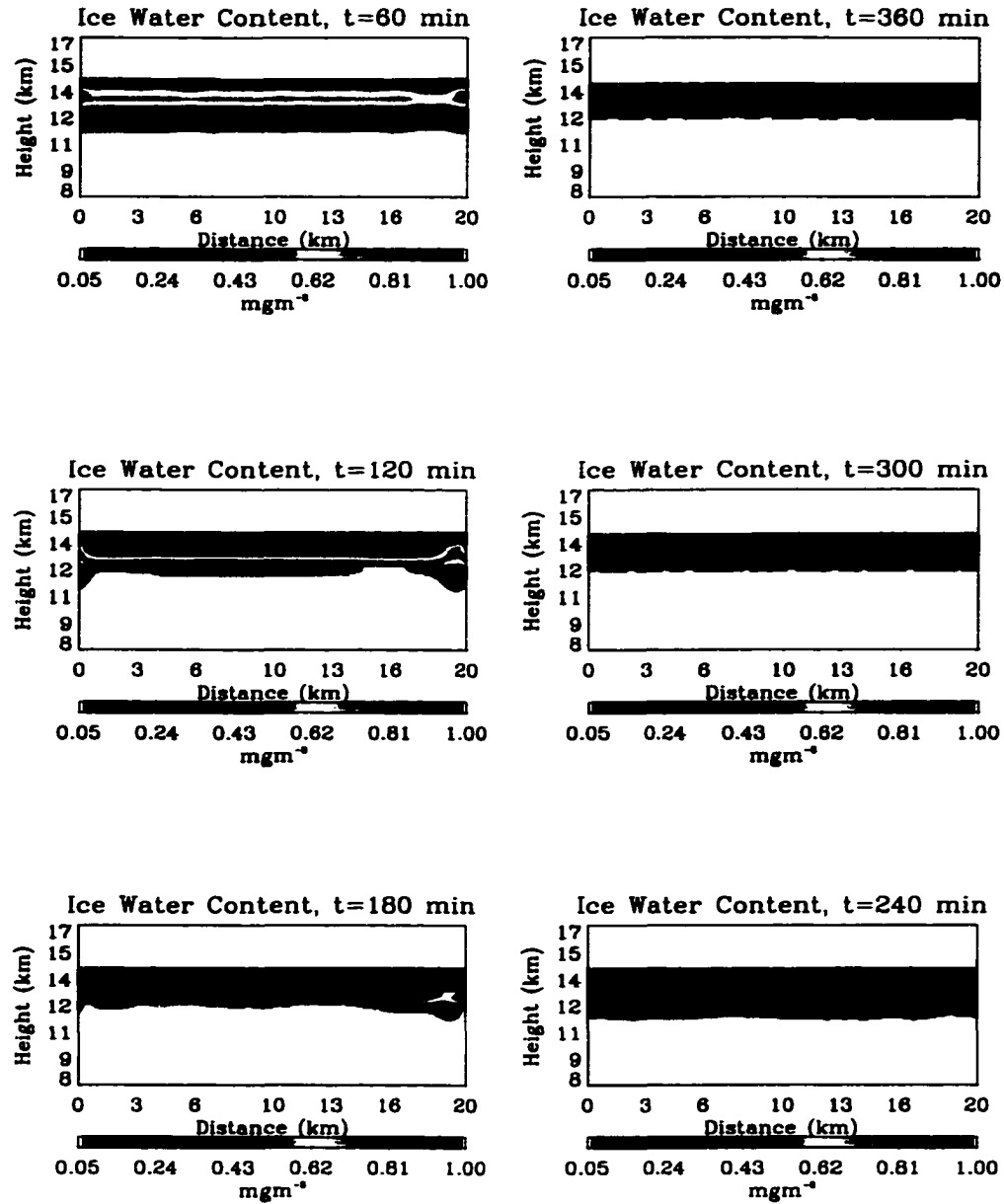


Figure 3.7: Same as figure 3.6, but for the cold cirrus simulations.

the large-scale uplifting which requires further investigation. The root mean square of zonal and vertical velocities shows a relatively active cloud and some turbulence developing at a later stage of the simulation, right after the large-scale forcing is turned off.

Lower panels show in-cloud potential temperature (left) and broadband radiative fluxes at cloud top and base (right). Note the decrease in mean potential temperature due to the superimposed large-scale cooling. Some impact on radiative fluxes due to presence of the cloud is noticeable, especially in the warm cirrus case due to higher IWP: the upward flux at cloud top appears slightly depressed at first due to thickening of the cloud over time⁵, then increases to its original value due to cloud dissipation. The opposite behavior is observed in the downward flux at cloud base.

3.3.3 Sensitivity studies

Sensitivity studies were performed using RAMS version 3b as well in a 2D configuration. Initial temperature, pressure, wind and relative humidity profiles are taken from the idealized warm cirrus test case. The model was run for three hours, and no forcing was implemented. Changes in IWP and cloud base height were considered as a measure of model response (what here it is referred to as model sensitivity) to change in parameterizations, as in the WG2 intercomparisons.

⁵Upward and downward broadband fluxes in the presence of a single cloud layer can be described by

$$\begin{aligned} F^\uparrow(z) &= F^\uparrow(z_b)(1 - \epsilon^\uparrow) + \epsilon^\uparrow \sigma T_c^4 \\ F^\downarrow(z) &= F^\downarrow(z_t)(1 - \epsilon^\downarrow) + \epsilon^\downarrow \sigma T_c^4 \end{aligned}$$

where z_b , and z_t represent cloud bottom and top respectively, and ϵ represents the cloud emissivity. The cloud emissivity can be parameterized in terms of IWP as

$$\epsilon = 1 - \exp(-\beta k IWP)$$

where β is the *diffusivity* and k is an *absorption* coefficient. When the emissivity increases, there is a corresponding decrease (increase) in upward (downward) flux at cloud top (bottom). In the particular case shown, there is the additional effect of a decrease in cloud temperature T_c , due to the superimposed cooling that lowers it over the time interval the forcing is active (Stephens, 1994).

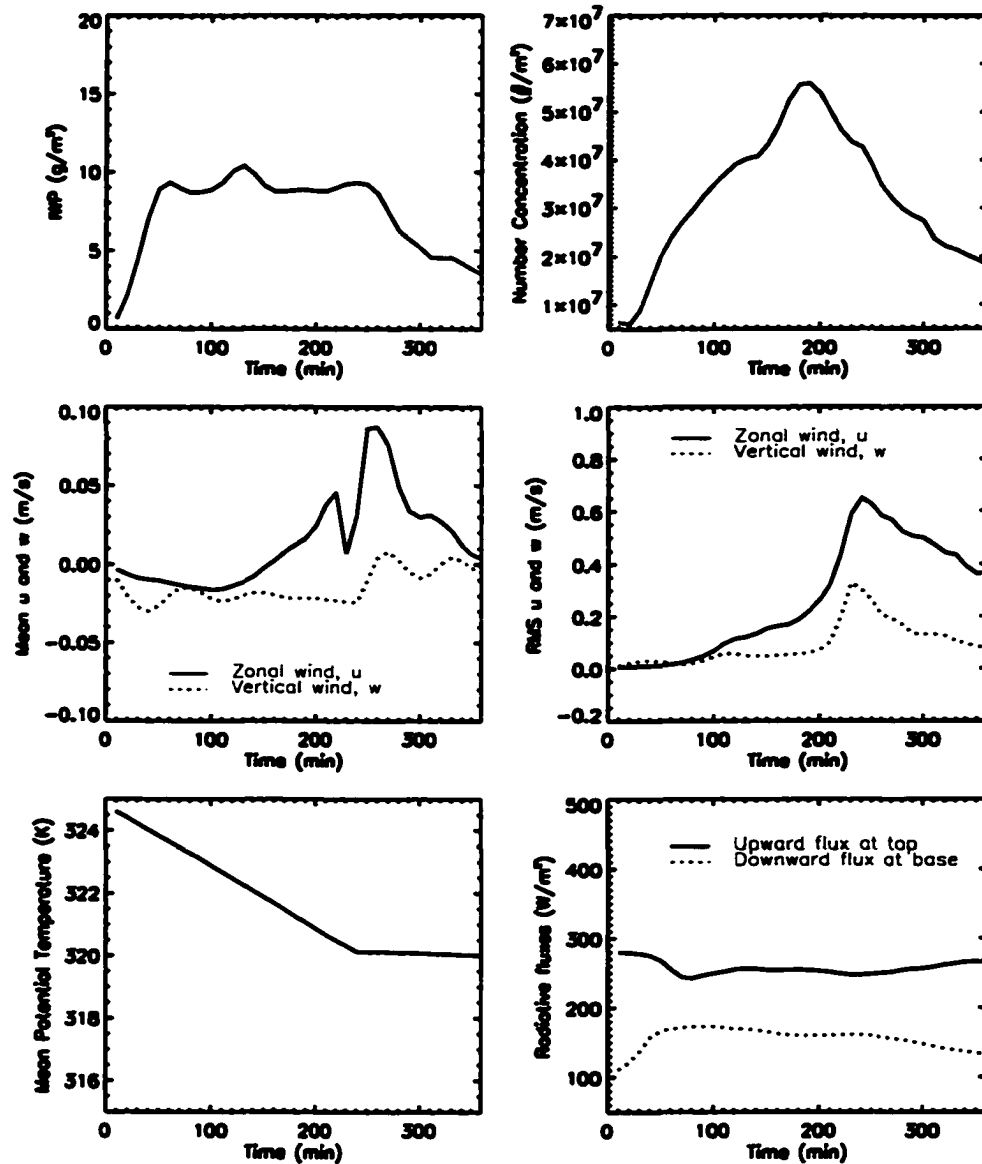


Figure 3.8: Warm cirrus statistics. Upper left panel shows IWP (gm^{-2}) as a function of time (min); upper right panel shows vertically integrated particle number concentration ($\#\text{m}^{-2}$); left central panels shows mean zonal (solid line) and vertical (dotted line) velocities (m/s) and right panels shows root mean square zonal (solid line) and vertical (dotted line) velocities vertically averaged over cloud generation layer; bottom left panel shows environment potential temperature (K), and bottom right panel shows upward radiative flux (Wm^{-2}) at cloud top (solid line), and downward radiative flux at cloud base (dotted line).

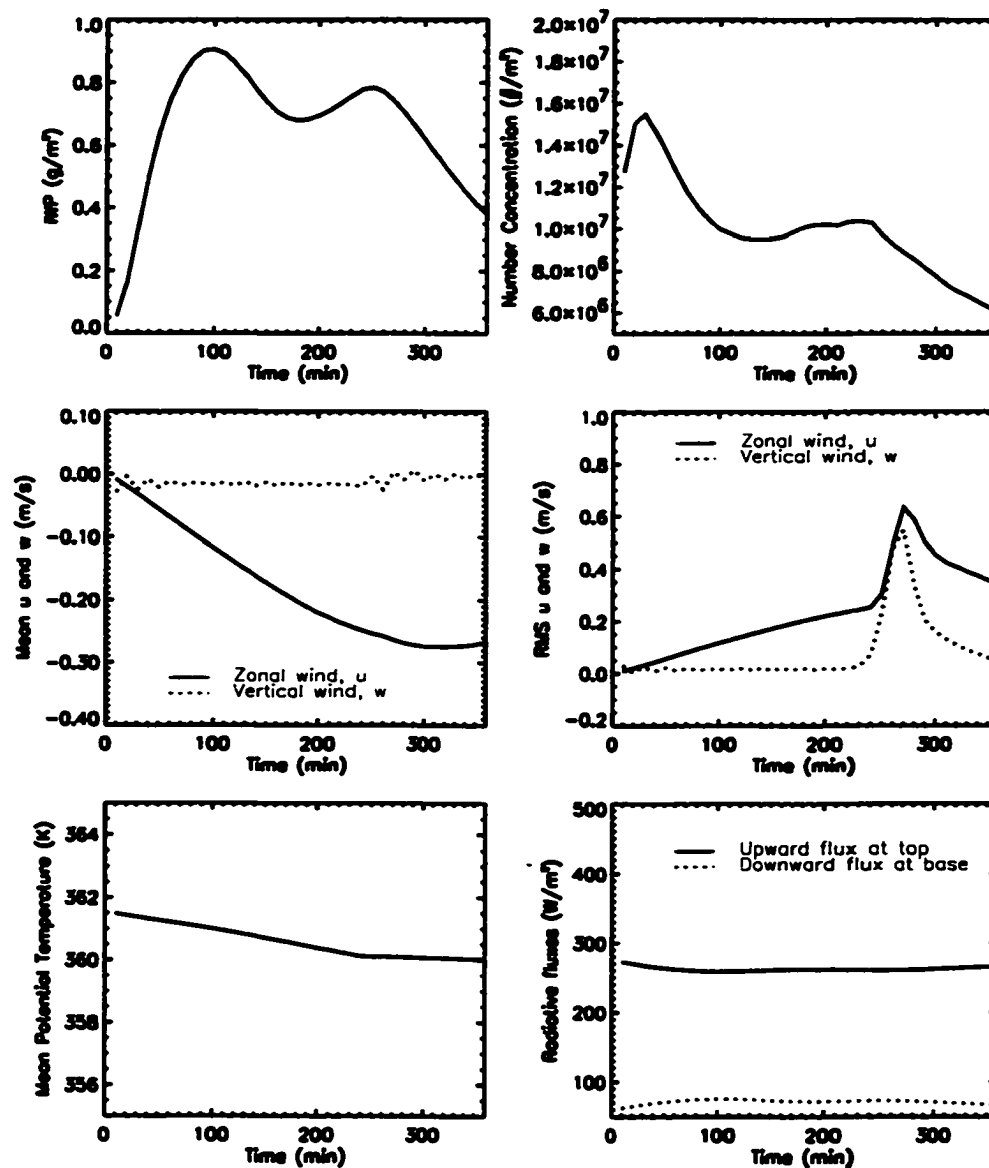


Figure 3.9: Same as figure 3.8, but for “cold” cirrus.

Sensitivity to ice microphysics

In the RAMS, the user can select the hydrometeor categories of interest, and turn off computations relative to other ones. This is done on a case-by-case basis to save computational time. For example, in a cirrus simulation it is unlikely to have formation of graupel or hail, since they are more likely associated with deep convec-

tive systems. The cirrus simulation will be faster if computations relative to those categories are not performed.

In what follows it was assumed that ice categories relevant to cirrus formation and evolution were pristine ice, snow and aggregates. Sensitivity to the inclusion/exclusion of these three categories was investigated. This corresponds indirectly to studying the influence and relative importance of various growth mechanisms. For instance allowing the cirrus cloud to only be constituted of pristine ice crystals and snow effectively imposes that only vapor depositional and riming growth are active, and that only conversions between these two categories of *unaggregated* crystals are possible, while aggregation growth is totally inhibited. This is an assumption which might be closer to reality in cirrus clouds developing at colder temperatures and less realistic in cirrus forming in warmer environment. Nevertheless, it is interesting to understand the impact of assumptions made in the choice of model configuration, and how those assumptions affect the final results. The following model configurations are considered:

- all relevant ice categories (pristine ice, snow and aggregates – baseline case)
- pristine ice only (no snow or aggregates)
- pristine ice and snow (no aggregates)
- pristine ice and aggregates (no snow)

A run was made to confirm that the assumption that no other ice typology such as hail or graupel formed during the cirrus simulation. Those ice categories were found to have no relevance in cirrus composition.

Figure 3.10 shows a time series of IWP for runs including different ice categories. The run with pristine ice only presents an IWP that increases steadily with time and tends to asymptote toward the end of the simulation. The run with aggregates and pristine ice shows a similar behavior, but the IWP starts to flatten out at an earlier time, and then slightly decreases. The IWP associated with the baseline case and the case with no aggregates shows a different trend; the initial increase is not as large as in the previous two cases, and the subsequent decrease in time is more

marked. Assumptions about ice microphysical processes have a direct influence on total ice production and vertical distribution. When large crystals, both aggregated and unaggregated, are inhibited from forming, the removal by sedimentation is less effective, and the IWP of the cloud is larger (i.e. the ice "hangs" around longer). When the aggregation process is activated (pristine+aggregates case), there is a decrease in IWP. However, due to the absence of large crystals (snow), the aggregation process is not as efficient. The case with no aggregates shows that the principal agent of sedimentation removal is snow. The fact that the baseline case shows similar IWP indicates that the aggregates mixing ratio is much lower than snow+pristine ice mixing ratio, and aggregates contribute only slightly to sedimentation removal. These results are pertinent to the specific cirrus case under consideration. In general the impact of inclusion/exclusion of ice species might have a different outcome depending on the condition of cloud formation and the external environment. However, these results are relevant in showing the impact on cloud properties of assumptions regarding microphysical processes.

This impact is also illustrated in the plot of cloud base height at various times during the simulation (figure 3.11). Cloud base is much higher for the pristine ice only case, consistent with the lower fall speed of small ice crystals. This should imply a lower IWP, since IWP is proportional to cloud geometrical thickness. However, ice removal by sedimentation is less efficient in this configuration, such that the total IWC at cloud levels is actually higher, compensating for the lower geometrical thickness.

Sensitivity to fall speed parameterizations

Further studies of the model response to different choices in parameterizations are presented in this section. The interest in fall speed parameterizations is motivated by the GCSS-WG2 intercomparison results, since there is a general consensus that the spread in fall speed parameterizations is responsible for major discrepancies between models. A survey conducted by P. Brown revealed that there are many different fall speed parameterizations currently used in NWPs, GCMs, and CRMs.

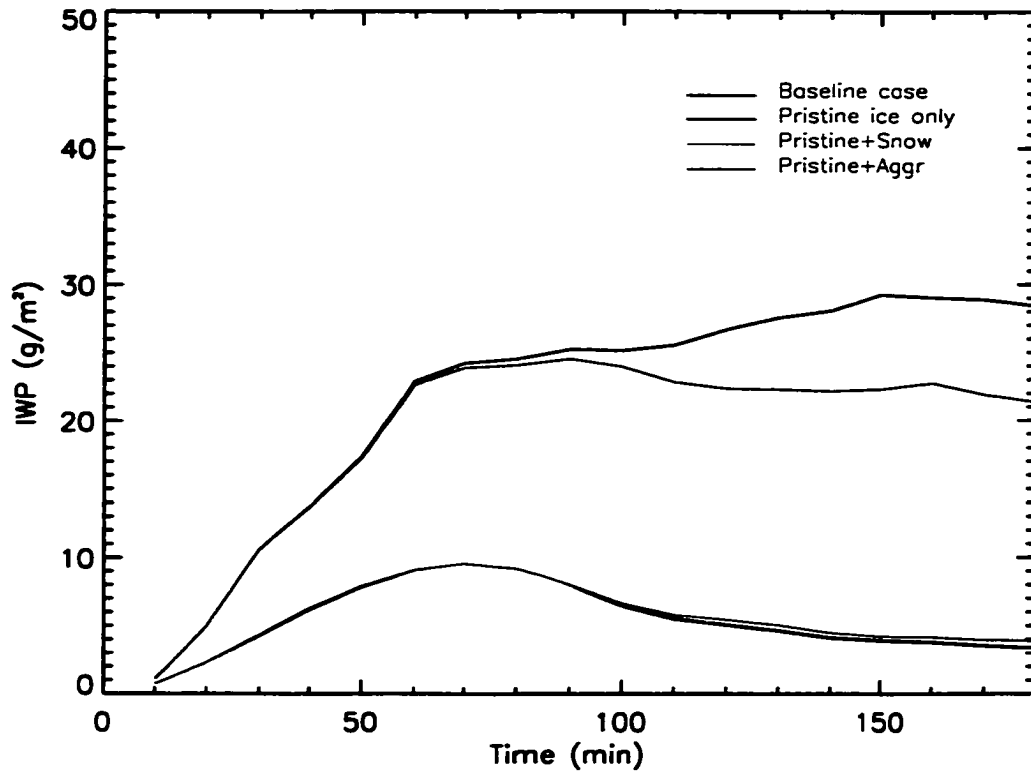


Figure 3.10: IWP versus time for various model configurations (see text for explanations).

Three main categories are identified:

1. parameterizations based on particle size, $v = aD^b$
2. parameterizations based on IWC. $v = cIWC^d$
3. parameterizations that link terminal velocity to particle mass, projected area and maximum dimension (e.g. Mitchell (1996), hereafter M96).

In general, parameterizations of type 1 are preferred in Cloud Resolving Models which generally assume a size distribution, and hence evaluate explicitly mean particle size from mixing ratio and number concentration at each timestep. A degree of complexity to this simple parameterization is added by the fact that coefficients a and b depend on ice crystal shape (e.g. Mitchell, 1988).

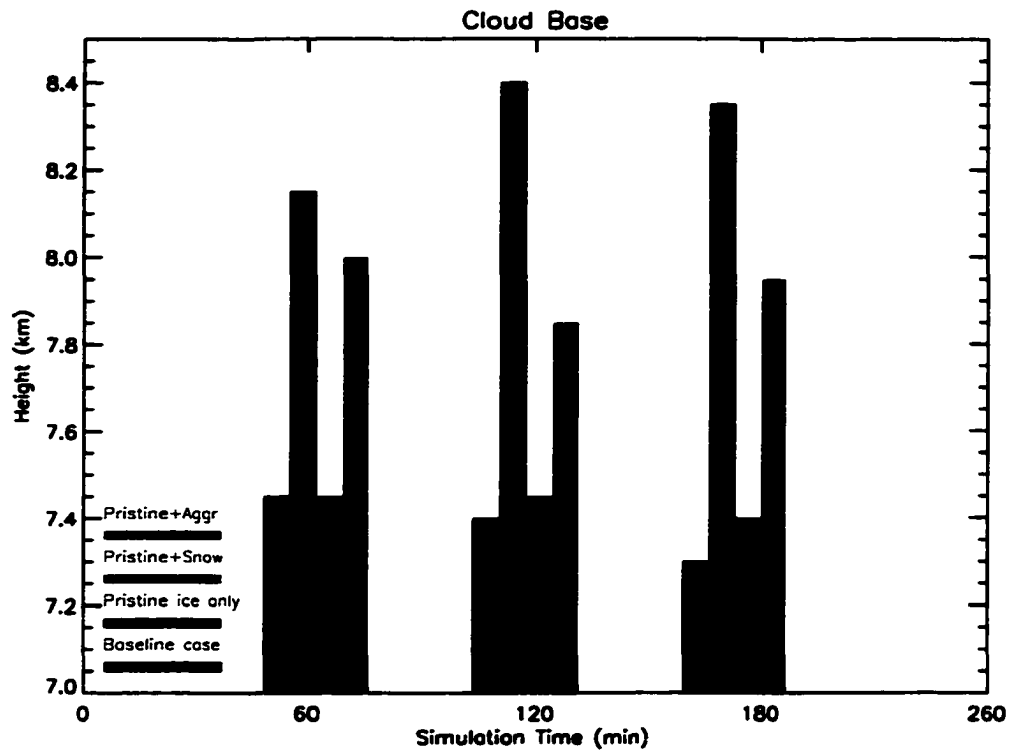


Figure 3.11: Cloud base at selected times for various model configurations (see text for explanations).

Parameterizations of type 2 are generally preferred in NWP and GCMs which do not explicitly predict particle size but rather predict IWC.

Parameterizations of type 3 are more general since they are derived from first principles (i.e. starting from the general expression for the aerodynamic drag force on a particle⁶), and depend only weakly on particle shape (through the dependence on

⁶Following M96 we can write the aerodynamic drag force, F_D , as:

$$F_D = \frac{1}{2} \rho v^2 A C_D \quad (3.4)$$

where ρ is the air density, v is particle velocity, A is the particle's area projected normal to the flow, and C_D is the drag coefficient.

To solve for particle terminal velocity, equation (3.4) is equated to the force of gravity ($= mg$, where m is particle mass and g is the gravity acceleration). An expression for the terminal velocity

particle projected area perpendicular to the flow), which makes them particularly attractive for description of cirrus crystal fall speeds.

Figure 3.12 shows results from 1998 survey of ice water fall speed relationships used in GCSS cirrus cloud models, conducted by P. Brown of the United Kingdom Meteorological Office. Shown are the relationships between mass weighted ice water fall speed and ice water content used in various models. For reference, the Heymsfield and Donner (1990) cirrus parameterization is shown as is the one used in Starr and Cox (1985) model (indicated with the legend SCM1). Values of terminal velocities can differ by up to two orders of magnitude between different parameterizations for the same IWC.

Default particle terminal velocities in RAMS 3b (the model version used in this part of the investigation) are computed via a combination of parameterization of type 1 and type 2. The average fall velocity, v_t , is obtained by integrating over the assumed mass distribution, $m(D)$, and weighted according to the IWC as follows:

$$\bar{v} = \frac{\int v(D)m(D)dD}{IWC}. \quad (3.8)$$

In more recent versions of RAMS (4.2 and higher), v_t is calculated from a bin representation using directly $v(D)$. A plot of RAMS fall speeds for pristine ice, is derived as:

$$v_t = \left(\frac{2mg}{\rho AC_D} \right)^{1/2} \quad (3.5)$$

From (3.5), it can be seen that the fall speed depends on the particle mass and projected area. In practice, equation 3.5 is hard to use, due to the presence of C_D , and the terminal velocity is found by empirically relating the Reynolds number, defined as

$$Re \approx \frac{v_t D}{\nu}, \quad (3.6)$$

which depends on v_t , ν , the kinematic viscosity, and size of the particle, D (radius for a spherical particle), to the Best number defined as,

$$X = C_D Re^2 = \frac{2mg\rho D^2}{A\nu^2} \quad (3.7)$$

containing m and A , and independent of v_t . For more details on this type of fall speed parameterizations see M96.

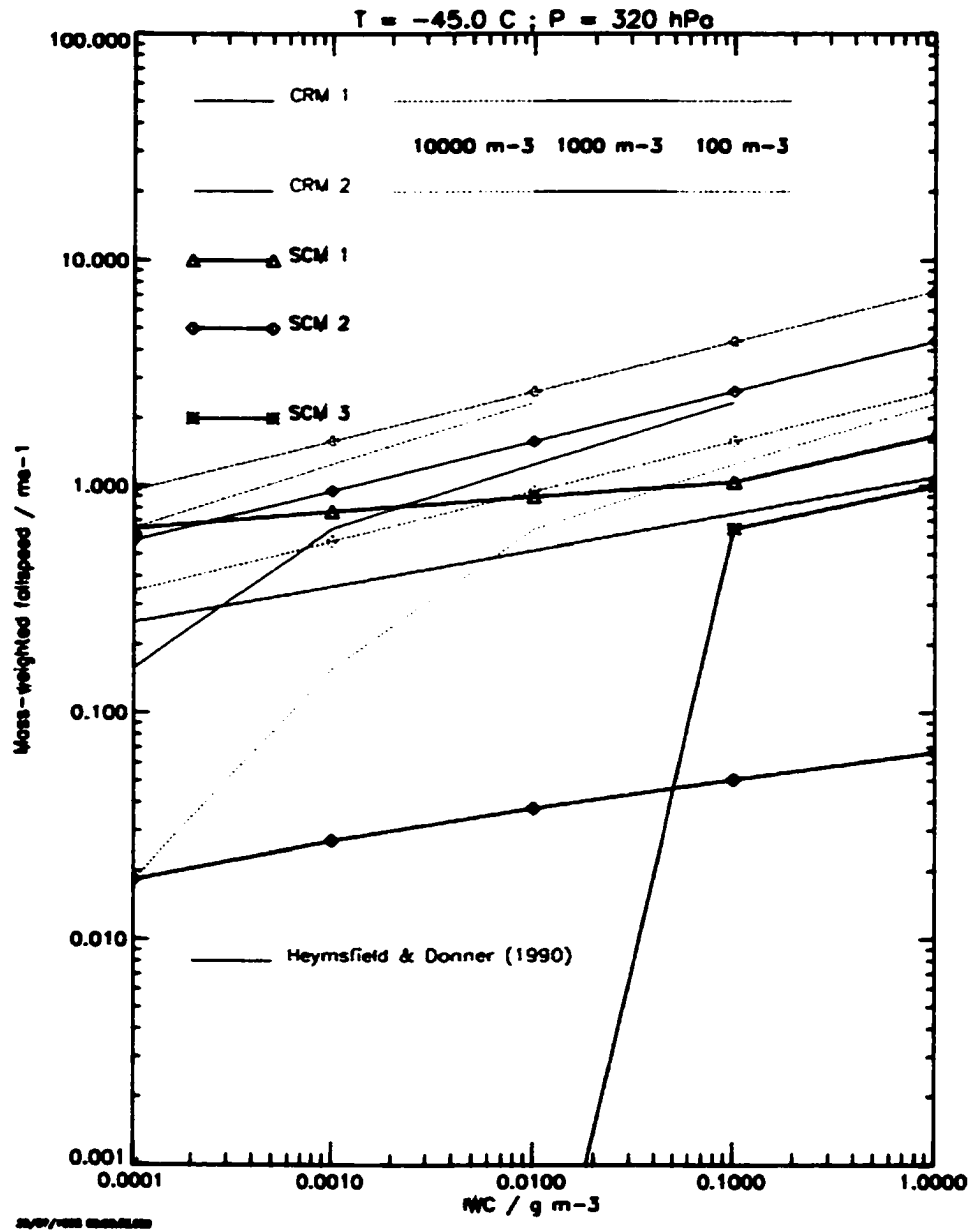


Figure 3.12: Fall speed velocities (ms^{-1}) vs IWC (gm^{-3}). x and y axis scale is logarithmic. Courtesy of P. Brown, U.K. Meteorological Office.

snow and aggregates is shown in figure 3.13. Coefficients a and b for the various ice categories are given in table 3.1.

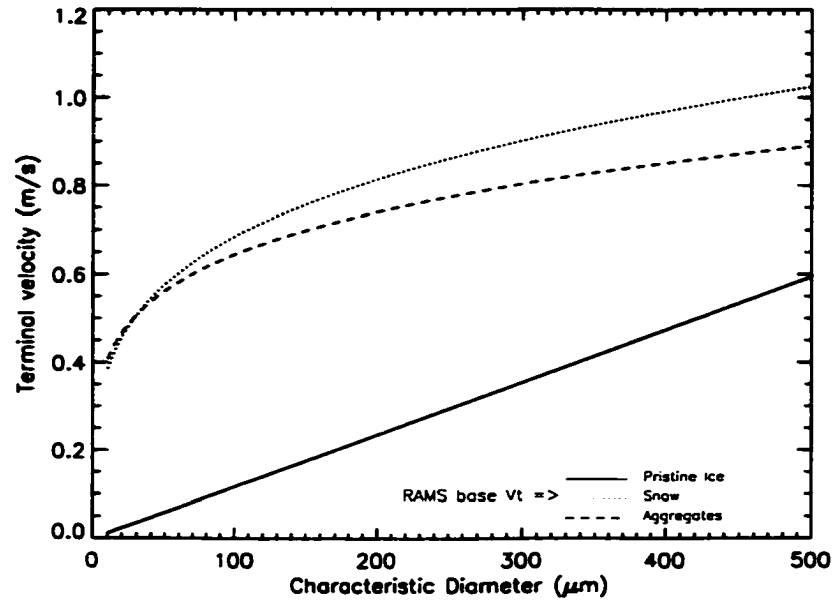


Figure 3.13: Fall speed velocities (ms^{-1}) vs size (μm) for the three RAMS ice categories pristine ice, snow, and aggregates.

Table 3.1: Fall speed coefficients for RAMS ice categories (MKS units).

Ice category	Coefficient a	Coefficient b
Pristine ice	316	1.01
Snow	4.836	0.25
Aggregates	3.084	0.20

In this section, we compare a baseline run with default RAMS terminal velocities with runs in which the parameterization of terminal velocities was changed. In particular, we considered five parameterizations: three are taken from the above-mentioned survey, corresponding to a fall speed treatment of type 2 as illustrated in figure 3.14. The terminology comes from the fact that the two extreme parameterizations (upper v_t and lower v_t), and an intermediate one provided by Heymsfield and Donner (1990) (referred to as H&D (1990) in the figures) were considered.

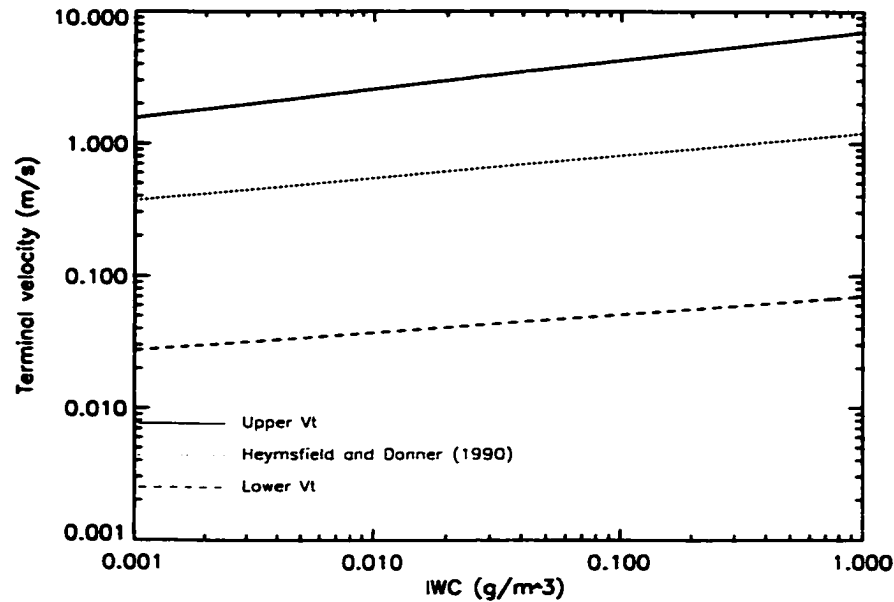


Figure 3.14: Selection of fall speed velocities–IWC parameterizations chosen for this study (see text for explanations)

The other two parameterizations included one in which the terminal velocity is assumed constant for all IWCs and sizes, and one of type 3 provided by M96. IWP results for the different runs are shown in figure 3.15. There is not a large impact of changing fall speed parameterization on the total IWP other than the higher IWP for the run with “lower” v_t . This result, in part unexpected, can be explained with the apparent tendency of the RAMS model to create some compensating average vertical motion at cloud levels that might be responsible for the similarities in the IWPs between the different fall speed cases. However if we look at cloud base height statistics (figure 3.16), large differences emerge. Type 1 parameterizations (RAMS baseline) tend to place the cloud base much higher, implying the presence of numerous small crystals that tend to sediment slowly. On the other hand, parameterizations of type 2 have lower cloud base heights, which imply a faster sedimentation rate of ice crystals. The largest difference is cloud base height (≈ 400 m) is observed two hours after the beginning of simulation between the baseline case and the run with upper v_t .

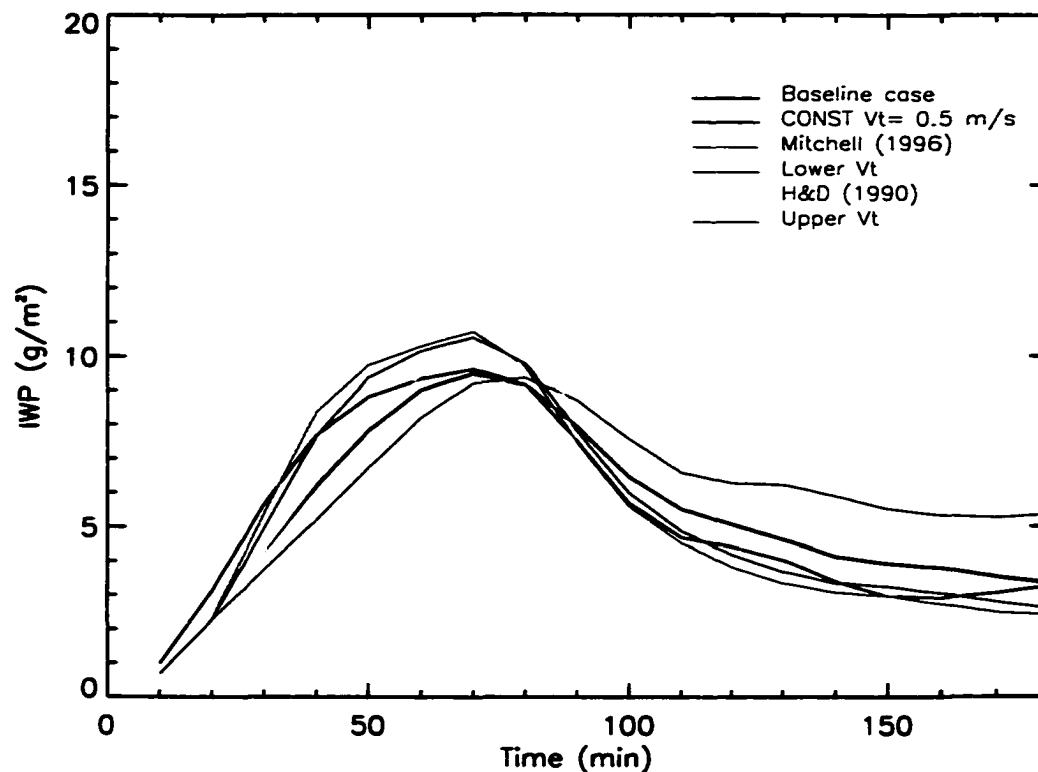


Figure 3.15: IWP versus time for various terminal velocities (see text for explanations).

In both cases the two effects—higher IWC, but lower vertical cloud extension for type 1, and lower IWC but higher cloud geometrical thickness for type 2 parameterizations—tend to compensate each other so that the total IWP is not overly affected. The v_t from M96 produces results curiously similar to the constant v_t case, indicating that there is not much variation in drag force parameters over time in this simulation.

While the vertically-integrated mass characteristics of the cloud are unchanged, the vertical distribution of IWC is dramatically affected by the choice of parameterization. The impact of the different parameterizations would mainly show in the vertical distribution of diabatic (latent and radiative) fluxes which may in turn affect atmospheric lapse rate, and the stability properties of the atmosphere.

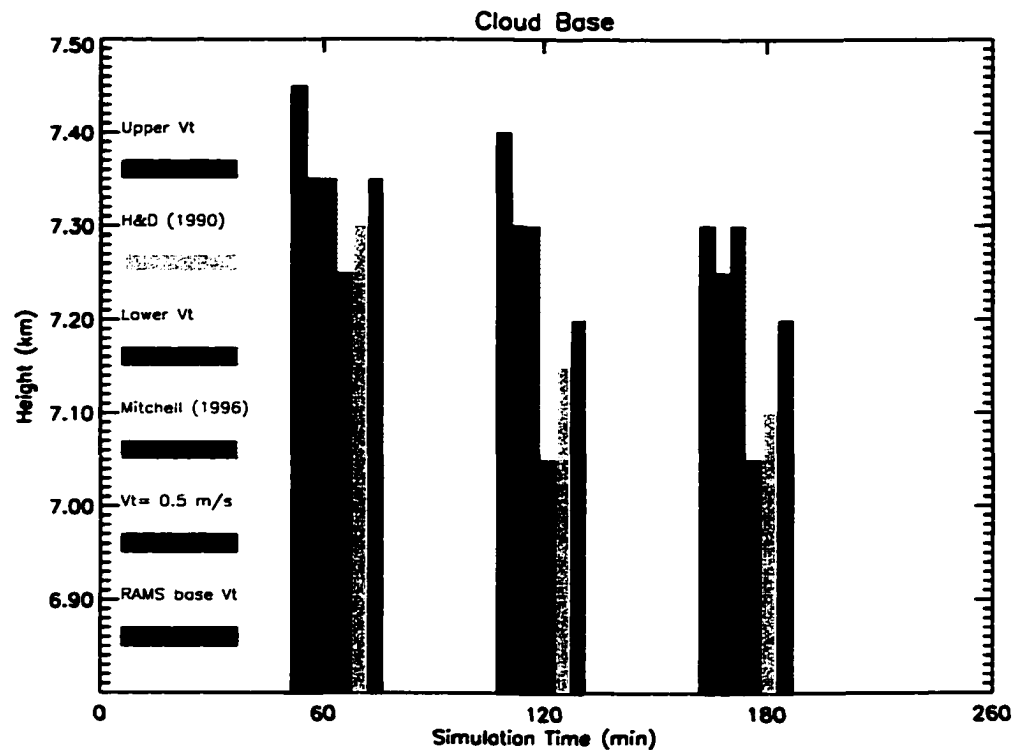


Figure 3.16: Cloud base at selected times for various terminal velocities (see text for explanations).

3.3.4 Concluding remarks on the sensitivity study

We presented cirrus simulations for an idealized test case, and studied model response to changing microphysical parameterization. In particular, we focused on the sensitivity of model output to the choice of bulk ice categories to describe the cloud ice phase, and on the sensitivity to the choice of terminal velocity schemes. These types of studies provide some understanding of model behavior under different conditions. The ultimate test remains comparison with real measurements, but sensitivity analyses help narrow down the relevant parameters that influence the representation of the modeled cloud. These studies are also useful as a guideline in choosing the type of parameterization that minimizes the discrepancies with real world cases. Awareness of model sensitivity also helps to set error boundaries on

model results.

3.4 Real case simulations

While important knowledge can be gained by looking at idealized test case in which forcings and initial conditions are controlled, the evaluation of model skill ultimately must involve comparisons with real data. In the next sections we present results from 3D-channel and 3D simulations of a cirrus cloud observed off the coast of Kauai, Hawaii (22N–160W). This cirrus case was part of a series of observations carried out in the context of the ARM–UAV Spring 1999 field campaign.

3.4.1 The ARM–UAV Spring 1999 Experiment

The ARM–UAV program established in 1991 has been pivotal in demonstrating how measurements from unmanned aircraft platforms can contribute to our understanding of cloud microphysical and radiative processes. As part of this program, a field campaign took place over the Pacific Ocean during Spring 1999 (Stephens *et al.*, 2000b). The campaign was operated from the Pacific Missile Range Facility, located in Kauai, Hawaii (22N–160W). It was designed as a two aircraft experiment to obtain accurate and multi-instrument measurements of cirrus clouds. The stratospheric aircraft Altus II was flown above cloud top (at an altitude of approximately 15 km) in formation with the Twin Otter flown below cloud base. The Altus II payload provided measurements of spectral and broadband radiative fluxes, spectral radiances from the Scanning Spectral Polarimeter (SSP), and lidar backscattering from the Cloud Detection Lidar (CDL). The Twin Otter provided similar radiometric measurements as well as radar reflectivity from the NASA JPL/University of Massachusetts's 94 GHz Airborne Cloud Radar (ACR). Data were collected from April 28 to May 18.

3.4.2 Synoptic overview of April 30 cirrus case

The cirrus case chosen for this study consisted of a single layer cloud observed on April 30, 1999. On that day, an upper level low located about 1500 km north of

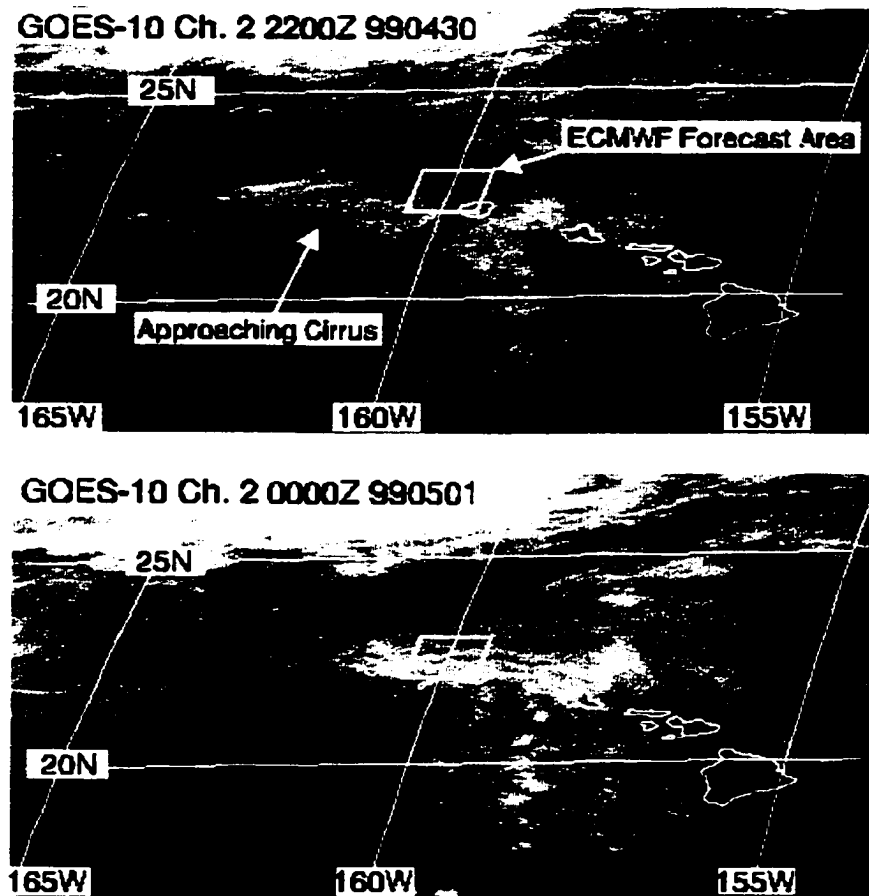


Figure 3.17: GOES 10 visible images at 2200UTC on April, 30, 1999 (upper panel) and at 000UTC on May, 1, 1999 (lower panel) showing the advection of the cirrus layer over the experiment area. Images are courtesy of Steve Miller.

Honolulu was deepening and moving slowly to the southwest. A trough extended from the upper level low southwest to a second low near the dateline. Isolated thunderstorm activity was associated with the second low, with the cirrus cloud being advected over the observation area in response to the west to a southwest wind aloft. The advection of the cirrus layer is visible in the satellite images from GOES 10 at 2200UTC and 000UTC respectively (fig. 3.17).

The aircraft lidar/radar measurements indicated a relatively thick layer of cirrus between 7 and 12 km, deepening over time between 2200 and 000UTC (figure 3.18),

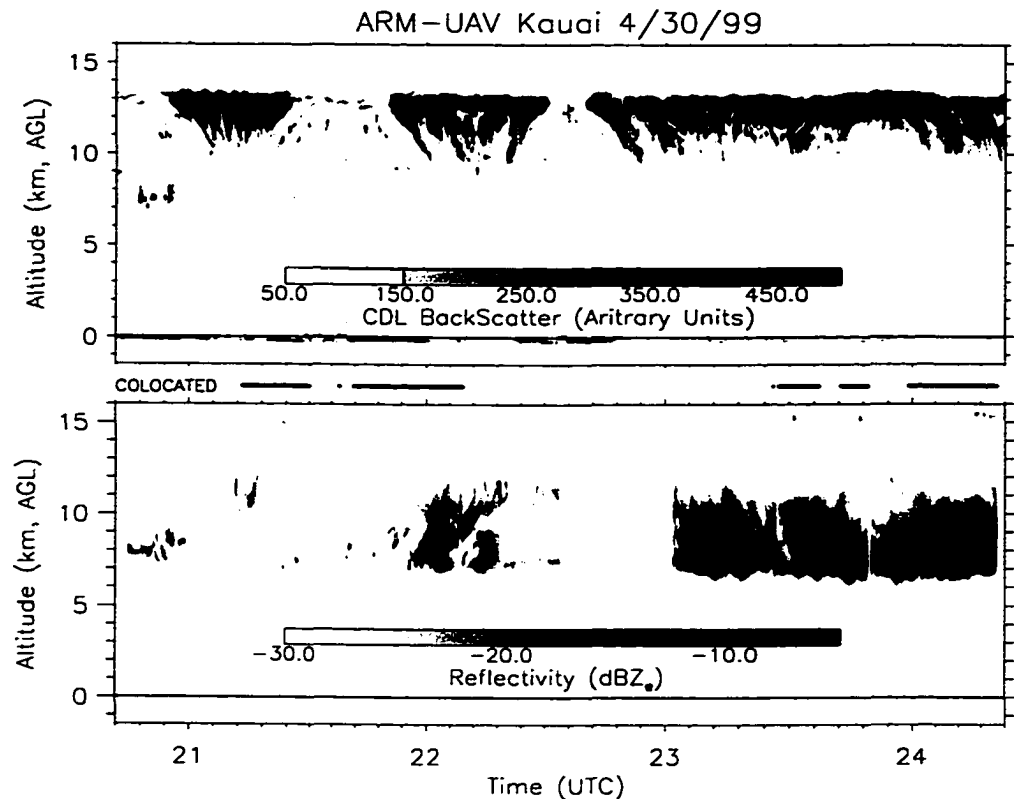


Figure 3.18: Co-located lidar and radar cirrus observation for April, 30, 1999.

congruent with the synoptic analysis.

It is interesting to note that lidar and radar measurements perfectly complement each other. The radar misses the signal from small ice crystals residing at cloud top, which is captured by the lidar. On the other hand the lidar signal gets attenuated rapidly going downward from cloud top, due to the presence of larger crystals and aggregates.

3.4.3 Overview of analysis method

We used two versions of RAMS, version 3b in a 3D-channel configuration and version 4.2 in a 3D configuration. Main differences between model versions were discussed in section 3.3.1. Model predicted ice water content is converted into radar reflectivity and compared with observations along the flight track. Horizontal averages are also

compared. Conversion to radar reflectivity is explained in the next section. Results from the 3D-channel simulations are presented in section 3.4.6. Results from the 3D simulations are presented in section 3.5.3 along with the RAMS error analysis.

3.4.4 Synthetic radar reflectivities from model fields

The mixing ratios of the various ice species for the model domain are saved at selected times throughout the simulation, along with thermodynamic variables (temperature, pressure, vapor mixing ratio). The total ice mixing ratio is computed as the sum of the ice species mixing ratios. Temperature and pressure are used to convert mixing ratios into total IWC by computing the value of air density at each grid-point and time. The IWC is then converted into radar reflectivity, Z . The conversion to radar reflectivity is straightforward if we assume a power law relationship between IWC and Z ,

$$Z = cIWC^d, \quad (3.9)$$

where coefficients c and d depend on the radar wavelength. Values for a 94 GHz radar, empirically estimated from in situ measurements, are taken from Liu and Illingworth (2000) (hereafter, LI00). Relation (3.9) can also be inverted to obtain a simple estimate of IWC from measured radar reflectivities. In LI00 the accuracy of such an estimate is reported to be around 20–30%. However, the authors point out that the error on an individual value of IWC derived from specific measurements of Z can be as high as 50–100%. To improve the estimate of IWC based on power law relationships, they suggest the use of size or temperature dependent coefficients. The uncertainty might thereby be reduced to 30–50% error on any single IWC value.

Results are presented for both radar reflectivity and IWC fields, bearing in mind that equation (3.9) establishes a perfect equivalence between the two fields. IWC can also be obtained from radar reflectivities combined with other types of observations (for example radiometric measurements) as shown in appendix A.

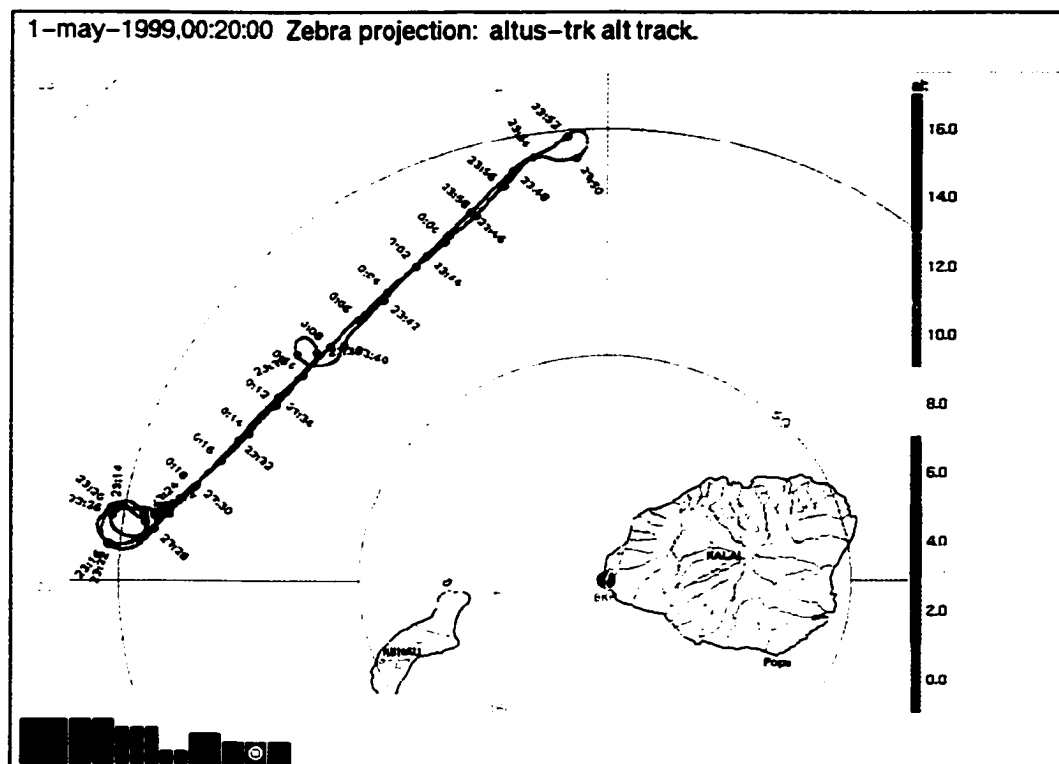


Figure 3.19: Example of Altus II trajectory off the coast of Kauai, on April, 30 1999.

3.4.5 Model implementation for the 3D channel-type LES runs

The CRM domain consisted of a single grid, centered at 22.5N and 160.25W, 80 km long and 4 kilometers wide oriented from east to west, representative of a cross section along the study area at an angle of roughly 45 degrees with the flight track, shown in figure 3.19.

The model resolution was 400 m in the horizontal and variable in the vertical from 400 m in the Boundary Layer to 50 m at cloud levels. The configuration of the model run was 3D, but due to the the fact that the total number of grid points in the zonal (x-) direction was much greater than the number of grid points in the meridional (y-) direction, it is more appropriate to consider the run a LES 3D-channel simulation. It is thus impossible to establish a one-to-one comparison between observations and model fields. Nevertheless, if we assume horizontal homogeneity of the cloud

field, it is still meaningful to compare the simulated fields with the measurements, especially in a domain average sense.

A rigid lid ($w=0$) was assumed at the top of the domain, and a damping method to absorb wave energy was used to avoid possible reflections at the upper boundary. At the lower boundary, the impermeability condition ($w=0$) was enforced; the lateral boundary conditions were modeled as cyclic. Ocean surface temperature was assumed constant. The initial fields (pressure, temperature, specific humidity and wind speed) were specified from the ECMWF forecast over the study area. The initial profiles were assumed to be horizontally homogenous. A large scale forcing of roughly $+3 \text{ cm s}^{-1}$, constant in height, representative of a positive average vertical velocity as given by the ECMWF analysis, was implemented throughout the simulation. This forcing featured the large-scale advection of cloudy air mass over the model domain, otherwise absent in the CRM, due to the fact that the ECMWF fields were only ingested at the initial time and were horizontally homogeneous, i.e. one profile was applied to the whole domain, and not updated during the course of the simulation. Without this forcing it was verified that the CRM produced a much shallower cloud which did not thicken over time as the observations were showing.

3.4.6 Results for the 3D-channel runs

Portions of the flight level data shown in figure 3.18 are reproduced in figures 3.20 along with the simulated radar reflectivity.

As anticipated, model simulations of cloud variability do not match the observations in any quantitative detail. Besides the fact that the model cross-section and the data cross-section are not directly comparable on a one-to-one basis as mentioned in the previous section, there are a number of other reasons why the structure of the predicted cloud differs from the observed cloud. The model initialization is horizontally homogeneous as previously noted, and, as such, it lacks large-scale structure. Despite these obvious deficiencies, encouraging qualitative agreement of the modeled and the observed fields was found.

More meaningful is the comparison of the domain averaged model results with the

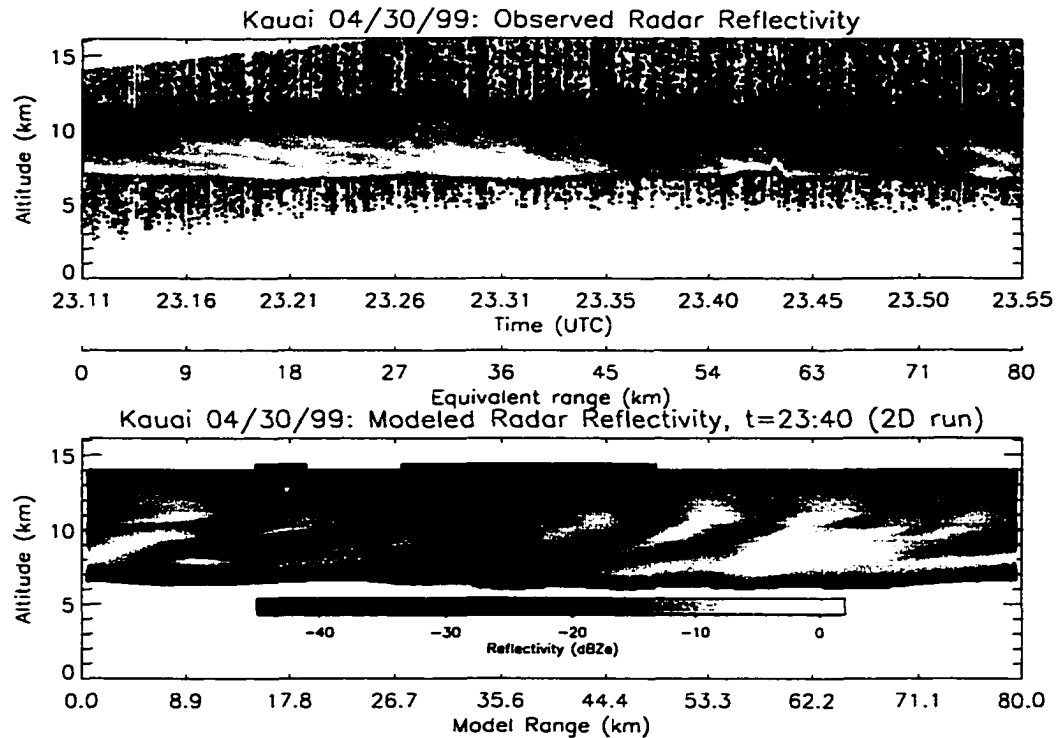


Figure 3.20: Observed (upper panel) versus modeled (lower panel) radar reflectivity for the 3D-channel runs.

average of the observations. Figure 3.21 shows the vertical profile of the measured radar reflectivity, averaged along the flight track over time period from 2311 and 2389UTC, along with average profiles derived from the ECMWF and the CRM Ice Water Content. The comparison shows similarity between the CRM profile and the profile measured by the ACR. However, the simulated cloud is too thick at upper levels compared with the observations.

The CRM in the two-dimensional configuration is only capable of capturing gross average features of the observed cirrus, but no real quantitative agreement is found. The lack of similarity with the observations is mostly due to the rudimentary implementation of the external forcing.

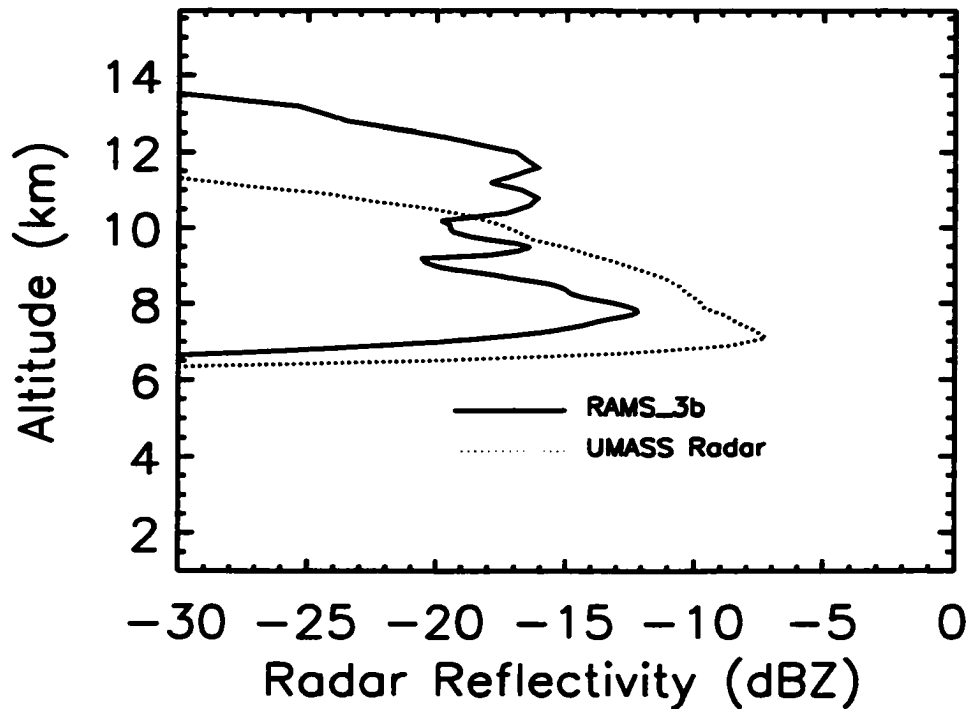


Figure 3.21: Domain averaged radar reflectivity: comparison between observations and CRM results.

3.5 RAMS error analysis⁷

3.5.1 Estimating model bias

By direct comparison of model derived radar reflectivities with observations, it is possible to investigate the performance of the model for the specific case study, and identify any model biases within the uncertainties of the relationship used to convert IWC to radar reflectivity (or vice versa). It is usual to provide a qualitative assessment of the model performance by verifying that the model succeeded in producing a cloud of the correct geometrical thickness and in the right position.

⁷This section is part of a paper entitled: “*Characterization of errors in cirrus simulations from a Cloud Resolving Model for application in Ice Water Content Retrievals*” by A. Benedetti and G.L. Stephens, accepted for publication in *Atmospheric Research*, 2001.

A more quantitative assessment is obtained by comparing average fields, such as the domain-average IWC. One-to-one comparisons are also possible, although it is more difficult to obtain a quantitative agreement when model fields are compared with observed fields at each observation point. In this section, examples of both methods (i.e. point-by-point and average field analysis) are shown. The procedure to compute model bias statistically is outlined in what follows. The model fields from the 3D simulations are interpolated at the observation locations. For each data point, the difference between measured and model value (hereafter referred to as the observation increment, Daley, 1991) is computed using:

$$\epsilon_{ik} = Z_{obs}^{ik} - Z_{mod}^{ik}, \quad (3.10)$$

where k represents the index for the vertical levels, and i represents the index for the horizontal grid points along the flight track, and the indices obs and mod refer to observed and modeled reflectivities, respectively. The observation increment can also be defined in terms of IWC, making use of equation (3.9) to convert the observed radar reflectivity into IWC.

To compute the expected value of the random variable $\bar{\epsilon}_k$, which represents the average bias between model and observations at every vertical level (i.e. the model bias if observations are assumed bias-free), one can evaluate the probability density, $p(\bar{\epsilon}_k)$, using all available horizontal data points. The definition of expected value can then be applied:

$$\langle \bar{\epsilon}_k \rangle = \int_{-\infty}^{\infty} \bar{\epsilon}_k p(\bar{\epsilon}_k) d\bar{\epsilon}_k. \quad (3.11)$$

The symbol $\langle \rangle$ is generally used to indicate the expected value of any generic function of the random variable. Assuming the observations are unbiased, if the model were unbiased, the expected value in (3.11) would be zero. A nonzero value indicates that the cloud forecast has a bias which, for a mesoscale model, can be due either to the external large-scale initialization and forcing, or to physical parameterizations intrinsic to the model, as discussed in section 3.2.

The average bias at each level can also be estimated by taking the mean of all

observation increments, ϵ_{ik} , over the total number of horizontal points, N_h :

$$\bar{\epsilon}_k = \frac{1}{N_h} \sum_{i=1}^{N_h} \epsilon_{ik}. \quad (3.12)$$

In the error analysis that follows, both definitions were used and the same result was obtained for the average bias profile.

3.5.2 Error Covariance Matrix

Using equations (3.10) and (3.11), the concept of the Error Covariance Matrix is now introduced. The covariance of two variables is defined as:

$$cov(s, q) = \langle (s - \langle s \rangle)(q - \langle q \rangle) \rangle = \langle sq \rangle - \langle s \rangle \langle q \rangle \quad (3.13)$$

If s and q are totally uncorrelated then $\langle sq \rangle = \langle s \rangle \langle q \rangle$, and their covariance is zero. To compute the covariance from the general definition (3.13), it is necessary to know $p(sq)$, the probability density of the product of the two random variables, which enters in the evaluation of $\langle sq \rangle$ via the analogue of (3.11). In our application, the two correlated variables are the biases at different vertical levels, $\bar{\epsilon}_k$ and $\bar{\epsilon}_l$. As described in the previous section, it is in principle possible to empirically derive $p(\bar{\epsilon}_k \bar{\epsilon}_l)$, and use definition (3.13) to obtain the error covariance. Otherwise, an estimate of the error covariance can be computed by:

$$C_{kk'} = \frac{1}{N_h} \sum_{i=1}^{N_h} (\epsilon_{ik} - \bar{\epsilon}_k)(\epsilon_{ik'} - \bar{\epsilon}_{k'}) \quad (3.14)$$

(e.g., L'Ecuyer and Stephens, 2001). The diagonal elements of \mathbf{C} represent the error variance at each level, and are, by definition, positive. Their square root is the standard deviation associated with the profile of horizontally averaged model fields. Off-diagonal elements describe cross correlations between uncertainties at different vertical levels. The amplitude of these cross correlations falls off as the vertical separation between levels increases.

The error covariance matrix \mathbf{C} characterizes model errors assuming that the observational error is uncorrelated in the vertical and is known, so that it can be

subtracted off. For mesoscale cloud forecasts, it is safe to assume that observational errors are smaller than forecast errors, considering that there are multiple potential sources of model errors (large-scale initialization and forcing, parameterizations, etc.). However, this might not be true for uncalibrated sparse observations. If these assumptions cannot be made, \mathbf{C} represents a sum of both observational and model forecast error covariances. Here we consider that \mathbf{C} represents the model error covariance, subject to the specific (and arbitrary) choice of initial and boundary conditions and of mapping used to transform IWC into radar reflectivity. In the analysis that follows, equation (3.14) was used to compute the ECM. Examples of calculated ECMs are shown in section 3.5.3.

3.5.3 Error analysis results

Model implementation for the 3D nested-grid runs

For the 3D simulations, RAMS 4.2 was used. The CRM domain consisted of three nested grids, situated at 22.5N and 160.25W, center of the field experiment area. For the largest grid, 42x42 points and a horizontal resolution of 16 km were chosen. For the second grid, the same number of grid points was maintained, and the horizontal resolution increased to 4 km. The inner grid was designed to cover the measurement area. It consisted of 82x82 horizontal points with a horizontal resolution of 1 km. The vertical resolution for all three model grids was 500 m and the vertical domain extended from the surface up to 15 km. The impermeability condition ($w=0$) at the bottom and rigid lid at the top of the vertical domain were implemented as in the previous 3D-channel run.

The main difference between 3D-channel and 3D nested-grid runs consisted in the implementation in the latter of a large-scale forcing on the boundaries of the outermost grid through the nudging technique described in chapter 2. Thermodynamics and dynamics fields were specified from the ECMWF forecast over the study area, as in the 3D-channel simulations, but initialization was horizontally-nonhomogeneous. Ambient variables were assimilated over the total length of the simulation (6 hours) with input of ECMWF data every 3 hours. Since model equa-

tions are continuous and the nudging term has to be implemented at every model timestep, large-scale model fields are linearly interpolated in time between two subsequent intakes of data. In the 3D-channel simulations, it was seen that to form a cloud from horizontally-homogenous initial conditions and periodical lateral boundary conditions, it was necessary to superimpose a large scale vertical velocity of roughly 3 cm s^{-1} throughout the simulation (as in the WG2 simulations presented in the previous sections).

The nudging of the CRM with the ECMWF forecast playing the role of “observations”, insured the proper large-scale forcing necessary for the initial formation and subsequent persistence of the cirrus layer. However, the model was found to be sensitive to different nudging assumptions. In fact, strength and location of the nudging are controlled by the user via the relaxation coefficient, as discussed in section 2.2.1. A range of recommended values, mostly dictated by experience and practice, is given in the RAMS 4.2 manual. In this study, the radar observations were used as a guide to select location and strength of the nudging, in order to reduce the degree of arbitrariness in the choice of the coefficient, and in an attempt to minimize the model bias. Despite this effort, the CRM still showed a positive bias with respect to radar measured IWC as shown in the next section. This points to the need for a more optimal way than nudging to incorporate large-scale information into RAMS. A possible answer to this need is represented by the 4D-Var system described in chapter 5.

Direct comparisons with observations

The observed reflectivities are shown in figure 3.22 (top panel) along with the model results (central panel) for a portion of the flight track. The 3D configuration of the model run permitted horizontal interpolation of model fields from the inner grid along the flight track, and comparison of the observations and model fields at each point.

The bottom panel of the same figure shows the observational increment, ϵ_{ik} , defined by equation (3.10). The synthetic reflectivities appear more horizontally

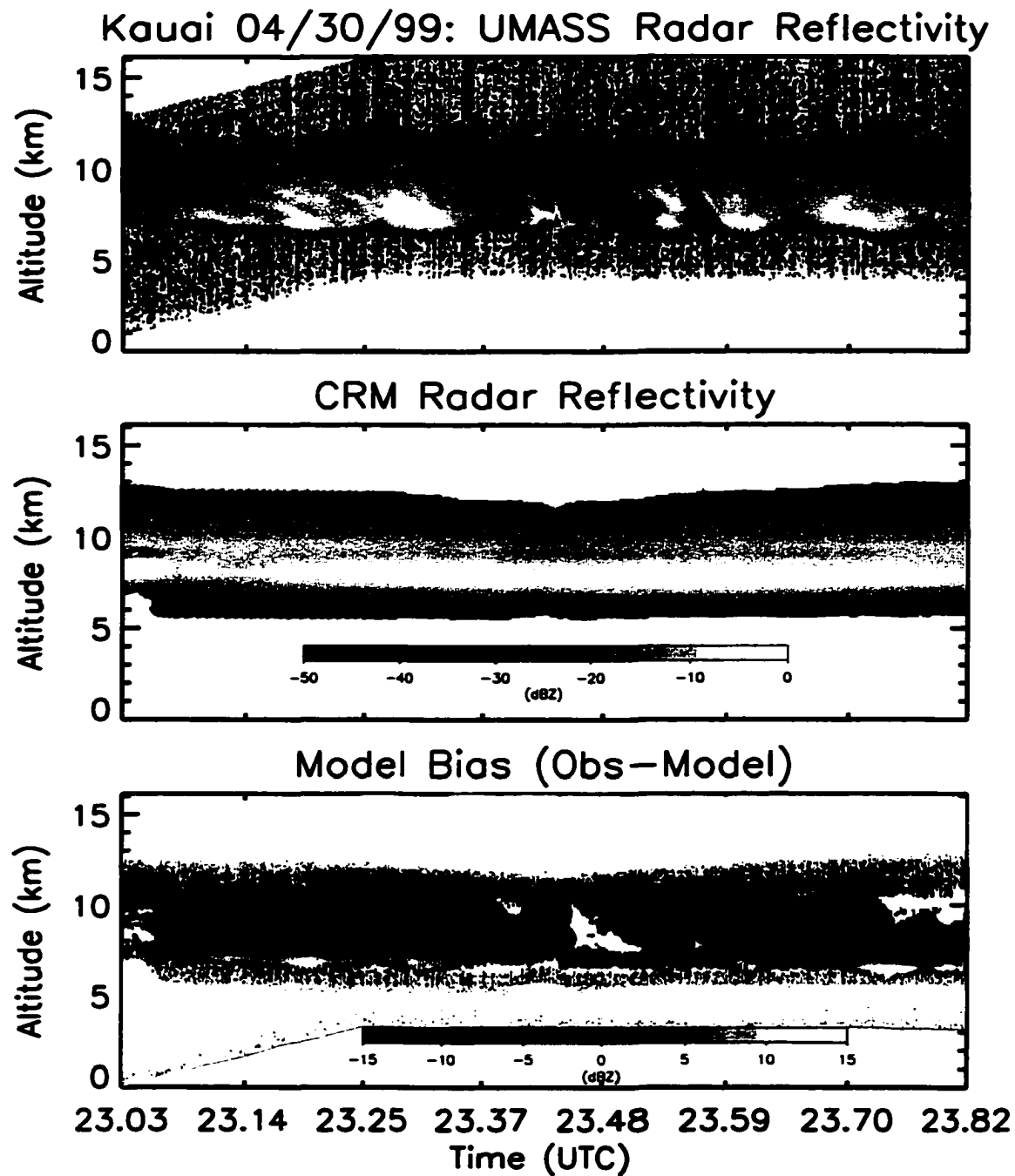


Figure 3.22: Observed (top) and model radar reflectivity (center) for one flight leg. Model bias is shown in lower panel.

homogeneous than the observed reflectivities, due to the coarse (compared to the radar) model resolution. Since the observation increment is defined as the difference between modeled and observed fields, the small scale discrepancies between model and measurements are emphasized in the ϵ_{ik} plot. However, the model seems able to capture the bulk features of the observed cloud, such as vertical position and domain average radar reflectivity. A similar plot is also shown for the IWC field (figure 3.23). Horizontally averaged profiles of model and measured reflectivity and IWC are shown in figure 3.24 and figure 3.25 respectively.

The definitions of expected value introduced in section 3.5.1 were used to compute the average model bias, $\bar{\epsilon}_k$. The probability distribution function (pdf), $p(\bar{\epsilon}_k)$, is shown in figure 3.26 at selected vertical levels.

This pdf was used to compute the expected value of the observation increment according to equation (3.11). Figure 3.27 shows the average bias for the IWC computed both as average and as expected value over the pdf. Both evaluations yielded the same result.

The bias plot shows a general wet bias in the model, i.e. the model IWC is larger than the observed IWC, except around 6.5 km where the situation is slightly reversed. This is also apparent in the average IWC profiles of figure 3.25. We believe that this bias is not related to the model physics, rather to the model response to large scale forcing via the choice of nudging coefficients or to errors in the forcing. However, both factors contribute to cloud model biases, and to understand the relative importance of one versus the other, more model realizations are necessary together with *ensemble* statistics from a number of different observational cases.

Error covariance matrix results

Following the procedure outlined in section 3.5.2, the error covariance matrices were computed. Figure 3.28 and 3.29 show the ECM for the average reflectivity and the average IWC, respectively.

The largest variance is shown around 6–8 km. This is where values of IWC (and reflectivity) are also large. A good indicator of model uncertainty is the relative error

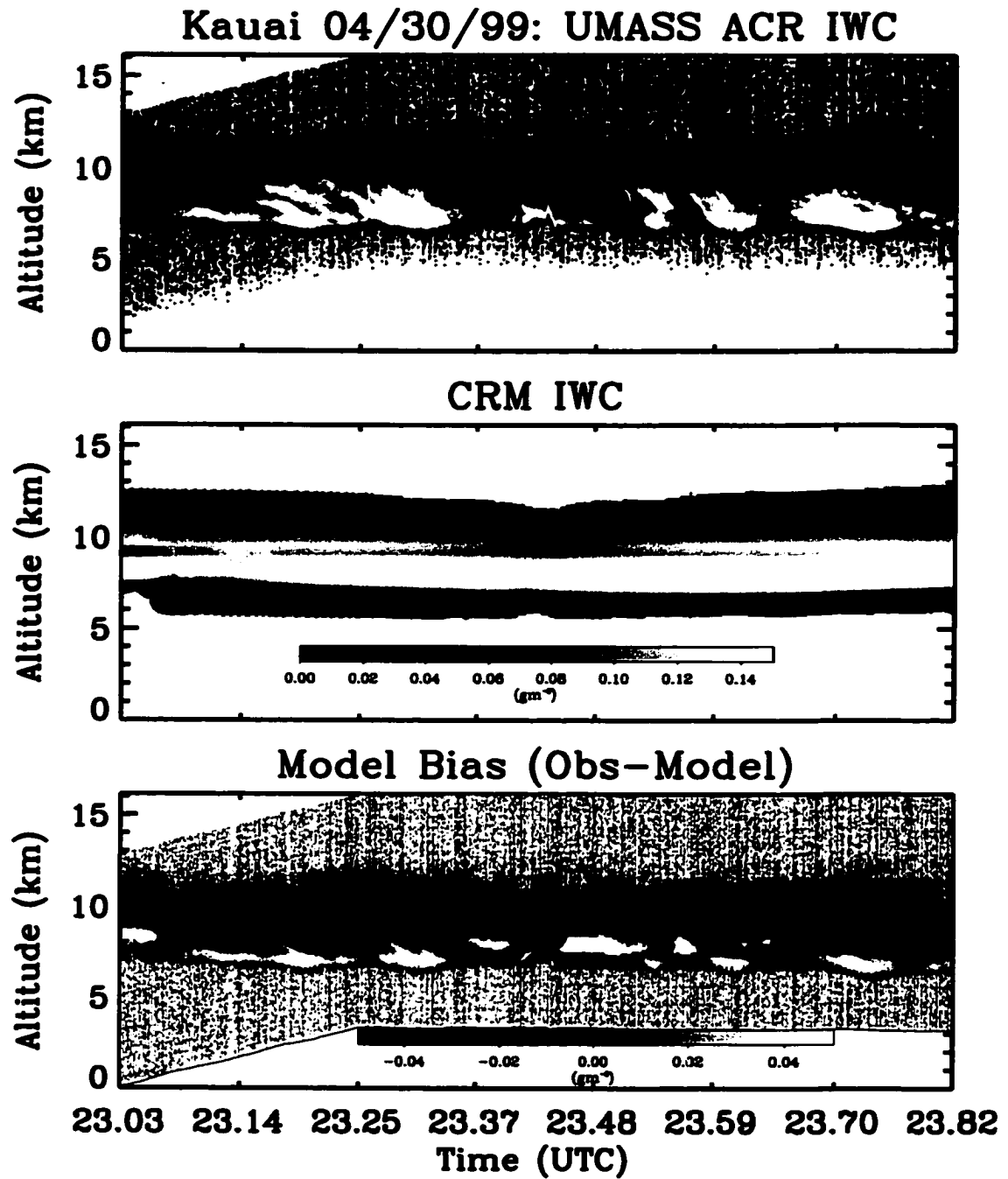


Figure 3.23: As in figure 3.22, but for the Ice Water Content.

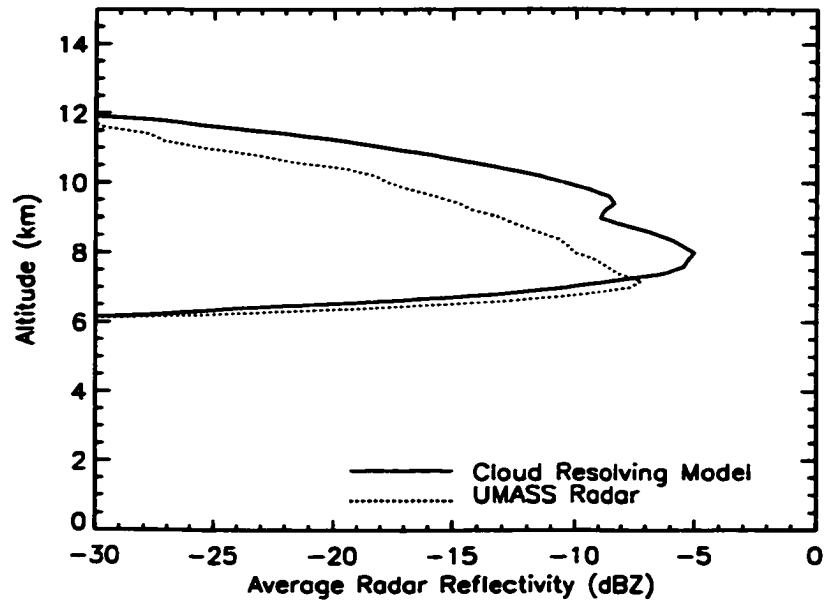


Figure 3.24: Domain average model and observed radar reflectivity.

which can be obtained by taking the square root of diagonal elements and dividing by

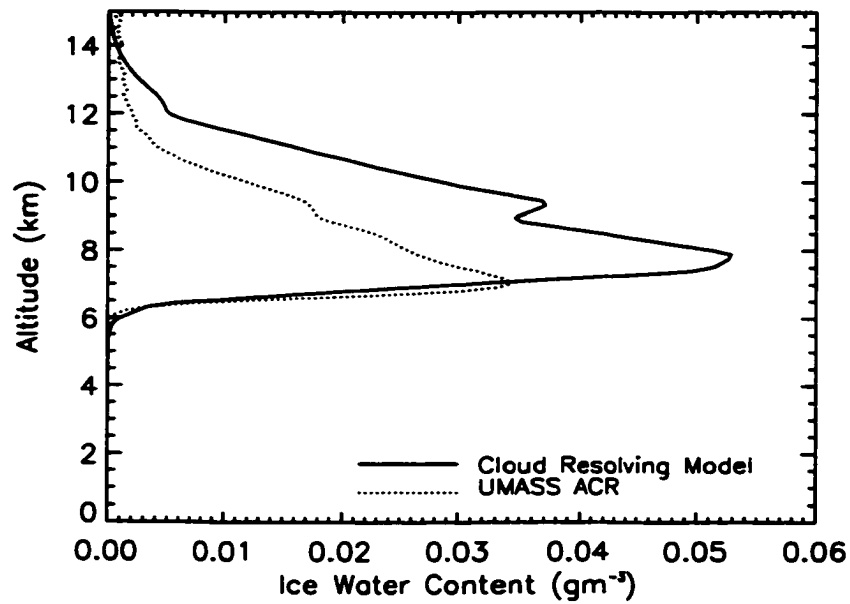


Figure 3.25: As in figure 3.24, but for the domain average IWC.

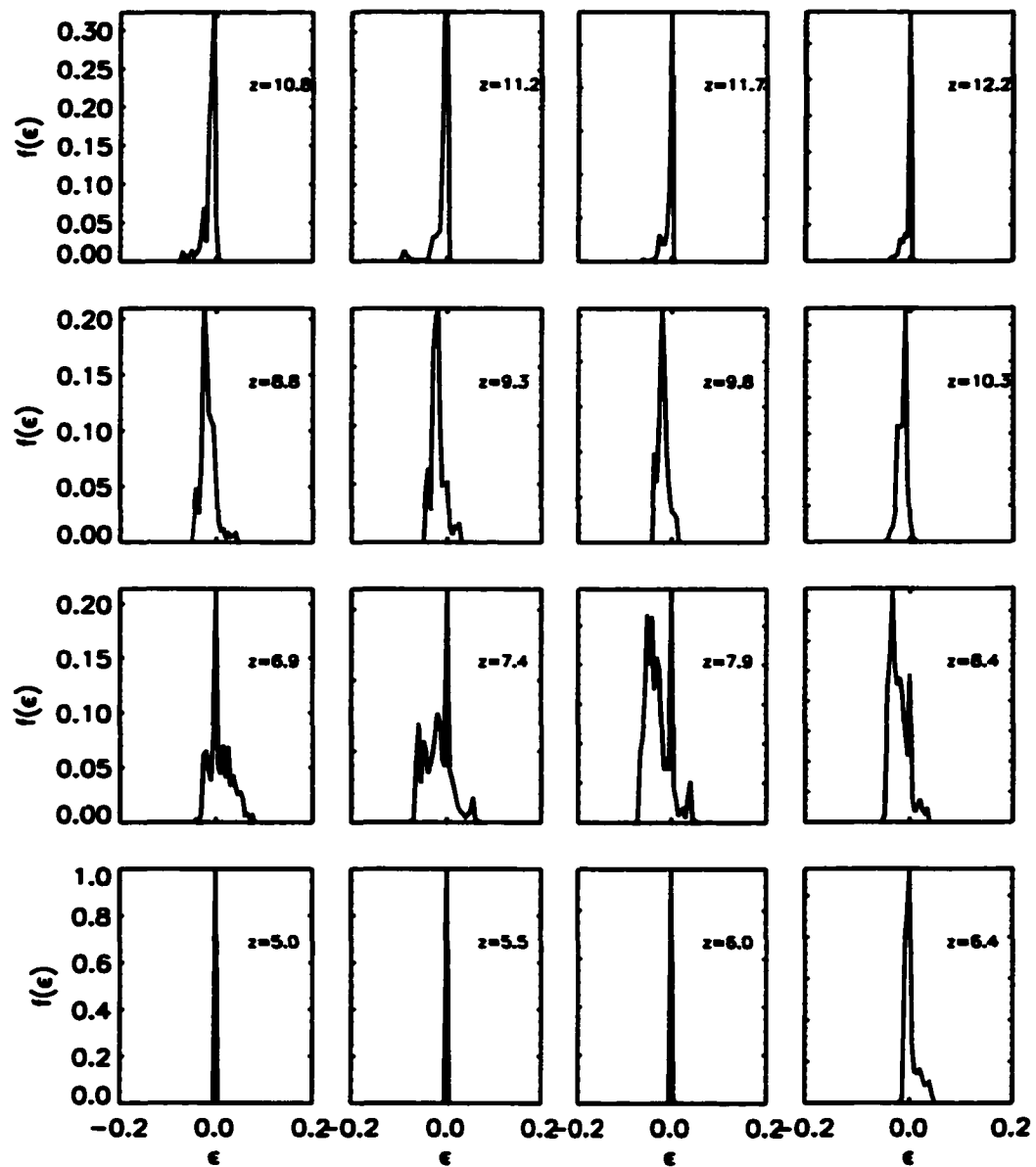


Figure 3.26: Model IWC bias probability distribution function at selected vertical levels.

the value of average model field at the given vertical level. Figure 3.30 shows a profile of model percent uncertainty as a function of altitude for the IWC. The error is huge right above cloud base, and toward cloud top, indicating a mismatch between model IWC and observations which could be caused by the coarse vertical resolution used

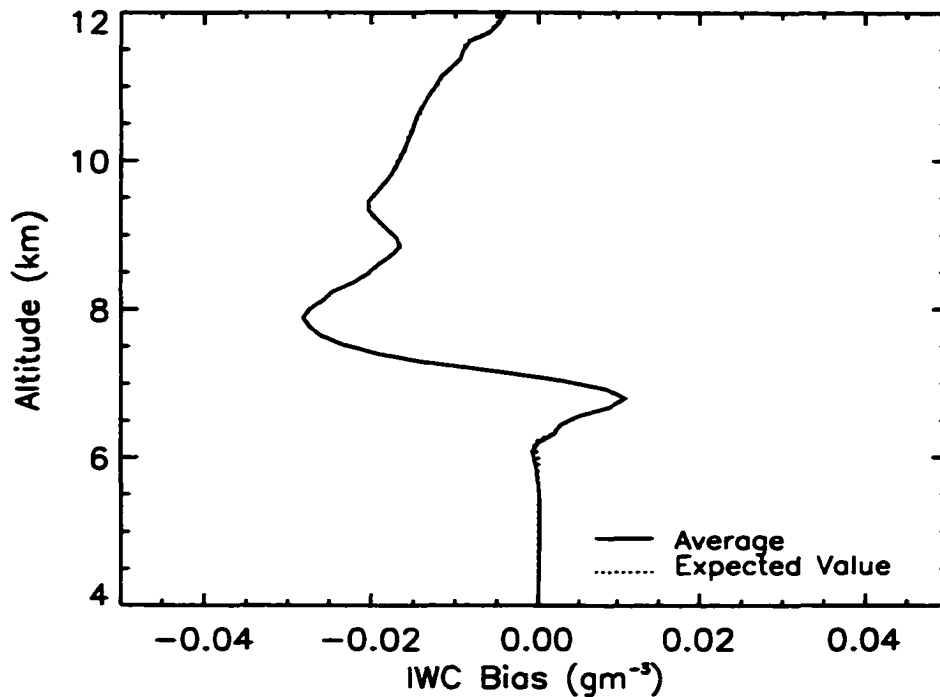


Figure 3.27: Profile of average model bias. Bias is expressed in IWC units (gm^{-3}).

in the simulation (500 m) and/or errors in the ECMWF analysis forcing, resulting in a limited CRM capability to resolve fine vertical structures. This discrepancy, which is also disclosed upon accurate scrutiny of the average profile of figure 3.25, is revealed clearly in the error analysis. At all other levels, the relative error is around 50%.

The model average IWC profile, corrected for the bias, along with its error covariance can be used as *a priori* information to help/insure convergence of an estimation-based retrieval algorithm, in a similar way as the background is treated in data assimilation problems. Appendix A illustrates this application of the computed ECM.

3.6 Summary

In this chapter, many different aspects of cirrus modeling were discussed. Model behavior in simulating ice cloud processes was tested with various tools: controlled

Reflectivity Covariance Matrix

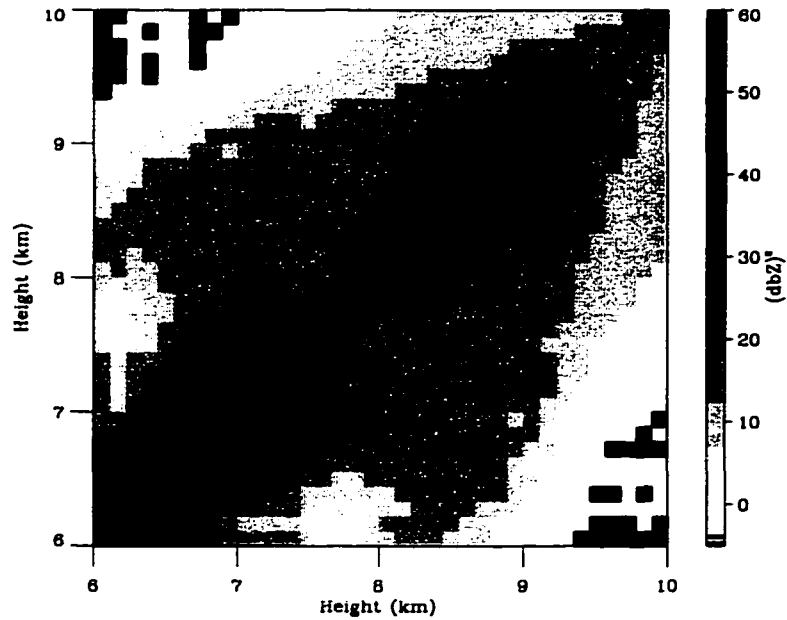


Figure 3.28: Radar reflectivity Error Covariance Matrix.

IWC Covariance Matrix

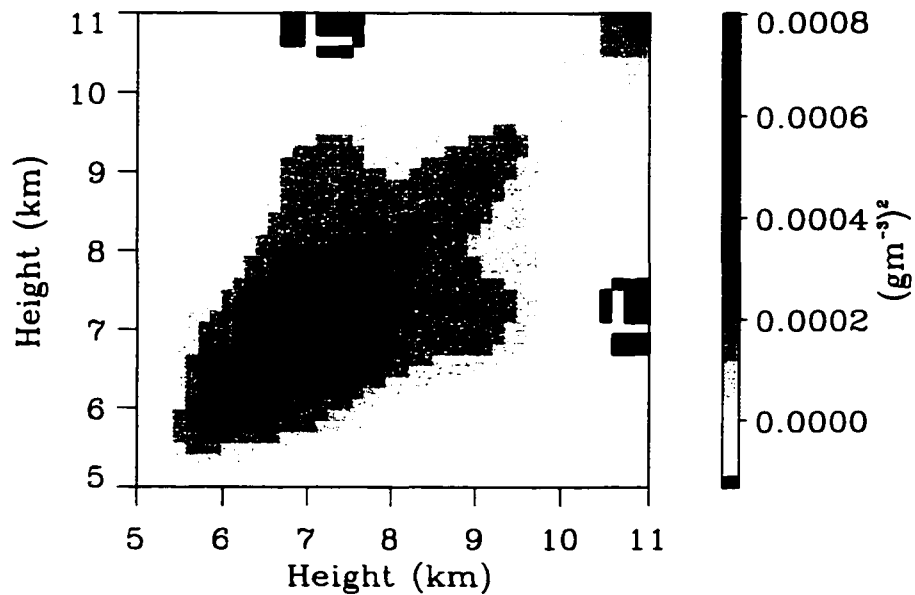


Figure 3.29: IWC Error Covariance Matrix.

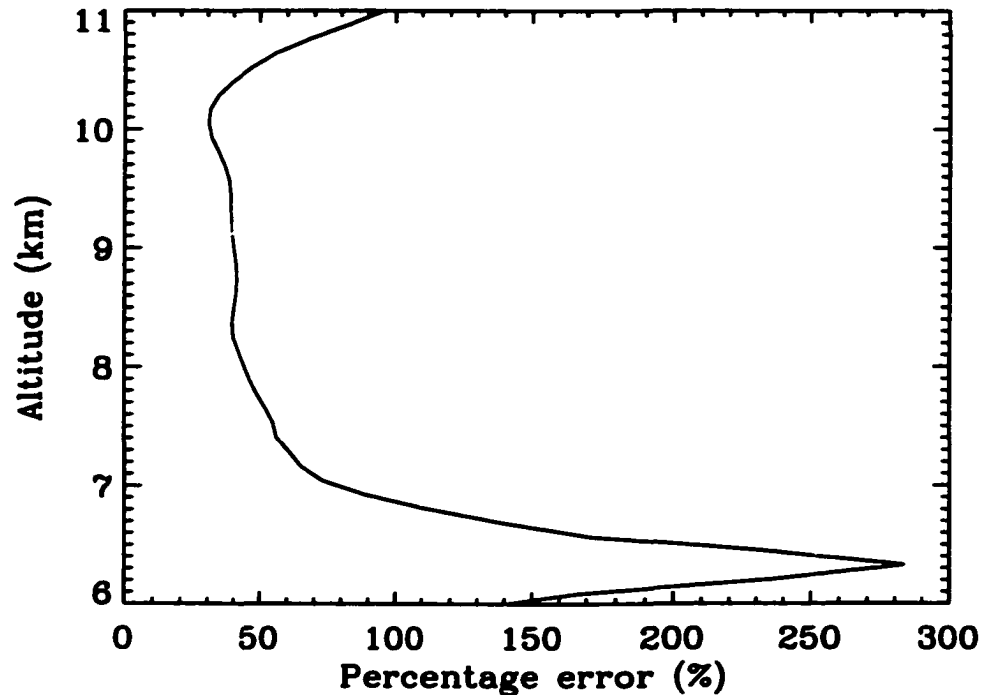


Figure 3.30: Profile of relative error on horizontally averaged IWC, expressed as percentage.

experiments, sensitivity analysis, and real case simulations. From the first two, it was learned how model errors can be introduced by specific choices that are made regarding model parameterizations and description of microphysical processes. From comparison with observations, the impact of model initialization was explored. It was found that only under specific conditions does the model reproduce a cloud that resembles that which has been observed, and these conditions include the introduction of large-scale information. A nudging technique was used in this chapter to implement the assimilation of external data. Due to the arbitrary choice of nudging strength, this is a suboptimal way of initializing and forcing the cloud model, yet model performance is totally dependent on it, at least for the cirrus case examined. It is concluded that exploration of other ways of *assimilation* of external information, such as variational methods, are necessary. This in turn eliminates, or at least minimizes, cloud model errors derived from initial condition specification (i.e. from

a NWP forecast) which are not directly quantifiable, permitting a focus on errors related to model physics.

The next chapter discusses variational experiments performed with a one-dimensional ice growth model and serves as an overture for the assimilation of radar data into cloud models, which is the main focus of this research.

Chapter 4

Cloud radar data assimilation

4.1 Rationale for the use of a simplified model

In this chapter, a one-dimensional time-dependent ice growth model is introduced and used in the context of variational retrievals and data assimilation. The simplicity of the model facilitates learning since it allows for in-depth exploration of various aspects of the data assimilation problem, while offering the advantages of a direct control over model implementation, from the equations-on-paper stage to the computer-coding stage. Multiple experiments to probe different solutions to the assimilation problem are possible because the entire assimilation procedure is computationally inexpensive.

While more sophisticated models, such as RAMS, offer the benefits of a coupling between all physical processes,– dynamical, microphysical and radiative, and hence adduce a more realistic representation of the “real world”, they are also harder to implement. For instance, the job of developing a tangent linear and adjoint model (which are strictly model-dependent!) for a complex nonhydrostatic cloud resolving model might take months, even with the aid of automatic differentiation software. For a less complex model, the task of developing the tangent linear and adjoint might still not be trivial, but it is more tractable.

In the following sections, different applications of the ice growth model and its adjoint are shown. The main applications are the use of the adjoint for sensitivity studies as a way of examining the model, and the use of a variational technique to explore the potential of assimilation of radar data. Various ramifications of this investigation are explored and general implications are discussed.

4.1.1 Specific goals of this modeling effort

The goals of this chapter are summarized as follows:

- Develop an ice growth model which describes cirrus cloud vertical and temporal evolution from basic principles using physically-based parameterizations while maintaining an overall (but necessarily nonlinear) simple form;
- evaluate the model behavior by comparison of the model with radar-derived cloud profiles;
- construct an adjoint based on the tangent linear version of the nonlinear model to be used in sensitivity studies;
- use the adjoint in variational data assimilation of synthetic and real radar reflectivities.

These activities are aimed at demonstrating the impact of radar data assimilation in cloud model prediction. The steps taken in developing the Cirrus Ice Growth Model and Adjoint (CIGMA) system are general, and the procedure is applicable to more complex cloud models. Conclusions drawn from both sensitivity studies and assimilation experiments are relevant to the broader assimilation contexts (such as to the assimilation of cloud radar data into NWP models).

4.2 General overview of procedure

The microphysical model is based on the work of Mitchell (1988, 1991), Passarelli (1978a, 1978b) and Drake (1972). Mitchell's formulation has been extended to obtain two coupled prognostic equations to predict the time and vertical evolution of two parameters of a gamma size distribution of fixed width, the characteristic diameter (D_n) and the total number concentration of ice crystals (N_t).

The model is initialized with profiles of diameter and number concentration derived from temperature, specific humidity and pressure profiles specified from a NWP forecast or from a sounding, making use of selected parameterizations currently implemented in other models. At each time step during the time integration,

the updated profiles of D_n and N_t are mapped into radar reflectivity. At selected assimilation times, the quadratic cost function introduced in chapter 2 is computed and optimized with respect to the control variables. Two sets of optimum variables are used: the initial conditions for D_n and N_t (cloud variables) and the initial profiles of temperature, pressure and specific humidity (environment variables). It was found that some control variables yield more robust results than others. This conclusion was supported by sensitivity studies conducted with the adjoint of the model. The main findings of note are the extreme model sensitivity to initial conditions and high nonlinearity with respect to the environmental variables.

The results of the optimization are used to perform a new model prediction until convergence is reached. The optimization process uses the gradient of the cost function with respect to the control variables which is computed using the adjoint of the cloud model, as discussed in chapter 2.

The procedure is tested with synthetic data in order to insure its correct implementation. The use of synthetic data allows us also to study the problem under “controlled” settings. The model is run with an arbitrary initial condition and radar reflectivities are computed at given assimilation times. These “bogus” reflectivities are used in the optimization as measurements. The initial profile is then perturbed by a given percentage, and the optimization process is started. Various tests are conducted to gain a better understanding of the radar data assimilation system as a whole, and its response to various factors such as changing error statistics, including/excluding a background, adding/subtracting information. Finally, real observations are also used to study the impact of inclusion of radar reflectivity profiles on cloud model prediction.

The results are encouraging. Assimilation of radar data into cloud models is feasible, and radar data do contain valuable information on both cloud variables and the environment. However, results are not always optimal and convergence is not always achieved, especially when using real measurements. This problem might be due to the use of a 1D model lacking an explicit treatment of relevant physical processes. The use of a more complex model and the inclusion of observations in addition to radar reflectivity is an obvious extension of this research. However,

since the CIGMA model makes use of parameterizations that are also used in more complex models, some issues encountered with it are likely to be experienced even with a different model.

The outline of the chapter is as follows. Section 4.3 presents a detailed description of the ice growth model, along with parameterizations and numerical procedures introduced to solve model equations. Model validation using radar observations is also presented. A sensitivity study using the adjoint is presented in section 4.4. Section 4.5 describes and analyzes all components of the CIGMA assimilation system specific to radar data assimilation. In sections 4.6 and 4.7 numerical results from assimilation experiments that introduce both the synthetic data and actual observations are presented and discussed. Section 4.8 summarizes the chapter and elaborates on the major findings of this particular piece of research. Aspects of this research are contained in the paper “Exploring the possibility of assimilation of radar reflectivities in a cirrus model using a variational approach”, submitted for the P.Wagner Memorial Award to the Desert Research Institute of Reno, NV.

4.3 The microphysical module

The model presented here is a Lagrangian time-dependent 1D model that predicts the evolution of ice particle size spectra in cirrus clouds. The assumption is that given a state of the atmosphere that is favorable to the formation and maintenance of an ice cloud, i.e. a profile of relative or specific humidity that is supersaturated with respect to ice, we can predict the time evolution of cloud ice particles subject to growth by vapor diffusion and aggregation (riming and break-up processes are not included). The model does not predict the time rate of change of any thermodynamic or dynamic variables, which are assumed fixed and are specified from a NWP forecast or from a sounding. This is an obvious limitation to the degree of realism of the model. However, if it can be assumed that the time scale over which the large-scale atmospheric state changes is longer than the time scale over which cloud processes take place, it is reasonable to treat the atmospheric state as fixed. This hypothesis may more likely hold for long-lived clouds that are formed in connection to large-

scale uplifting or as residuals of convection, such as some cirrus clouds, since the time scale involved in their evolution ranges from a few hours up to a day (Boehm et al., 1999).

The model can also be regarded as a bulk microphysical module that can be incorporated into a more complex model which includes other microphysical processes (liquid water clouds, precipitating, convective clouds) as well as the necessary feedbacks on dynamical and radiative processes.

4.3.1 Model equations

Model equations are derived from the full continuity equation in flux form:

$$\frac{\partial n(m, x, y, z, t)}{\partial t} + \nabla \cdot [n(m, x, y, z, t)\underline{u}] = -P - VD + AG - AL \quad (4.1)$$

where $n(m, x, y, z, t)$ represents the mass distribution of ice crystals and \underline{u} is the vector velocity. The various terms in the RHS represents respectively:

- P = loss of particles of mass m by precipitation: $v(m) \frac{\partial n(m, x, y, z, t)}{\partial z}$
where $v(m)$ is the fall speed of a particle of mass m ;
- VD = loss of particles of mass m by vapor diffusion growth: $\frac{\partial \dot{m}n(m, x, y, z, t)}{\partial m}$
where \dot{m} is the mass time derivative;
- AG = gain by stochastic aggregation (production of particles of mass m due to aggregation of particles of mass $(m - m')$ and m'):
 $\frac{1}{2} \int_0^m n(m - m', x, y, z, t)n(m', x, y, z, t)K(m - m', m')g(T)dm'$
- AL = loss due to stochastic aggregation of particles of mass m and m' :
 $n(m, x, y, z, t) \int_0^\infty n(m', x, y, z, t)K(m, m')g(T)dm'$
where $K(m, m')$ represents the aggregation kernel.

Equation (4.1) can be explicitly rewritten as

$$\begin{aligned} & \frac{\partial n(m, x, y, z, t)}{\partial t} + \frac{\partial un(m, x, y, z, t)}{\partial x} + \frac{\partial vn(m, x, y, z, t)}{\partial y} + \frac{\partial [w - v(m)]n(m, x, y, z, t)}{\partial z} = \\ & - \frac{\partial \dot{m}n(m, x, y, z, t)}{\partial m} + \frac{1}{2} \int_0^m n(m - m', x, y, z, t)n(m', x, y, z, t)K(m - m', m')g(T)dm' \\ & - n(m, x, y, z, t) \int_0^\infty n(m', x, y, z, t)K(m, m')g(T)dm'. \end{aligned} \quad (4.2)$$

Assuming horizontal homogeneity and including the precipitation term in the vertical advection, the problem simplifies to:

$$\begin{aligned} & \frac{\partial n(m, z, t)}{\partial t} + \frac{\partial [w - v(m)]n(m, z, t)}{\partial z} = \\ & - \frac{\partial \dot{m}n(m, z, t)}{\partial m} + \frac{1}{2} \int_0^m n(m - m', z, t)n(m', z, t)K(m - m', m')g(T)dm' \\ & - n(m, z, t) \int_0^\infty n(m', z, t)K(m, m')g(T)dm'. \end{aligned} \quad (4.3)$$

Introducing the first and second moment of the mass distribution along with their respective fluxes:

first moment (Ice Water Content),	$\chi(z, t) = \int_0^\infty mn(m, z, t)dm$
Ice Water Content flux,	$\chi_f(z, t) = \int_0^\infty mv(m)n(m, z, t)dm$
second moment (\propto Radar Reflectivity),	$\zeta(z, t) = \int_0^\infty m^2n(m, z, t)dm$
second moment flux,	$\zeta_f(z, t) = \int_0^\infty m^2v(m)n(m, z, t)dm,$

Equation (4.3) can be used to write prognostic equations for the two moments of the mass distribution. The procedure is to multiply by m and m^2 respectively, and integrate over the mass distribution between 0 and ∞ . The resulting prognostic equations for χ and ζ are:

$$\frac{\partial \chi(z, t)}{\partial t} + \frac{\partial [w\chi(z, t) - \chi_f(z, t)]}{\partial z} = \int_0^\infty \dot{m}n(m, z, t)dm \quad (4.4)$$

$$\begin{aligned} \frac{\partial \zeta(z, t)}{\partial t} + \frac{\partial [w\zeta(z, t) - \zeta_f(z, t)]}{\partial z} &= \int_0^\infty 2m\dot{m}n(m, z, t)dm \\ &+ \int_0^\infty \int_0^\infty mm'n(m, z, t)n(m', z, t)K(m, m')g(T)dmdm' \end{aligned} \quad (4.5)$$

where the following relation from Drake (1972) was used:

$$\begin{aligned}
 & \int_0^\infty m^\gamma \left[\frac{1}{2} \int_0^m n(m-m', z, t) n(m', z, t) K(m-m', m') g(T) dm' - \right. \\
 & \left. n(m, z, t) \int_0^\infty n(m', z, t) K(m, m') g(T) dm' \right] dm = \\
 & = \frac{1}{2} \int_0^\infty \int_0^\infty [(m+m')^\gamma - m^\gamma - m'^\gamma] K(m, m') n(m, z, t) n(m', z, t) g(T) dm dm'.
 \end{aligned} \tag{4.6}$$

In the next section, equations (4.4) and (4.5) are used to derive prognostic equations for the parameters of a *size* distribution.

4.3.2 Model parameterizations

Equations for the first and second moment of the mass distribution derived in the previous section, can be reformulated through the introduction of a size distribution. The “modified gamma” distribution (Stephens et al., 1990) was chosen for this task, since it describes with relative realism observed size spectra in cirrus clouds and it has nice analytical properties. This distribution is also used in RAMS bulk microphysics scheme as described in section 3.3.1. The mathematical form of the modified gamma distribution is reported again here for convenience:

$$n_\gamma(D) = N_t \frac{1}{\Gamma(\nu)} \left(\frac{D}{D_n} \right)^{\nu-1} \frac{1}{D_n} e^{-(D/D_n)} \tag{4.7}$$

As mentioned, $n_\gamma(D)$ is fully described by the parameters N_t , D_n , and ν . For simplicity, it is assumed that the width of the distribution ν is a constant and does not vary with time and height. Following Dowling and Radke (1990) we set it equal to 2. An important property of the gamma-distribution is the following:

$$\int_0^\infty D^P n_\gamma(D) dD = N_t \frac{\Gamma(P + \nu)}{\Gamma(\nu)} D_n^P. \tag{4.8}$$

Some parameterizations are introduced in order to derive governing equations for the evolution of characteristic diameter and number concentration. The mass m

of an ice crystal and its fall speed $v(D)$ are parameterized in terms of the crystal diameter D :

$$m = \alpha D^\beta \quad (4.9)$$

$$v = a D^b \quad (4.10)$$

where α , β , a and b are parameters that depend on the shape of the ice crystals (e.g. Mitchell, 1988 and 1991).

Using these parameterizations and recognizing that $n(m)dm = n(D)dD$ for any size and mass distribution, we can express χ , χ_f , ζ and ζ_f in terms of D_n and N_t , using the property of the gamma distribution given in equation (4.8):

$$\chi = \alpha N_t D_n^\beta \frac{\Gamma(\beta + \nu)}{\Gamma(\nu)} \quad (4.11)$$

$$\chi_f = \alpha a N_t D_n^{\beta+b} \frac{\Gamma(\beta + b + \nu)}{\Gamma(\nu)} = a D_n^b \frac{\Gamma(\beta + b + \nu)}{\Gamma(\beta + \nu)} \chi \quad (4.12)$$

$$\zeta = \alpha^2 N_t D_n^{2\beta} \frac{\Gamma(2\beta + \nu)}{\Gamma(\nu)} \quad (4.13)$$

$$\zeta_f = \alpha^2 a N_t D_n^{2\beta+b} \frac{\Gamma(2\beta + b + \nu)}{\Gamma(\nu)} = a D_n^b \frac{\Gamma(2\beta + b + \nu)}{\Gamma(2\beta + \nu)} \zeta \quad (4.14)$$

Parameterizations also need to be introduced for the terms that appear as integrals in the RHS of equations (4.4) and (4.5). Starting with the growth by diffusion term, it can be shown from thermodynamical considerations that the growth rate by vapor diffusion is a power law in the diameter of the growing particle:

$$\dot{m} = f(T, p, q) D^\delta \quad (4.15)$$

where \dot{m} is the mass time rate change, δ is a coefficient that takes into account ventilation effects (i.e. for $\delta = 1$ no ventilation effects are included) and $f(T, p, q)$ is a thermodynamic coefficient that depends on temperature pressure and specific humidity. Factors that are related to the shape of the ice crystal which determine its capacitance,¹ are also included in f . In what follows, it is assumed that the

¹An analogy between diffusion of water vapor toward a growing particle and diffusion of charge toward a conductor can be made. The capability of an object to accumulate charge is described by

dominating shape for vapor diffusional growth processes is hexagonal plate, in which case $f(T, p, q)$ is simply:

$$f(T, p, q) = 4(S_i - 1)/(A + B) \quad (4.16)$$

where S_i is the supersaturation with respect to ice and A and B are constants related to properties of air, such as vapor diffusivity and heat conductivity (e.g., Pruppacher and Klett, 1997).

Using equation (4.15) and equation (4.8), the evaluation of the vapor diffusion integrals is straightforward:

$$\int_0^\infty f(T, p, q) D^\delta n(D) dD = f(T, p, q) N_t \frac{\Gamma(\nu + \delta)}{\Gamma(\nu)} D_n^\delta \quad (4.17)$$

$$\int_0^\infty 2f(T, p, q) \alpha D^{\delta+\beta} n(D) dD = 2f(T, p, q) \alpha N_t \frac{\Gamma(\nu + \delta + \beta)}{\Gamma(\nu)} D_n^{\delta+\beta} \quad (4.18)$$

Similarly the aggregation integral that appears in the RHS of equation (4.5) can be evaluated by assuming that the aggregation kernel has the form:

$$K(D, D') = \frac{\pi}{4} (D + D')^2 E_c |v(D) - v(D')| \quad (4.19)$$

with E_c , the collision efficiency. Set the aggregation efficiency² E_a equal to $g(T)E_c$, where $g(T)$ is a temperature dependent coefficient (Mitchell, 1994) and use (4.19) to get:

$$\begin{aligned} & \int_0^\infty \int_0^\infty mm' n(m, z, t) n(m', z, t) K(m, m') g(T) dm dm' = \\ & \int_0^\infty \int_0^\infty \alpha^2 D^\beta D'^\beta n_\gamma(D) n_\gamma(D') \frac{\pi}{4} (D + D')^2 E_a a |D^b - D'^b| dD dD' = \\ & \alpha^2 \frac{\pi}{4} E_a a \int_0^\infty \int_0^\infty \alpha^2 D^\beta D'^\beta n_\gamma(D) n_\gamma(D') (D + D')^2 |D^b - D'^b| dD dD' \end{aligned} \quad (4.20)$$

its *capacitance*. Similarly the capability of an ice crystal to accumulate water vapor is described by a capacitance which is shape related as in electrostatic theory. The capacitance of a spherical particle is simply equal to its radius; for an hexagonal plate, $C = \frac{2r}{\pi}$, where r is the ice crystal radius.

²Uncertainties in collision and aggregation efficiency are very large, especially for ice clouds. Direct measurements are practically inexistent; parameterizations are made difficult by the dependence of these quantities on habit of colliding crystals, turbulence, and ambient conditions.

Change variables to obtain:

$$\alpha^2 \frac{\pi}{4} E_a a \int_0^\infty \int_0^\infty \alpha^2 D^\beta D'^\beta n_\gamma(D) n_\gamma(D') (D + D')^2 |D^b - D'^b| dD dD' = \quad (4.21)$$

$$\alpha^2 \frac{\pi}{4} E_a a \frac{1}{\Gamma(\nu)^2} D_n^{2\beta+2+b} \int_0^\infty \int_0^\infty \alpha^2 \xi^{(\beta+\nu-1)} \zeta^{(\beta+\nu-1)} (\xi + \zeta)^2 |\xi^b - \zeta^b| e^{-(\xi+\zeta)} d\xi d\zeta = I_b$$

The integral above can be computed either numerically or analytically (Pasarelli, 1978a) as a function of the given parameters β , ν and b . However, it just represents a constant indicated in what follows with the symbol I_b .

Equations (4.4) and (4.5) can now be rewritten in the following form:

$$\frac{\partial \chi}{\partial t} + \frac{\partial [w\chi(z, t) - \chi_f(z, t)]}{\partial z} = f(T, p, q) N_t \frac{\Gamma(\nu + \delta)}{\Gamma(\nu)} D_n^\delta \quad (4.22)$$

$$\frac{\partial \zeta(z, t)}{\partial t} + \frac{\partial [w\zeta(z, t) - \zeta_f(z, t)]}{\partial z} = 2f(T, p, q) \alpha N_t \frac{\Gamma(\nu + \beta + \delta)}{\Gamma(\nu)} D_n^{\beta+\delta} + \frac{1}{\Gamma(\nu)^2} \alpha^2 N_t^2 \frac{\pi}{4} E_a a D_n^{2\beta+2+b} I_b \quad (4.23)$$

By using relations (4.11)–(4.14) in equations (4.22) and (4.23), a system of two coupled nonlinear prognostic equations for D_n and N_t is derived as follows. Divide equation (4.22) by χ and equation (4.23) by ζ , and simplify the RHS using (4.11) and (4.13) to get:

$$\frac{1}{\chi} \frac{\partial \chi}{\partial t} + \frac{1}{\chi} \frac{\partial [w\chi(z, t) - \chi_f(z, t)]}{\partial z} = \frac{f(T, p, q)}{\alpha} k_3 D_n^{\delta-\beta} \quad (4.24)$$

$$\frac{1}{\zeta} \frac{\partial \zeta(z, t)}{\partial t} + \frac{1}{\zeta} \frac{\partial [w\zeta(z, t) - \zeta_f(z, t)]}{\partial z} = 2 \frac{f(T, p, q)}{\alpha} k_4 D_n^{\delta-\beta} + \frac{1}{k_5} N_t G_1 a D_n^{2+b} \quad (4.25)$$

where $G_1 = \frac{\pi}{4} E_a I_b$ and:

$$k_1 = \frac{\Gamma(\nu+b+\beta)}{\Gamma(\nu+\beta)}, \quad k_2 = \frac{\Gamma(\nu+2\beta+b)}{\Gamma(\nu+2\beta)}, \quad k_3 = \frac{\Gamma(\nu+\delta)}{\Gamma(\nu+\beta)}, \quad k_4 = \frac{\Gamma(\nu+\beta+\delta)}{\Gamma(\nu+2\beta)}, \quad k_5 = \Gamma(\nu)\Gamma(\nu+2\beta)$$

Writing the two moments as:

$$\ln \chi = \ln \left(\alpha \frac{\Gamma(\beta + \nu)}{\Gamma(\nu)} \right) + \ln N_t + \beta \ln D_n$$

$$\ln \zeta = \ln \left(\alpha^2 \frac{\Gamma(2\beta + \nu)}{\Gamma(\nu)} \right) + \ln N_t + 2\beta \ln D_n$$

and assuming α , β , a and b are constant with height and time, it follows that:

$$\begin{aligned}\frac{\partial \ln \chi}{\partial t} &= \frac{\partial \ln N_t}{\partial t} + \beta \frac{\partial \ln D_n}{\partial t} \\ \frac{\partial \ln \zeta}{\partial t} &= \frac{\partial \ln N_t}{\partial t} + 2\beta \frac{\partial \ln D_n}{\partial t}\end{aligned}$$

and similarly

$$\begin{aligned}\frac{\partial \ln \chi}{\partial z} &= \frac{\partial \ln N_t}{\partial z} + \beta \frac{\partial \ln D_n}{\partial z} \\ \frac{\partial \ln \zeta}{\partial z} &= \frac{\partial \ln N_t}{\partial z} + 2\beta \frac{\partial \ln D_n}{\partial z}\end{aligned}$$

If one is to include a time or height dependence for the model parameters, the above expressions have to be modified to include their derivatives with respect to t or z . Finally, rearranging the terms which contain the first and second moment fluxes, χ_f and ζ_f :

$$\begin{aligned}\frac{1}{\chi} \frac{\partial \chi_f}{\partial z} &= ak_1 D_n^b \left(b \frac{\partial \ln D_n}{\partial z} + \frac{\partial \ln \chi}{\partial z} \right) \\ \frac{1}{\zeta} \frac{\partial \zeta_f}{\partial z} &= ak_2 D_n^b \left(b \frac{\partial \ln D_n}{\partial z} + \frac{\partial \ln \zeta}{\partial z} \right)\end{aligned}$$

and combining all of the above expressions together, two *coupled* differential equations for D_n and N_t are obtained:

$$\begin{aligned}\frac{\partial \ln N_t}{\partial t} + \beta \frac{\partial \ln D_n}{\partial t} + w \frac{\partial \ln N_t}{\partial z} + w\beta \frac{\partial \ln D_n}{\partial z} + \frac{\partial w}{\partial z} \\ - ak_1 D_n^b \left(b \frac{\partial \ln D_n}{\partial z} + \frac{\partial \ln N_t}{\partial z} + \beta \frac{\partial \ln D_n}{\partial z} \right) &= \frac{f(T, p, q)}{\alpha} k_3 D_n^{\delta-\beta} \quad (4.26) \\ \frac{\partial \ln N_t}{\partial t} + 2\beta \frac{\partial \ln D_n}{\partial t} + w \frac{\partial \ln N_t}{\partial z} + w2\beta \frac{\partial \ln D_n}{\partial z} + \frac{\partial w}{\partial z} \\ - ak_2 D_n^b \left(b \frac{\partial \ln D_n}{\partial z} + \frac{\partial \ln N_t}{\partial z} + 2\beta \frac{\partial \ln D_n}{\partial z} \right) &= 2 \frac{f(T, p, q)}{\alpha} k_4 D_n^{\delta-\beta} \\ &+ \frac{1}{k_5} N_t G_1 a D_n^{2+b}. \quad (4.27)\end{aligned}$$

By subtracting equation (4.26) from equation (4.27) and dividing through by β , an equation for the time and height evolution of the characteristic diameter is

derived:

$$\begin{aligned} \frac{\partial \ln D_n}{\partial t} + w \frac{\partial \ln D_n}{\partial z} - \frac{a}{\beta} D_n^b (k_2 - k_1) b \frac{\partial \ln D_n}{\partial z} - a D_n^b (2k_2 - k_1) \frac{\partial \ln D_n}{\partial z} &= \quad (4.28) \\ &= \frac{f(T, p, q)}{\alpha \beta} D_n^{\delta-\beta} (2k_4 - k_3) + \frac{1}{k_5 \beta} N_t G_1 a D_n^{2+b} - \frac{a}{\beta} D_n^b (k_1 - k_2) \frac{\partial \ln N_t}{\partial z}. \end{aligned}$$

Similarly, by multiplying equation (4.26) by 2 and by subtracting equation (4.27) from the product, an equation for the number concentration is obtained:

$$\begin{aligned} \frac{\partial \ln N_t}{\partial t} + w \frac{\partial \ln N_t}{\partial z} - a D_n^b (2k_1 - k_2) \frac{\partial \ln N_t}{\partial z} + \frac{\partial w}{\partial z} &= \\ &= 2 \frac{f(T, p, q)}{\alpha} D_n^{\delta-\beta} (k_3 - k_4) - \frac{1}{k_5} N_t G_1 a D_n^{2+b} + a 2\beta D_n^b (k_1 - k_2) \frac{\partial \ln D_n}{\partial z} \\ &\quad - ab D_n^b (2k_1 - k_2) \frac{\partial \ln D_n}{\partial z}. \quad (4.29) \end{aligned}$$

Selection of a maximum value for D_n , D_{max} , and for N_t , N_{max} , allows to introduce new variables:

$$x = \ln \left(\frac{D_n}{D_{max}} \right) \quad (4.30)$$

$$y = \ln \left(\frac{N_t}{N_{max}} \right) \quad (4.31)$$

and, upon rearrangement, to derive a new set of equations in the nondimensional variables x and y :

$$\begin{aligned} \frac{\partial x}{\partial t} + \left[w - \frac{a}{\beta} b D_{max}^b (k_2 - k_1) e^{bx} - a (2k_2 - k_1) D_{max}^b e^{bx} \right] \frac{\partial x}{\partial z} &= \\ &= \frac{f(T, p, q)}{\alpha \beta} (2k_4 - k_3) D_0^{\delta-\beta} e^{(\delta-\beta)x} + \frac{N_0 e^y}{k_5 \beta} G_1 a D_0^{2+b} e^{(2+b)x} \\ &\quad - \frac{a}{\beta} (k_1 - k_2) D_0^b e^{bx} \frac{\partial y}{\partial z} \quad (4.32) \end{aligned}$$

$$\begin{aligned} \frac{\partial y}{\partial t} + \left[w - a D_0^b e^{bx} (2k_1 - k_2) \right] \frac{\partial y}{\partial z} &= \\ &= 2 \frac{f(T, p, q)}{\alpha} (k_3 - k_4) D_0^{\delta-\beta} e^{(\delta-\beta)x} - \frac{N_0 e^y}{k_5} G_1 a D_0^{2+b} e^{(2+b)x} \\ &\quad + a D_0^b e^{bx} [b(2k_1 - k_2) + 2\beta(k_1 - k_2)] \frac{\partial x}{\partial z} - \frac{\partial w}{\partial z}. \quad (4.33) \end{aligned}$$

Equations (4.28)-(4.29) or their equivalent (4.32)-(4.33) constitute the system of equations that describe the vertical and temporal evolution of D_n and N_t .

4.3.3 Numerical implementation

The numerical implementation is performed using a discrete version of the model on a time and height grid constituted of equally spaced points. The numerical integration uses the leap-frog scheme with the Robert time filter Asselin (1972) to eliminate the computational mode. The maximum timestep is determined by the Courant criterion for numerical stability. It was found that the plain leap-frog scheme yielded a numerically unstable solution for equations (4.32) and (4.33). The leap-frog method with LAX correction offered a stable solution, but introduced too large of a numerical dissipation (Press et al., 1992). A compromise was reached by adding a dissipation term proportional to an arbitrary coefficient to equation (4.32). The value for the coefficient was chosen to be equal to the smallest positive number that allowed a stable solution. No dissipation was added to the prognostic equation for the number concentration (4.33).

Coefficients that appear in the mass-diameter and fall speed-diameter parameterizations are taken from Mitchell (1988). Specific values used in the model are relative to a mixture of unrimed plates, side planes, bullets and columns, and are given here in MKS units:

$$\alpha = 18.3010^{-3}, \quad \beta = 1.9, \quad a = 11.7, \quad b = 0.41$$

In what follows, we will assume that these parameters are not functions of time and height. This is not true in reality because as the ice crystals grow and change in shape and size, the parameters that define mass-diameter and fallspeed-diameter relationships also change. In addition the ice growth rate has a dependence on crystal shape as pointed out in a study by Mitchell and Chai (1998). In section 4.4, model output sensitivity to choice of model parameters is discussed in detail.

The initial condition for characteristic diameter and number concentration can be specified from various options:

1. Constant characteristic diameter and number concentration computed from supersaturation with respect to ice, using a parameterization introduced by Meyers et al. (1992). This parameterization describes heterogeneous ice nucleation processes in RAMS, and prescribes the number of nucleated crystals as

an exponential function of supersaturation with respect to ice (S_i). It must be observed, though, that a more realistic treatment of ice crystal nucleation processes must include also homogeneous freezing of droplets and haze particles which can be dominant at cold cirrus temperatures.

$$N_{IN} \approx e^{-0.639+0.1296S_i}, \quad (4.34)$$

2. Number concentration computed as in step 1., and characteristic diameter derived from a parameterization used in the ECMWF model, cubic in the difference between ambient temperature (T) and reference temperature ($T_0 = 273$ K):

$$D_n \approx 326.3 + (T - T_0)(12.42 + (T - T_0)(0.197 + (T - T_0)0.0012)). \quad (4.35)$$

3. Number concentration computed as in step 1., and characteristic diameter derived from N_i and IWC, parameterized in term of vertical velocity and temperature, following Heymsfield (1977) as reported in Cotton and Anthes (1989):

$$IWC \approx 0.072w^{0.78}e^{(-0.01w^{0.186}(-T^{1.59w^{0.04}}))}. \quad (4.36)$$

It is assumed that the cloud layer is confined in height and no cloud is present at the lowest and uppermost level of the z -grid which extends from the surface to the upper troposphere. Boundary conditions are thus simplified.

4.3.4 Observational operators

The radar mapping

In chapter 2, the notion of an observational operators was discussed. In this section, the general concepts are applied to assimilation of radar data. The model equivalent

radar reflectivity in the Rayleigh approximation can be obtained from the following definition:

$$Z = \int_0^{\infty} n_{\gamma}(D) D^6 dD \quad (4.37)$$

where the size distribution is assumed to be a modified gamma distribution (see equation (4.7)), as done in the development of the cloud model equations. Note that the underlying assumption in the derivation of (4.37) is that the ice particles behave as Rayleigh³ scatterers at the radar wavelengths of interest (i.e., 3.19 or 8.57 mm for cloud radars), hence the dependence on the sixth power of the diameter. We hence expect much higher sensitivity of radar reflectivity to diameter than to number concentration (roughly sixfold).

By integrating with respect to D , we obtain:

$$Z = \frac{\Gamma(\nu + 6)}{\Gamma(\nu)} N_t D_n^6 \quad (4.38)$$

where the general property of the modified-gamma distribution was used again:

$$\int_0^{\infty} n_{\gamma}(D) D^P dD = \frac{\Gamma(\nu + P)}{\Gamma(\nu)} N_t D_n^P.$$

For larger crystals, the Rayleigh radar reflectivity overestimates the reflectivity, due to its extreme dependence on particle diameter. To prevent this from happening, a correction based on exact Mie calculations for the radar reflectivity which depends on particle size was implemented.

Radar observations are usually given in equivalent radar reflectivity, which is the normalized power that a population of water droplets would backscatter. For comparison purposes, it is necessary to transform model ice reflectivities into equivalent reflectivities. In order to do so, account must be taken of the variation of the refractive index of ice with ice particle density. This is especially important for larger aggregates: an ice aggregate of a given size and with a low density (i.e.

³Particles are said to be Rayleigh scatterers if they are much smaller than the wavelength λ of the incident radiation. Introducing the size parameter, $x = \frac{2\pi r}{\lambda}$ where r is the radius of the particle, the Rayleigh approximation applies if $x \ll 1$.

with air pockets, and holes in it) will backscatter less than a solid ice particle of the same size. In the comparisons with the radar observations, a density correction to reflectivity was applied (courtesy of Dr. S. Matrosov, private communication).

Optical depth

The total optical depth represents a path integrated quantity:

$$\tau = \int_{z_b}^{z_t} \sigma_{ext}(z) dz, \quad (4.39)$$

where $\sigma_{ext}(z)$ represents the extinction coefficient profile and the limits of the integral are cloud base z_b and cloud top z_t . At visible wavelengths the extinction coefficient is given by:

$$\sigma_{ext} = \int_0^{\infty} n_{\gamma}(D) \frac{\pi}{4} D^2 Q_{ext}(D) dD. \quad (4.40)$$

By assuming that cirrus crystals are much larger than the wavelength so that the extinction efficiency, $Q_{ext} \rightarrow 2$, equation (3.14) is integrated to:

$$\sigma_{ext}(z) = \frac{\pi}{2} \frac{\Gamma(\nu + 2)}{\Gamma(\nu)} N_t D_n^2. \quad (4.41)$$

From equation (4.41) it can be seen that the dependence of σ_{ext} on D_n is quadratic as opposed to the sixth power dependence of radar reflectivity. We expect the sensitivity of τ to the characteristic diameter to be twice as large in magnitude with respect to the sensitivity of τ to the number concentration. The different sensitivities of τ and Z on particle size and number concentration are also directly exploited in the IWC retrieval schemes introduced in the appendix.

4.3.5 Some comparison with model and observed cirrus spectra

In its original formulation (Mitchell, 1988), the model predicts ice crystal concentrations and sizes of ice particles given an input IWC. Comparisons show favorable agreement of model versus measured ice particle spectra as discussed in Mitchell (1988) and (1994), and more recently in Lawson et al. (1998) and Mitchell and Chai (1998). However, the formulation presented here largely differs from the original

one since no information on the cloud IWC is provided as input, and the cloud is formed from a supersaturated profile. In order to verify the degree of realism of the model, it is necessary to compare model results with observations. This is an essential part of the model skill assessment process discussed in chapter 2. To this end, two case studies were considered. One is the ARM-UAV Spring '99 campaign, which is described in chapter 3. The other one is part of the cirrus observations carried out at the ARM Southern Great Plain site during an Intensive Operational Period (IOP) in Spring 1998. This latter case study is introduced below. A third case study from the Fall 1996 ARM-SGP IOP is presented in appendix B. Data from the ARM-SGP IOPs were provided by Dr. G.Mace.

Case study 1: Oklahoma, 8 May 1998

The cloud observed on May 8 appeared homogeneous and thick, as shown in the time series of the Millimeter Wave Cloud Radar ⁴ (MMCR) radar backscattering. Figure 4.1 shows a vertical cross-section of radar reflectivity derived from the MMCR signal along with a time series of cloud optical depth derived from Raman lidar data (courtesy of Zhien Wang, University of Utah). The black solid line represents the flight track of the University of North Dakota (UND) Citation, carrying several in situ instruments. Microphysical quantities such as number concentration and mean size were measured with a 2D-C probe. These observations are used in conjunction with the radar profiles to evaluate the performance of the ice growth cloud model.

Figure 4.2 shows the time series of temperature, relative humidity with respect to ice and aircraft altitude as recorded on board of the UND Citation. Figure 4.3 presents the time series of particle number concentration and mean size measured by the 2D-C probe. The average mean size is around $250 \mu m$ and the average number concentration is around 20 particles per liter with peaks of $80 l^{-1}$.

Initial profiles of pressure, temperature, and specific humidity to initialize the

⁴The MMCR is a zenith-pointing radar that operates at a frequency of 35 GHz. The main purpose of this radar is to determine cloud boundaries (e.g., cloud bottoms and tops). This radar also reports radar reflectivity (dBZ) of the atmosphere up to 20 km. The radar possesses a Doppler capability that allows the measurement of cloud constituent vertical velocities.

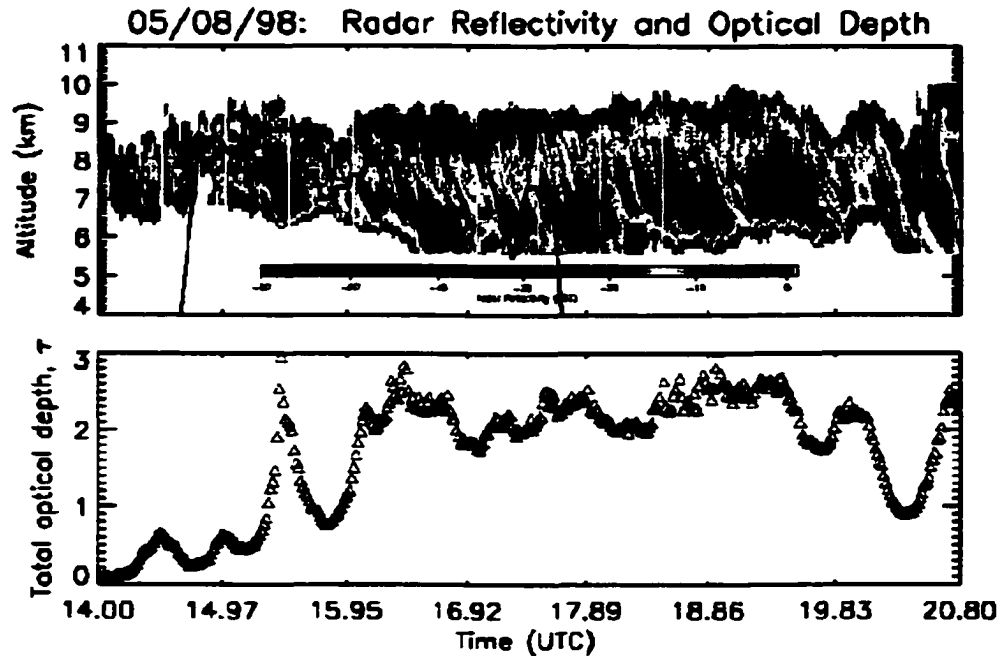


Figure 4.1: Radar reflectivity and total optical depth time series for a thick cirrus layer observed over the ARM SGP site, Oklahoma, on May 8, during the Spring 1998 Cloud IOP.

model are from radiosonde measurements which are routinely taken over the ARM central facility. The sounding we used, taken at 17:30UTC, is shown in figure 4.4. Cloud temperatures range from -17°C at cloud base to -37°C at cloud top.

The relative humidity profile shows a relatively dry layer right below 5.6 km, and moist layer between 6 and 12 km. This sounding is used to provide initial environmental fields to the the ice growth model. Type 1 initialization for cloud variables was used, and the initial constant characteristic diameter was set equal to $100 \mu\text{m}$. For comparison purposes, the mean size for the gamma distribution is computed from the following integral:

$$\bar{D} = \frac{\int_0^{\infty} Dn(D)dD}{\int_0^{\infty} n(D)dD} = \nu D_n. \quad (4.42)$$

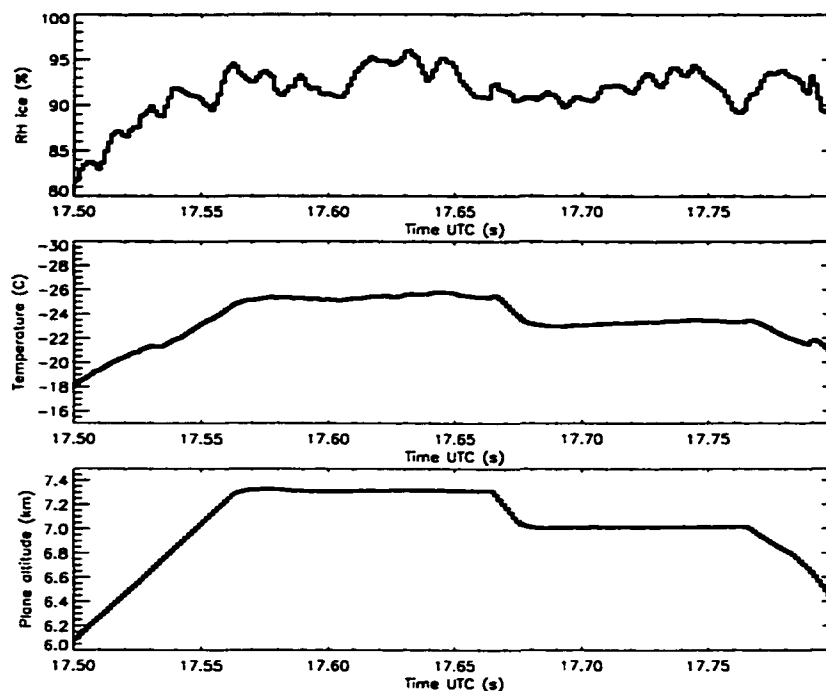


Figure 4.2: Relative humidity with respect to ice (upper panel), temperature (middle panel), and plane altitude (lower panel) recorded between 17:30 and 17:48 UTC on board of the UN Citation on May 8, 1998.

In situ observations of ice particle number concentration and mean size were compared with the corresponding model fields over approximately a half hour interval from 17:30 UTC to 18:00 UTC. Model-predicted reflectivities were also compared with MMCR profiles averaged over the same time interval.

Figure 4.5 shows the microphysical parameters predicted by the model versus the in situ values, the latter are rebinned to match model vertical resolution (100 m). As it appears, although the shape of the profiles is reasonable, the model number concentrations are much too low with respect to the in situ concentrations. The discrepancy is as large as two orders of magnitude. Note that this discrepancy can also be due to the lack of homogeneous nucleation processes in the model. Better agreement is found for values of mean size. This is reflected in the comparison between model and observed radar reflectivities (see figure 4.6).

The model predicts a radar reflectivity that is lower than the observed reflectivity due to the underprediction of the number of ice crystals. The difference is, however, less marked than the difference in number concentration, due to the fact that the mean size vertical structure is reasonably well represented, and radar reflectivity is more sensitive to size than number concentration (this will become evident in discussions in section 4.3.4). We notice that the reflectivity profile is slightly misplaced with respect to the observed profile. Since the base of a cloud is directly related to the fall speed of the cloud particles, as shown in chapter 3, this problem is “fixed” by using a lower coefficient a for the power law $v = aD^b$ (more on this in section 4.4). The result is a better coincidence of observed and model cloud base (see figure 4.7). The decreased value of the parameter ($a = 5.8$) is used throughout the rest of this case study comparison. This result suggests that a is an important model parameter that can be optimized using the observations.

The main reason for the discrepancy between observed and model number con-

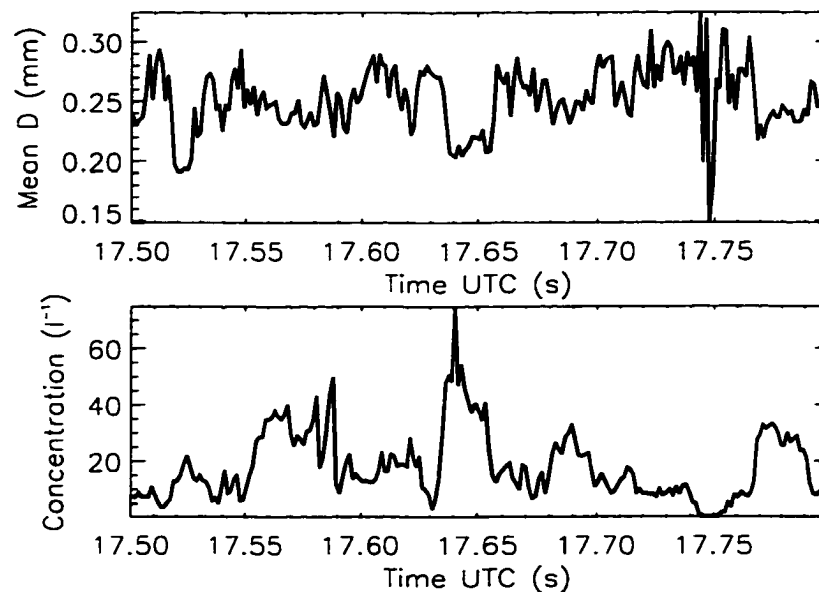


Figure 4.3: Ice particle mean size (upper panel), and number concentration (lower panel) measured between 17:30 and 17:48 UTC by the 2D-C probe on board of the UN Citation on May, 8 1998.

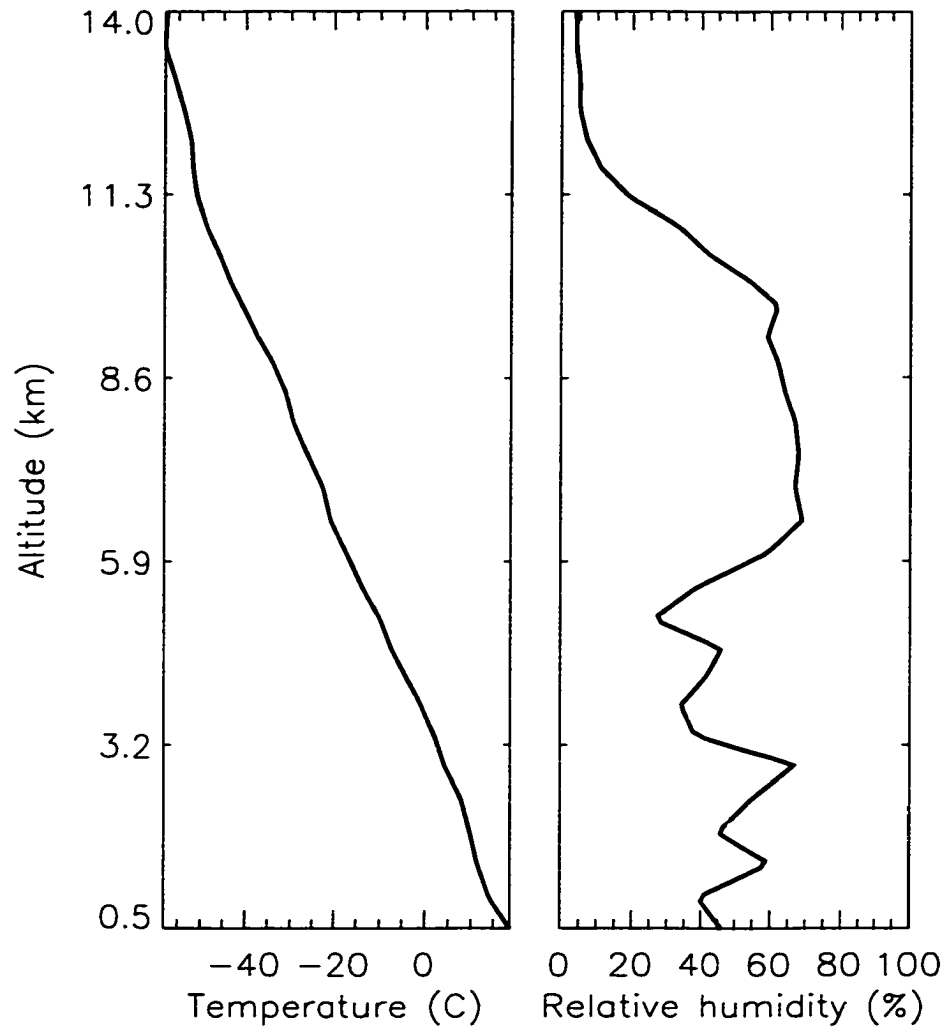


Figure 4.4: Temperature (upper panel) and relative humidity (lower panel) at 17:30 UTC on May, 8 1998. Sounding launched from the ARM-SGP central facility. Data were obtained from the ARM Program sponsored by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Environmental Sciences Division.

centrations resides in the fact that the initialization for N_t requires local supersaturation with respect to ice, but the initial sounding of figure 4.4 is not supersaturated with respect to ice at cloud levels. This is a problem often encountered when using radiosonde measurements which are known to lack accuracy in the humidity fields,

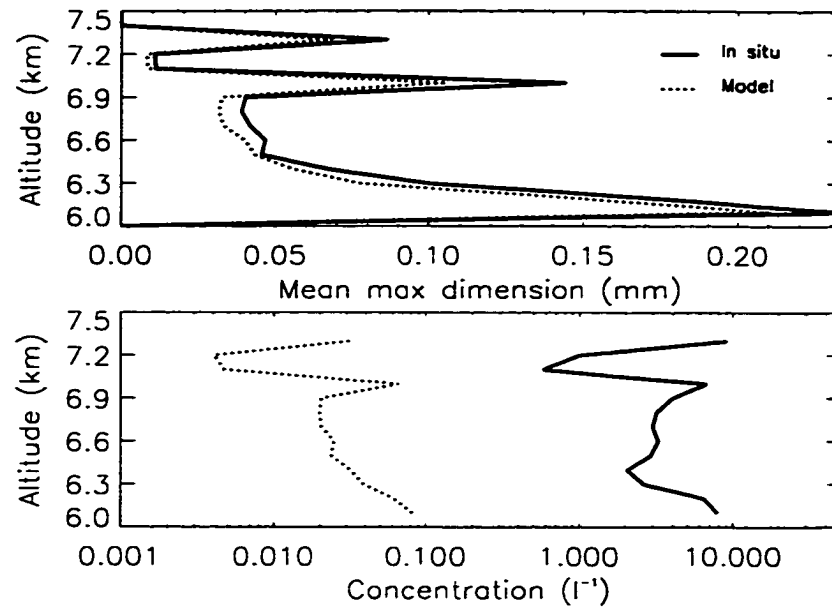


Figure 4.5: Comparison between observed (solid line) and model number concentrations and mean sizes (dashed line) averaged over a half hour interval.

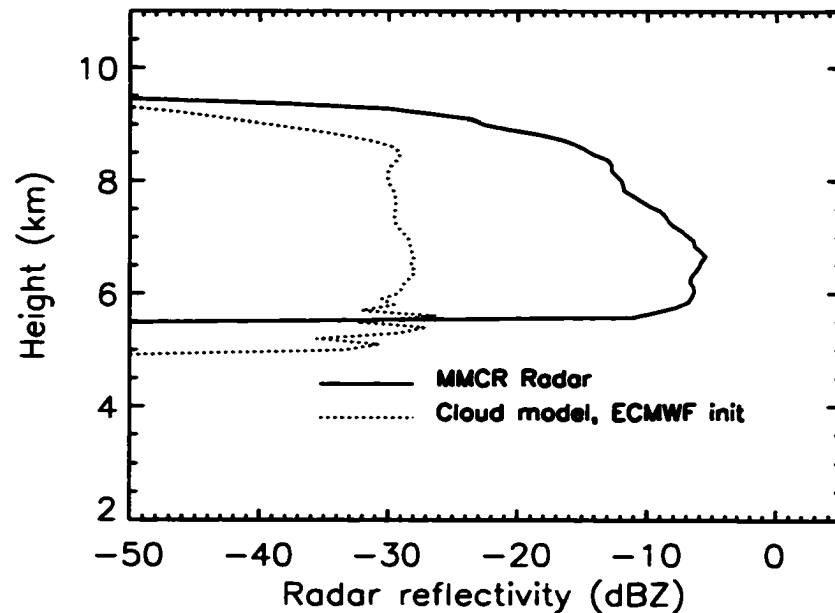


Figure 4.6: Comparison between observed (solid line) and model reflectivities (solid line) averaged over a half hour interval.

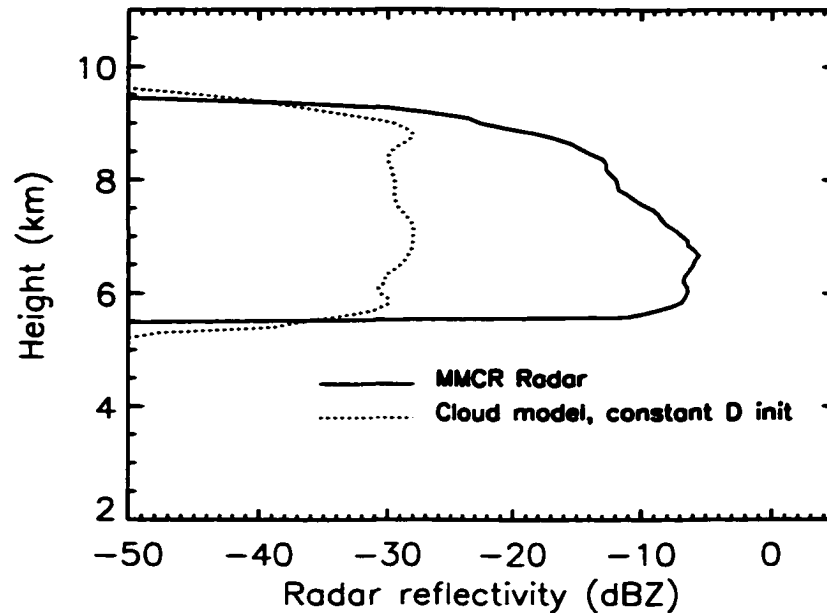


Figure 4.7: Same as in figure 4.6. Decreased value of a parameter (see text for explanations).

especially at upper tropospheric levels (Zipser and Johnson, 1998). Investigation showed that most radiosonde soundings presented a dry bias due to technical problems with the humidity sensor as reported in (Miller et al., 1999a). When using the relative humidity profile from radiosonde to initialize cloud models, a very long spin-up time for cloud formation in the model is often observed. Another reason for this discrepancy is the lack of a parameterization for homogeneous nucleation. However, in this particular case, the latter is not the dominating source of error since the cloud temperatures are comparatively warm, and do not go below the threshold of -38°C for homogenous nucleation (De Mott et al., 1994).

In situ measurements for the levels between 6 and 7.5 km do not show supersaturation with respect to ice either, but it must be remembered that they are taken *after* the formation of the cloud. The particle must have been exposed to large ice supersaturations in order to nucleate, but the growth processes act as supersaturation sinks, and unless new supersaturation is created by, for example vertical uplifting induced, for example, by radiative-convective overturning, the environment becomes

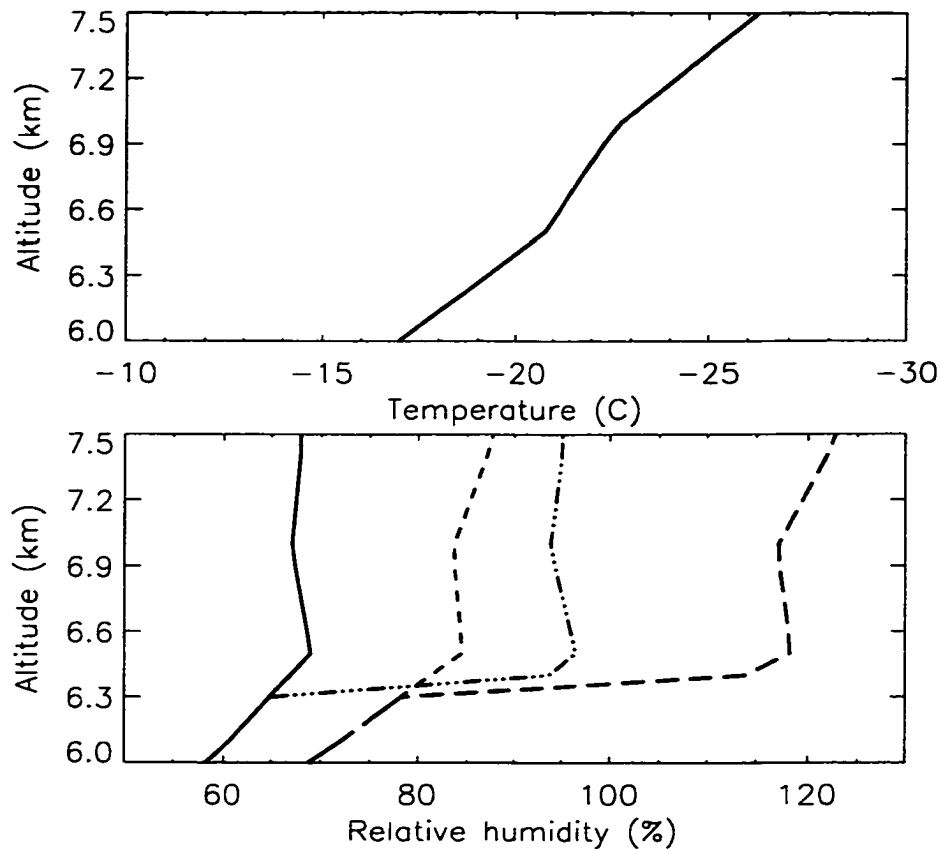


Figure 4.8: Atmospheric state variable profiles between 6 and 7.5 km. Temperature is shown in the upper panel. Relative humidity (RH) in the lower panel. The various profiles represents respectively: original RH wrt water (solid line), original RH wrt ice (short dashed), increased RH wrt water (dot-dashed), and increased RH wrt ice (long dashed).

undersaturated especially around cloud base. To verify this hypothesis, the initial relative humidity was arbitrarily increased to reach a level close to water saturation, thus producing a supersaturation with respect to ice that is large (figure 4.8).

This figure shows the relative humidity with respect to ice profile before and after the arbitrary increase, along with the relative humidity with respect to water. Note that the cloud levels are now close to being water saturated, and are ice supersaturated. If we allow higher ice supersaturations to occur, the resulting number

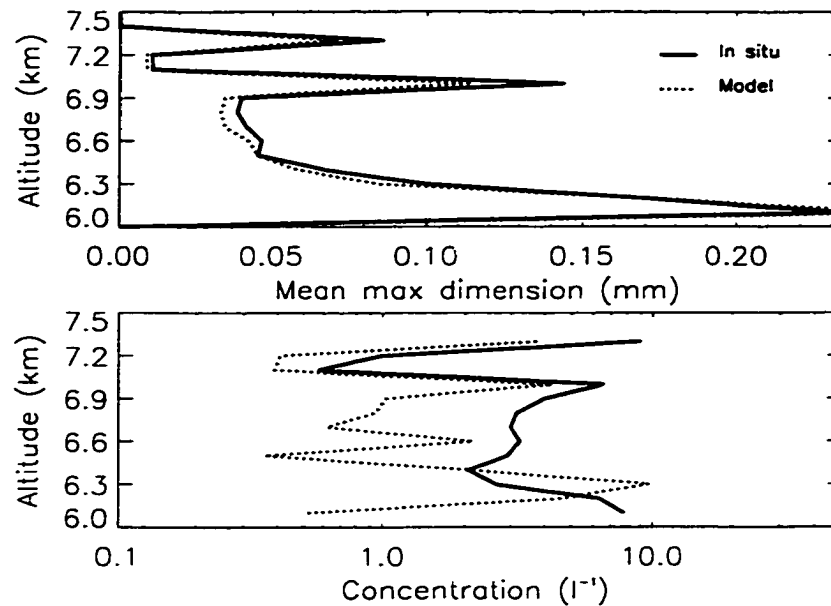


Figure 4.9: Comparison between observed (solid line) and model number concentrations and mean sizes (dashed line). Meyers et al. (1992) initialization for N_t and constant initialization for D_n ($100\mu m$). Increased value of RH wrt water (see text for explanations).

concentration becomes closer to the in situ observations, as exemplified in figure 4.9. The improvement in agreement between model and observed reflectivity is also noticeable (see fig. 4.10).

These results suggest that the model is capable of reproducing realistic fields, but due to the specific high sensitivity of the initial number concentration to supersaturation with respect to ice, the model can only operate under “controlled”, i.e. supersaturated, conditions. A positive aspect of this otherwise undesirable but expected high sensitivity is that there is information contained in the number concentration field regarding the initial supersaturation experienced by the cloud upon its formation. Since radar reflectivity is sensitive to number concentration, even if to a lesser degree than it is sensitive to characteristic size, some information may be extracted from radar measurements regarding the initial moist field that “forces” the cloud.

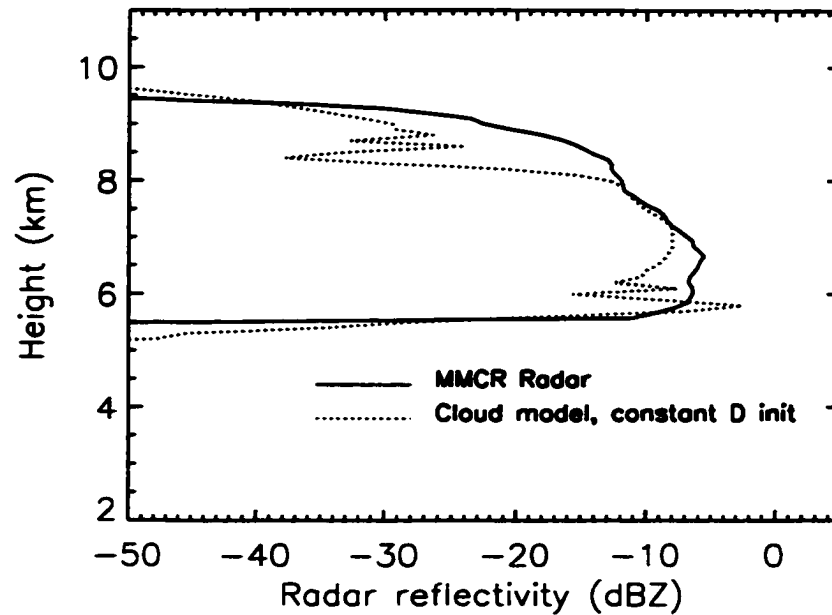


Figure 4.10: Same as figure 4.9, but for radar reflectivity.

When other types of initializations for D_n were used (different value of constant D_n , and ECMWF parameterization, see section 4.3.2), similar behaviors were observed as seen in figures 4.12 and 4.11. and 4.13 and 4.14.

Case study 2: Kauai, April 30 1999

Our second case study is the already familiar cirrus case observed on April 30, 1999 during the ARM-UAV Spring 1999 campaign. For details on this set of cirrus observations, we refer to section 3.4.1. For this case study, there was no availability of in situ measurements, and the comparison is carried out only between model predicted and observed radar reflectivities profiles. However, RAMS and ECMWF model results for the same case were also available, and were included in the comparison.

Initial atmospheric profiles (shown in figure 4.15) and vertical motion were taken from the ECMWF forecast over the study area. Cloud temperatures range from -17°C at cloud base to -52°C at cloud top. Relative humidity is low in the lower middle troposphere ($\approx 2.5\text{km}$), and higher in the middle troposphere. The profile

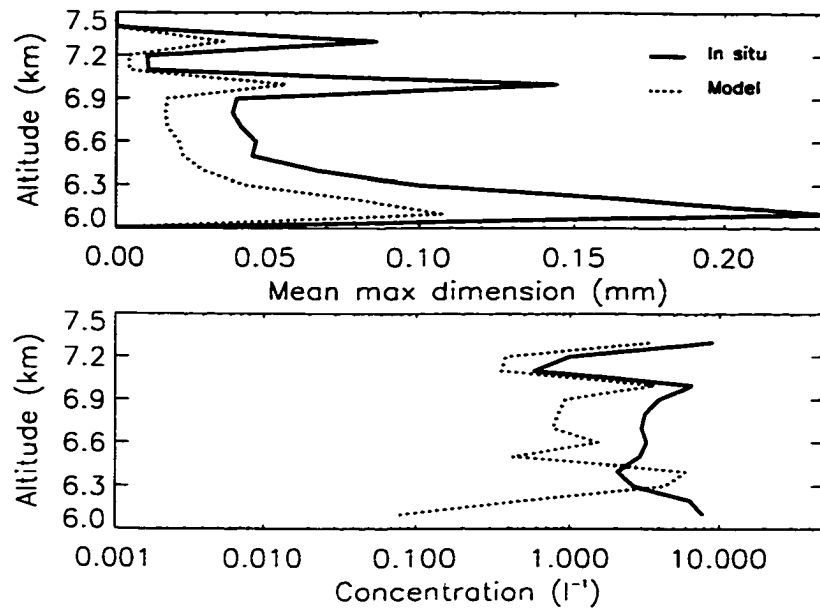


Figure 4.11: Same as figure 4.9, but for initial value of $D_n = 50\mu m$.

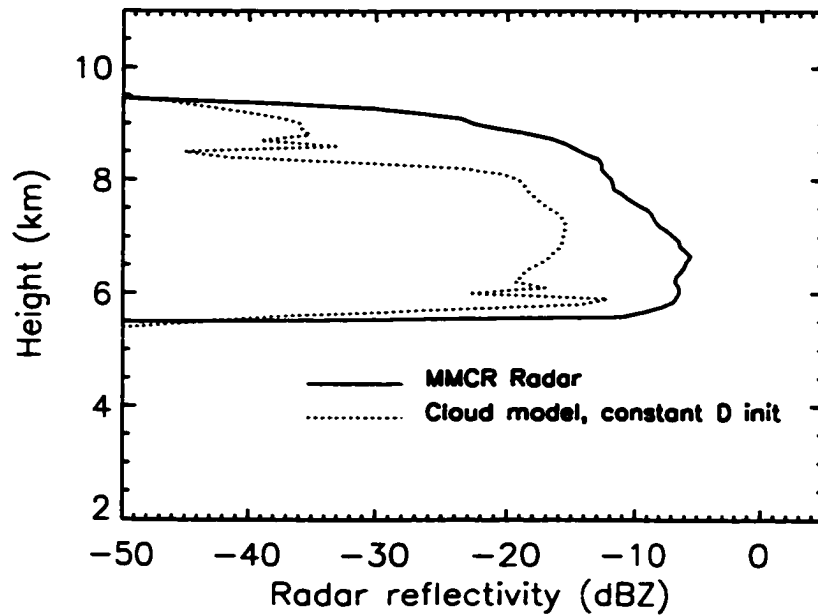


Figure 4.12: Same as figure 4.10, but for initial value of $D_n = 50\mu m$.

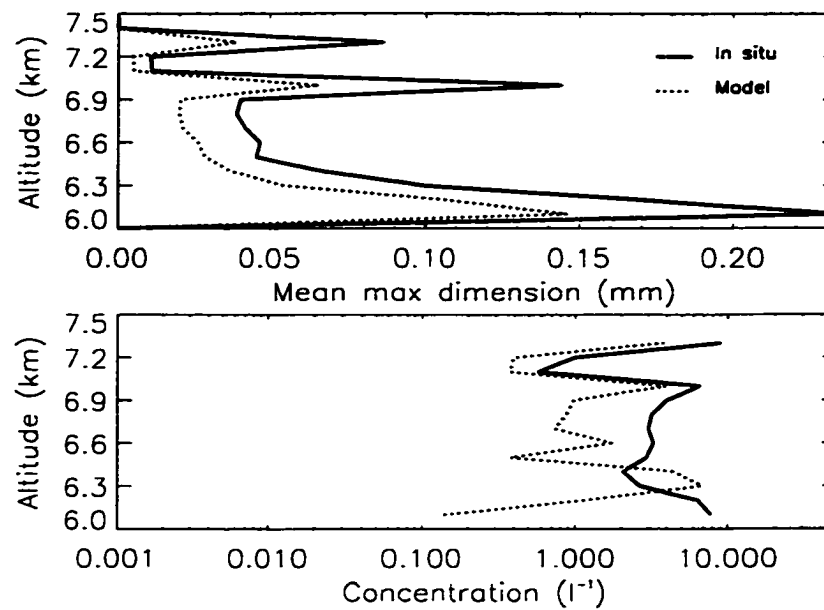


Figure 4.13: Same as figure 4.9, but for ECMWF initialization for D_n .

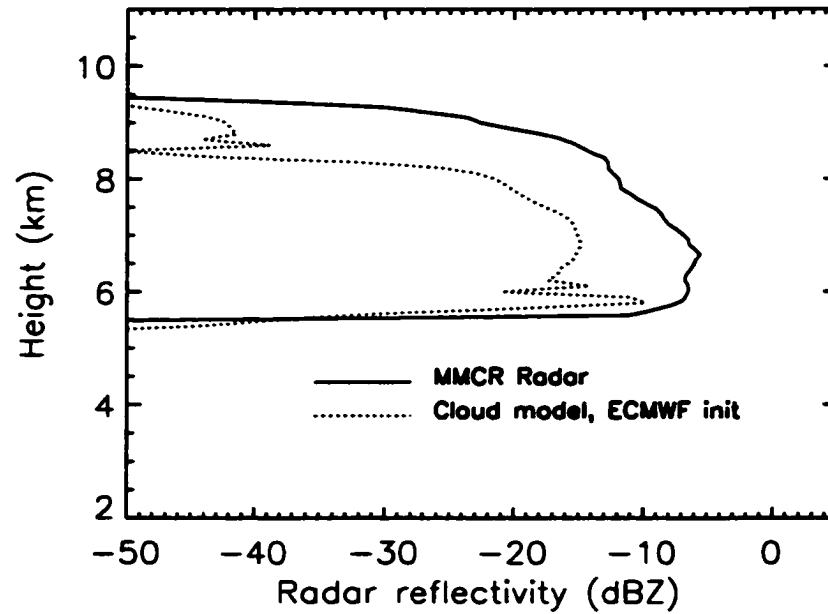


Figure 4.14: Same as figure 4.10, but for ECMWF initialization for D_n .

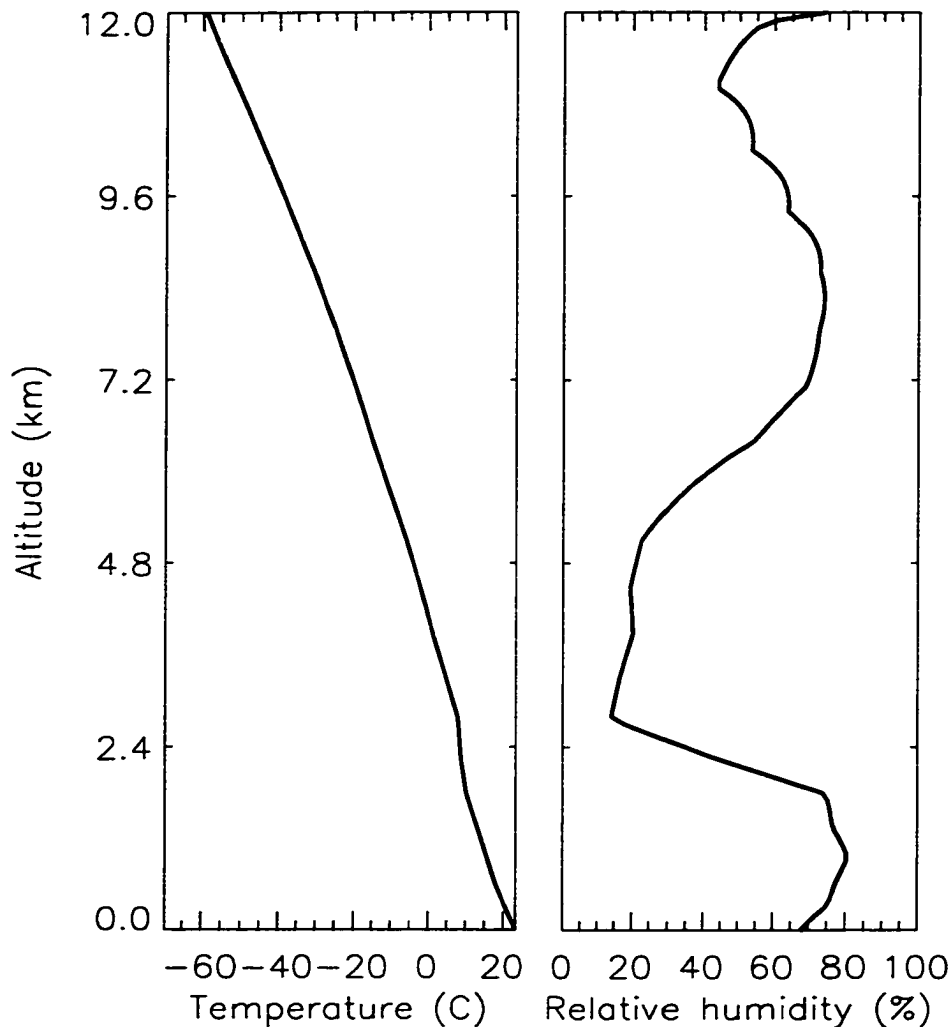


Figure 4.15: Temperature (left) and relative humidity (right) at 23:20 UTC on April, 30 1999, Kauai, Hawaii.

is not saturated with respect to water at cloud levels.

The same issues encountered in the SGP case study when using ice unsaturated profiles were found in this case study (see figure 4.16). The model radar reflectivity appears to be too low compared to observed values. In this particular case, the lack of homogeneous nucleation could be a large source of error, since cloud temperatures are low. This factor in addition to the subsaturated relative humidity profile contributes to the discrepancy between model and observed reflectivities. When

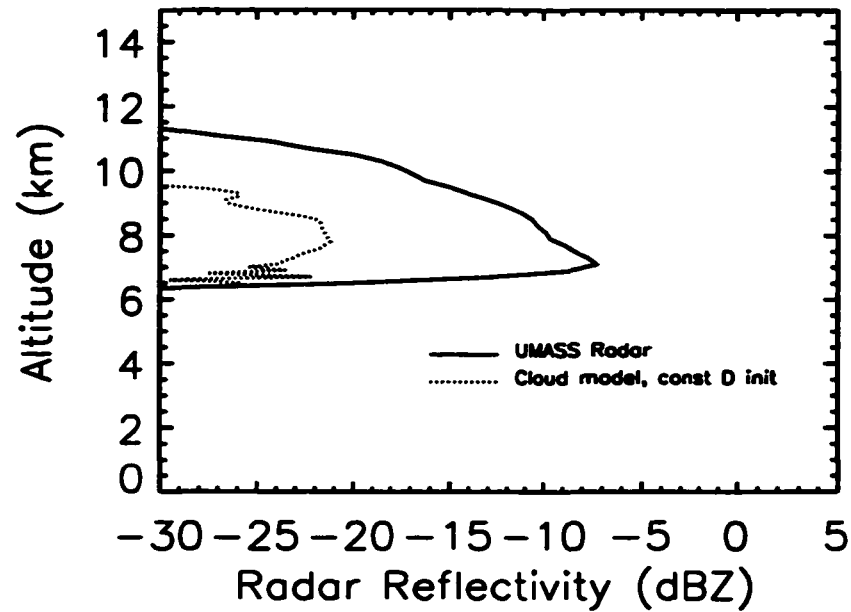


Figure 4.16: Comparison between observed (solid line) and model reflectivities. Meyers et al. (1992) initialization for N_t and constant initialization for D_n ($100\mu m$).

increasing the relative humidity at cloud levels (6–12 km) as shown in figure 4.17, the agreement is improved for all different characteristic diameter initializations (see figures 4.18, 4.19, 4.20, and 4.21).

Comparisons with other models is presented in figure 4.22.

The reflectivity profile for the model presented in this section is the one from the simulation with increased relative humidity and constant initial characteristic diameter equal $100\mu m$. Radar reflectivities for ECMWF and RAMS runs, are computed from the model derived IWC using an IWC–Z relationship, as explained in section 3.4.4. Simulations from the two different version of RAMS are included. Version 4.2 produces an average radar reflectivity profile which is in better agreement with observations. For more details on the RAMS simulations, and explanations of the differences between the two model versions, we refer the reader to chapter 3. In general all models do a reasonable job in reproducing the observed cloud. The major discrepancies, examined model by model, are as follows:

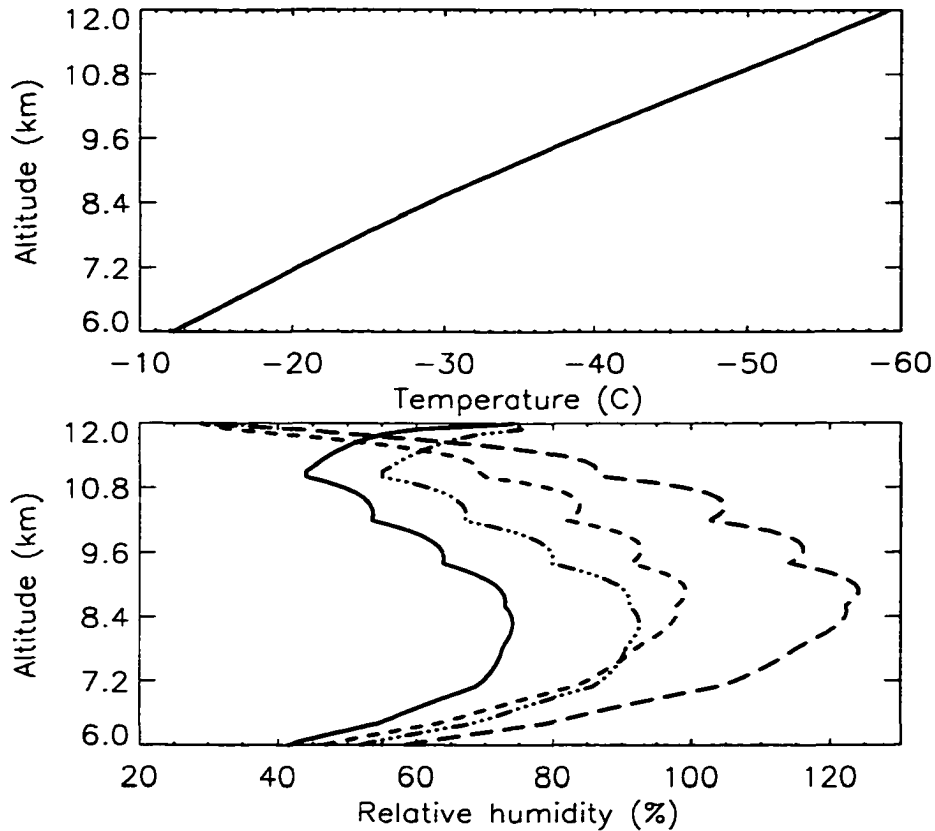


Figure 4.17: Atmospheric state variable profiles between 6 and 12 km. Temperature is shown in the upper panel. Relative humidity (RH) is shown in the lower panel. The various profiles represent respectively: original RH wrt water (solid line), original RH wrt ice (short dashed), increased RH wrt water (dot-dashed), and increased RH wrt ice (long dashed).

RAMS 3b reflectivity mostly overestimates observed value, except right above cloud base.

RAMS 4.2 places the cloud in the right position, but has a bright reflectivity bias throughout the cloud, and a slight dark bias right above cloud base, similar to RAMS 3b.

ECMWF has the right reflectivity near cloud base, but no cloud above 10.5 km,

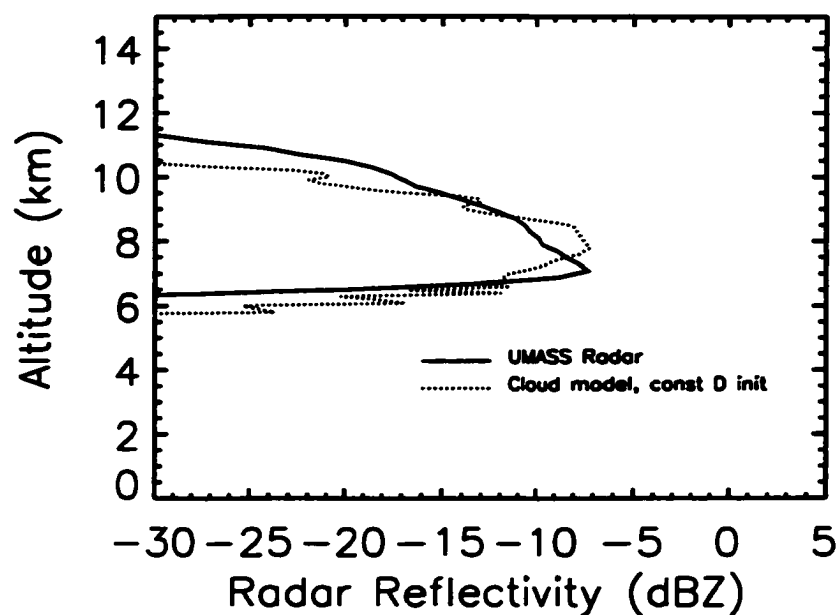


Figure 4.18: Same as figure 4.16, but for the increased relative humidity profile of figure 4.17.

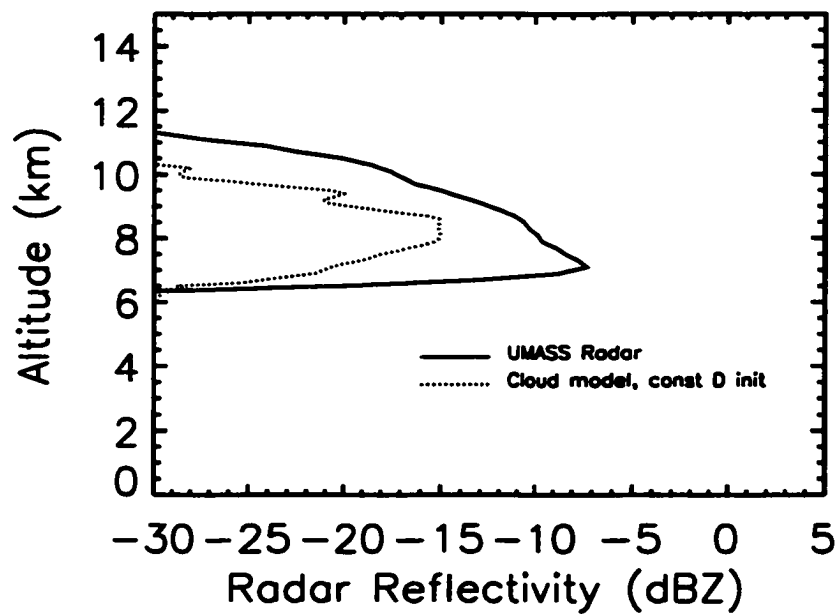


Figure 4.19: Same as figure 4.18, but for $D_n = 50\mu m$.

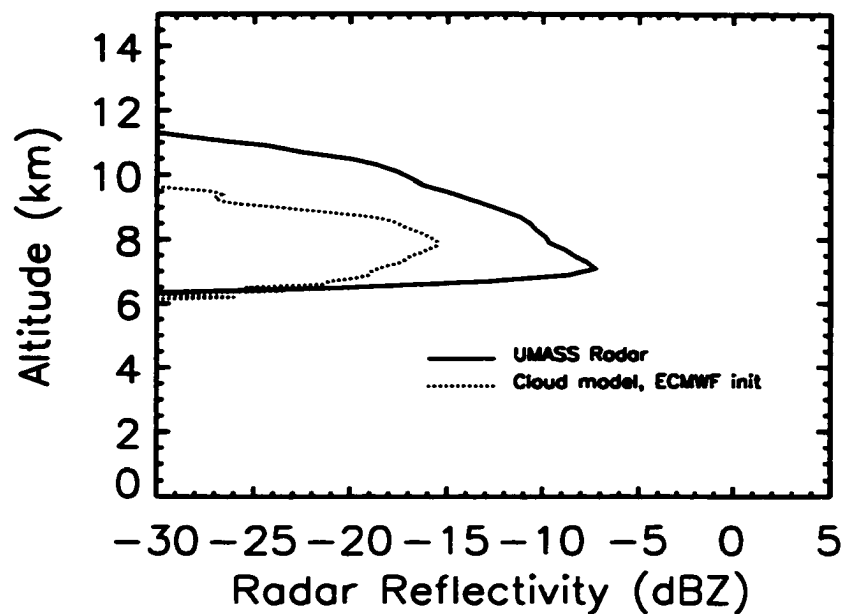


Figure 4.20: Same as figure 4.18, but for ECMWF D_n initialization.

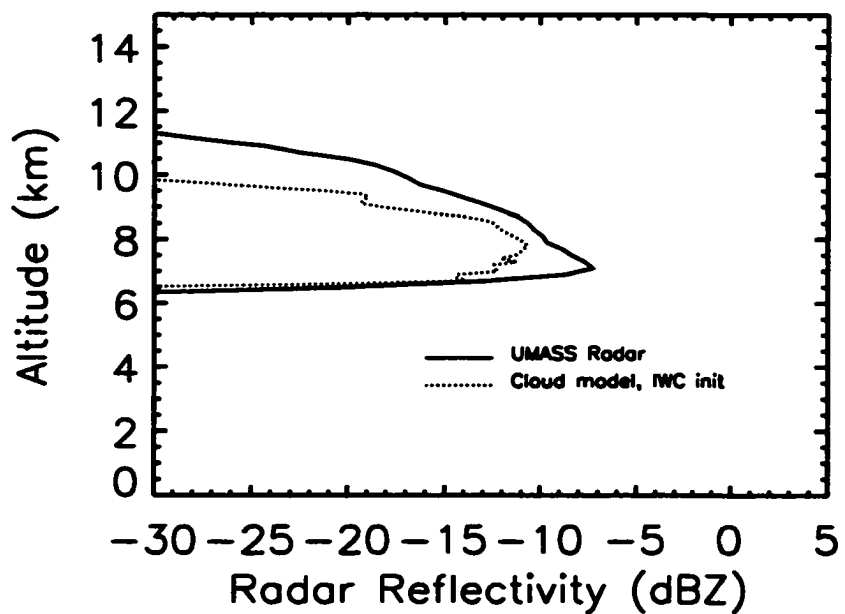


Figure 4.21: Same as figure 4.18, but for IWC D_n initialization.

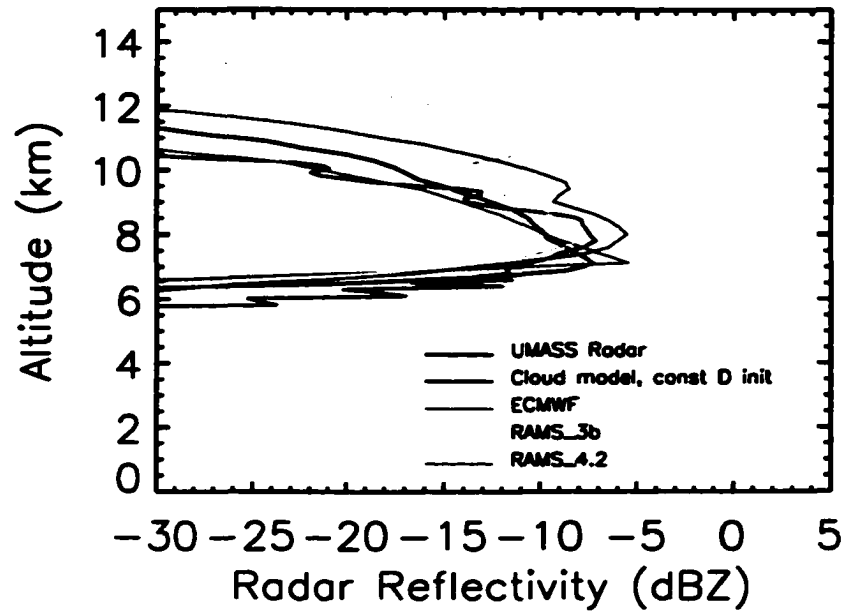


Figure 4.22: Comparison of radar reflectivity fields for a cirrus case observed in Kauai, Hawaii, on April 30, during the Spring 1999 ARM-UAV campaign.

whereas radar observations show a top at 12 km.

CIGMA shows a similar behavior to the ECMWF profile, except for an overestimation of reflectivity at cloud base.

Conclusions on model evaluation

The comparisons presented above show that the model, when initialized with relative humidity profiles close to water saturation (hence supersaturated with respect to ice at cloud levels), is capable of reproducing realistic ice clouds, if not perfectly in the magnitude of the fields, at least in the correct shape of the profiles. Some issues remain concerning the applicability of this simple model under variable environment conditions. A major deficiency hinted at previously is the lack of a prognostic equation for supersaturation that would take into account local changes in supersaturation induced by cloud processes (evaporation as a source and depositional growth as a sink of moisture). The lack of radiative/convective overturning,

crucial in cirrus dynamics as discussed in Wu et al. (2000) is also a major limitation of the model. Highly desirable would also be the inclusion of other microphysical processes such a parameterization for homogeneous nucleation processes.

However, even using a model with obvious limitations, it is possible to learn much about how radar data assimilation might impact our understanding of the model and of processes important to cirrus clouds in general. The remainder of this chapter focuses on showing that radar reflectivity measurements contain information regarding both cloud parameters and state parameters of the environment in which the cloud formed. This potentiality is fully explored in the next sections, starting from the study of model sensitivity to external inputs and model internal parameters.

4.4 Sensitivity studies

In this section we discuss the adjoint model and present results for the sensitivity studies. Based on the description of the adjoint of a linear model provided in chapter 2, it can be seen that the adjoint solution provides the sensitivity of the final model state (or a function of the model state) to an arbitrary change in initial condition or input parameters. The sensitivity is linear since it is based on a tangent linear approximation of the nonlinear model. For small perturbations, this is generally a good approximation. For large perturbations, or for highly nonlinear models, the linear sensitivity analysis might not be appropriate, as discussed in Saltelli (1999), since it does not consider higher order terms that account for interactions between different inputs. Given this limitation, the linear sensitivity analysis remains a powerful tool to learn about the model and assess its skills.

A study of sensitivity is necessary in order to establish whether the output variable, radar reflectivity in this application, is sensitive to the chosen control variables. Two opposite situations can be found: the output variable is not sensitive enough to the input or the output is too sensitive. In the former case, even large perturbations do not affect the final result, and it will not be possible to recover the optimal value of the control variable due to the lack of information regarding the control in the observable quantity. The second case is more delicate; on the one

hand it is desirable to have sensitivity to initial condition, but if the model is highly nonlinear this might turn out to be detrimental to convergence. Small perturbations can change the output by large amounts and the optimization algorithm will not be able to find the real minimum of the cost function. The use of linear (adjoint) analysis identifies *a priori* variables to which the model is most sensitive, providing those clues about the model performance.

The use of the adjoint model in sensitivity studies has a long-established tradition. Since its early application to investigate the sensitivity of a nonlinear radiative-convective climate model (Hall et al., 1982), the adjoint analysis has been applied in the study of cyclogenesis (Vukićević and Raeder (1995); Vukićević (1998)), atmospheric blocking (Zou et al., 1993a), tracer transport (Vukićević and Hess, 2000), and CO₂ seasonal cycle (Kaminski et al., 1996). As an advantage over the so called *forward model sensitivity* in which the model is perturbed for every grid point and each time in space, the *adjoint sensitivity* provides a computation of the model sensitivity to a given variable (or set of variables, including initial or boundary conditions, physical model parameters, etc.) in only one integration, since the adjoint is defined for all model points and times. The adjoint technique requires a solution of a differential equation, and the accuracy of the method is subject only to the accuracy of the linear assumption invoked.

4.4.1 Construction of the adjoint model

Adjoint construction can be summarized by the following practical steps (Errico and Vukićević, 1992):

1. Develop a Tangent Linear (TL) version of the nonlinear model under consideration (if model is already linear, this step is obviously not necessary);
2. in the TL model find every equation (diagnostic or prognostic) where a particular dependent variable occurs on the right-hand side. For discrete models, variables that have distinct gridpoint indices are considered distinct.
3. For each variable found in step 2, the equation that determines its corre-

sponding adjoint variable has the same terms as the TL equation on the RHS, but here these are coefficients that multiply whatever dependent variable was originally on the left-hand side.

4. If particular TL equations must be solved in a certain order, then for the adjoint they must be solved in the *reverse* order. For example, in CIGMA the computation of radar reflectivity is performed after the prediction of diameter and number concentration; in the adjoint it is exactly the opposite, i.e. the adjoint of Z is computed before that of D_n and N_t .
5. Diagnostic variables continue to be determined by diagnostic equations. Diagnostic variables are computed *backward* in time.

Automatic adjoint construction is based on these few steps. More details on the actual implementation and efficient coding of these operations can be found in Giering and Kaminski (1998). There are several automatic differentiation packages available to the scientific community, such as the Tangent linear and Adjoint Compiler (TAMC) software (Giering, 1999), ODYSSEÉ (Faure and Papegay, 1997), ADIFOR (Bischof et al., 1995). For our application, TAMC was used to derive the adjoint of the discretized version of the model described by equations (4.32) and (4.33). The software is easily accessible, and provides a rapid assembling of the adjoint, provided the input and output variables are all specified in the call to the main model subroutine. The resulting adjoint code is not optimized for speed, and some modifications are necessary in order to increase its computational efficiency. In order to do that, the tangent linear version of the model was also derived using the same automatic differentiation code. The tangent linear model derived with TAMC is usually accurate, and can be used to identify the coefficients (first derivatives) that need to be saved from a run of the nonlinear model and then used in the adjoint calculation. If all coefficients are saved in the forward run and read in direct access mode, the adjoint computation is greatly sped up. The tangent linear code also offers a way of verifying the correct functioning of the adjoint. By comparing the adjoint and tangent linear solution for a given perturbation, the validity

of the adjoint is established if a perfect (within numerical uncertainties) equality is recovered.

Initially the adjoint was built using cloud variables, such as initial characteristic diameter, initial number concentration and the model parameters α , β , a and b introduced in section 4.3.2 as control variables. A second version of the adjoint was then constructed with respect to environmental variables (vertical velocity, temperature, pressure and specific humidity). When changing control variables, the adjoint of the model has to be rederived since it is strictly dependent on the choice of the variable with respect to which the variation is taken. Once the adjoint is derived, the linear sensitivity of the model to a given input variable can be efficiently computed.

4.4.2 Sensitivity to model cloud-related variables

In this first part of the sensitivity study, we focus strictly on the cloud-related variables. We consider a 5% change in the input parameters: percentage variation makes it easier to compare perturbations in variables with different units. For vector variables, the percentage perturbation is taken over the whole profile.

Plots of radar reflectivity absolute variation (in dBZ) are shown in figure 4.23 and an interpretation of the results follows. The sensitivity of Z to D_0 and N_0 is straightforward to interpret: an increase in either field induces an increase in radar reflectivity. As anticipated, the radar reflectivity is more sensitive to the initial characteristic diameter profile, than to the initial number concentration profile. More exactly the radar reflectivity is almost six times more sensitive to initial characteristic diameter than to initial number concentration. It hardly seems worthwhile going to the trouble of deriving an adjoint to reach this obvious conclusion. However, the sensitivity of model radar reflectivity to other variables is far less intuitive. As an example, the central panels of figure 4.23 show the sensitivity to fall speed-diameter relationship parameters. An increase in a of 5% leads to an increase of 15 dBZ in Z at cloud base. This means that if the fall speed velocity of ice crystals is increased, the bigger crystal will concentrate preferentially at cloud base before hitting the subsaturated region, i.e. before starting to evaporate, hence enhancing the radar re-

flectivity at cloud base. Recalling the comparisons presented in section 4.3.5, it was noticed that the model put the cloud base too low (see figure 4.6), and a reduction in the fall speed parameter a , gave a better agreement with observations. This is explained clearly by the sensitivity plot. Conversely, if the coefficient b is increased, ice crystals fall at a lower speed (remembering that $b < 1$), resulting in a deficit of radar reflectivity at lower levels. It is worth noticing, however, that sensitivities to parameters a and b are large throughout the cloud depth, indicating how important a correct representation of the particle fall speed is for cirrus modeling. Results shown in the lower two panels indicate a large sensitivity to the mass–diameter relationship parameters below cloud base and above cloud top, and little sensitivity inside the cloud.

The sensitivity analysis was also performed on the cloud–layer optical depth. Results are presented in table 4.1. As expected, the sensitivity of τ to D_0 is two–fold the sensitivity of τ to N_0 , reflecting the quadratic dependence of τ on characteristic diameter, and linear dependence on number concentration. Varying the fall speed coefficients a and b had a small (almost negligible) negative impact on the total variation of optical depth. The mass–diameter relationship parameters have opposite impact with respect to each other. An increase in the parameter α , which is related to the density of the nonspherical ice crystals, decreases the optical depth,⁵ while an increase in the parameter β increases the cloud optical depth.

4.4.3 Sensitivity to environmental fields

Further sensitivity studies were conducted with a different version of the adjoint of the cloud model derived using environment variables as control. A 0.5% change in the input fields, temperature pressure, specific humidity and vertical velocity, was considered. The response of radar reflectivity to changing these variables is large in

⁵Visible optical depth can be approximated by the following relation from Stephens (1978)

$$\tau \approx \frac{3IWP}{2\rho_i r_e} \quad (4.43)$$

where ρ_i is the density of ice and r_e is the effective radius.

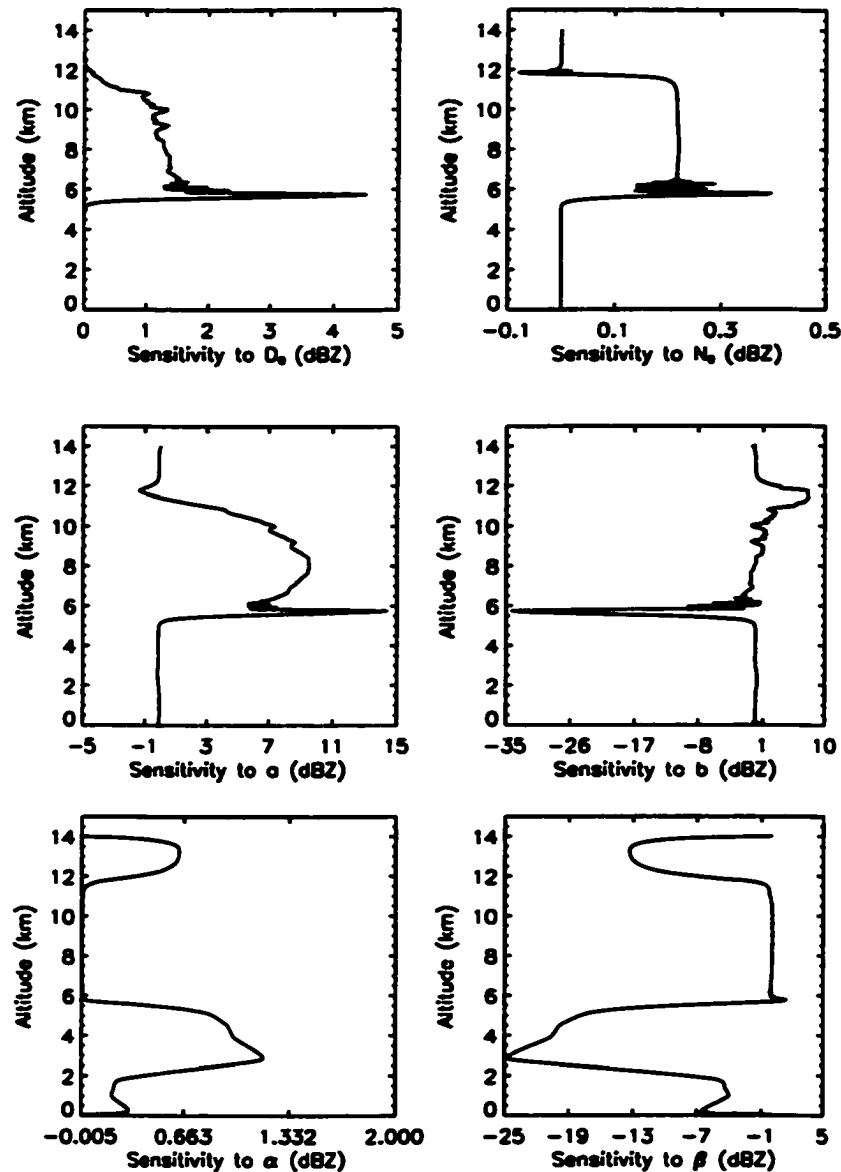


Figure 4.23: Sensitivity of radar reflectivity to 5% change in input parameters. In order, starting from the upper left corner: sensitivity of Z to initial characteristic diameter profile (left top panel); sensitivity to initial characteristic number concentration profile (top right panel); sensitivity to fall speed parameter (middle left panel), a ; sensitivity to fall speed parameter, b (middle right panel); sensitivity to mass parameter, α (bottom left panel); sensitivity to mass parameter, β (bottom right panel). See text for explanations.

Table 4.1: Sensitivity of cloud visible optical depth to a 5% change in input variables. Value of optical depth for unperturbed conditions is $\tau = 0.2338$.

Parameter	$\Delta\tau$	Change in τ (%)
Initial diameter profile, D_0	0.0241660271	10
Initial number concentration profile, N_0	0.0126533825	5.5
Fall speed parameter, a	-6.25227694E-05	-0.01
Fall speed parameter, b	-0.00216991128	-3.5
Mass-Diameter relationship parameter, α	-0.00123489264	-0.5
Mass-Diameter relationship parameter, β	0.0012982504	0.6

amplitude, except the one for vertical velocity (see figure 4.24).

This particularly high sensitivity to temperature is related to the Meyers et al. (1992) parameterization used to initialize ice crystal number introduced in section 4.3.2, which we transcribe here for convenience:

$$N_{IN} \approx e^{-0.639+0.1296S_i}, \quad (4.44)$$

Since this is an exponential function of S_i , which, in turn, is a nonlinear function of temperature, small changes in temperature can lead to a large difference in the initial condition for the number concentration. This, in turn, dominates the time evolution of the particle number concentration in the cloud. The resulting extreme sensitivity to the temperature profile appears to be an intrinsic characteristic of the physical system, amplified by our choice of parameterization to describe the system. Results indicate that an increase (decrease) in temperature leads to a large decrease (increase) in radar reflectivity. The explanation follows: when T increases, the saturation vapor pressure with respect to ice decreases according to the following relationship directly derived from the Clausius-Clayperon equation:

$$e_{si} \approx 6.11 \exp \left[\frac{L}{R_v} \left(-\frac{1}{273} + \frac{1}{T} \right) \right] \quad (4.45)$$

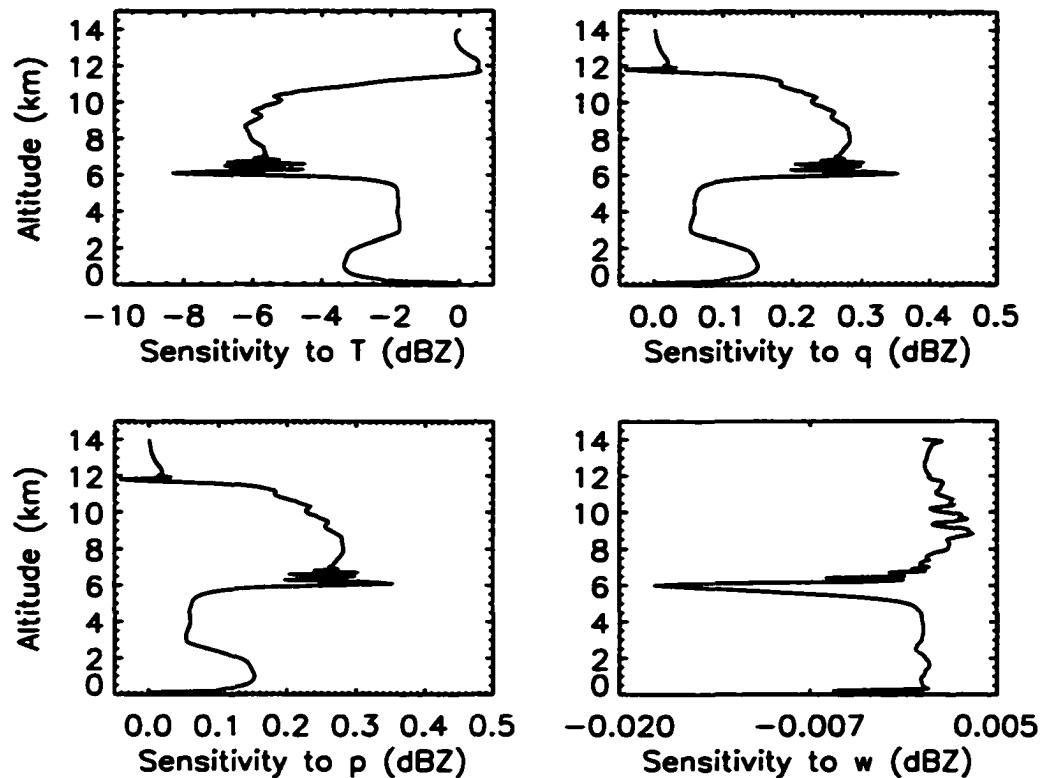


Figure 4.24: Sensitivity of radar reflectivity to 0.5% change in atmospheric variables. In order, starting from the upper left corner: sensitivity of Z to initial temperature profile (left upper panel); sensitivity to initial specific humidity profile (right upper panel); sensitivity to pressure (left lower panel); sensitivity to vertical velocity (right lower panel).

where $T < 273K$. Consequently the supersaturation with respect to ice diminishes and the number of ice nuclei that heterogeneously nucleate decreases. Since the radar reflectivity is a linear function of the number concentration, it can be concluded that an increase in temperature will decrease the radar signal. As mentioned, this seems to be a property of the physical system under study, but might not necessarily be a characteristic of all ice cloud systems⁶. However, a similar behavior may also

⁶For example, in mixed phase clouds, the sharp increase in radar reflectivity around the melting layer, also known as *bright band* (Stephens, 1994), reflects a different behavior of the physical system

be encountered when studying the sensitivity to temperature of more sophisticated models that make use of parameterizations similar⁷ to equation (4.44).

With regard to the model reflectivity sensitivity to a change in the specific humidity, it is observed that an increase (decrease) in specific humidity increases (decreases) the radar signal. This is also related to the change in supersaturation as discussed in section 4.3.5. Sensitivity to pressure has a similar magnitude and shape as sensitivity to specific humidity. Sensitivity to vertical motion is especially pronounced at cloud base, although much smaller than the sensitivity to other state variables. When the vertical velocity is increased, the variation of Z is particularly emphasized and negative at cloud base, indicating that a stronger updraft (w is positive upwards) results in a smaller number of large particles at cloud base, and hence in a reflectivity deficit. This result might have been different had the model included a prognostic equation for supersaturation. An increase in vertical velocity creates supersaturation in the lifted parcel by adiabatic cooling, and hence would have a larger impact on the final reflectivity profile.

The sensitivity of the optical depth to changes in the state variables is summarized in table 4.2. According to this analysis, a colder cloud is optically thinner than a warmer cloud, i.e. an increase in temperature leads to a dramatic decrease in cloud optical depth. This is once again related to the decrease in supersaturation, and in the consequent reduction in initial number concentration. An increase in either pressure or specific humidity, hence in supersaturation, induces an increase in optical depth. A vertical velocity increase has a small negative impact on optical depth.

4.4.4 Concluding remarks on the sensitivity study

It is important to note that the sensitivity analysis is based on a linearization of the nonlinear cloud model, and, as such, might not correctly represent nonlinear feedbacks that a change in one or more parameters could induce on the output fields with respect to temperature variations.

⁷A commonly used parameterization to estimate the number concentration of ice nuclei is also exponential in temperature and is found in Fletcher (1962).

Table 4.2: Sensitivity of cloud visible optical depth to a .5% change in atmospheric state variables. Value of optical depth for unperturbed conditions is $\tau = 0.01764$.

Parameter	$\Delta\tau$	Change in τ (%)
Temperature, T	-0.0263474584	-150
Specific humidity, q	0.00109758112	6
Pressure, p	0.00109898427	6
Vertical velocity, w	-1.41718516E-07	-0.0008

and on the other model parameters. Moreover, only a subset of the parameter space is sampled in the linear analysis, and its validity is limited to small perturbations around the baseline case. Nevertheless this analysis allows understanding of model response to perturbations in input parameters, and detection of nonphysical model behaviors (for example, the disproportionate sensitivity to temperature or lack of sensitivity to vertical velocity).

The knowledge gained from these sensitivity studies is applied in the next sections that explore more directly the impact of assimilating radar data into the ice growth model described in the first part of the chapter.

4.5 Toward variational assimilation of radar data

The variational experiments described here fall under the category of 1+1D-Var assimilation (one spatial and one temporal coordinate). All the concepts introduced in chapter 2 are applied and illustrated using concrete examples. These include the use of synthetic data to investigate the possibility of recovering the “true” initial conditions (controlled experiments), and the use of true observations to study the potential of radar information for improving cloud prediction.

Experiments designed to evaluate the CIGMA assimilation system developed in this study comprise:

- use of different control variables;

- sensitivity to the magnitudes of mapping and observational errors;
- inclusion/exclusion of background term;
- inclusion/exclusion of additional information (specifically optical depth).

In the next section, these aspects are briefly discussed from a theoretical point of view to provide further background material to the reader. A detailed overview of the optimization procedure is also presented.

4.5.1 On the choice of control variables

Based on the sensitivity study presented in section 4.4 we used two different sets of control variables: initial conditions on cloud variables (initial characteristic diameter, D_0 , and initial number concentration, N_0) and initial conditions on environmental variables. For the latter, we experimented with temperature and specific humidity, two variables commonly optimized in operational assimilation systems. Due to the high sensitivity of the model to temperature, it was difficult to use temperature as a control variable. Even a small perturbation in temperature profile induces a large change in radar reflectivity for the reasons discussed in section 4.4. Under these circumstances, it was verified that the optimization procedure was unstable, and the algorithm did not converge to the minimum. Thus, it was decided to use only specific humidity as a control variable. The sensitivity to the initial profile of this variable is still substantial, but not as dramatic as for temperature. Another reason to use specific humidity is that it is easier to justify a large correction to in-cloud specific humidity on a physical basis than it is to justify a large correction to in-cloud temperature. For example, specific humidity can be advected and undergo a substantial increase over a short time scale. Some experiments of parameter optimization using the variational technique for the parameter a of the fall-speed velocity relationship were also conducted. Results are presented in section 4.6.3.

4.5.2 Discussion of radar observational and mapping error

In chapter 2 we discussed the importance of error statistics in the data assimilation system. The matrix \mathbf{W}^{-1} that appears in the cost function (2.4) is the inverse of the error covariance matrix. There are two contributions to this error: the contribution that comes from the measurements and the contribution that comes from the mapping. In general the first contribution is the smallest, since measurements of radar reflectivities are rather accurate (within 0.5–1.0 dBZ which corresponds to approximately 20% accuracy⁸); the error in the mapping is harder to define. Assuming that the forward dynamical model has no error, then the mapping error is only due to errors in the parameterizations used, for example in the value $\nu = 2$, chosen for the width of the gamma distribution, or the approximation made in using the Rayleigh radar reflectivity instead of performing an exact calculation. If the mapping is analytical, an attempt to estimate this type of error can be made from the derivative of the mapping with respect to specific parameters. The radar mapping of section 4.3.4 depends on the parameter ν , and it follows that

$$\frac{\partial Z}{\partial \nu} = N_t D_n^6 \frac{\Gamma(\nu + 6)}{\Gamma(\nu)} [F(\nu + 6) - F(\nu)] \quad (4.46)$$

where $F(x) = \frac{\Gamma'(x)}{\Gamma(x)}$, proportional to the first derivative of the gamma function, is the *digamma* function (Abramovitz and Stegun, 1974). If the mapping is nondifferentiable with respect to the parameter, first derivatives can be obtained by perturbing the parameter value and taking the finite difference. Similarly, if the adjoint of the mapping with respect to the parameter is available, first derivatives can be obtained by an adjoint computation. The absolute error is then written as:

$$\Delta Z = \frac{\partial Z}{\partial \nu} \Delta \nu. \quad (4.47)$$

⁸Radar reflectivity is usually measured in dBZ which is a logarithmic scale. The mathematical definition is $Z_{dBZ} = 10 \text{Log}_{10}(Z)$, where the units of Z are $mm^6 m^{-3}$. When propagating the error on Z to Z_{dBZ} we have to differentiate the above expression with respect to Z to get: $\Delta Z_{dBZ} = \frac{\partial Z_{dBZ}}{\partial Z} \Delta Z = \frac{10 \text{Log}_{10}(e)}{Z} \Delta Z$. If $\Delta Z = 0.2Z$ (20% error on reflectivity), then $\Delta Z_{dBZ} \approx 0.86 \text{dBZ}$. The accuracy quoted in the text is relative to the MMCR radar as reported in the ARM web site.

In what follows, this mapping error was not included, nor was the dynamical model error. Instead the sensitivity to the error assumptions was investigated by arbitrarily increasing the magnitude of the elements of \mathbf{W} .

4.5.3 Adding information

It is possible to add other types of observations, if available, to the measurement vector \vec{y}_{obs} . This can potentially improve the constrained optimization procedure, provided that the new information is sensitive to the control variables. However, when new information is added, convergence might be slower since the requirements the optimum variables have to satisfy are stricter, i.e. the search in the parameter space is limited by the introduction of the additional constraint. To test this, the visible total cloud optical depth, τ , was used as a further piece of information in the synthetic data assimilation with cloud variables as control. The mapping for optical depth from model number concentration and characteristic size is described in section 4.3.4. This measurement is expected to be available along with radar reflectivity in future aircraft and satellite experiments.

4.5.4 Including the background

In our application using synthetic measurements and specific humidity as a control, a fictional background was defined as a perturbation of the true (known) profile with a preassigned uncertainty associated with it. The influence of the inclusion of the background and of the magnitude of the background error on the performance of the assimilation system was then investigated. It was found that if no information is contained in the measurements regarding the control variable, then the optimization converges back to the background, and this convergence depends on the relative magnitude of specified observational and background errors, as discussed in section 2.3.4. In the assimilation of real measurements, no background was used, and the inversion was done purely using observations.

4.5.5 Optimization procedure

The optimization procedure is started with an initial guess for the control variables. The nonlinear cloud model is run, and the model equivalent radar reflectivity is obtained. The cost function is evaluated using equation (2.4) which we rewrite here in discrete form:

$$J = \frac{1}{2}(Z - Z_{obs})^T \mathbf{W}^{-1}(Z - Z_{obs}). \quad (4.48)$$

The adjoint forcing is given by the derivative of J with respect to the model output variables, and has to be computed at each evaluation of the cost function. Explicitly:

$$\frac{\partial J}{\partial \boldsymbol{\kappa}} = \left(\frac{\partial Z}{\partial \boldsymbol{\kappa}} \right)^T \frac{\partial J}{\partial Z} \quad (4.49)$$

where $\boldsymbol{\kappa} = (x, y)$ is the model output vector, and x and y represent the two non-dimensional variables introduced in section 4.3.2. $\frac{\partial Z}{\partial \boldsymbol{\kappa}}$ represents the Jacobian of the radar reflectivity with respect to the model variables. We rewrite $\frac{\partial J}{\partial \boldsymbol{\kappa}}$ as two vectors:

$$\frac{\partial J}{\partial x} = \left(\frac{\partial Z}{\partial x} \right)^T \frac{\partial J}{\partial Z} \quad (4.50)$$

$$\frac{\partial J}{\partial y} = \left(\frac{\partial Z}{\partial y} \right)^T \frac{\partial J}{\partial Z}. \quad (4.51)$$

Recalling that

$$\begin{aligned} x &= \ln\left(\frac{D_n}{D_{max}}\right) \\ y &= \ln\left(\frac{N_t}{N_{max}}\right), \end{aligned}$$

Z can be rewritten in terms of x and y :

$$Z = \frac{\Gamma(6 + \nu)}{\Gamma(\nu)} N_{max} e^y D_{max}^6 e^{6x} \quad (4.52)$$

hence

$$\frac{\partial Z}{\partial x} = 6Z \quad (4.53)$$

$$\frac{\partial Z}{\partial y} = Z \quad (4.54)$$

and the adjoint forcing is simply:

$$\frac{\partial J}{\partial x} = 6Z\mathbf{W}^{-1}(Z - Z_{obs}) \quad (4.55)$$

$$\frac{\partial J}{\partial y} = Z\mathbf{W}^{-1}(Z - Z_{obs}) \quad (4.56)$$

which is dimensionally consistent with units of \mathbf{W} which are mm^6m^{-3} . The “forced” adjoint equation is integrated backward in time in order to compute the gradient of the cost function wrt the control variables needed by the optimizer. The initial guess for the control variables, the value of the cost function and its gradient with respect to the control variables are passed to the optimization routine which computes the value for the new guess. Several line search routines are available. This application uses a quasi-Newtonian algorithm developed by Zou et al. (1993b). A new value of the initial condition is calculated, and the whole process repeated until convergence is reached. The convergence criterion is fixed internally in the optimizer via user-specified optimization parameters, i.e. minimum value of the norm of the gradient, minimum and maximum step-lengths. etc. Once the predefined convergence criterion is satisfied, the “optimal” initial condition is used to run the nonlinear model and produce the optimal model solution. In section 4.6 we will show some examples of optimized initial conditions using synthetic and real data.

A schematic of the optimization procedure is presented in figure 4.25.

4.6 Results of optimization using synthetic measurements

The optimization procedure described above was tested using synthetic measurements. These tests are useful for verifying the correct implementation of the procedure, as well as for understanding the nature of the optimization problem at hand, as a step toward the assimilation of real measurements.

From an arbitrary initial condition (“truth”), the radar forward model was used to calculate synthetic radar reflectivities. Similarly, synthetic optical depths were also computed using the forward model described in section 4.3.4. Unperturbed temperature, pressure and specific humidity profiles from the ECMWF forecast for the April, 30 1999 cirrus case were used. Biases and errors due to initial conditions

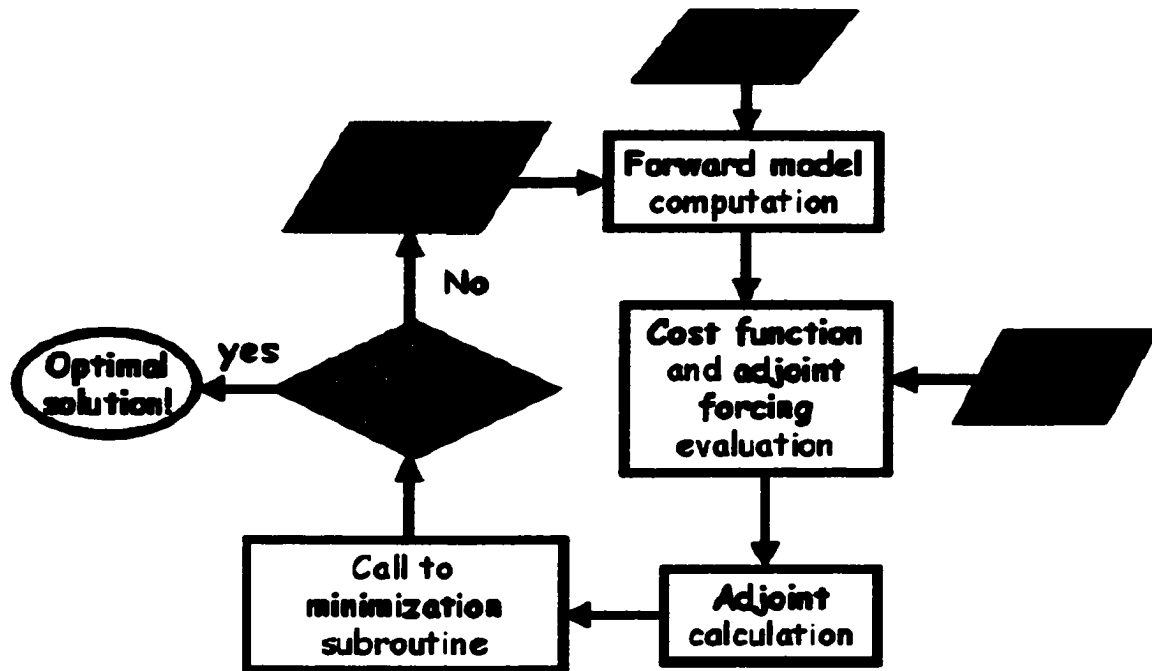


Figure 4.25: Flowchart of the optimization procedure, beginning with the forward model computation from a first guess initial condition.

and model assumptions are not a concern here because the same model is run to produce synthetic measurements and for the optimization process. The initial value of the characteristic diameter is derived using the ECMWF parameterization (type 2. initialization). The reflectivity threshold for cloud detection is chosen to be -30 dBZ. Cloud radar can detect equivalent cloud signals as low as -50 dBZ (Clothiaux et al., 1995). The reason for our choice is related to the spaceborne 94-GHz radar that will be on board of CloudSat whose sensitivity is estimated to be around -28/-30 dBZ. Once the synthetic measurement dataset is created, and synthetic measurements are made available at the given assimilation times, the model is run with a perturbed initial condition and the optimization is performed following the step procedure outlined in section 4.5.5.

4.6.1 Cloud variables as control variables

From preliminary tests, it was found that it is better to work with dimensionless control variables, since this ensures a better behavior of the optimization algorithm.

The optimization was then performed with respect to x_0 and y_0 , defined as:

$$\begin{aligned} x_0 &= \ln\left(\frac{D_0}{D_{max}}\right) \\ y_0 &= \ln\left(\frac{N_0}{N_{max}}\right) \end{aligned} \quad (4.57)$$

where D_0 and N_0 represent the initial profiles of characteristic diameter and number concentration. The following sections present results for the synthetic data assimilation experiments performed with these control variables.

Sensitivity to magnitude of perturbation

Initial profiles of characteristic diameter and number concentration were perturbed by a variable percentage, and the behavior of the optimization with respect to recovering the true initial profiles was monitored. The magnitudes of the elements of the error covariance matrix \mathbf{W} were fixed to approximately 25% of the “bogus” reflectivity value at each level (1dBZ). The assimilation system was run until the change in the cost function from one iteration to the next was no more than a prescribed small increment ($\epsilon = 10^{-6}$). This was achieved in approximately 15–20 iterations depending on the specific case. The criterion for testing the convergence is to compare the initial and the final values of the cost function, and to look at profiles of recovered control variables as functions of iteration. Intuitively, we expect the convergence to be better for small perturbations.

Table 4.3 offers an overview of the behavior of the cost function. The amplitude of perturbation and the initial and final values of the cost function are listed.

For negative perturbations, it is indeed true that the larger the perturbation the more the algorithm has problems with the convergence. For example, with a perturbation in both control variables of -30%, the final cost function decreases to 4% of its original value whereas for a perturbation of -10% the cost function decreases to 0.18% of its original value. For positive perturbations, however, the opposite

Table 4.3: Initial and final values of the cost function versus initial condition perturbation.

Perturbation (%)	J_0	J_{FIN}	$\frac{J_{FIN}}{J_0}$ (%)
-30	7204	294	4.
-20	3517	14	0.3
-10	1978	3.5	0.18
+10	860	1.8	0.2
+20	3224	6.7	0.2
+30	6783	13	0.19
+50	16618	27	0.16

behavior is observed (larger perturbations lead to larger percentage decreases in the cost function). It is also interesting to notice that the assimilation system recovers the true profiles better when starting from positive perturbed initial conditions, that is to say that it performs better in *reducing* the amplitude of the initial fields rather than in *increasing* it. This characteristic is illustrated in figure 4.26 and figure 4.27 in which it can be seen that the analysis profile matches more closely the true profile of reflectivity in the case of a positive perturbation than in the case of a negative perturbation. This can be summarized by the statement that the algorithm is more effective at dissipating the cloud than at creating it. This results is also confirmed in section 4.6.2, where we investigate the use of specific humidity as control variable. A limitation of the assimilation system related to the low sensitivity of radar reflectivity to initial number concentration also appears evident from the analysis of both figures 4.26 and 4.27: regardless of the sign of the perturbation, the initial condition for the characteristic diameter is better recovered than the initial condition for the number concentration, whose profile remains close to the initial guess. However, at upper levels, where there is no radar signal, the final profile of the characteristic size coincides with the initial guess.

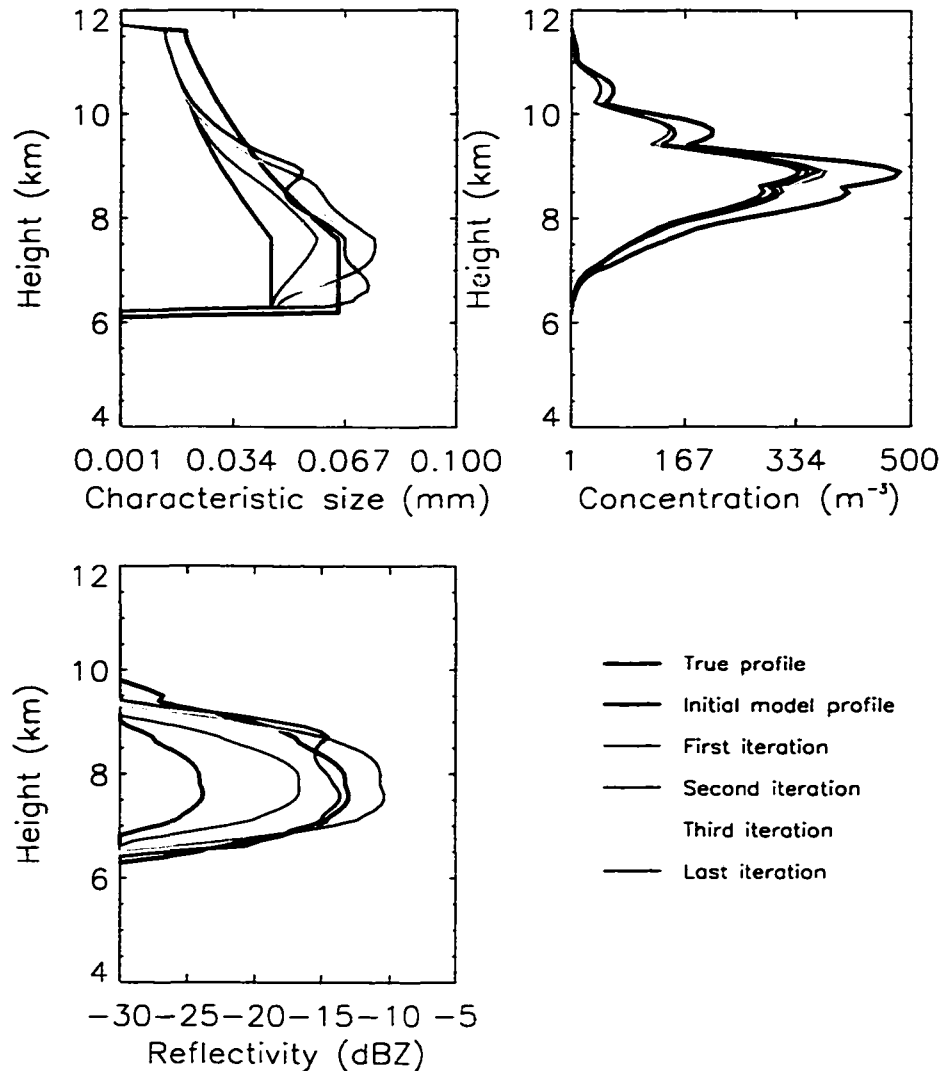


Figure 4.26: Characteristic diameter (top left panel), number concentration (top right panel) and synthetic reflectivity at final time (bottom left panel) as functions of iteration. The initial perturbation is -30%.

Sensitivity to vertical displacement perturbations

In this section, the focus is on ability of the assimilation system to recover true profiles from a perturbation that shifts the system in the vertical direction. Various perturbations were explored. Results are shown in figure 4.28 for one case where

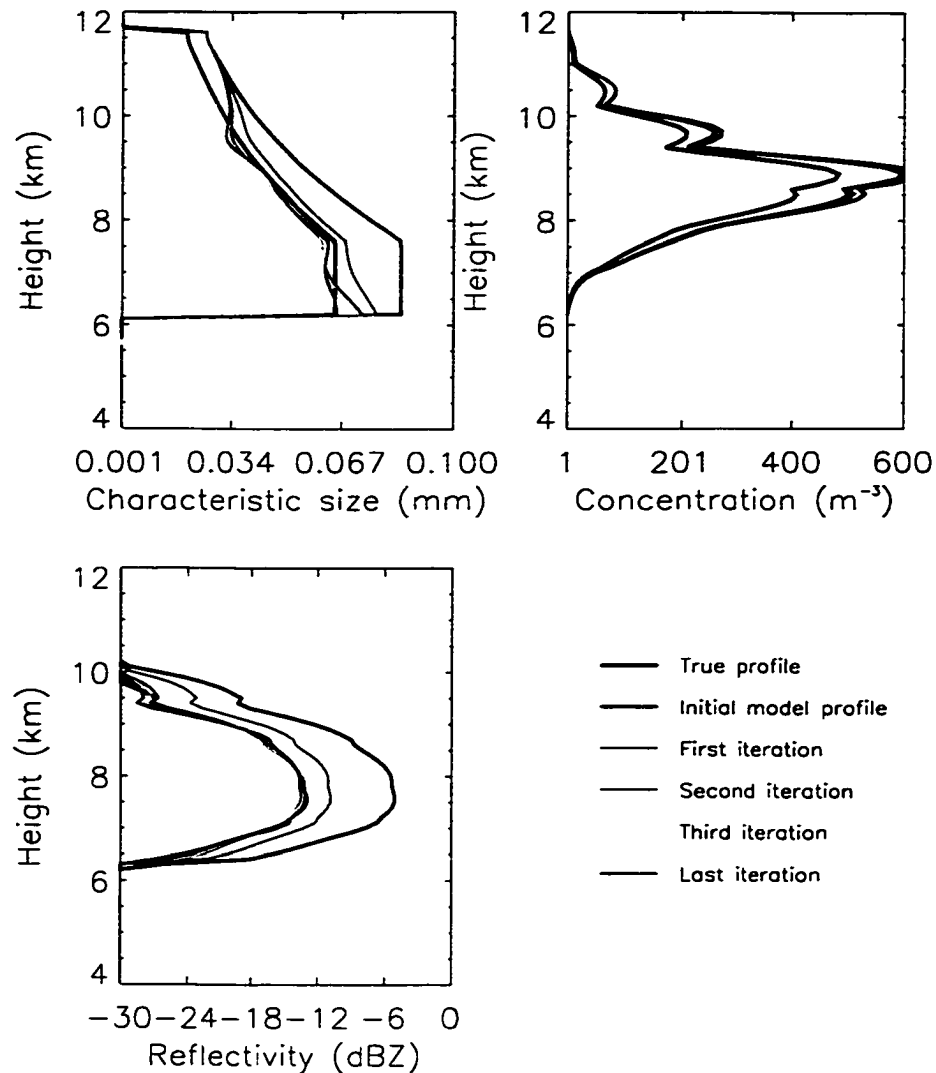


Figure 4.27: Same as figure 4.26. but for an amplitude perturbation of +30%.

the initial condition misplaced the cloud 1 km higher than the true position, and the magnitude of the characteristic diameter and number concentration profiles was also increased by 20%.

The model is not able to recover true initial conditions below cloud base where, according to the model, there is no radar signal. Once again this shows, that there is no possibility of creating a cloud where there is no cloud to begin with by adjusting only the initial cloud profiles. However, the assimilation is successful in *reducing* the

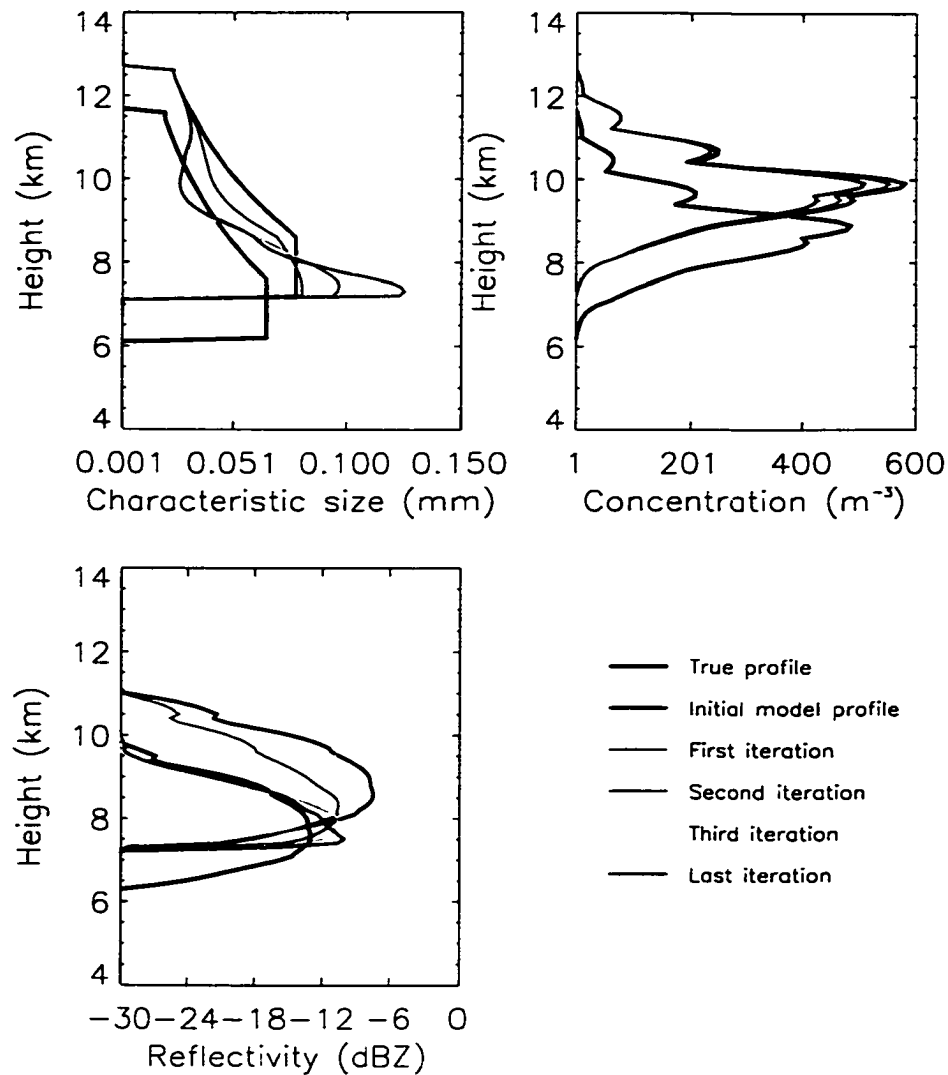


Figure 4.28: Characteristic diameter (top left panel), number concentration (top right panel) and synthetic reflectivity at final time (bottom left panel) as functions of iteration. The initial profile is perturbed by 20% and misplaced by 1 km with respect to true profile.

initial excessive reflectivity at upper levels, mainly by reducing the characteristic size.

Table 4.4: Initial and final values of the cost function versus magnitude of observational plus mapping errors.

Error (dBZ)	J_0	J_{FIN}	$\frac{J_{FIN}}{J_0}$ (%)
0.5 ($\approx 10\%$)	12900	27	0.21
1 ($\approx 23\%$)	3224	6.7	0.21
2 ($\approx 46\%$)	806	1.7	0.21
3 ($\approx 69\%$)	358	0.74	0.20
5 ($\approx 115\%$)	129	0.27	0.20

Sensitivity to error assumptions

The response of the CIGMA assimilation algorithm to varying magnitudes of errors is investigated in this section. The perturbation of the initial profiles is fixed to a given value (i.e. 20%), and the error covariance matrix for observations and mapping is varied. The results are summarized in table 4.4 using the criterion introduced in the previous section (comparison of initial and final values of the cost function).

The relative decrease of the cost function is approximately three orders of magnitude, regardless the magnitude of observational and mapping errors. This implies that, even when it is assumed that the measurements and the mapping are close to perfect, there is an intrinsic limitation on the system with respect to recovering the true profiles of cloud variables imposed by the statistical description of the state variables and respective errors.

Figures 4.29 and 4.30 show optimum profiles of size, concentration and reflectivity for 1dBZ and 5dBZ errors, respectively.

The magnitude of observation plus mapping errors does not seem to influence the convergence. This is a meaningful result in the sense that information can be extracted even when model mapping and/or measurements are not accurate. Since assimilation is performed with the same nonlinear model used to produce the synthetic data set, model biases are irrelevant. The data set used in the experiments

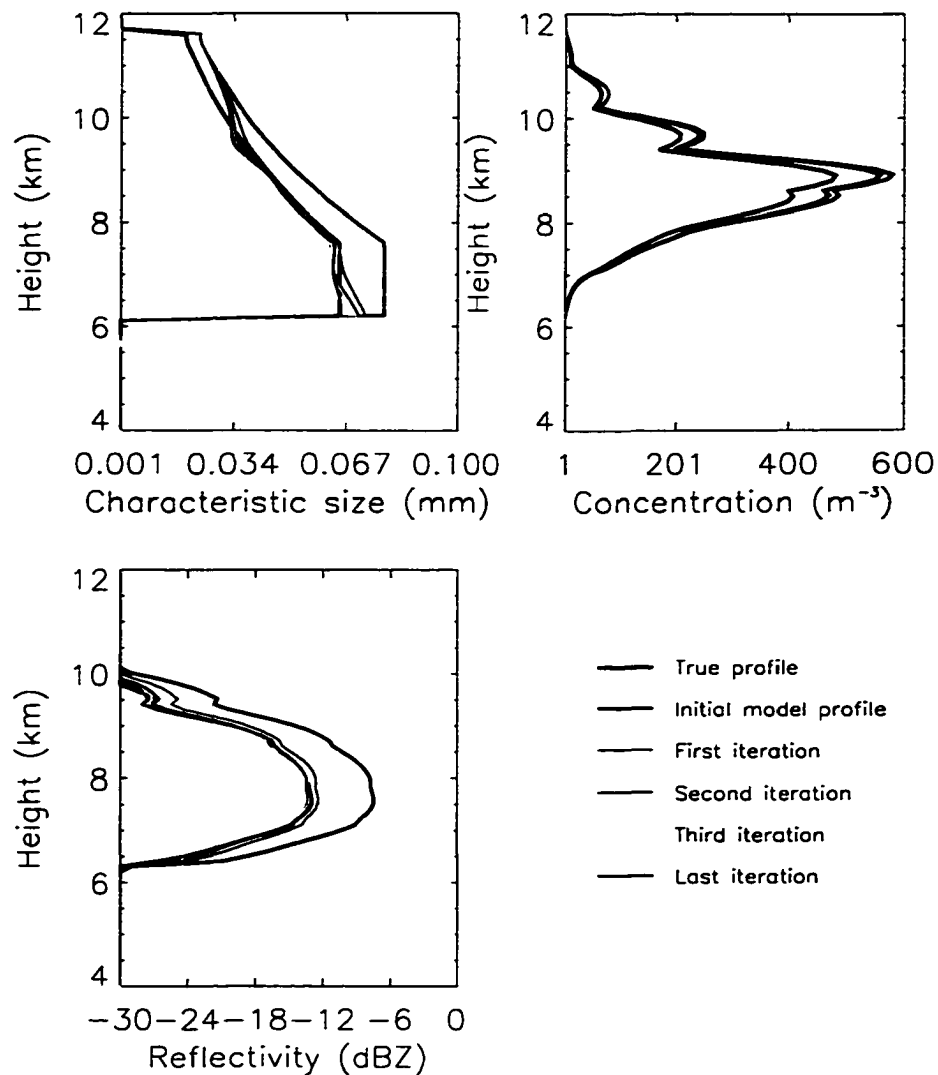


Figure 4.29: Characteristic diameter (top left panel), number concentration (top right panel) and synthetic reflectivity at final time (bottom left panel) as functions of iteration. The initial perturbation is 20%. Observation plus mapping error is 1 dBZ ($\approx 23\%$).

is thus ideal for the specific model. This is obviously not true when the model is confronted with actual observations, and this conclusion might need to be reconsidered.

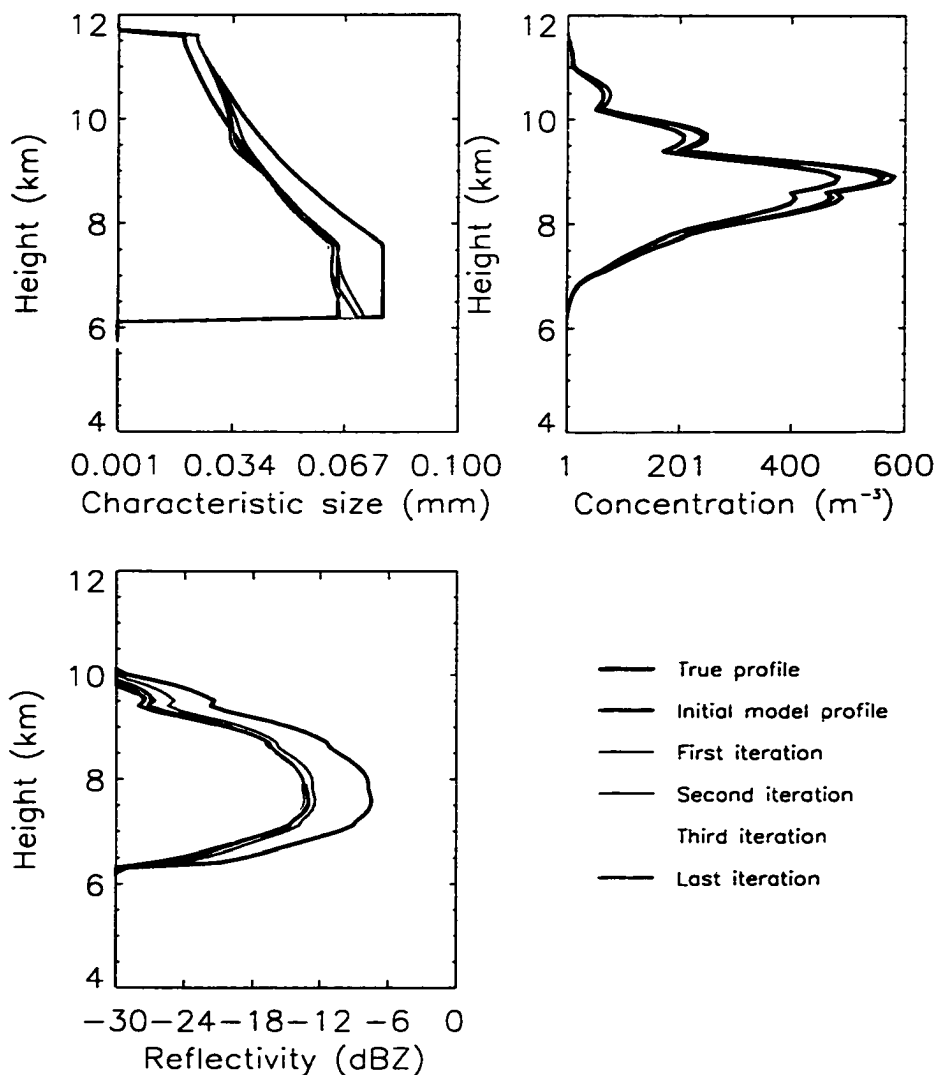


Figure 4.30: Same as figure 4.29, but for an error of 5dBZ.

Inclusion of optical depth

The effect of adding further information to the assimilation system is now examined. Optical depth information can be derived from a variety of active and passive instruments. Miller et al. (2000) introduced an estimation-based technique for the retrieval of cloud optical depth from satellite radiances (examples were shown in chapter 3). The optical depth shown in figure 4.1 was derived from Raman lidar with a technique based on Young (1995).

The additional constraint should ensure a more physical solution, and hence should improve the analysis, provided that the information added is meaningful, and the error in the new piece of information is not degrading the convergence. In particular, since the total optical depth is linearly sensitive to the number concentration, it is anticipated that adding this information will lead to a recovery of the true initial concentration. However, this improvement did not occur, and it was found that the inclusion of the optical depth information did not have a strong impact on the analysis. This is shown in figure 4.31. If these results are compared to the same case with no optical depth information, shown in figure 4.29, no substantial differences are evident.

An hypothesis to explain this behavior is that total optical depth alone does not provide enough additional independent information with respect to radar reflectivity to dramatically improve the algorithm. The sensitivity of optical depth to size is still dominant over its sensitivity to concentration. The inclusion of profile information such as cloud extinction as a function of height (as obtained by, i.e., lidar or other profiling instruments) could be of greater benefit to the assimilation. This can be easily tested with synthetic data, but is beyond the scope of this research, which focus on assimilation of radar measurements. Investigation of the assimilation of information other than radar reflectivity will be pursued in future research.

Discussion of results

From these synthetic assimilation experiments with cloud variables, two main conclusions can be drawn: the assimilation system performs relatively well in recovering the true initial condition for the characteristic diameter under various conditions for perturbation magnitude and mapping plus observation errors. However, the initial condition for the number concentration is not recovered as effectively, reflecting the fact that radar reflectivities contain more information about the size of the cloud particles than the number concentration. Inclusion of optical depth information does not lead to any improvement.

The algorithm is also found to be less effective in recovering true profiles when

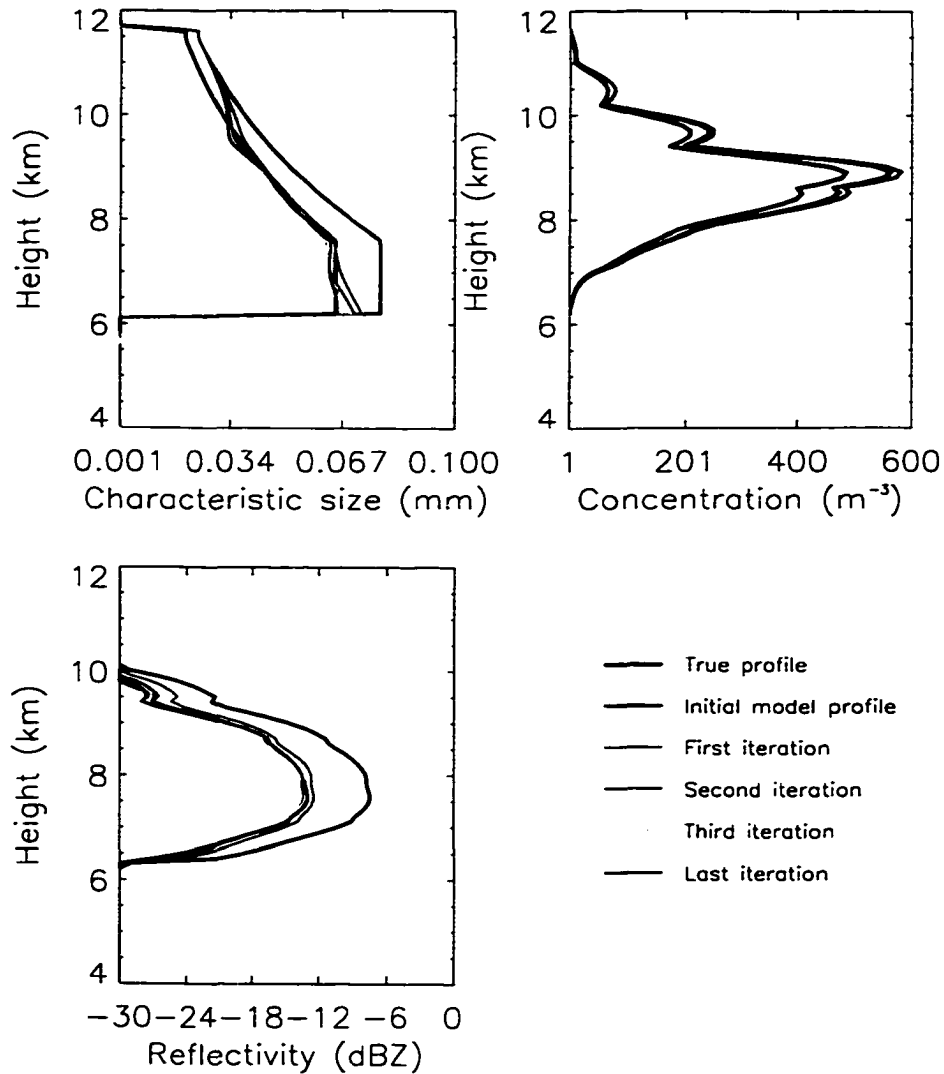


Figure 4.31: Same as figure 4.29, but with the inclusion of total optical depth information. The accuracy of the optical depth is same as the accuracy in radar reflectivity (23%).

the initial conditions are misplaced. This represents a limiting factor for the assimilation of real measurements, since it seems that the model is not able to adjust the cloud variables in a region where no cloud is present to begin with. Adjusting environmental variables might be the solution to this problem.

4.6.2 Specific humidity as a control variable

In this series of experiments, the dimensionless specific humidity, defined as

$$q^* = \log \left(\frac{q}{q_{max}} \right) \quad (4.58)$$

was used to ensure the positivity of the corrected value of q and better behavior of the optimization procedure.

Sensitivity to the magnitude of the perturbation

This section explores the capability of the assimilation system to “recover” the true profile from the perturbed profiles. Wet (specific humidity too high) and dry (specific humidity too low) biased profiles are specified. The perturbation is applied between 5 and 11 km, and the error on observations plus mapping is fixed to 25%. Table 4.5 summarizes the results. As in previous sections, the criterion for comparison between different perturbations is the relative decrease in the cost function.

Figures 4.32 and 4.33 show sample profiles of specific humidity and radar reflectivity as functions of iteration for perturbation amplitudes of -20 and +20%

Table 4.5: Initial and final values of the cost function versus the magnitude of the perturbation to the specific humidity profile.

Perturbation (%)	J_0	J_{FIN}	$\frac{J_{FIN}}{J_0}$ (%)
-20	14700	781	5
-15	6942	292	4
-10	2989	260	9
-5	1018	318	31
-5	1058	5	0.4
+10	3363	35	1
+15	7440	155	2
+20	13463	132	0.9

Table 4.6: Initial and final values of the cost function as functions of error magnitude.

Error (dBZ)	J_0	J_{FIN}	$\frac{J_{FIN}}{J_0}$ (%)
0.1 ($\approx 2.3\%$)	1342282	13198	0.98
1 ($\approx 23\%$)	13422	132	0.98
2 ($\approx 46\%$)	3535	33	0.93
3 ($\approx 69\%$)	1491	15	1.00
5 ($\approx 115\%$)	537	5.3	0.98

respectively. For all perturbations, the assimilation system was able to recover true specific humidity profiles, but only at cloud levels. However, as already found in the experiments with cloud variables as control variables, the assimilation system is more efficient in reducing rather than increasing the amount of ice, and hence the radar reflectivity.

Sensitivity to error assumptions

The sensitivity to error assumptions is investigated, as in section 4.6.1, by fixing the perturbation to a given value (i.e. 20%), and by varying the error covariance matrix \mathbf{W} . The results are summarized in table 4.6.

For a given perturbation to the initial condition, no sensitivity to the error magnitude is found in the relative decrease of the cost function from its initial value to its final value. Example of profiles of specific humidity and radar reflectivity recovered from assimilations with a value of the error amplitude equal to 1 dBZ (corresponding to approximately 23%) are shown in figure 4.34.

Inclusion of background

Results from the optimization without a background term in the definition of the cost function indicate that the observations contain no information about the specific

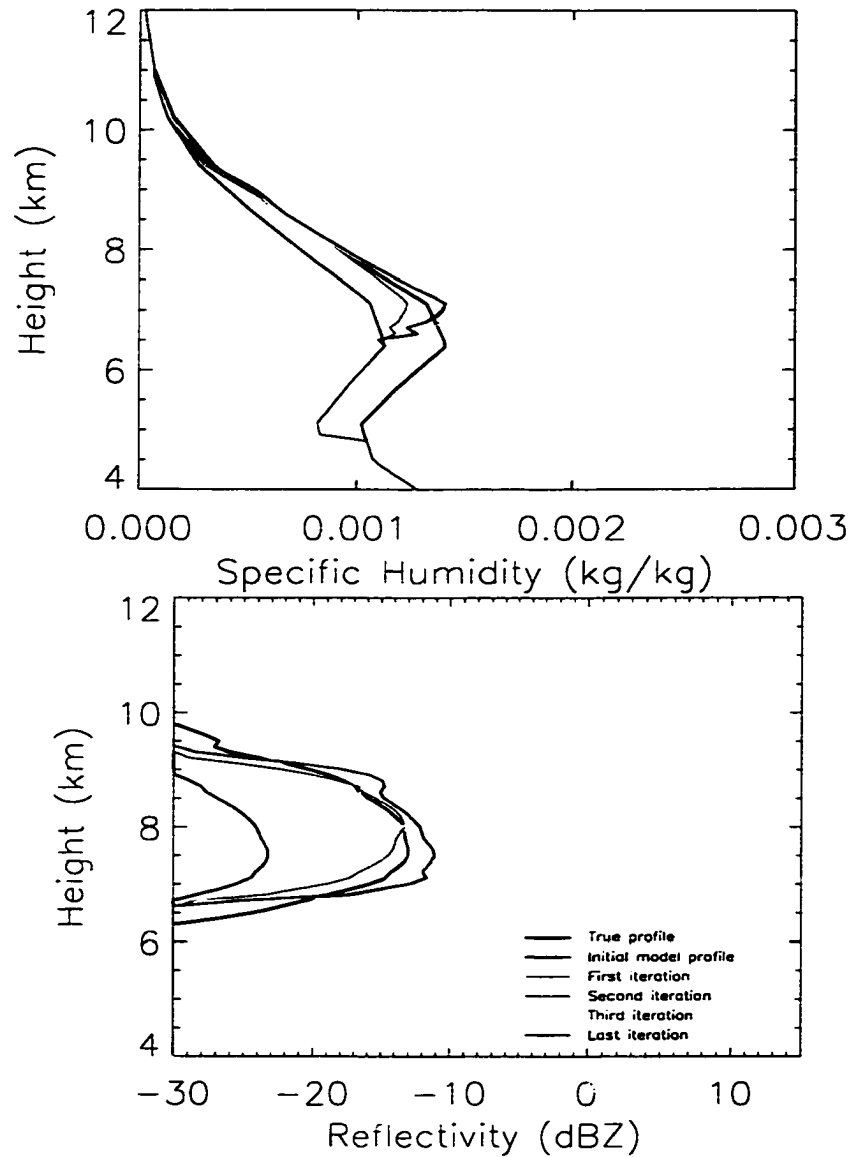


Figure 4.32: Specific humidity (top panel) and synthetic reflectivity at final time (bottom panel) as functions of iteration. The initial perturbation is -20%.

humidity profile below cloud base. The optimization procedure is not capable of correcting the initial guess, and there is only partial convergence to the true solution. The implication is that without additional information on specific humidity below cloud levels, it is not possible to recover the profile from the radar reflectivity alone.

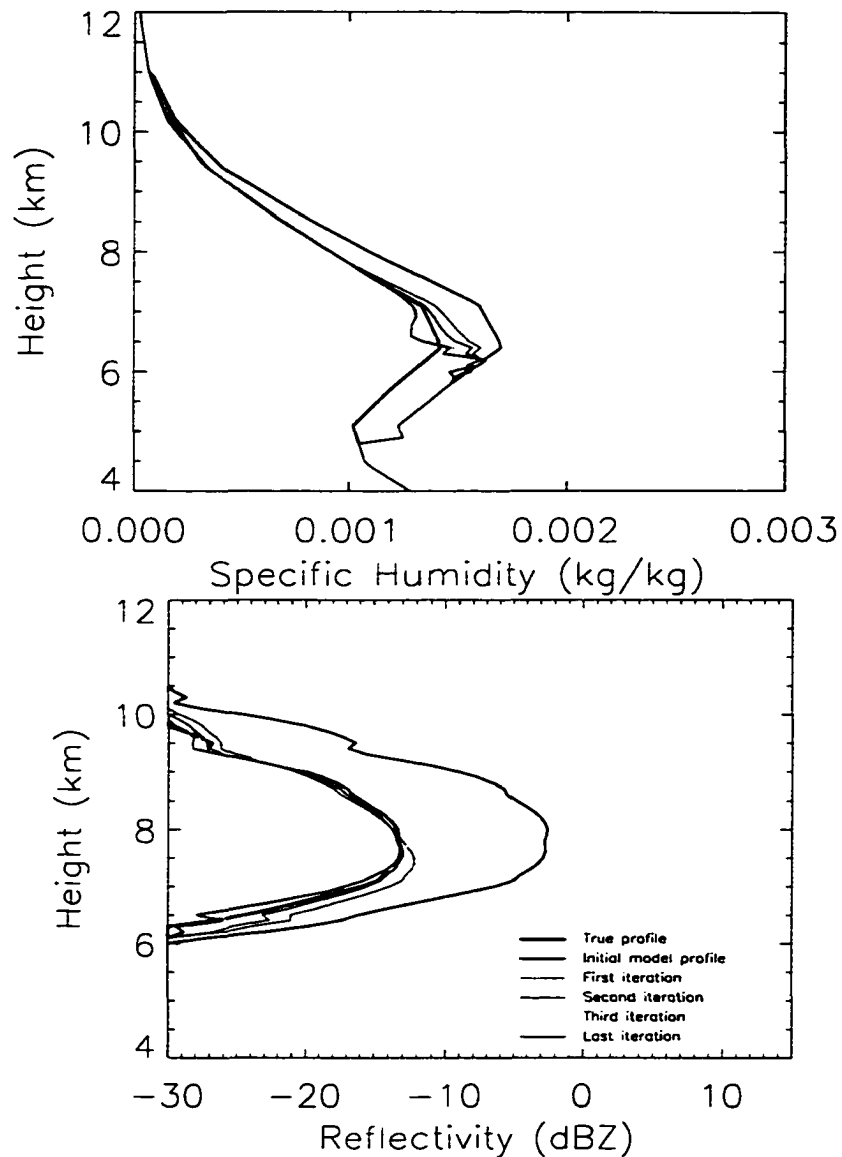


Figure 4.33: Same as fig 4.32, but for an initial perturbation of +20%.

This can be further examined by introducing an arbitrary background humidity profile with a prescribed error covariance matrix, and by observing the behavior of the optimization as a function of the accuracy of the background relative to the magnitude of the measurements plus mapping errors.

For these experiments, the cost function was modified to include the background

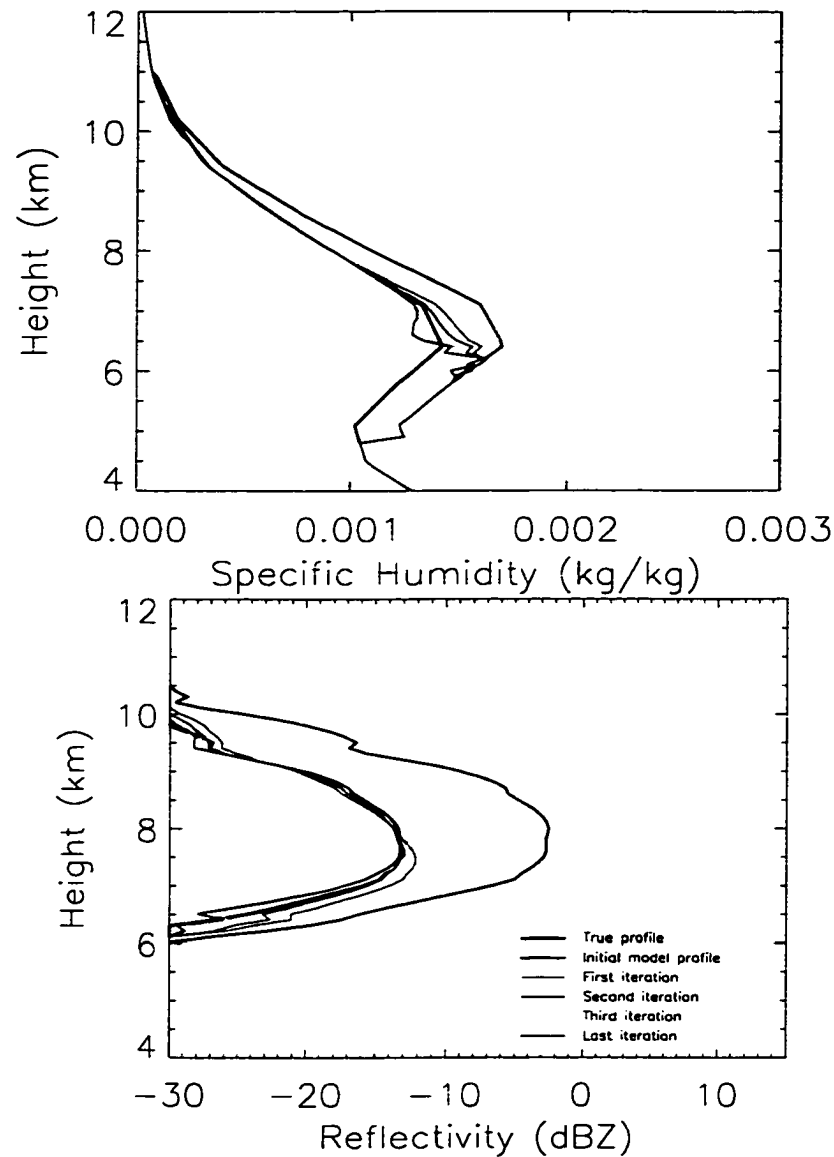


Figure 4.34: Specific humidity (top panel) and synthetic reflectivity at final time (bottom panel) as functions of iteration. Observational plus mapping error is 1 dBZ corresponding to $\approx 23\%$.

term and rewritten as:

$$J = \frac{1}{2}[(Z - Z_{obs})^T \mathbf{W}^{-1}(Z - Z_{obs}) + (q_0^* - q_b^*)^T \mathbf{B}^{-1}(q_0^* - q_b^*)] \quad (4.59)$$

where q_0^* is the initial profile of the dimensionless specific humidity, and q_b^* is the

specific humidity background profile. The gradient of the cost function with respect to q_0^* also includes the background term, $\mathbf{B}^{-1}(q_0^* - q_b^*)$. The initial perturbation was fixed to 20% of the true profile between 5 and 11 km, as in the previous experiments. The background is chosen to be -20% of the true profile between 5 and 11 km.

Two extreme cases are considered: for the first one, the background error is much larger than the observational error (50% error on background versus 1% error on observations); in the second case the situation is reversed (1% error on background versus 50% error on observations). Figure 4.35 applies to the first case, and it shows that, when higher confidence is put in the measurements, the convergence toward the true profile of specific humidity (upper panel) and radar reflectivity (lower panel) is good everywhere except below cloud base. This is to be expected: the radar measurements are not sensitive to changes in specific humidity where the radar does not see a cloud, i.e. where the signal is below the threshold for radar detection (here arbitrarily fixed to the value of -30 dBZ). Moreover, the model does not have other prognostic or diagnostic equations that can relate ice production/dissipation, and hence the radar reflectivity, to the water vapor field. Only so much information can be extracted from this combination of observations and model fields. The background information is not able to compensate for this lack of information in the measurements since low confidence is placed on the former information. On the other hand, when low confidence is given to the measurements, and high confidence is attributed to the background, the profiles will converge to the background as shown in figure 4.36. In this figure, it is also shown that the final result is not close to the true profile, due to the fact that the chosen background was poor and a high weight was undeservedly assigned to it. This result underscores the critical nature of placing proper error assumptions on the various components of the assimilation system. If there are reasons to completely trust the background information, and the observations are only intended to perform a slight adjustment on the background, the error covariance matrices can be defined such as to reflect this confidence. Otherwise, if the observations have a high information content with respect to the control variable, and the background information cannot be trusted—or there is no background information available—the system can be set up so as to use

only measurements. For assimilation of radar data, a situation of compromise is desirable.

From the synthetic examples, it is clear that the radar observations contain significant information about the true profile of specific humidity at cloud levels, but we cannot hope to infer the specific humidity at lower levels by using just the microphysical model as a constraint. By having a background that has some influence below cloud base, it is possible to compensate for the lack of information of the measurements. Another possibility is to complement radar observations with other types of measurements that contain information on the specific humidity at low levels or to use a more complex model.

Discussion of results

Results from this set of experiments with specific humidity as a control variable show that information is contained in radar reflectivities about the water vapor field within the cloud. Some limitations that are intrinsic to the nature of the assimilation problem with the CIGMA model are highlighted:

- radar data assimilation is more successful in suppressing an excess in the cloud rather than in creating a cloud;
- radar measurements contain information only on the specific humidity field at cloud levels, due to the constraint imposed by the model.

The first of these conclusions seems to be robust and consistent with findings from the precipitation assimilation studies discussed in the introductory chapter. Both Hou et al. (2000) and Marécal and Mahfouf (2000), using different models, came to similar conclusions that the optimization was more successful in correcting pixels with an excess of precipitation rather than in increasing precipitation where the model underpredicts. A strict requirement for the assimilation of precipitation or cloud data is that the model has to at least succeed in forecasting some precipitation or some cloud, so that there is a nonzero sensitivity of these fields to the control variable, thus allowing the observations to impact the model prediction. It is not

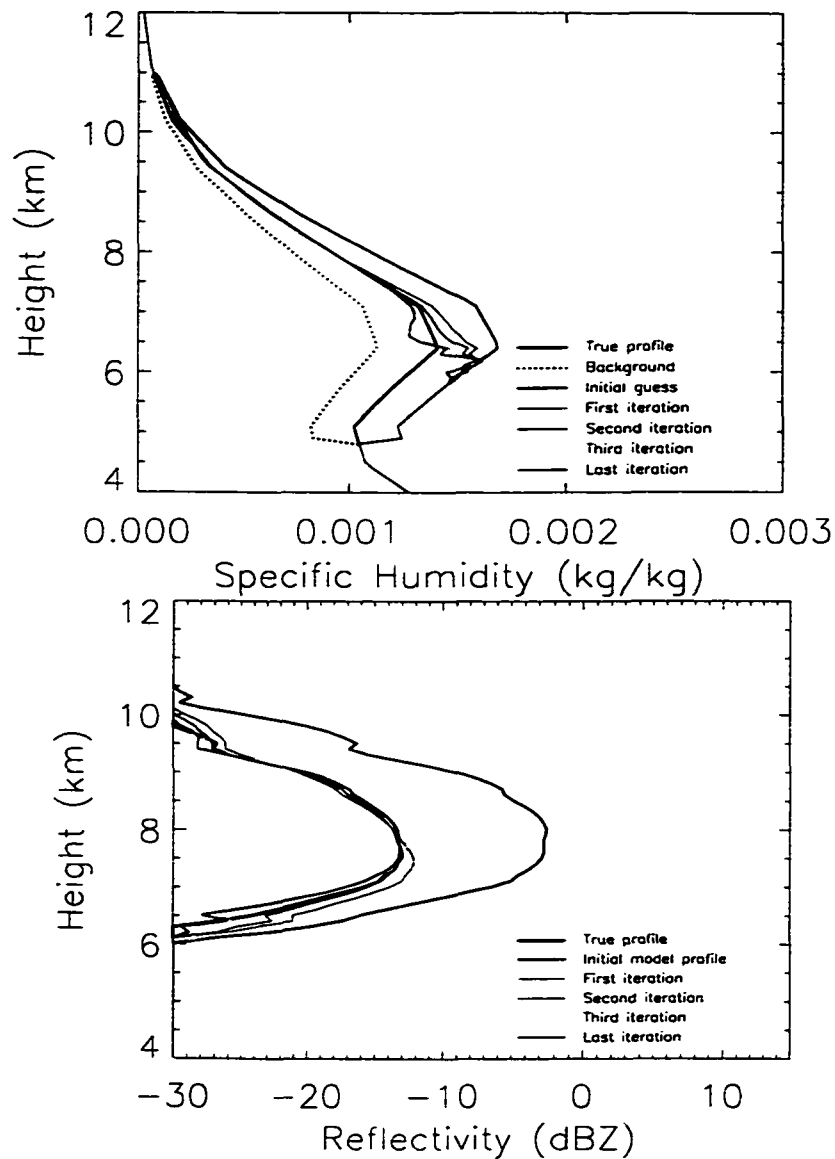


Figure 4.35: Specific humidity (top panel) and synthetic reflectivity at final time (bottom panel) as functions of iteration. Background error is 50%, observational error is 1%.

possible for the assimilation to recover from a situation in which no cloud or no precipitation has been forecast by the model, but where a cloud or precipitation has been observed. This represents a real limitation for a successful variational

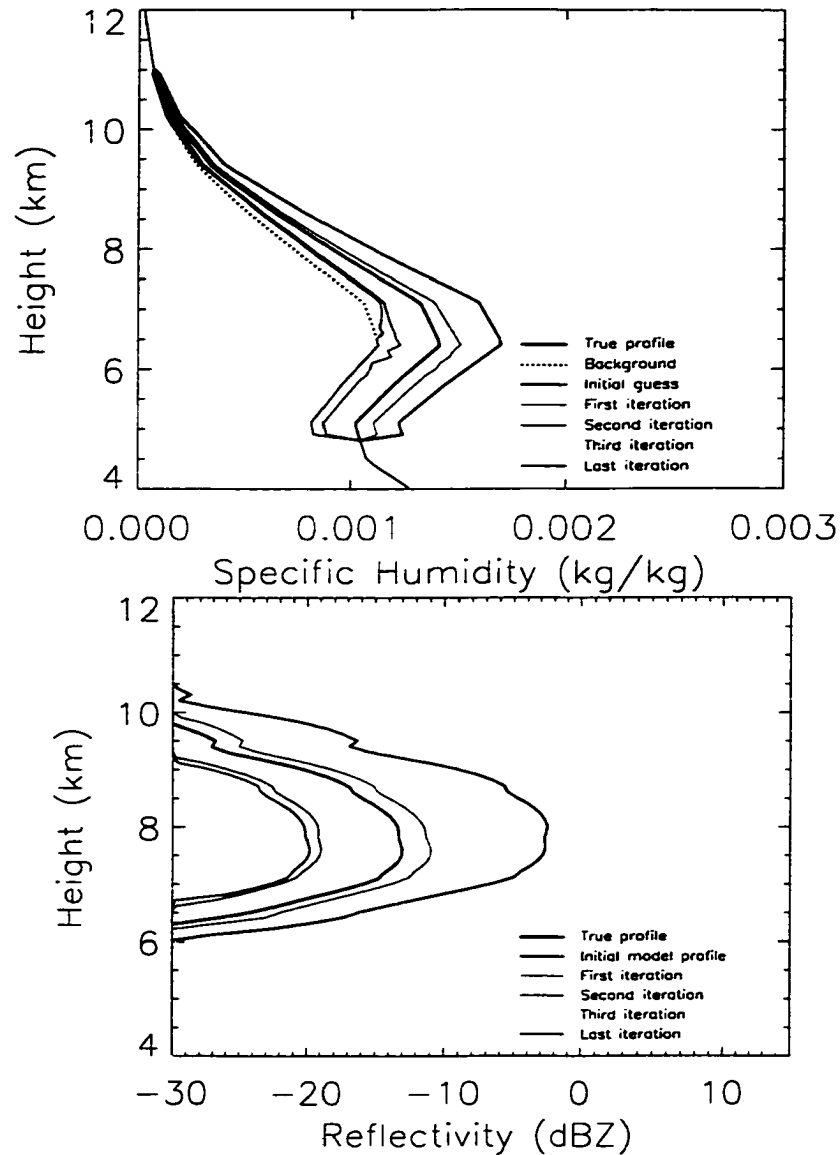


Figure 4.36: Same as fig 4.35, but background error is 1%, and observational error is 50%.

assimilation of cloud and precipitation data since both these fields are susceptible to being under-predicted by the model more than over-predicted.

The second conclusion is more pertinent to the ice growth model used. In chapter 5, it will be shown that assimilation of cloud radar reflectivities in a different

cloud model also has an impact on the humidity fields at levels outside the cloud boundaries.

4.6.3 Model parameters as control variables

The goal of this section is to show that radar reflectivity measurements can be used in variational data assimilation to optimize model parameters on which the solution depends. The same set of synthetic data was used to probe the capability of the assimilation system in recovering optimal parameter values used in the “creation” of the dataset. The focus is on the parameter a that appears in the fall velocity–size parameterization ($v = aD^b$). From the sensitivity study, it was found that radar reflectivity is sensitive to this parameter, especially in the vicinity of the cloud base. In the section dedicated to the validation of the forward model, we saw that a reduction in a could lead to a better agreement between model and observed reflectivities, especially concerning the positioning of the cloud. For this assimilation experiment we varied the parameter a by a given percentage and attempted to recover its true value through the assimilation. Results are shown in table 4.7. Even if perfect convergence was not achieved, the iteration procedure was truncated when the decrease of the cost function and its gradient were reasonably large. This allowed us to retrieve a final value for the variable a which is accurate to within 3% with respect to the true value, i.e. the assimilation of radar data was able to correct for approximately 97% of the initial parameter mismatch.

Figure 4.37 shows the initial and final profiles of the radar reflectivity compared to the true profile for the optimization of the parameter a . The difference between the initial guess and the true value is minimal. It seems that the optimized profile of the reflectivity does not follow the true profile any more closely than the initial guess. However, it can be noticed by looking right above cloud base that the optimization has performed a correction at cloud levels where the sensitivity of radar reflectivity to a is the highest (see figure 4.23). This indicates that the variational assimilation of reflectivity data is capable of correcting the model prediction in the right direction subject to the specific model sensitivity characteristics. Assimilation

Table 4.7: Value of a as a function of iteration. True value of $a= 11.7$ Initial perturbation -50%.

Iteration #	a	J
0	5.85	860
1	6.85	646
2	7.34	581
3	9.31	218
4	17.2	1819
5	11.39	6

of measurements that are more directly related to particle fall speed velocity, such as Doppler radar reflectivities, is likely to offer better results in this type of model parameter optimization. To do this will require the development of a Doppler radar mapping operator and relative adjoint, and this is a task that will be part of future research.

4.7 Results of optimization using real measurements

From the use of synthetic measurements, a lot was learned about the radar data assimilation problem. However, synthetic measurements are no substitute for real observations, whose assimilation is the ultimate goal. In this section we experiment with assimilation of *real* radar data for the two cases presented in section 4.3.5 (ARM-UAV Kauai, April 30, 1999 and ARM-SGP May, 8 1998). As noticed in that section, it was not possible to get a reasonable agreement between the model and observations without adjusting the profile of relative (or specific) humidity to values close to water saturation within the cloud. Nonetheless, since a cloud was generated, and we know that sensitivity to change in initial conditions is large, it is possible to examine how the system behaves when assimilating real measurements, and whether the patterns found in the synthetic data assimilation experiments are

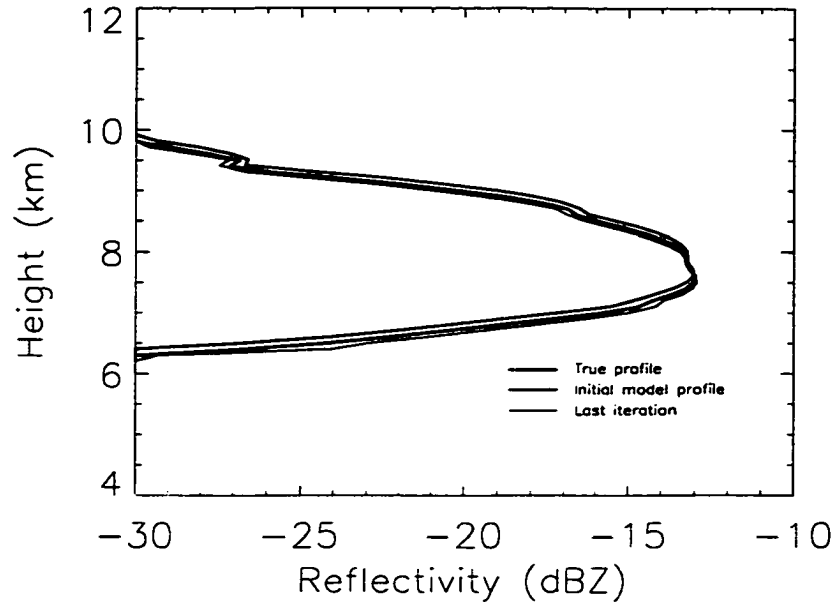


Figure 4.37: Comparison of true synthetic reflectivity at final time (black) with reflectivity computed initial value of a (blue) and optimized value of a (red).

also reproduced in the experiments with real data.

The type of experiments conducted are divided in two categories: experiments with original state variable profiles, and experiments with increased relative (or specific) humidity. Experiments of the former type were performed to determine the possibility of improving the initial humidity field experienced by the cloud through variational assimilation, instead of “manually” increasing the in-cloud relative humidity to a value close to water saturation as described in section 4.3.5. In this context, the manual increase of humidity to match model and observed values, can be considered a (very) rudimentary form of data assimilation. It is desirable to use a more objective and reliable approach. Experiments of the second type only serve the purpose of determining whether or not a further improvement in the increased humidity profile is possible so as to obtain a better agreement between model predicted and observed reflectivities. The remainder of the section presents results from this series of experiments. Only radar reflectivities were used since the addition of optical depth did not particularly improve the synthetic data assimilations. No

background was included. Results show that some of the conclusions derived from experiments with synthetic data still hold for assimilation of real data.

4.7.1 Cloud variables as control variables

Assimilation experiments using cloud variables (characteristic diameter and number concentration) as control variables were only partially successful. Some improvement in model predicted radar reflectivity upon assimilation of observed reflectivities was obtained. However, due to the higher sensitivity to size, the model tried to correct the initial condition on diameter rather than on both diameter and number concentration. This behavior, already observed in the synthetic runs, prevents the assimilation from correcting for erroneous initial conditions on the number concentration. As a result, the cost function stalls after a few iterations, and its decrease is limited. The improvement in model prediction is also modest. Moreover, the fact that the model underpredicted the radar reflectivity reproduces the worst case scenario of the synthetic measurements. Recalling the discussion about the sensitivity of the assimilation system to the amplitude of the perturbation, it was in fact observed that the optimization recovered true profiles more effectively for positive rather than negative perturbations.

Figure 4.38 shows results from the optimization of initial particle size and number concentration using a constant diameter initial profile. The atmospheric profiles are the original ones given by the ECMWF forecast at 23:20 UTC over the Kauai area (shown previously in figure 4.15). Instantaneous observed reflectivities are assimilated over a half hour time interval with a frequency of 5 minutes. The impact of a more frequent assimilation of radar data over the forecast window on the results of the optimization was not investigated. This will be a subject of future work. Comparison is made between the model-predicted profile at final time and the observed profile at final time as a function of the iteration. Note that the instantaneous profiles show more variability than the domain-averaged profiles shown in the comparison section 4.3.5 (figure 4.16). The figure shows dramatically that almost all the correction in the radar reflectivity profile is accounted for by a

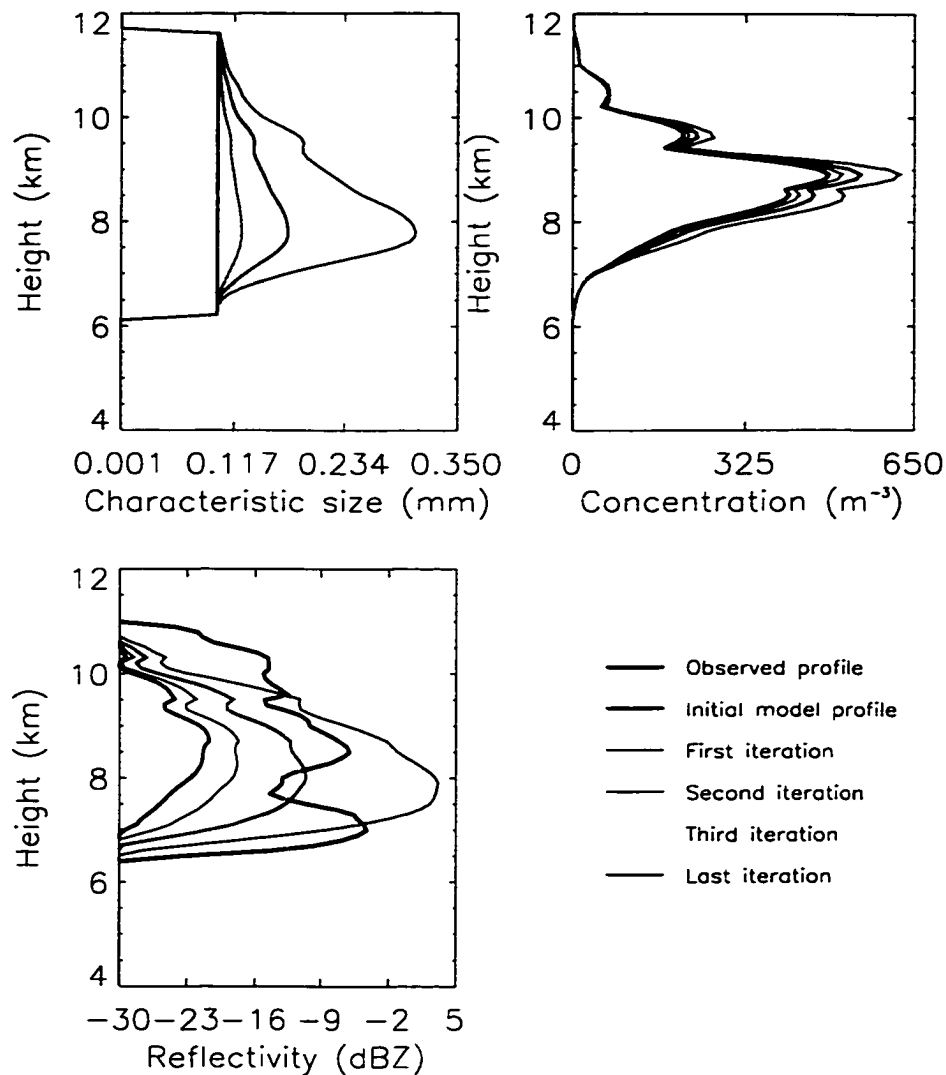


Figure 4.38: Characteristic diameter (top left panel), number concentration (top right panel) and synthetic reflectivity at final time (bottom left panel) as functions of iteration for assimilation of real measurements. Observation plus mapping error is 2 dBZ ($\approx 46\%$). Constant initialization for the characteristic diameter.

correction on initial particle size. The number concentration is increased only by a small amount, if at all.

Figures 4.39 and 4.40 show optimum initial conditions on the size and number concentration for an initial condition for the specific humidity close to water satura-

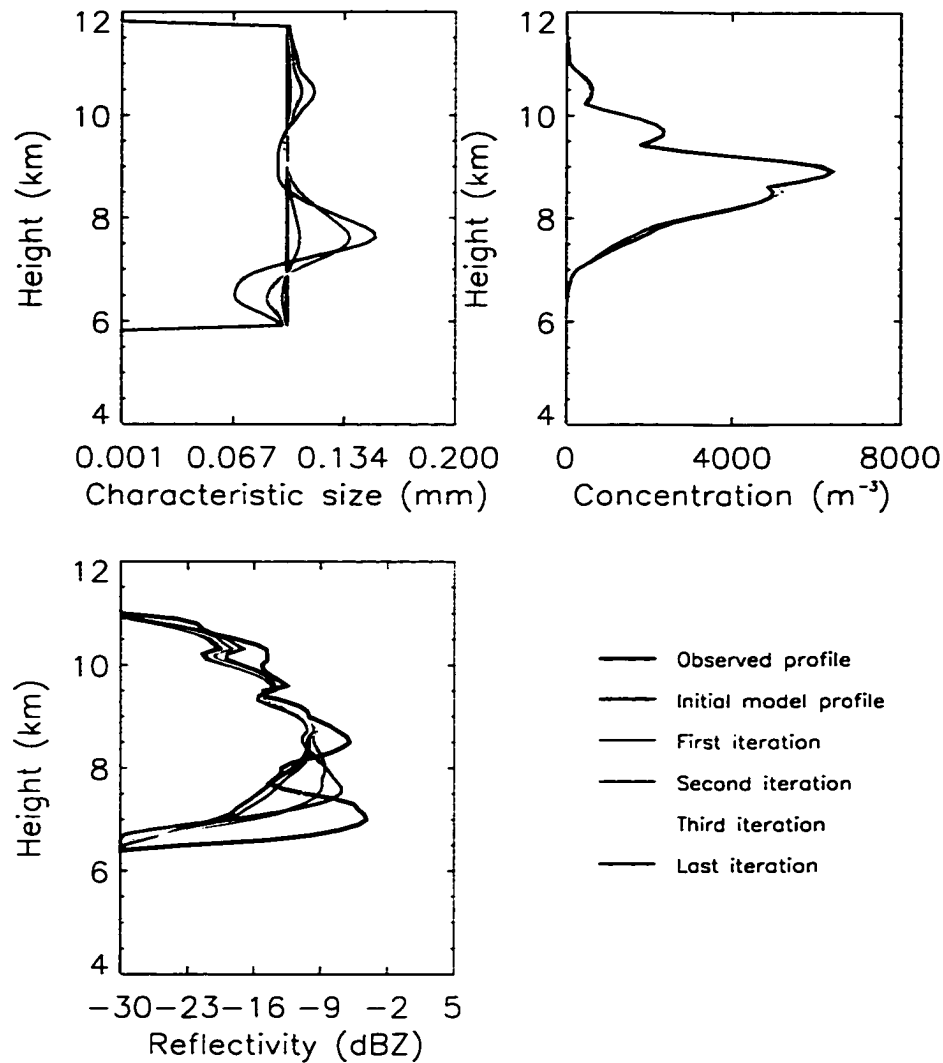


Figure 4.39: Same as figure 4.38, but with increased initial specific humidity. Constant initialization for the characteristic diameter.

tion (see figure 4.17). The difference between these two figures is the initialization chosen for the characteristic diameter.

Experiments with the ARM-SGP indicated a similar behavior. In both examples, there is definitely a large impact of the radar data assimilation on the initial profiles of characteristic size, and the cloud predicted with optimized fields is geometrically thicker than the first guess cloud and better resembles the instantaneous

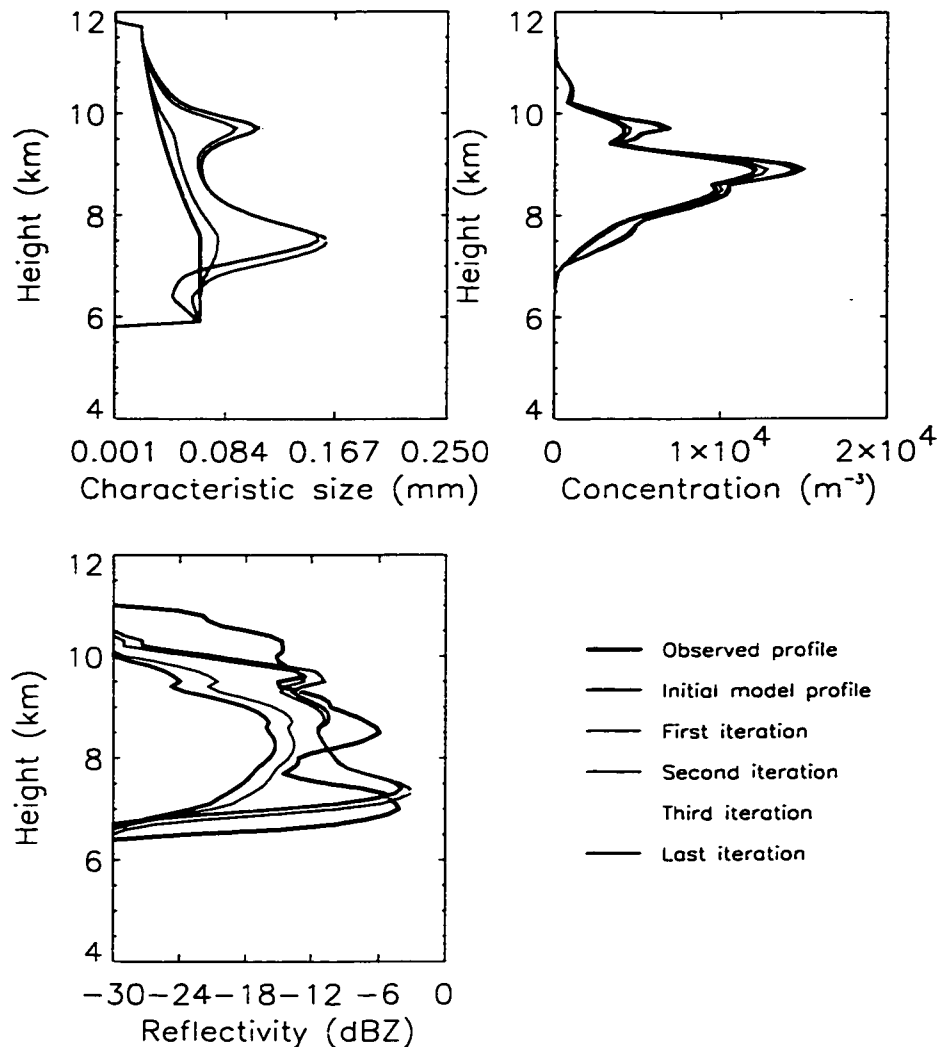


Figure 4.40: Same as figure 4.38, but with increased initial specific humidity. ECMWF initialization for the characteristic diameter.

observed profile. However, this is not sufficient, since the initial condition for the number concentration is almost unaffected by the inclusion of radar data. This reinforces the conclusion drawn from the synthetic data experiments: it is possible to correct for an initial condition for the diameter by using real observations, but it is not possible to correct the number concentration. Better results are obtained by correcting the specific humidity profile on which the number concentration depends (see the next section).

4.7.2 Specific humidity as a control variable

Assimilation experiments with real data using specific humidity as a control variable rather than N_0 and D_0 were more successful. Starting from the Kauai case with the original specific humidity profile, figure 4.41 shows the profile of specific humidity, and the corresponding initial number concentration as functions of iteration. Radar reflectivity is also shown. A small increase in specific humidity at upper levels can induce a huge increase of up to several orders of magnitude in the number concentration (note the logarithmic scale on the horizontal axis of the number concentration plot). This increase would be even larger if homogeneous nucleation were considered.

This result implies that the radar measurements can be used to infer information in the vapor field within the cloud. When the specific humidity is initially increased at cloud levels, the correction on specific humidity, and hence on initial number concentration, is much smaller (see figure 4.42). However, given the results from figure 4.41, this step of initially increasing the specific humidity is not necessary. The inclusion of radar reflectivities has no real impact near cloud base, where the initial guess for the specific humidity is not corrected. This behavior was also manifest in the assimilation of synthetic data.

For the cirrus case observed at the ARM-SGP on May 8 1998 the results are similar to the Kauai case. The assimilation of radar reflectivities improves the agreement between the model prediction and the observed cloud by optimizing the initial condition for the specific humidity (see figure 4.43).

4.8 Summary and conclusions

This chapter offered an overview of the ice growth model developed for variational assimilation studies of radar data. A full derivation of model equations was offered along with a detailed discussion of model parameterizations and assumptions. A detailed evaluation of model performance by comparison to cirrus observations was also given.

A linear sensitivity analysis to a range of variables (particle size and number

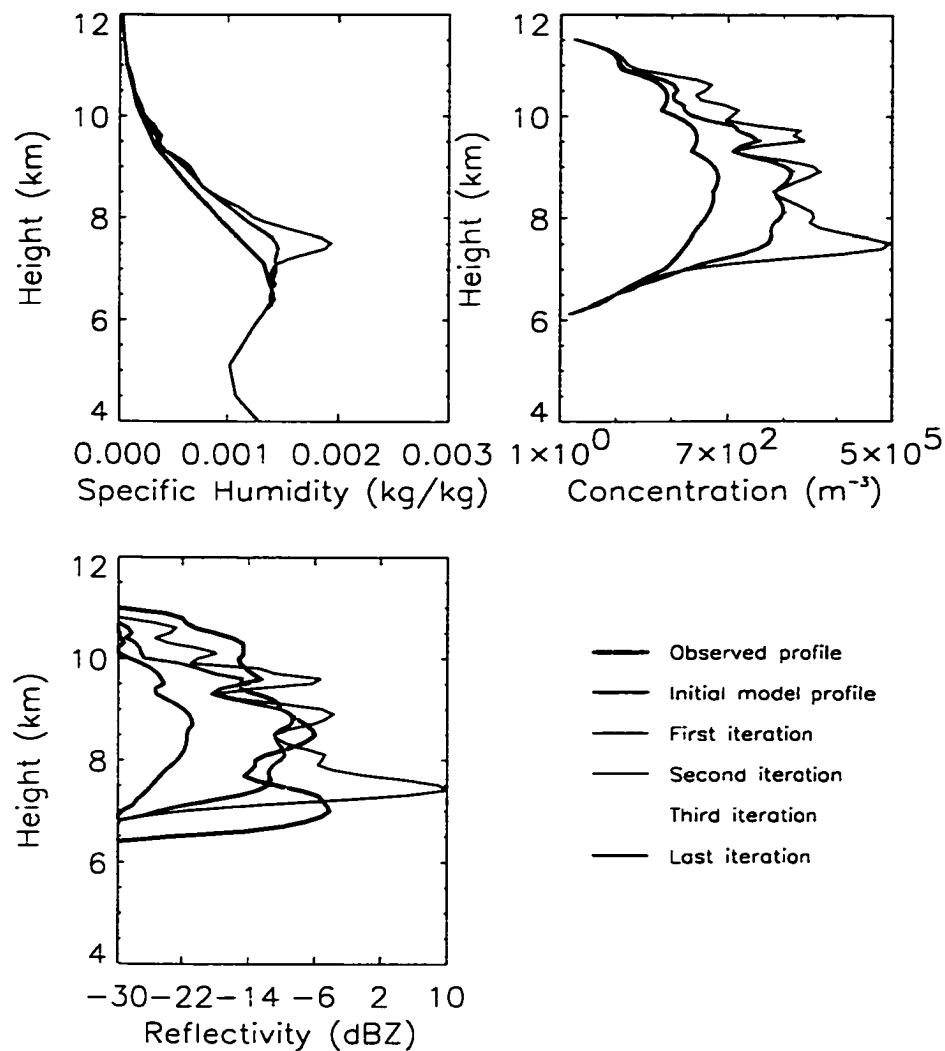


Figure 4.41: Specific humidity (top left panel), number concentration (top right panel) and synthetic reflectivity at final time (bottom left panel) as functions of iteration for assimilation of real measurements. Observation plus mapping error is 2 dBZ ($\approx 46\%$). Constant initialization for the characteristic diameter. Original specific humidity profile.

concentration initialization, model parameters, and atmospheric profiles) was also performed by introducing the *adjoint* of the model. The results highlight the acute sensitivity of the cloud model to certain key parameters.

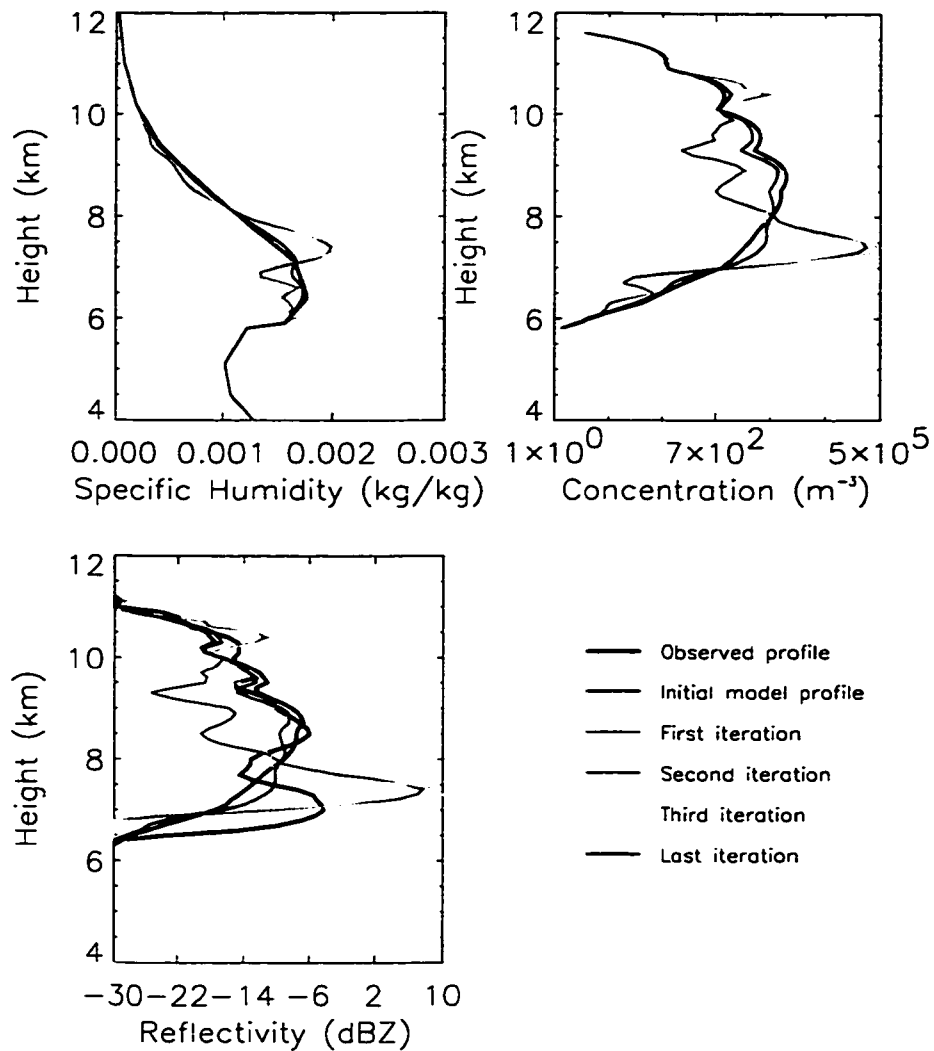


Figure 4.42: Same as figure 4.41, but for an increased initial specific humidity.

Variational data assimilation experiments were performed using both synthetic and true measurements. Responses to varying initial conditions and elements of the assimilation system, such as the observational plus mapping error covariance matrix, were investigated. The results demonstrate the feasibility of the assimilation of radar reflectivity and that such assimilation has real potential to improve cloud prediction. Experiments also highlight a number of serious limitations, some of which are related to the specific model used⁹, such as the lack of prognostic equations for

⁹Results from any assimilation exercise are necessarily model dependent, i.e. a different model

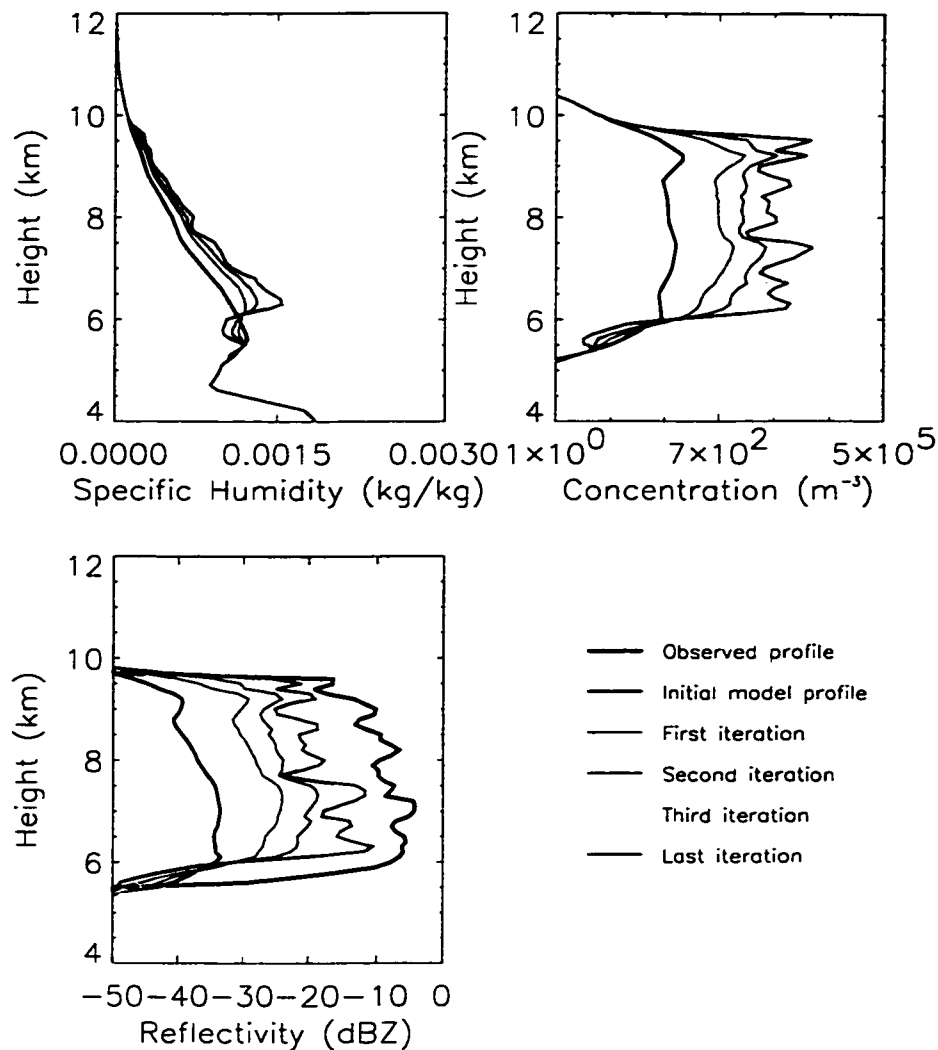


Figure 4.43: Same as figure 4.41, but for the SGP May 8 1998 cirrus case.

thermodynamical and dynamical fields, and the lack of representation of turbulent and radiative processes, which limit the applicability to real world clouds whereas others are more pertinent to the radar data themselves, such as the higher sensitivity of radar reflectivity to particle size than to number concentrations. We now attempt to address the first type of limitation using a more detailed cloud resolving model (RAMS) and pursue the assimilation of radar data using the adjoint of this with different physics and/or parameterizations will have peculiar sensitivities, and hence a specific response to the assimilation of radar fields.

model.

Chapter 5

Prelude to 4D-VAR assimilation using RAMS adjoint

5.1 Introduction

This chapter introduces some preliminary results obtained with the Regional Atmospheric Modeling Adjoint System (RAMAS, Vukićević et al. (2001)), which can be considered the *prelude* to four-dimensional data assimilation of radar reflectivities into the RAMS model.

The RAMAS software has been completed only recently, and it has not yet been tested under a wide range of conditions. Results shown in this chapter are hence highly experimental. They include the investigation of system sensitivity of the ice cloud amount to thermodynamic variables (potential temperature), moist variables (specific humidity), and dynamic variables (horizontal and vertical winds). The goal of the first part of this study is to examine whether the sensitivity of the radar reflectivity to changes in environmental conditions is as large as indicated by the CIGMA system in section 4.4 of Chapter 4. As pointed out, this disproportionate sensitivity could be either a model artifact or a feature common to both a simplified and a complex mesoscale model. RAMAS can help address this question.

The second set of experiments is more directly connected to four-dimensional variational assimilation. Various synthetic radar reflectivity datasets are created by perturbing a RAMS baseline forward run, and then used to compute the adjoint forcing for the cost function as in chapter 4. The forcing is applied to the adjoint equation, which is then solved in a time backward integration to get the system response to the assimilation of “bogus” reflectivities. Gradients of the cost function

with respect to ice variables (pristine ice and snow), water vapor, temperature, perturbation Exner function (related to pressure), and winds are obtained from the adjoint solution. These gradients at initial time represent what would be needed to implement the search routine and get the updated optimal initial or boundary conditions for the chosen control variables, i.e. specific humidity or temperature. This is the only missing step in the 4DVar process, which will be fully carried out as part of future work. Results presented in this chapter are analyzed and discussed under the 4DVar optimal model initialization perspective, but also under a new perspective, as instruments for understanding the model, and for extracting the largest quantity of information from cloud radar data by using the model to connect all atmospheric fields to ice cloud characteristics.

5.1.1 Variational data assimilation in regional models

In the introductory chapter of this thesis some of the previous variational assimilation studies performed with regional forecast models were briefly reviewed. In section 3.2 of Chapter 3, sources of uncertainties for regional models were concisely presented, and data assimilation was identified as a way to address errors related to sub-optimal choices of initial and boundary conditions, and by doing so, to minimize the mesoscale model forecast error. In this section, we expand on this topic with the goal of providing a more complete picture. A fundamental difference between global and regional models is the need of the latter for accurate lateral boundary conditions (LBC) in addition to initial conditions (IC). There is evidence that mesoscale forecasts are more sensitive to LBC than to IC specification (Errico et al., 1993). This is due to the fact that as time progresses features are advected from inflow boundaries into the interior of the domain, as interior features exit at the outflow boundaries. In general, boundary conditions for limited-area models are specified from some source external to the model, for example a global-scale NWP forecast or weather station sounding data, unless horizontal periodicity of the fields is imposed. This specification of boundary conditions requires some form of data assimilation. An example of this was offered in chapter 3 with the nudging of ECMWF fields into

RAMS where, based on the discussion presented in chapter 2, it was reemphasized that nudging is a sub-optimal data assimilation technique. For this reason, a valid alternative (or a complementary tool) to nudging is variational assimilation.

However, the value of variational assimilation goes beyond that of being an optimal model initialization technique. In fact, the power of the adjoint model which is at the core of variational systems can be exploited to learn about the behavior of a nonlinear models as illustrated in chapter 4. In addition, the optimal combination of observations and model fields opens the doors to extracting more information than either of the two components has by itself. Model prediction needs to be evaluated with real data. Observations on the other hand, offer only a partial representation of the atmosphere state, since they are only sensitive to a specific field, whether it is cloud variables, specific humidity or CO₂ concentrations. The model however can represent all the links between directly observed variables and non-observed variables, and it can bring out the full potential of the observations. An example of this application is presented in the following sections: cloud radar data alone do not provide information on the state of the whole 3D atmosphere. Yet it will be shown that, since radar reflectivities are sensitive to ice cloud amount, and the latter is a result of specific atmospheric conditions, it is possible to derive information about the ambient environmental conditions in which the cloud formed, through the use of the mesoscale model and its adjoint.

5.2 Requirements for VDA into limited-area models

Requirements for variational DA into mesoscale models include primarily the need for computational efficiency. If the model is run for real-time forecasts, the optimization of the cost function has to be as fast as possible. In general, optimization is truncated after a few iterations, and an approximate solution is used.

Methods to improve convergence through *preconditioning* of the control variable are also utilized extensively. Preconditioning implies a redefinition of the first guess for the control variable so to bring it closer to its optimal value and improve convergence. These methods are generally based on an approximation to the matrix of

the second derivatives of the cost function—the Hessian¹—(see for example Zupanski (1998)).

Alternatively, an expedient approach used in operational settings is to implement a *linearized* version of the nonlinear model in the optimization procedure as done in the ECMWF 4DVar system (e.g Mahfouf and Rabier, 2000). As shown in figure 4.25, each time a new value of the control variable is obtained as result of the optimization, a new call to the nonlinear model is performed before recomputation of the cost function. This call makes many applications computationally intractable, since any weather forecasting or cloud model is usually very expensive. This represents the main drawback of the variational technique, and the solution of this efficiency issue is a discipline in and of itself. However, it is obviously desirable to keep efficiency in mind while constructing the adjoint of the nonlinear model. Similarly advisable is the efficiency of the operational operator. This is especially crucial for assimilation of cloudy radiances, since the radiative transfer operator is rather complex, and computation can be quite slow. An efficient radiative transfer code is presented in

¹The fact that the Hessian matrix is useful in preconditioning can be illustrated very simply. Considering the cost function, $J(x_0)$, we wish to minimize with respect to the control variable x_0 , the necessary condition for a minimum is that its gradient, $\nabla J(x_0)$, is equal zero. If the gradient is expanded in Taylor series about a point x_0 near the minimum, x_0^* , we have

$$\nabla J(x_0^*) = \nabla J(x_0) + \nabla^2 J(x_0)(x_0^* - x_0) + o((x_0^* - x_0)^2) \quad (5.1)$$

where $\nabla^2 J(x_0)$ is the Hessian matrix. At the minimum $\nabla J(x_0^*) = 0$, which implies

$$\nabla J(x_0) + \nabla^2 J(x_0)(x_0^* - x_0) + o((x_0^* - x_0)^2) = 0 \quad (5.2)$$

Neglecting higher order terms and rearranging, the optimal value is given by

$$x_0^* = x_0 - [\nabla^2 J(x_0)]^{-1} \nabla J(x_0) \quad (5.3)$$

If the Hessian is known and its inverse is well-behaved, and the observational operator is linear, then the least square minimization problem can be solved in one step by applying equation (5.3). In general the Hessian matrix is too big to be computed and stored at all model points during the optimization procedure. For this reason, approximate estimates to the Hessian (for example only diagonal elements) are calculated beforehand and used in the optimization.

appendix C. For this reason when assimilating cloud information, radar data are ideal, since the mapping can be implemented very efficiently.

In the experiments with the CIGMA model, speed was not a real issue, due to the relative simplicity of the model. However, when dealing with a model such as RAMS, accuracy and speed are priorities. The next section briefly outlines specific characteristics of RAMS adjoint construction.

5.3 RAMS adjoint construction

Construction of an adjoint system even for a complex mesoscale model such RAMS follows the simple steps outlined in section 4.4.1 with the caveats about discontinuous processes listed in the introduction (section 1.1.3). The derivation of RAMAS was conducted mainly with the Automatic Differentiation software TAMC (Giering, 1999), also used in the CIGMA adjoint model construction. The validation of the adjoint code obtained with TAMC was performed by first comparing the Tangent Linear solution to the difference of two slightly perturbed nonlinear model solutions and then using the TL to validate the adjoint via the scalar product equality of equation (2.11). More details on RAMAS development and evaluation can be found in Ghemires et al. (2001) and Vukićević et al. (2001).

The structure of the RAMAS software conforms to the general RAMS structure. A main driver coordinates, at each time step, the calls to the different subroutines that treat various physical processes such as advection, turbulence, microphysics, and radiation. In the adjoint, these calls are performed in *reverse* order and the time stepping is backward from final to initial time. All main subroutines in the forward model have their corresponding adjoint subroutine. The initial condition on each adjoint variable is zero.

Efficiency of the adjoint integration, listed as a requirement in the previous section, is achieved quite simply but ingeniously in RAMAS by saving the basic state in the nonlinear model run and by reading it in during the evaluation of the adjoint solution. This technique was also used in CIGMA. One major limitation is represented by the fact that a huge quantity of memory is used to store the forward

trajectory. In this study, the adjoint integration had to be limited to one hour for this reason. One solution to the problem of memory space is the subdivision of the adjoint integration into small adjoint sub-problems for the construction of small pieces of trajectory and the parallelization of the adjoint code as suggested by Ghemires et al. (2001).

5.4 Experiment setup

As already discussed, a prerequisite for cloud variational assimilation is that a cloud must be present in the model, i.e. the model has to show some ability to predict at least the occurrence of the cloud, if not the correct position and microphysics characteristics, otherwise the inclusion of cloud data does not have any impact. As briefly mentioned in chapter 3 when discussing results from 2D simulations, it is very difficult to initiate cloud formation from ice undersaturated horizontally homogeneous initial conditions, regardless the degree of sophistication of the model. The intrinsic model spin up time can be long, and in certain cases the cloud might not form at all if an external forcing is not superimposed. We hence decided to use the idealized warm cirrus profile from the GCSS-WG2 intercomparison activity which, being supersaturated, ensured cloud formation². The single grid was chosen to cover an area of approximately 20x20 km in the horizontal, and it extended from the surface up to 17.5 km in the vertical. Horizontal resolution was fixed at 500 m. Vertical resolution was variable, lower at low levels and higher at cloud levels. Since the idealized profiles are generally representative of midlatitude average conditions, the grid was centered over the Oklahoma ARM site. This choice also minimized the

²The rationale for implementing the model in this configuration was to avoid additional external model forcing that could mask the effect of radar data assimilation. This is also the main reason why new cirrus simulations were performed and the 3D simulations presented in chapter 3 were not used. Those simulations were in fact heavily nudged, and the radar observations were assimilated indirectly by the manual tuning of the nudging coefficients to minimize model bias. Another major limitation to using multiple nested grid simulations such as the Kauai case is the fact that at present RAMAS does not have a multi-grid nesting capability in the adjoint formulation that mirrors RAMS'.

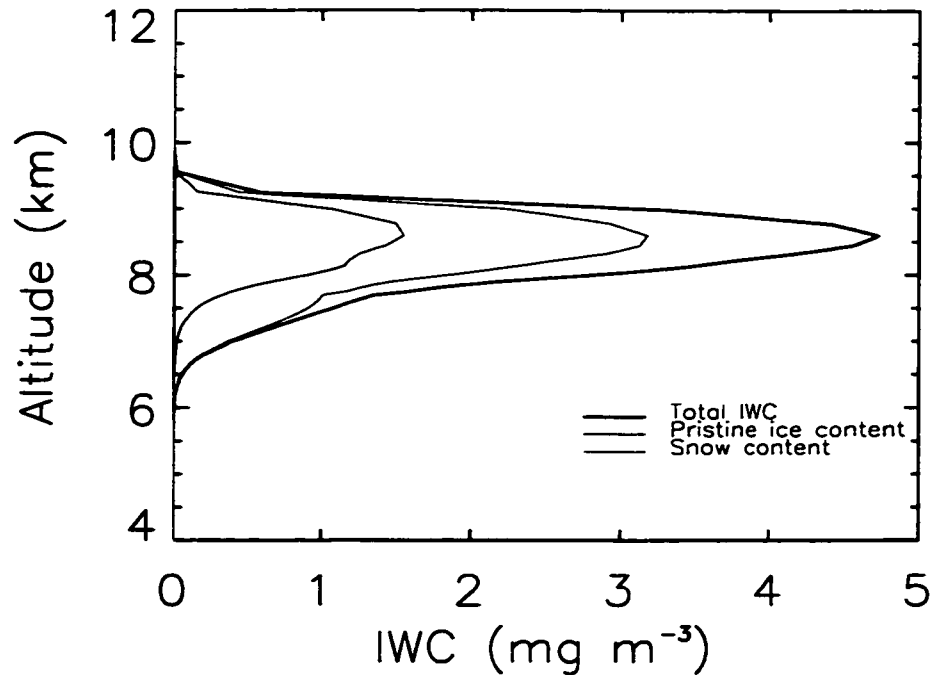


Figure 5.1: Domain averaged profile of IWC ($mg\ m^{-3}$). Relative contributions from pristine ice and snow are also shown.

impact of topography. The simulated cloud is similar to the one shown in figure 3.6 of chapter 3 with a peak ice water content of $5\ mg\ m^{-3}$ and a vertical extension of approximately 2.5 km, from 7 to 9.5 km, at one hour after the beginning of the simulation. The horizontally-averaged profile of IWC and the relative contribution of pristine ice and snow³ contents are shown in figure 5.1.

The presence of the initial cloud field allowed for investigation of the model response to forcings in the environment and in the ice cloud itself. Results from these experiments are shown in sections 5.5 and 5.6.

³As a reminder, pristine ice crystals represent unaggregated crystal of diameter smaller than $125\ \mu m$, whose main growth process is vapor deposition, while snow represents larger crystals grown by vapor deposition and riming, i.e. by collection of supercooled cloud droplets. The latter process is not relevant in cirrus clouds.

5.5 System response to change in environmental variables

This first set of experiments seeks to quantify the impact of a percentage change in environmental parameters on cloud ice and radar reflectivity. As illustrated in chapter 4, this impact can be quantified by analyzing the solution of the adjoint equation. The adjoint forcing for the variable cloud ice is chosen to be equal to unity. This choice is explained by the following simple considerations. In chapter 2, section 2.3.2, it was shown how the adjoint operator provides a means to compute the gradient, and hence the sensitivity, of *any* scalar function of the variable x with respect to any other variable that x depends on (y in our example) via the simple inner product of equation 2.15:

$$\nabla_y \mathcal{J} = \mathbf{F}'^\dagger \nabla_x \mathcal{J}. \quad (5.4)$$

where \mathbf{F}'^\dagger is the adjoint operator. Suppose that the functional we are finding the gradient for is the variable x itself, i.e. $\mathcal{J} = x$. In this case $\nabla_x \mathcal{J}$ is simply equal to 1, and $\nabla_y \mathcal{J} = \mathbf{F}'^\dagger$. In a way, this can be considered the most elementary choice of cost function. Assuming the variable x depends on a set of variables, i.e. y is a vector with components y_1, y_2, y_3 , etc., the total variation of x will be given by:

$$\Delta x = (\nabla_y x, \Delta y) \quad (5.5)$$

which explicitly is

$$\Delta x = \frac{\partial x}{\partial y_1} \Delta y_1 + \frac{\partial x}{\partial y_2} \Delta y_2 + \frac{\partial x}{\partial y_3} \Delta y_3 + \dots \quad (5.6)$$

The relative contributions are represented by the terms of the sum, which are the solutions of the adjoint equation with respect to the given variable. In what follows, x is first identified with the pristine ice mixing ratio (r_{pristine}) and then with the snow mixing ratio (r_{snow}). The variables y_1, y_2, \dots are identified with the environmental variables of potential temperature (θ), specific humidity (q), zonal

wind (u), meridional wind (v), vertical wind (w), and perturbation Exner function (Π'). Hence we have

$$\Delta r_{ice} = \frac{\partial r_{ice}}{\partial \theta} \Delta \theta + \frac{\partial r_{ice}}{\partial q} \Delta q + \frac{\partial r_{ice}}{\partial \Pi'} \Delta \Pi' + \dots \quad (5.7)$$

and similarly for snow mixing ratio.

From equation (5.7), the corresponding variation of radar reflectivity due to a change in ice content can also be computed using the following equation from chapter 3:

$$Z = c(IWC)^d, \quad (5.8)$$

where

$$IWC = (r_{pice} + r_{snow})\rho_a \quad (5.9)$$

ρ_a is the local air density, and c and d are coefficients that relate the radar reflectivity to the cloud ice water content and depend on the radar wavelength. In the runs used for these pre-assimilation experiments, aggregates were not included, based on the results from the sensitivity studies to inclusion/exclusion of ice categories presented in chapter 3. The variation in radar reflectivity is given by:

$$\Delta Z = \frac{\partial Z}{\partial r_{pice}} \Delta r_{pice} \quad (5.10)$$

where from equation (5.9) and (5.8):

$$\frac{\partial Z}{\partial r_{pice}} = \frac{\partial Z}{\partial (IWC)} \frac{\partial (IWC)}{\partial r_{pice}} = cd(IWC)^{(d-1)}\rho_a. \quad (5.11)$$

5.5.1 Sensitivity results

Environment fields provided by the nonlinear model integration and gradients provided by the adjoint integration were averaged over the horizontal domain to obtain vertical profiles. A perturbation of 1% of forward model fields was applied uniformly to the average profiles of ambient variables, and corresponding pristine ice and reflectivity variations at cloud levels, i.e. where average ice water content was not

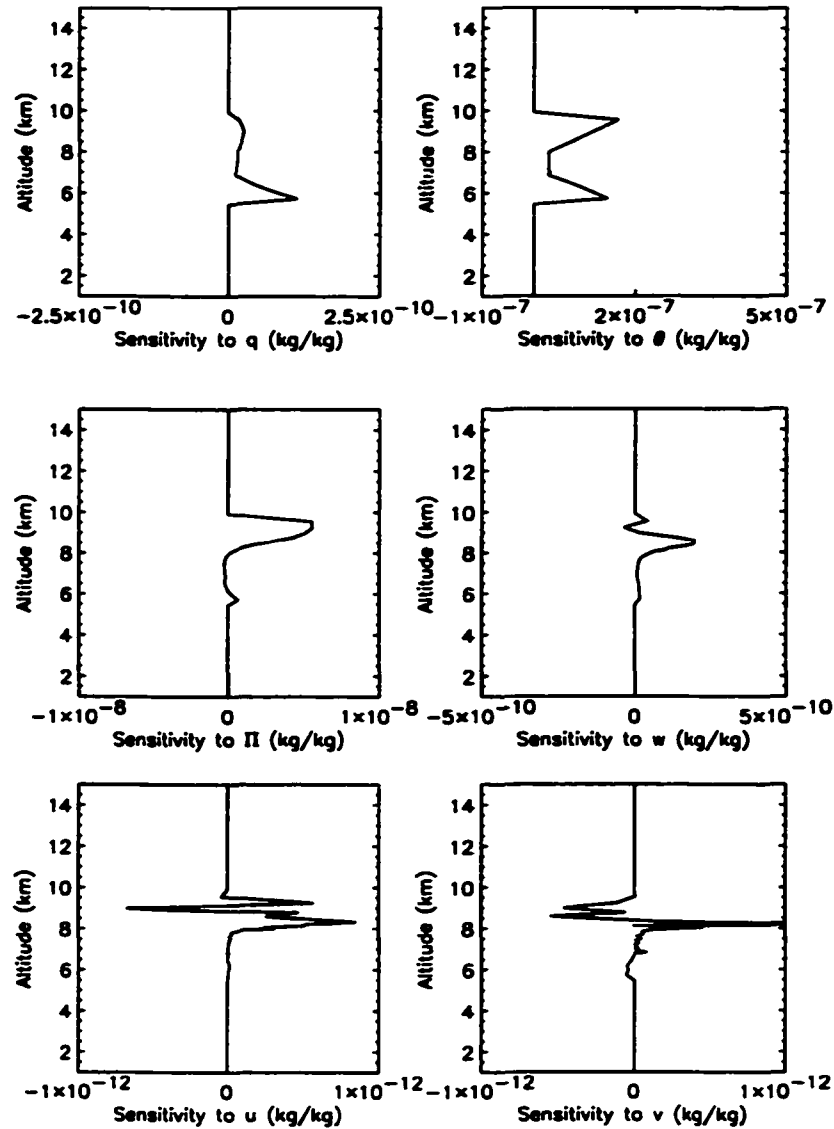


Figure 5.2: Sensitivity of pristine ice mixing ratio to 1% change in ambient variables. In order, starting from the upper left corner: sensitivity of r_{ice} to specific humidity, q (top left panel); sensitivity to potential temperature, θ (top right panel); sensitivity to Exner perturbation function, Π' (middle left panel); sensitivity to vertical wind, w (middle right panel); sensitivity to zonal wind, u (bottom left panel); sensitivity to meridional wind, v (bottom right panel). See text for explanations.

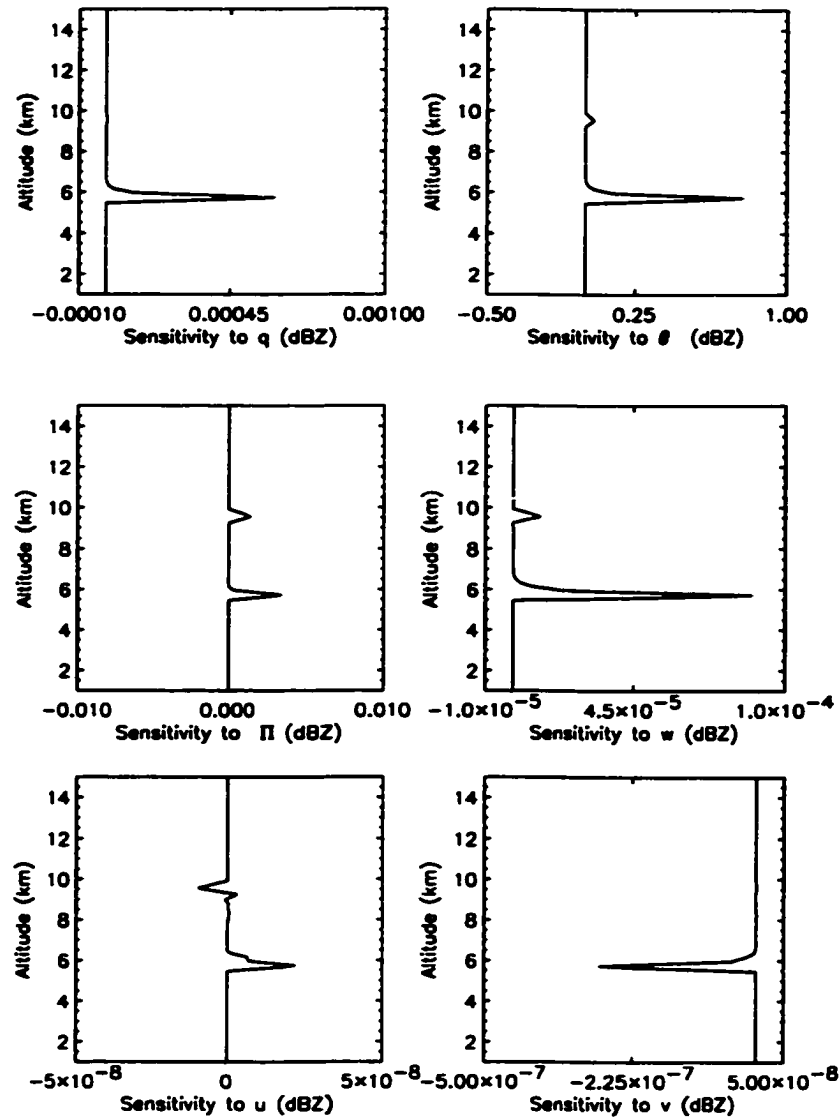


Figure 5.3: Same as figure 5.2, but for radar reflectivity.

zero, were computed. To facilitate comparison, all field sensitivities are expressed in pristine ice mixing ratio units (single terms of equation (5.7), that is the “absolute” sensitivities are shown). Results are presented in figures 5.2 and 5.3 for pristine ice.

A comparison of sensitivity to specific humidity and potential temperature (upper panels of figure 5.2), indicates that a higher sensitivity at cloud levels is explained by the pristine ice mixing ratio profile to the potential temperature field.

Considering an average value for pristine ice of $2 \cdot 10^{-06} \text{ kg/kg}$, corresponding to approximately 1 mg m^{-3} , a 1% increase in potential temperature at cloud levels induces a percentage increase in pristine ice of roughly 10%. Much lower is the change in pristine ice induced by a 1% variation in specific humidity ($\approx 0.01\%$). It must be observed, however, that specific humidity values at cloud altitudes from 6–10 km are quite small, and a 1% variation is not very significant. By contrast, a 1% variation in potential temperature at those levels corresponds to approximately 3K, which is a rather large perturbation in temperature. The potential temperature response is in phase with the response to a perturbation in the specific humidity. Since potential temperature and temperature are anticorrelated, i.e. an increase in the former corresponds to a decrease in the latter at constant pressure, this implies that the temperature response is out of phase with the response to a specific humidity perturbation. An *increase* in temperature at cloud levels, leads to a *decrease* in ice content. These RAMAS sensitivity experiments seem to confirm what was found with the ice growth model in chapter 4. The main difference is that the impact of a change in temperature is not as dramatic in the results presented here as in the results shown in section 4.4. This is due to the fact that temperature affects, in a complex manner, different processes related to ice cloud formation and evolution, such as ice nucleation. Also, there might be some compensation occurring in the regional model where the processes are all interdependent. The CIGMA model captures certain aspects of the ice cloud physics, but it does not provide a full picture of all cloud–ambient interactions.

The impact of perturbing the other ambient variables is very small, except for pristine ice variation due to perturbation of Π' which reaches a maximum of 1% toward cloud top. Another interesting feature is the larger sensitivity to vertical motion shown in the right middle panel, than for the horizontal motion shown in the bottom panels (note the change in scale in the horizontal axes). In particular, an increase in vertical motion increases supersaturation at upper levels, hence enhancing the amount of condensed ice. Nonlinearity of model dynamics and physics can also amplify small changes in forced vertical motion, contributing to the creation of supersaturation. This effect was not captured by the CIGMA system due to the

lack of a predictive equation for supersaturation. Radar reflectivity variation profiles (figure 5.3) also show a higher sensitivity to temperature, although the shape of the profiles is quite different. The impact is especially felt at the cloud boundaries where the values of the radar reflectivity are smaller and the sensitivity is higher.

Similar results for the snow mixing ratio and the relative radar reflectivity variation are shown in figures 5.4 and 5.5. The shape of these profiles appear very similar, although the absolute magnitudes of the variations are lower than for the corresponding variations in the pristine ice profile (note that same scale is used for the horizontal axes). The slightly large impact that a change in ambient variables has on pristine ice than on snow can also be interpreted as a sign that the balance at upper tropospheric levels, where pristine ice is preferentially concentrated, is more delicate, and more sensitive to perturbations in surrounding conditions, as well as being related to the nonlinearity of nucleation processes. The percent variation is, however, unchanged due to the lower average cloud snow mixing ratio (see figure 5.1).

Pristine ice and snow maximum and minimum gradients and absolute variations over the entire domain at cloud levels are shown in tables 5.5.1 and 5.5.1. Units for the gradients are: $\frac{kg/kg}{K}$ for the variable θ , $\frac{kg/kg}{kg/kg}$ for q , $\frac{kg/kg}{m/s}$ for the wind variables and $\frac{kg/kg}{J/(kgK)}$ for the Exner function. Units for $\Delta r_{pristine}$ and Δr_{snow} are kg/kg and for ΔZ_{dBZ} are dBZ (multiplicative factors take care of the conversion from radar reflectivity in mm^6m^{-3} to dBZ). The maximum positive and negative value of the gradients are indicated in the table with the symbols G^+ and G^- . Variations of 1% of environment parameters are considered with respect to average values at cloud levels. The results confirm what was shown by the average vertical profiles. A relevant feature discovered from the analysis of the results was that maximum (and minimum) values occurred preferentially at the boundary of the domain. This is a very promising result for the use of boundary conditions as control variables in the assimilation process. A similar behavior is also found in the RAMAS system for a cost function-like forcing as presented in the following sections.

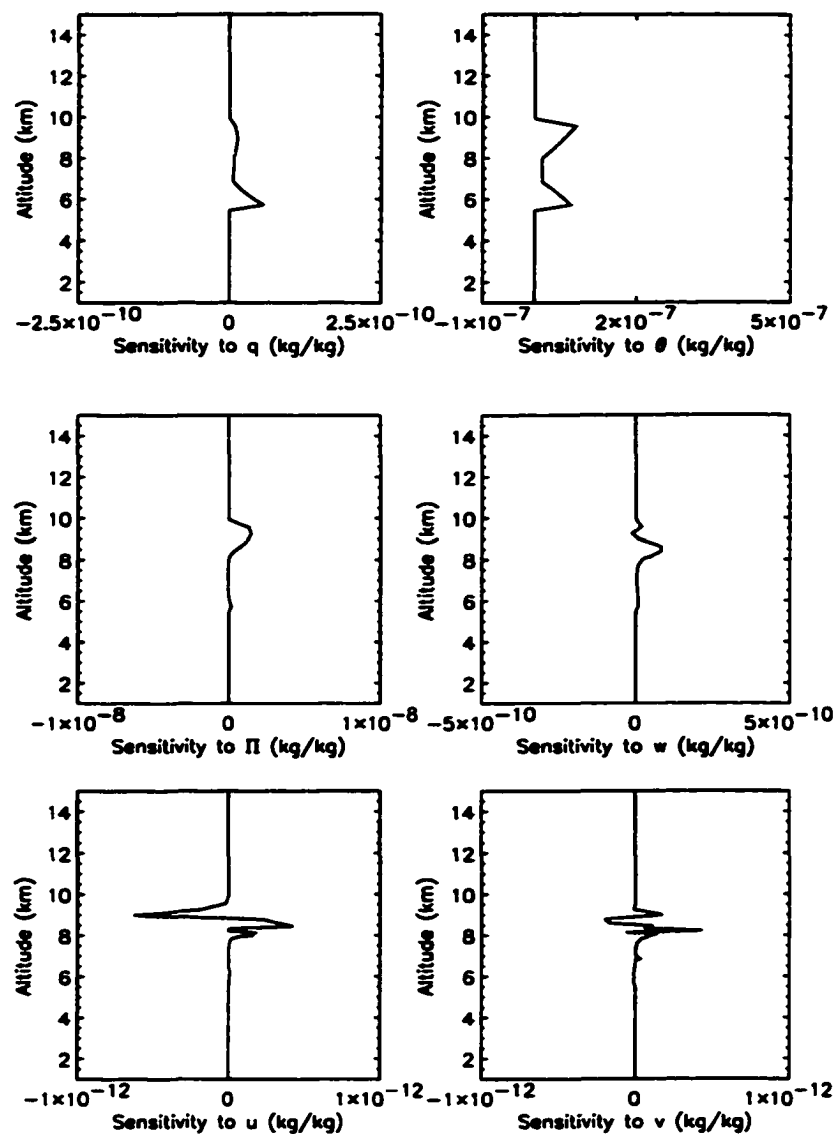


Figure 5.4: Same as figure 5.2, but for the snow mixing ratio.

5.6 System response to cloud synthetic data inclusion

Different synthetic datasets were created by perturbing the original model reflectivity. As discussed in the RAMS error analysis (chapter 3, section 3.5), model errors can be both systematic (biases) and random. The synthetic data sets were designed to account for different possible scenarios, and three cases were chosen:

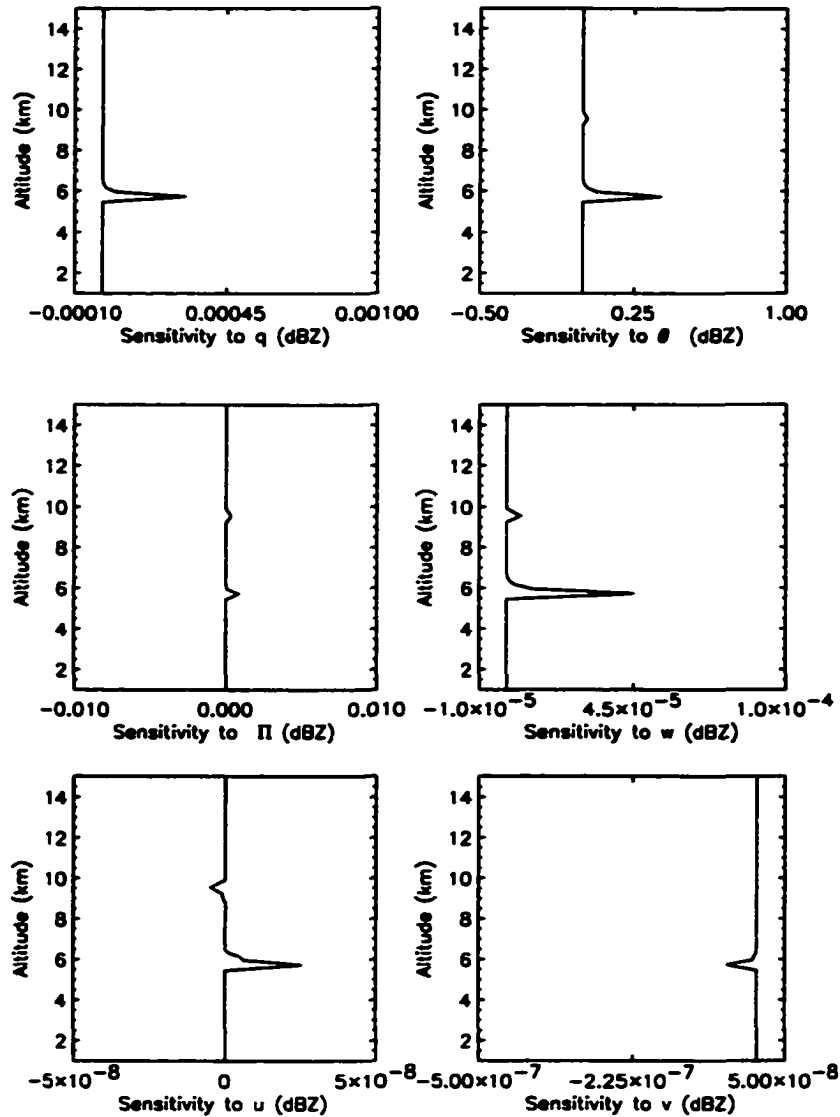


Figure 5.5: Same as figure 5.3, but for radar reflectivity variations induced by perturbations in the snow mixing ratio.

- (a) Pure random perturbation up to 50% of the local model reflectivity value on cloudy grid points only;
- (b) same as in (a) for cloudy grid points, but synthetic data are also defined for non-cloudy points (i.e. points with zero value for the radar reflectivity) as random perturbation up to 200% of an (arbitrary) reference value of IWC

Table 5.1: Maximum and minimum impact of 1% environmental parameter variation at cloud levels on pristine ice mixing ratio.

Variable	G^-	G^+	Δr_{ice}^-	Δr_{ice}^+	ΔZ_{dBZ}^-	ΔZ_{dBZ}^+
θ	0.000193	3.8e-05	0.00058	0.000114	0.39	0.077
q	0.108	0.022	1.08e-07	2.2e-08	7.3e-05	1.48e-05
u	0.0035	0.0034	3.5e-07	3.4e-07	0.0024	0.0023
v	0.0035	0.0034	3.5e-07	3.4e-07	0.0024	0.0023
w	0.00053	0.00011	5.3e-08	1.1e-08	3.58e-05	7.43e-06
Π'	0.0011	0.0013	2.14e-06	2.58e-06	0.0014	0.0017

Table 5.2: Same as table 5.5.1, but for snow mixing ratio.

Variable	G^-	G^+	Δr_{snow}^-	Δr_{snow}^+	ΔZ_{dBZ}^-	ΔZ_{dBZ}^+
θ	6.4e-05	1.6e-05	0.000192	4.8e-05	0.129	0.032
q	0.035	0.009	3.5e-08	9.0e-09	2.36e-05	6.07e-06
u	0.00011	0.00011	1.1e-08	1.1e-08	7.4e-06	7.4e-06
v	0.00011	0.00011	1.1e-07	1.1e-07	7.4e-05	7.4e-05
w	0.000176	4.5e-05	1.76e-08	4.5e-09	1.2e-05	3.0e-06
Π'	0.0035	0.0004	7.e-06	8.e-07	0.00472	0.00054

equal to 0.5 mg m^{-3} ;

(c) Same as in (b), but with added 20% bias.

The results from case (a) and case (b) were very similar, due to the fact that the cloud covered almost the whole study area. Results from case (b) and case (c) presented just slight differences in the magnitude of the gradients. For this reason, only results for case (c), which describes the most likely scenario are presented and discussed (in general model forecast might present a bias with respect to observations and definitely has a random error component).

The “synthetic” forcing is applied to the ice categories of pristine ice and snow. The forcing is derived from the cost function

$$J = \frac{1}{2}(Z - Z_{obs})^T \mathbf{W}^{-1}(Z - Z_{obs}), \quad (5.12)$$

and has the form

$$\frac{\partial J}{\partial r_{pice}} = \frac{\partial Z}{\partial r_{pice}} \mathbf{W}^{-1}(Z - Z_{obs}). \quad (5.13)$$

where Z represent the model-derived reflectivity, and Z_{obs} the “bogus” reflectivities that play the role of observations. \mathbf{W} is the observational and mapping error covariance matrix, already introduced and discussed in previous chapters. By using (5.11), equation (5.13) is explicitly rewritten as

$$\frac{\partial J}{\partial r_{pice}} = cd(IWC)^{(d-1)} \rho_a \mathbf{W}^{-1}(Z - Z_{obs}). \quad (5.14)$$

Similar equations are derived for the snow mixing ratio by simply substituting r_{pice} with r_{snow} .

Sensitivity to the magnitude of forcing, controlled by the magnitude of the matrix \mathbf{W} elements, was also investigated. Observational plus mapping errors were prescribed as a percentage of the bogus radar reflectivity value at every grid point. Results for errors equal to 25% (case c.1) and 50% (case c.2) are presented in the following sections. Comparison of case c.1 results is made between pristine ice and snow.

5.6.1 Experimental results for forced pristine ice cases

Case c.1: 25% observational+mapping error

Due to limitation in storage, and to the large amount of disk space required to save basic state fields at every timestep for the adjoint integration, the assimilation could only be performed for a one-hour time window. The forcing was hence applied at $t = 60$ minutes, and the adjoint integration was performed backward to time $t = 0$. The final time of the adjoint integration corresponds to initial time of nonlinear model integration. The adjoint solution at $t = 0$ corresponds to the negative gradient

of the cost function with respect to the particular control variable (see equation (2.25), chapter 2). If the 4DVar assimilation were to be carried out, this adjoint solution would be used to correct the initial guess on the control by the iterative procedure sketched in section 2.3.2. The sign of the correction is determined by the sign of the adjoint solution, since the minimization routine will follow the direction *opposite* to the gradient of the cost function to march toward the minimum. For instance, considering the initial condition on vertical motion as control variable and assuming that this initial condition is a state of rest, a *positive* adjoint vertical motion solution implies that a positive correction is applied to the initial vertical velocity field, and areas of ascending motions are created in the model domain as a result of the assimilation. When implementing the correction in a 4DVar operational setting, however, extra care is placed on filtering out, with the aid of penalty terms in the cost function, gravity waves that might be artificially excited by introducing an imbalance in the initial fields. However, this is more an issue in a hydrostatic model rather than in a nonhydrostatic model since in the latter the gravity waves that are excited are quickly shed and a new balance is achieved. The detailed description of how to operationally implement 4DVar assimilation in mesoscale models is beyond the scope of this chapter. However, the discussion about the sign of the adjoint solution corresponding to the sign of the correction can help interpret the results shown in the remainder of the chapter.

Figure 5.6 shows a west side three-dimensional plot of the adjoint pristine ice solution (representing minus the gradient of the cost function with respect to pristine ice) for experiment (c.1) at final time (upper panel) and fifteen minutes earlier (lower panel). A green cursor shows the relative vertical locations of adjoint fields. Figure 5.7 shows the same variable at $t = 30m$ and $t = 0m$. Note how the absolute magnitudes of the gradients decrease from final time to initial time. Light grey represents adjoint solution values smaller than $-3 \cdot 10^{07}$, whereas the orange isosurface represents adjoint solution values greater than $3 \cdot 10^{07}$: the random component of the forcing creates alternating areas of positive and negative perturbations.

Focusing our attention on the solution at $t = 0$ (lower panel of figure 5.7), let us consider how the correction would be implemented on the initial pristine ice field, if

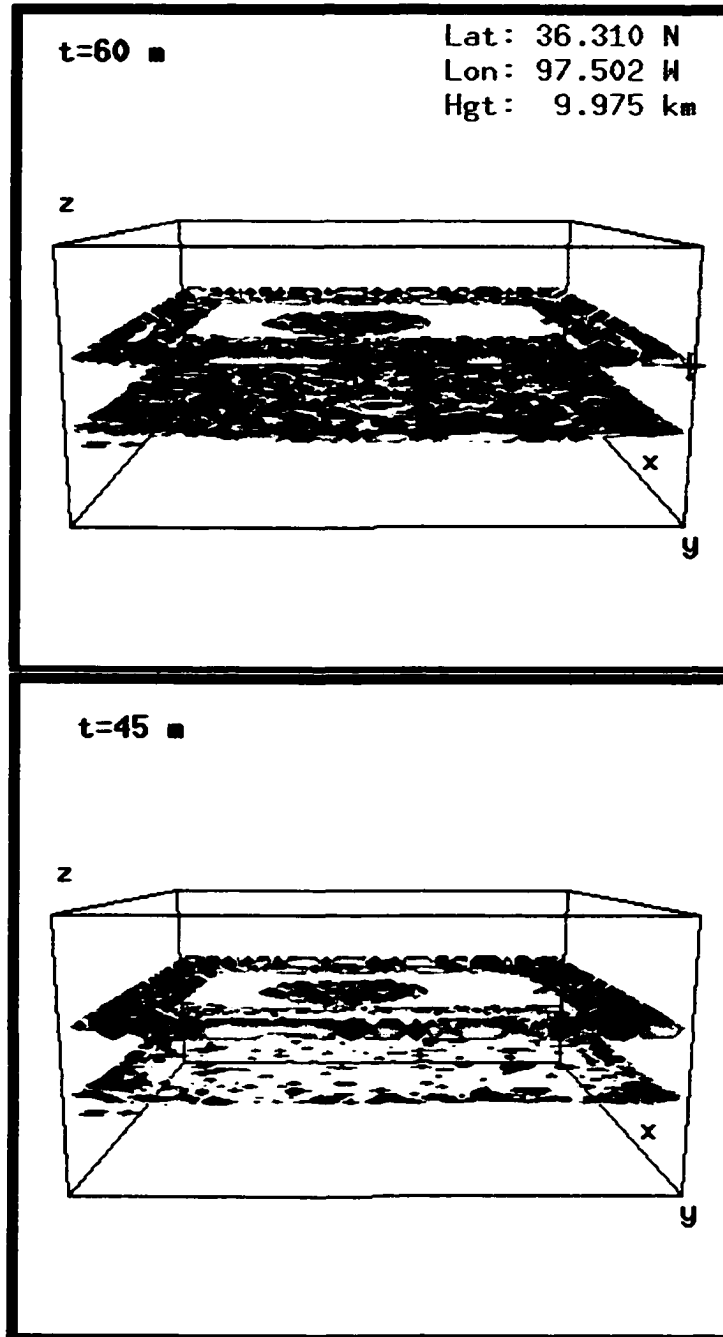


Figure 5.6: West view three-dimensional snapshots of positive (orange) and negative (light grey) adjoint pristine ice solution computed with RAMAS for case (c.1) forcing. Isosurface value is $3 \cdot 10^7$ for positive and $-3 \cdot 10^7$ for negative fields. Units are $\frac{1}{kg/kg}$.

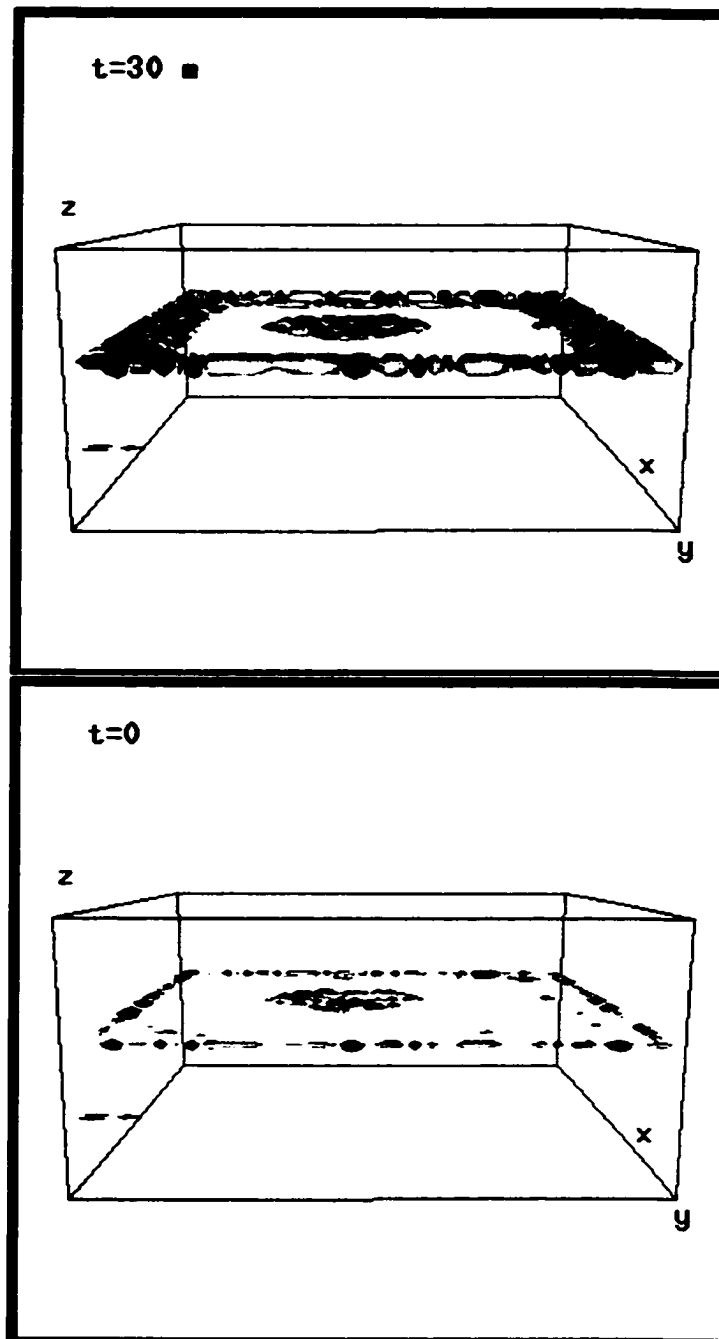


Figure 5.7: Same as figure 5.6, but for $t = 30m$ (upper panel) and $t = 0$ (lower panel).

the pristine ice mixing ratio were chosen as the control variable in the assimilation. The adjoint solution, although initially very large in magnitude, both toward cloud top and cloud base, and spread out over the whole domain area except for some “holes” at the initial time, is considerable only in sporadic areas around cloud top. The largest correction would hence be applied in selected regions of the domain and preferentially at upper cloud levels, where the forcing is mainly exercising its influence, and where the cloud is thicker (see figure 5.1). This finding confirms that radar data assimilation only corrects the cloud fields where the model originally had the largest amount of cloud, as already concluded in chapter 4 with the synthetic data assimilation using cloud variables as control variables.

The next set of figures (figures 5.8–5.12) shows the main result of this chapter which is the mesoscale model response in ambient conditions to cloud radar data assimilation. Figure 5.8 shows the gradient with respect to the humidity field. The impact of pristine ice forcing is shown both at cloud levels, and most importantly at levels away from cloud regions, especially at the domain boundaries and at upper tropospheric levels. This is a very interesting outcome, not only for optimal initialization of moisture fields, but, more generally, for the in-depth investigation of the relationship between high cirrus clouds and upper level humidity. Cirrus clouds are considered responsible for spatial redistribution of water vapor in the upper troposphere, both directly through particle growth at higher levels, which represents a sink of water vapor, and evaporation at lower levels, which represents a source of water vapor, and also indirectly through their radiative impact which affects the stability of the atmosphere creating water vapor gradients (Sherwood, 1999). The results presented here offer indirect evidence of this role played by high cirrus clouds.

Large response is also obtained with respect to potential temperature, as shown in figure 5.9. Note that the potential temperature correction is in phase with the specific humidity response, as already observed in the sensitivity study. Other thermodynamic fields such as the perturbation Exner function (figure 5.10) are also affected by radar data inclusion both at cloud levels and throughout the vertical domain, in particular at the boundaries.

Response in zonal wind (figure 5.11) is prevalently negative at lower levels and

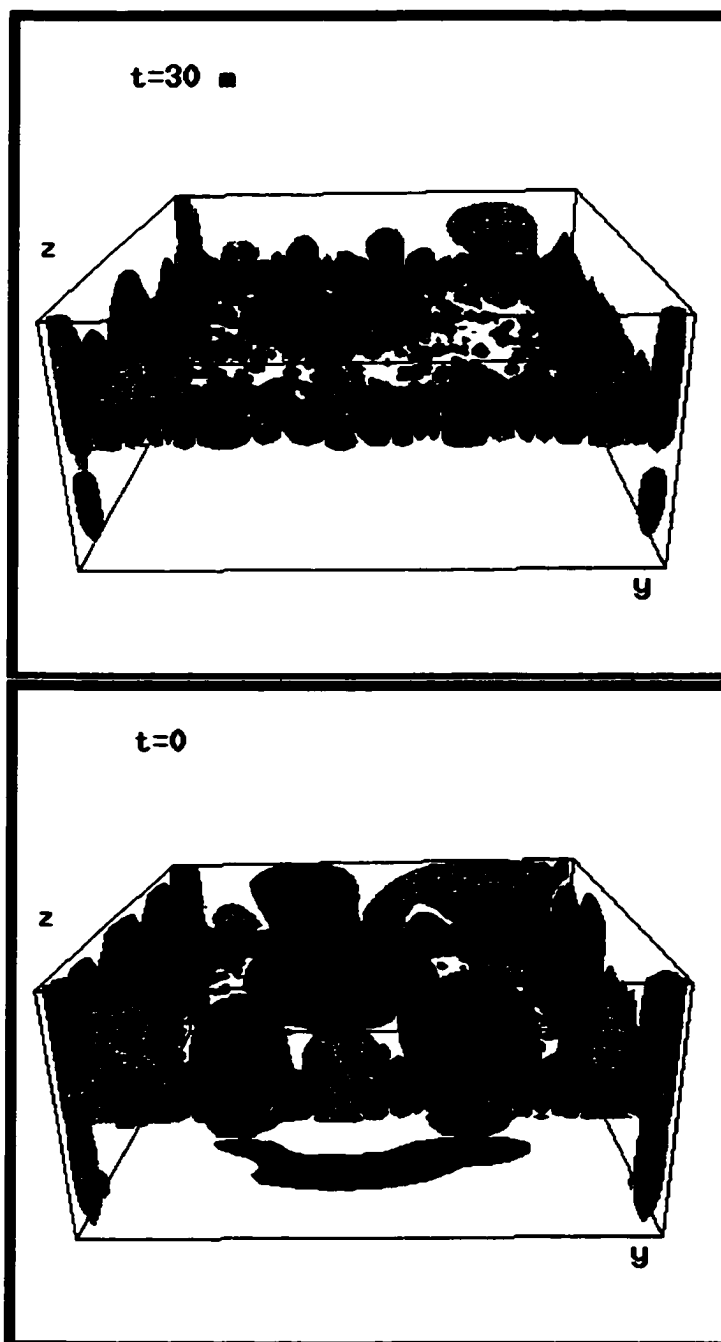


Figure 5.8: West view three-dimensional snapshots at $t = 30m$ (upper panel) and $t = 0$ (lower panel) of positive (dark yellow) and negative (green) specific humidity adjoint solution computed with RAMAS for case (c.1) forcing. Isosurface value is 10 for positive and -10 for negative fields. Units are $\frac{1}{kg/kg}$.

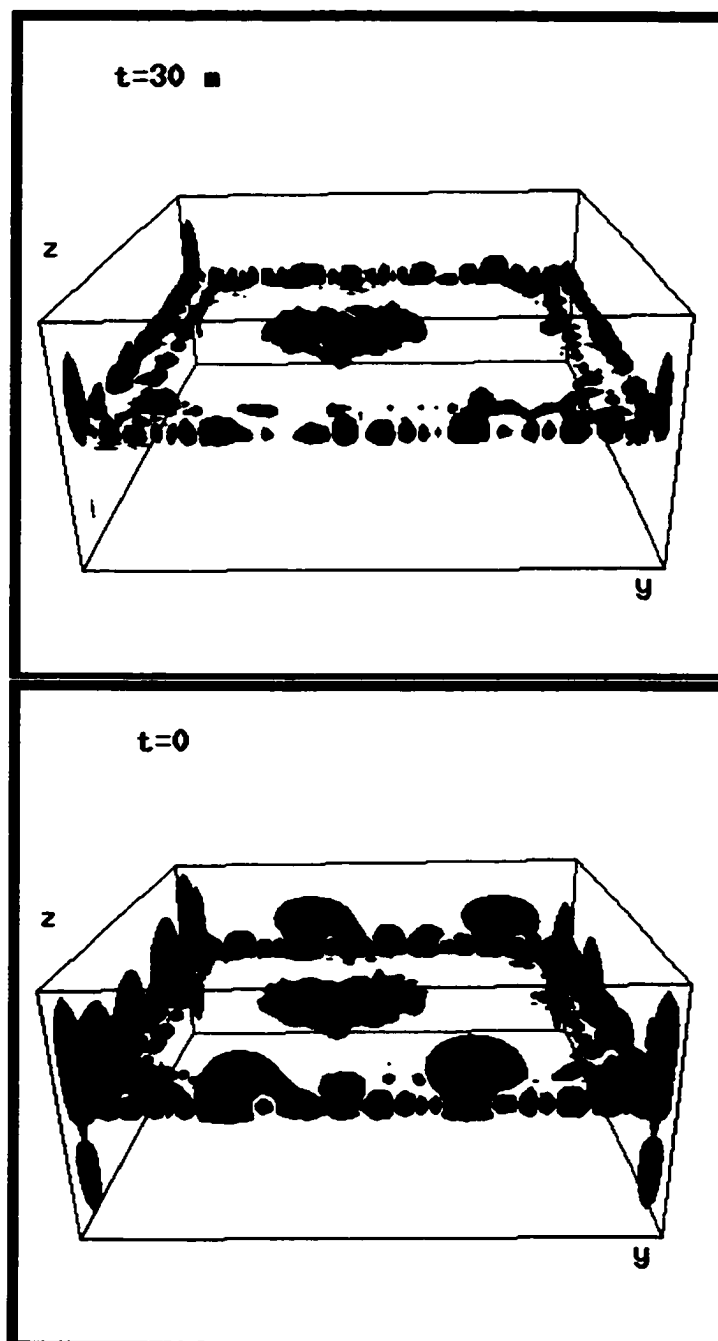


Figure 5.9: West view three-dimensional snapshots at $t = 30m$ (upper panel) and $t = 0$ (lower panel) of positive (red) and negative (blue) potential temperature adjoint solution computed with RAMAS for case (c.1) forcing. Isosurface value is 0.1 for positive and -0.1 for negative fields. Units are $\frac{1}{K}$.

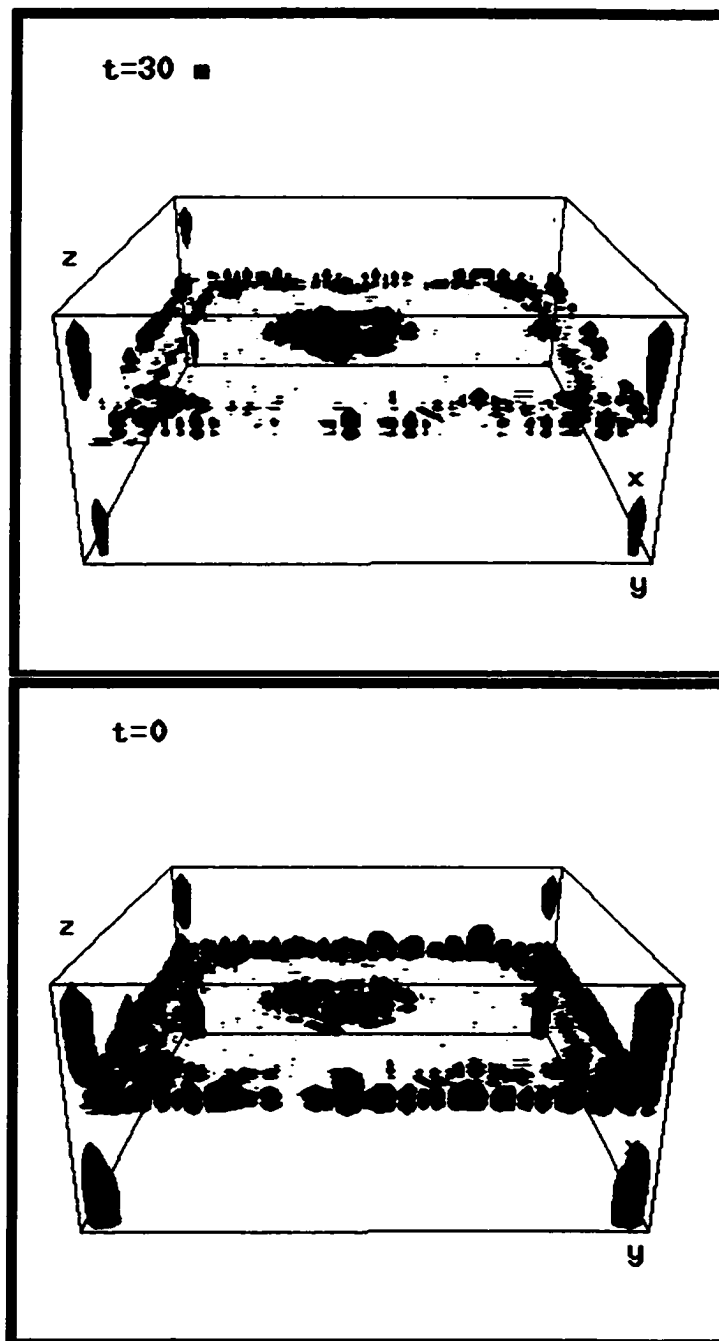


Figure 5.10: West view three-dimensional snapshots at $t = 30m$ (upper panel) and $t = 0$ (lower panel) of positive (light violet) and negative (light blue) perturbation Exner function adjoint solution computed with RAMAS (c.1) forcing. Isosurface value is 1 for positive and -1 for negative fields. Units are $\frac{1}{J/(kgK)}$.

positive at upper levels, whereas the response in vertical wind (figure 5.12) presents alternating positive and negative values, although positive perturbations are concentrated at the corners of the grid box. Assuming that the initial state is a state of rest (zero horizontal winds), this implies that the correction will create areas of moderate easterlies entering the lower east side of the grid box with associated westerlies aloft. For the vertical wind, areas of ascending motion will be preferentially created at model boundaries. Since dynamical processes are treated in the mesoscale model with physically consistent equations, dynamical responses will also be consistent amongst them. When implementing the correction, however, caveats mentioned earlier about the excitation of spurious gravity waves have to be taken into consideration.

Case c.2: 50% observational+mapping error

Increasing the amplitude of the errors, i.e. decreasing the forcing leads to smaller values for the gradients. This does not necessarily imply the impossibility of a successful data assimilation as seen in Chapter 4. Since the model is assumed perfect, however, the magnitude of observational and mapping errors does control how much the measurements are driving the model initialization. This result is evident in the plots of the cost function gradient for a higher error (50%) on the synthetic radar reflectivities. Gradients are generally smaller, indicating a lower impact of the assimilation on the variation of the cost function, and consequently a lower correction of control variables. For the sake of brevity, only results at the initial time for environmental variables are shown.

5.6.2 Experimental results for forced snow cases

The same cost function forcing implemented for pristine ice was applied to force snow mixing ratio and obtain adjoint solutions for snow ice fields and ambient parameters. Results are presented in the next section only for 25% magnitudes of errors (case c.1). Results for 50% magnitudes of errors (case c.2) are not shown, since increase in errors has the same impact as in the pristine ice case.

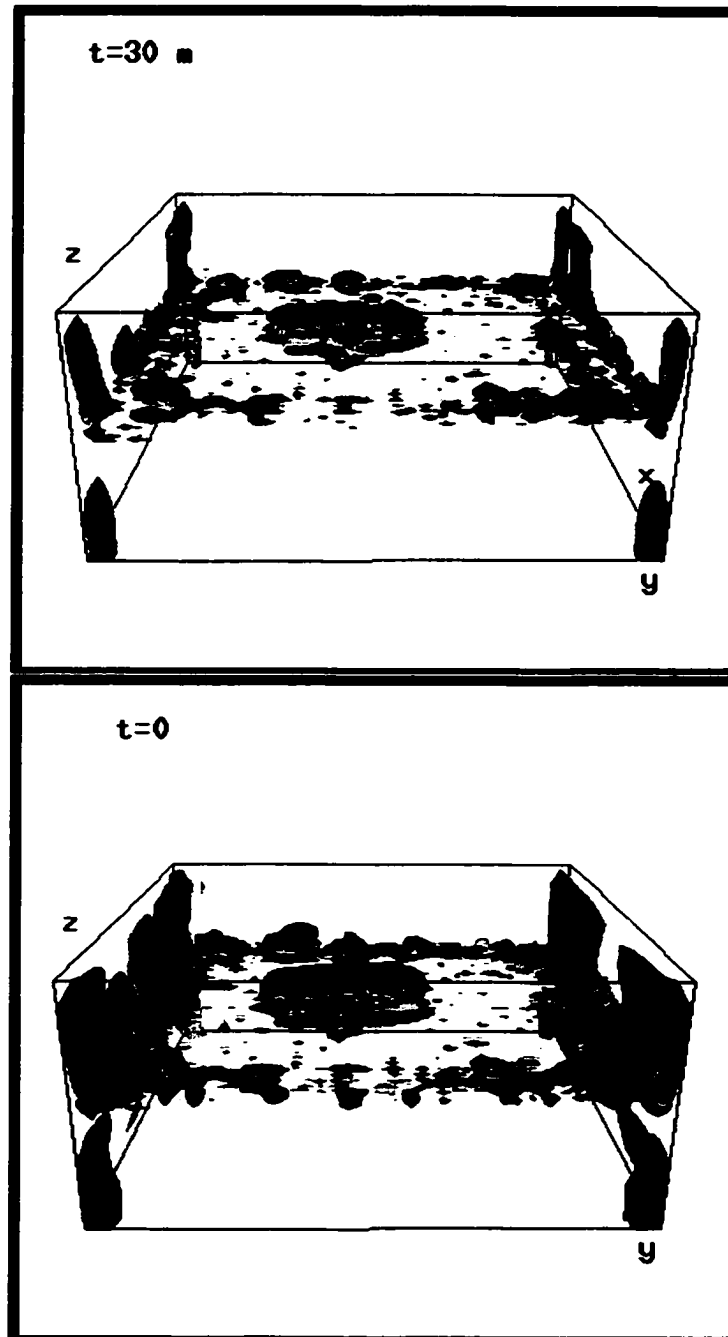


Figure 5.11: West view three-dimensional snapshots at $t = 30m$ (upper panel) and $t = 0$ (lower panel) of positive (light yellow) and negative (light green) zonal wind adjoint solution computed with RAMAS for case (c.1) forcing. Isosurface value is 0.1 for positive and -0.1 for negative fields. Units are $\frac{1}{m/s}$.

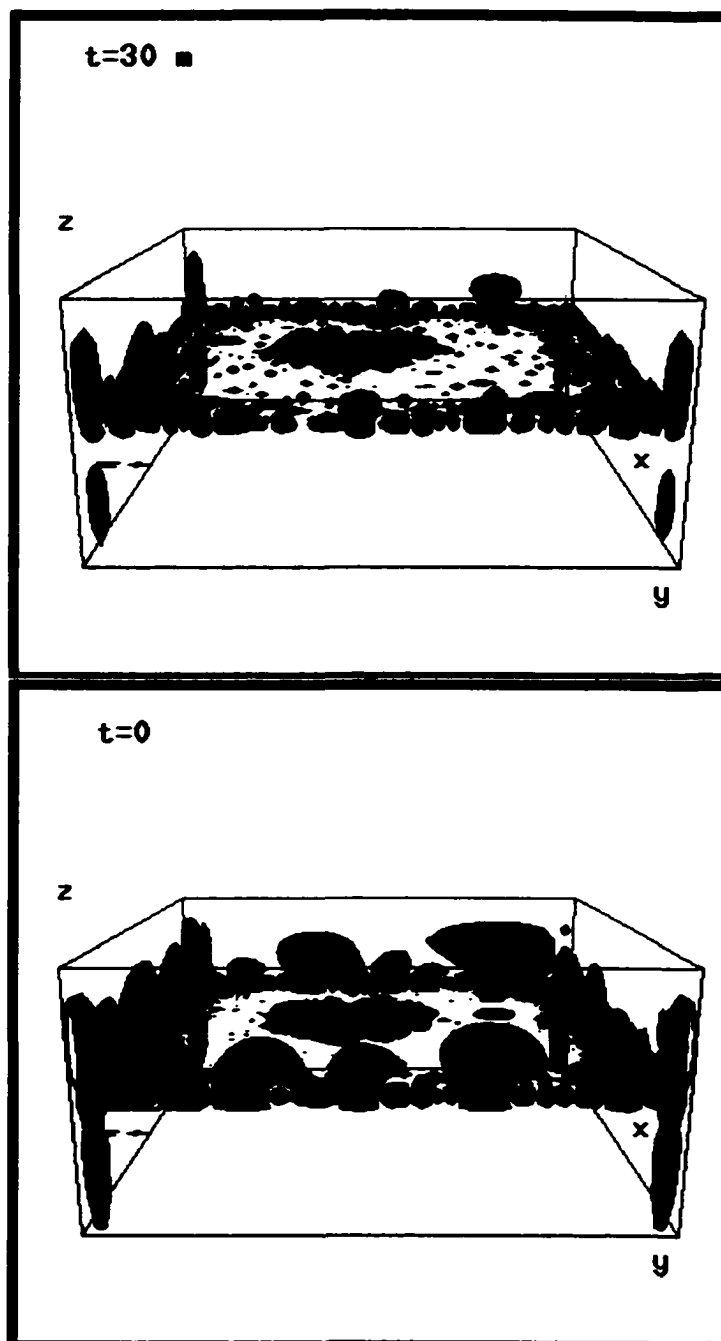


Figure 5.12: West view three-dimensional snapshots at $t = 30m$ (upper panel) and $t = 0$ (lower panel) of positive (light red) and negative (purple) vertical wind adjoint solution computed with RAMS adjoint for case (c.1) forcing. Isosurface value is 0.1 for positive and -0.1 for negative fields. Units are $\frac{1}{m/s}$.

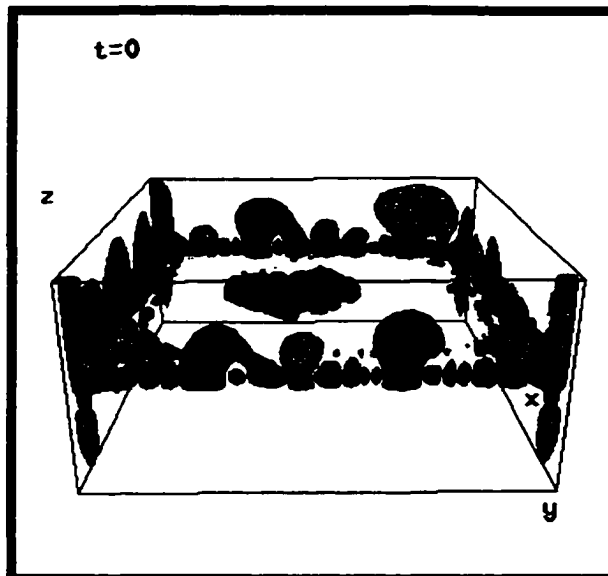


Figure 5.13: Same as figure 5.8 for the specific humidity adjoint solution at time $t = 0$, but for a 50% mapping+observation error (case c.2).

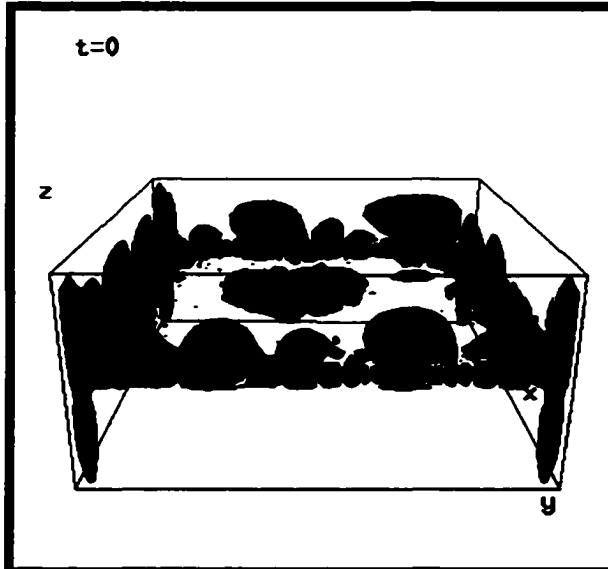


Figure 5.14: Same as figure 5.9 for potential temperature adjoint solution at time $t = 0$, but for a 50% mapping+observation error (case c.2).

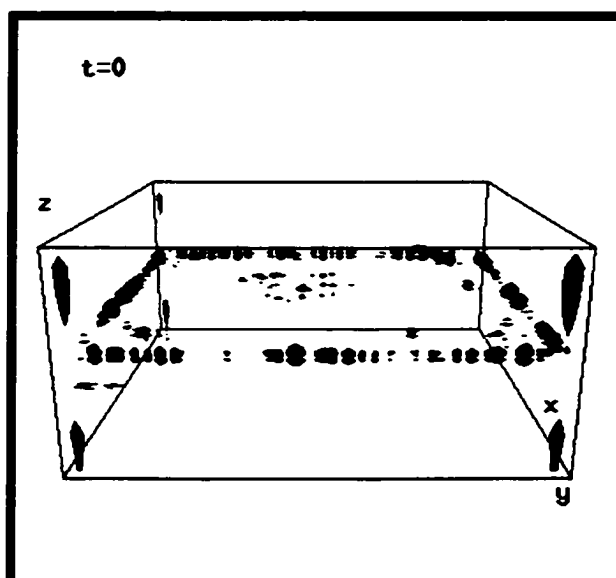


Figure 5.15: Same as figure 5.10 for perturbation Exner function adjoint solution at time $t = 0$, but for a 50% mapping+observation error (case c.2).

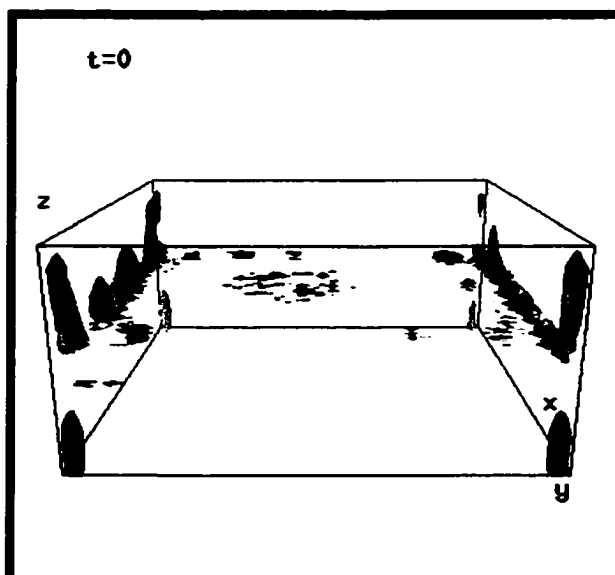


Figure 5.16: Same as figure 5.11 for zonal wind adjoint solution at time $t = 0$, but for a 50% mapping+observation error (case c.2).

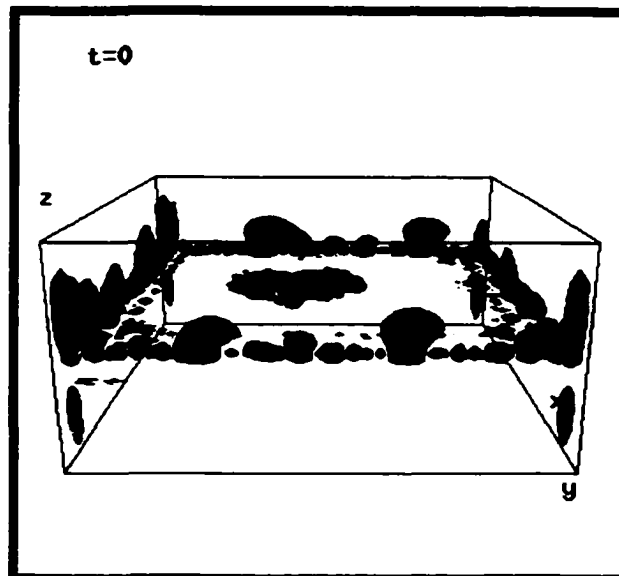


Figure 5.17: Same as figure 5.12 for vertical wind adjoint solution at time $t = 0$, but for a 50% mapping–observation error (case c.2).

Case c.1: 25% Observational+mapping error

Figures 5.18 and 5.19 show a times series of the cost function gradient with respect to snow mixing ratio. Isosurface contour levels are as in figure 5.6 to facilitate comparisons. Note that at final time (upper panel of figure 5.18) gradients are much larger at lower levels than at upper levels, indicating that the influence of snow forcing is mainly exerted at cloud base. The green cursor indicates the vertical position of the maximum sensitivity. In time, the gradient at lower levels tends to get much smaller, as illustrated in figure 5.19. In particular, at initial time only some sporadic areas can be seen, mainly at upper levels. This is not an indication that the assimilation of radar data would not work if the snow mixing ratio were to be used as a control variable, since the gradient is nonzero, but simply smaller than in the corresponding pristine ice case. Nonetheless, it raises a question regarding the frequency of data ingestion. Although, this question is not explicitly addressed in this study, the figures show that the impact of the assimilation of cloud data when the cloud variables are used as control decreases over time, i.e. the longer

the assimilation interval, the smaller the correction to the initial conditions. By extrapolation, if the interval is too long, there might be very little or no impact at all on the cloud variables which reiterates the conclusion already discussed in chapter 4 that cloud variables might not be the best candidates to be used as control variables. As already shown in the pristine ice cases, the situation for the ambient atmospheric variables is completely different. In fact, the magnitude of the gradient does increase going backward in time, and the highest influence is felt at the initial time. This is illustrated by figures 5.20–5.24. Another remarkable result is the fact that the gradient is also nonzero outside of the cloudy levels which reinforces what was found in the pristine ice cases, and offers a proof of the viability of ice cloud radar data assimilation to improve the model representation of the whole atmospheric state.

5.6.3 Comparative discussion of pristine ice and snow results

As already mentioned in section 5.5, the pristine ice forcing appears to induce a larger response in the ambient atmospheric variables. This can be due simply to the fact that the pristine ice contribution to the total ice water content is larger than the snow contribution, as shown in figure 5.1, leading to the higher impact of forcing the pristine ice adjoint solution. However, this could be related to the fact that pristine ice is more directly connected to the water vapor, temperature and vertical motion through nucleation and depositional growth processes. Upper tropospheric moisture balances might be more sensitive to radar data input due to the fact that the specific humidity fields at those levels are rather small, where the impact of observations is felt the most. However, when using real observations, a portion of the pristine ice crystals might be missed, thus reducing the impact of pristine ice forcing, due to the fact that radar reflectivity is more sensitive to large crystals than small crystals. Comparison between the two type of forcings (for pristine ice and snow) will have to be repeated with real measurements to ascertain whether the results are consistent with what was found using the synthetic data.

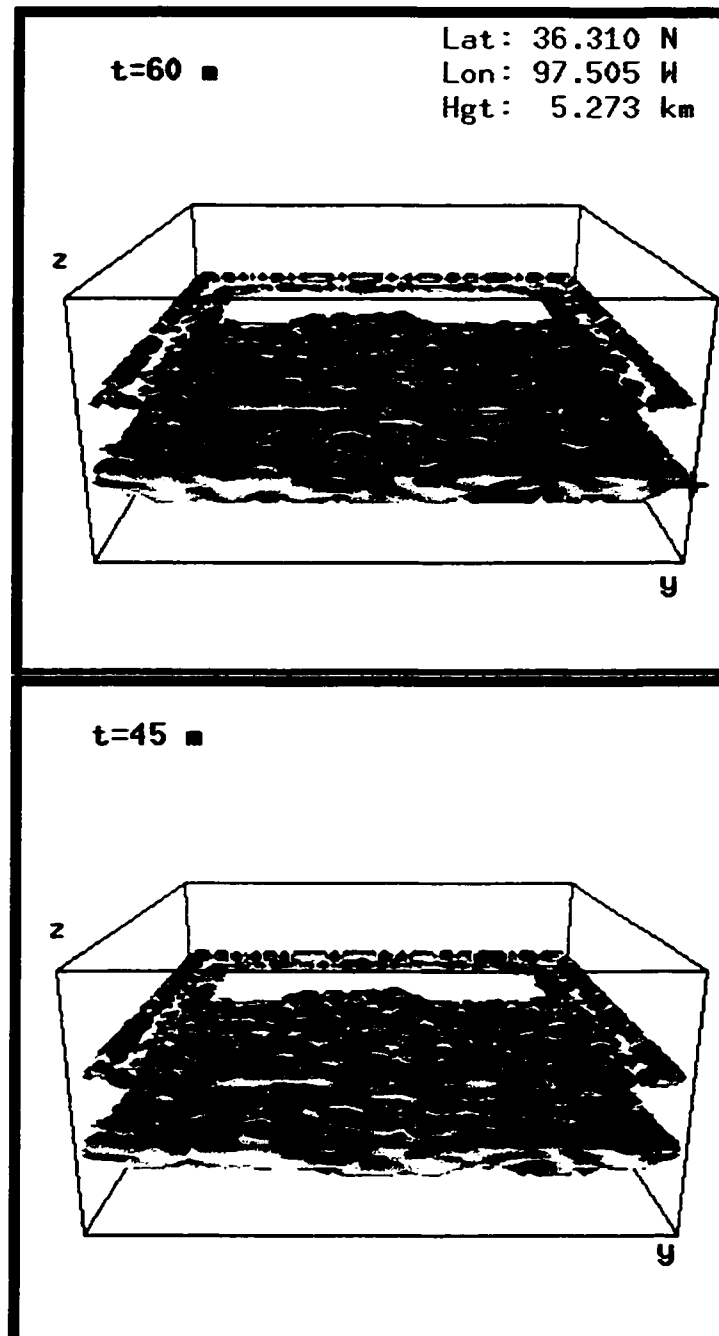


Figure 5.18: West view three-dimensional snapshots of positive (pink) and negative (very light yellow) snow adjoint solution computed with RAMAS for case (c.2) forcing. Isosurface value is $3 \cdot 10^7$ for positive and $-3 \cdot 10^7$ for negative fields. Units are $\frac{1}{\text{kg/kg}}$.

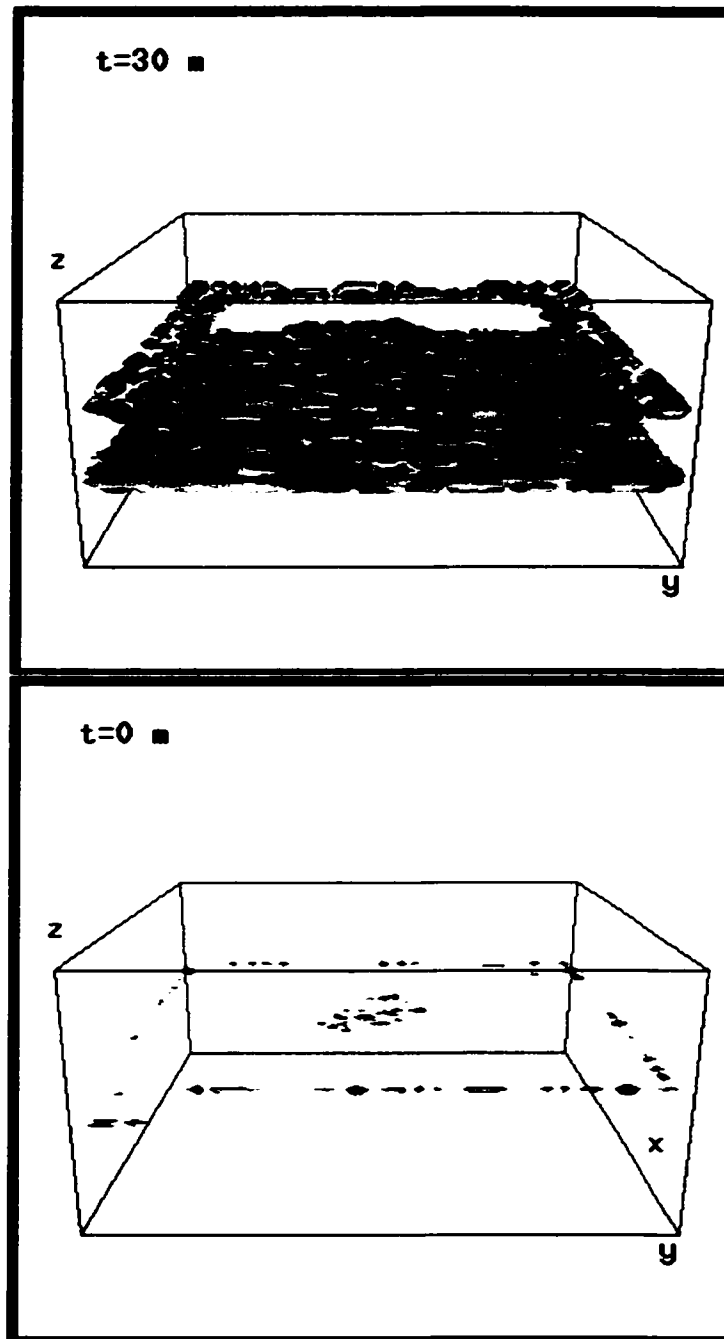


Figure 5.19: Same as figure 5.18, but for $t = 30m$ (upper panel) and $t = 0$ (lower panel).

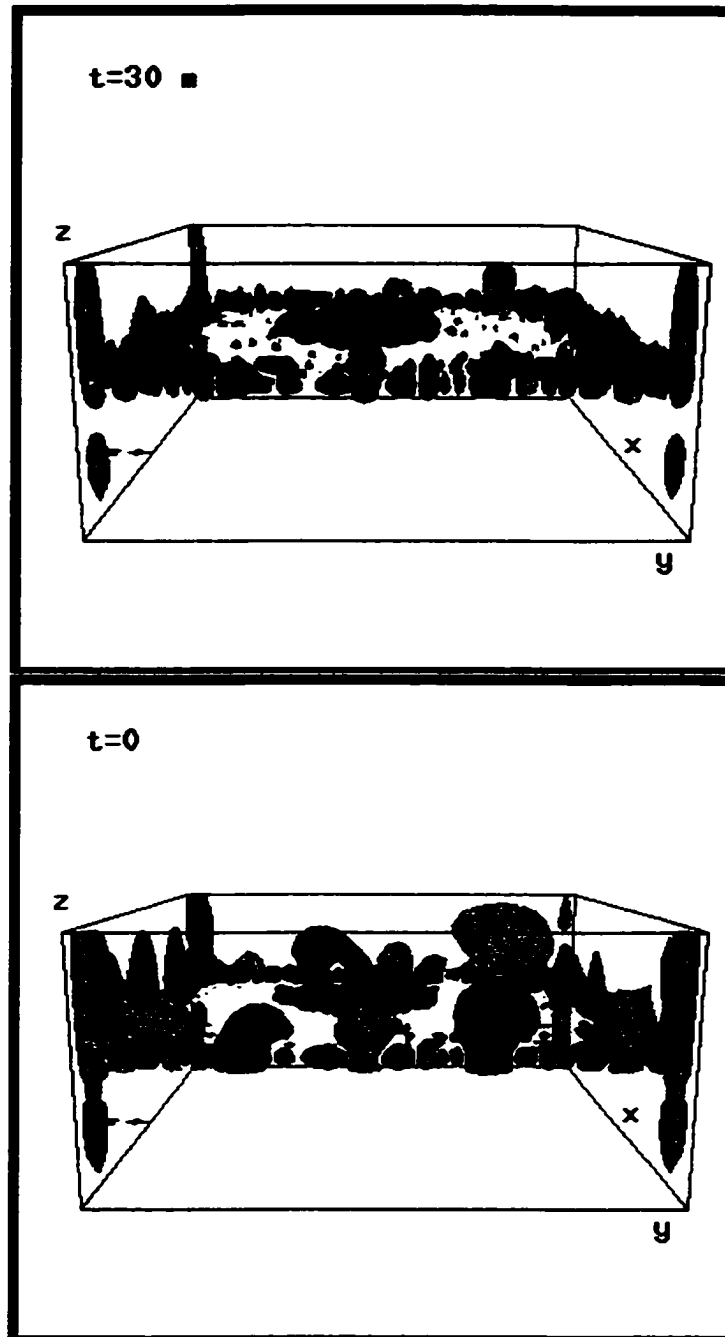


Figure 5.20: Same as figure 5.8 for the specific humidity adjoint solution, but for snow mixing ratio as the forcing variable.

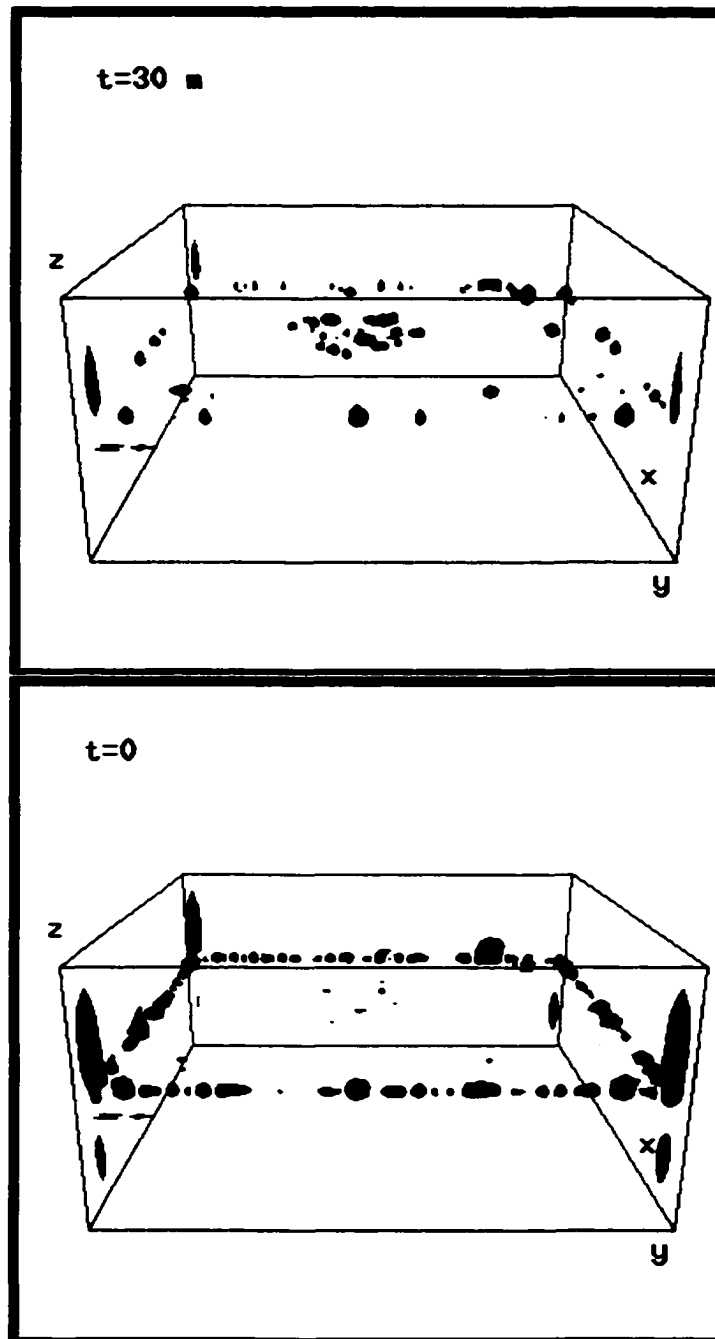


Figure 5.21: Same as figure 5.9 for the potential temperature adjoint solution, but for snow mixing ratio as the forcing variable.

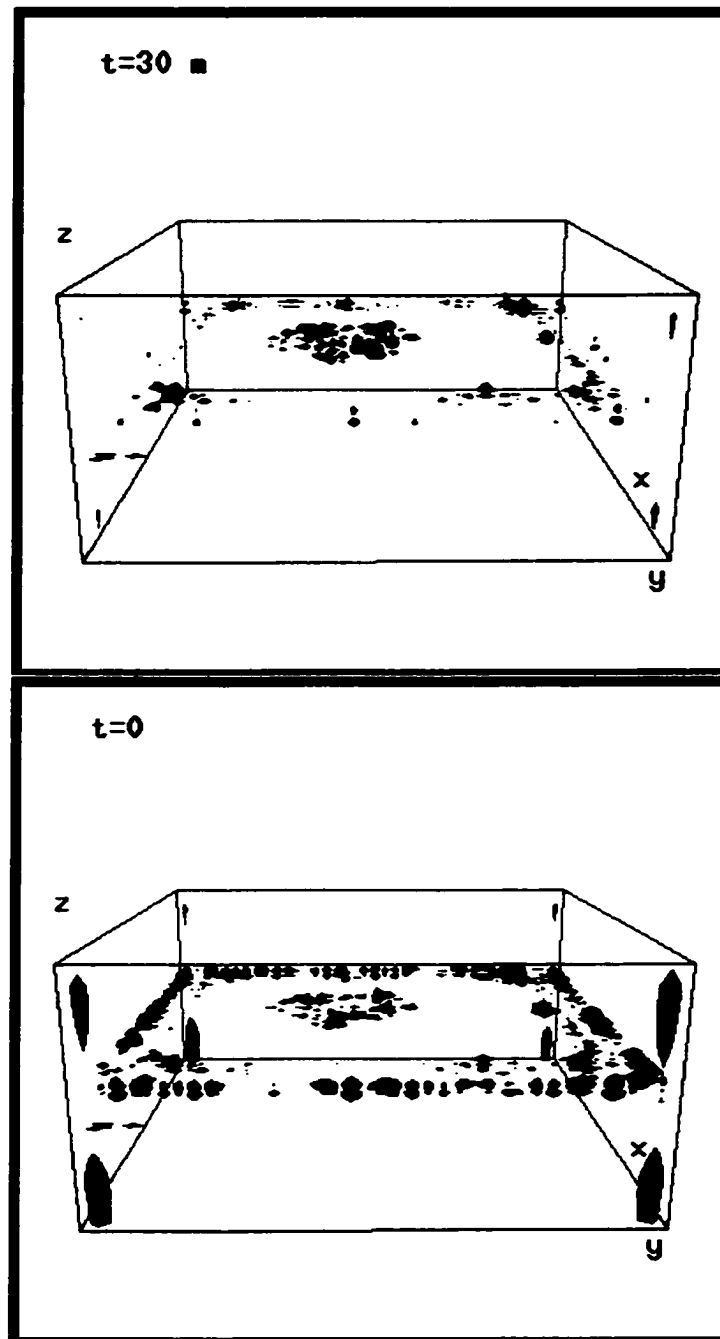


Figure 5.22: Same as figure 5.10 for the perturbation Exner function adjoint solution, but for snow mixing ratio as the forcing variable.

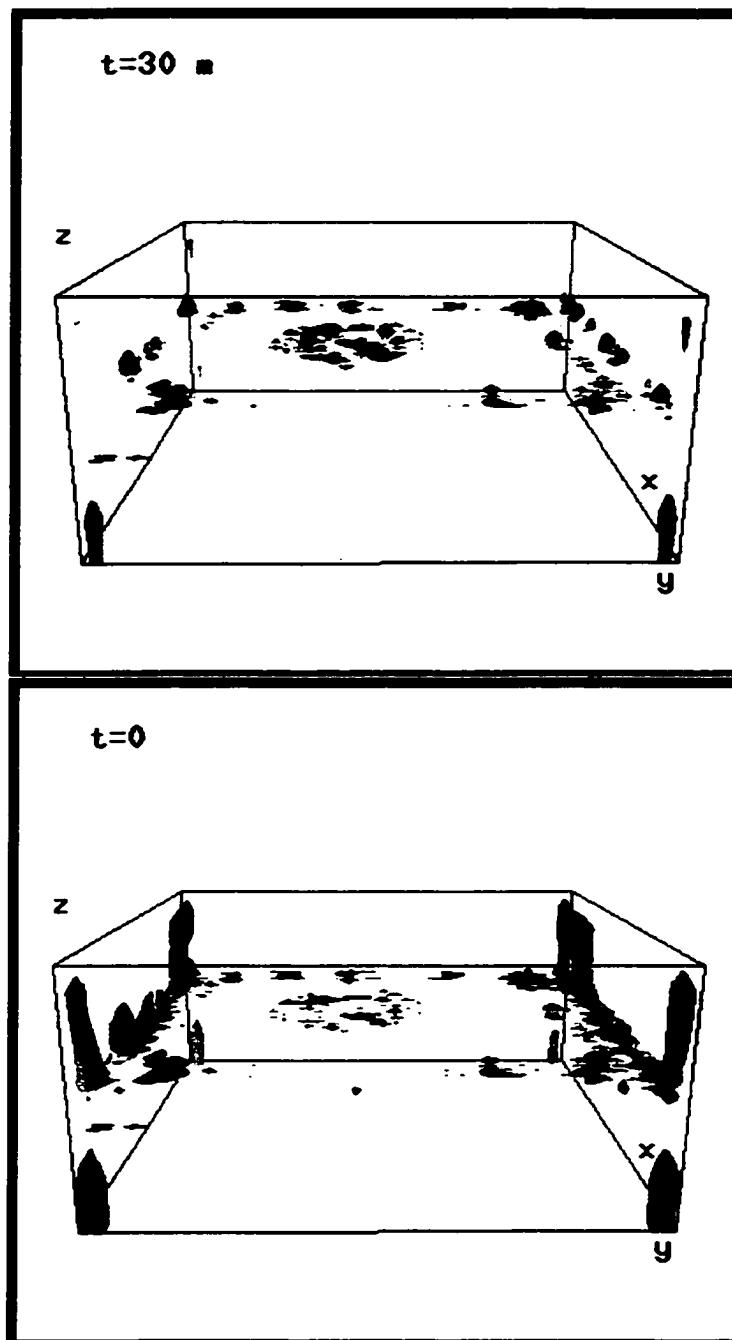


Figure 5.23: Same as figure 5.11 for the zonal wind adjoint solution, but for snow mixing ratio as the forcing variable.

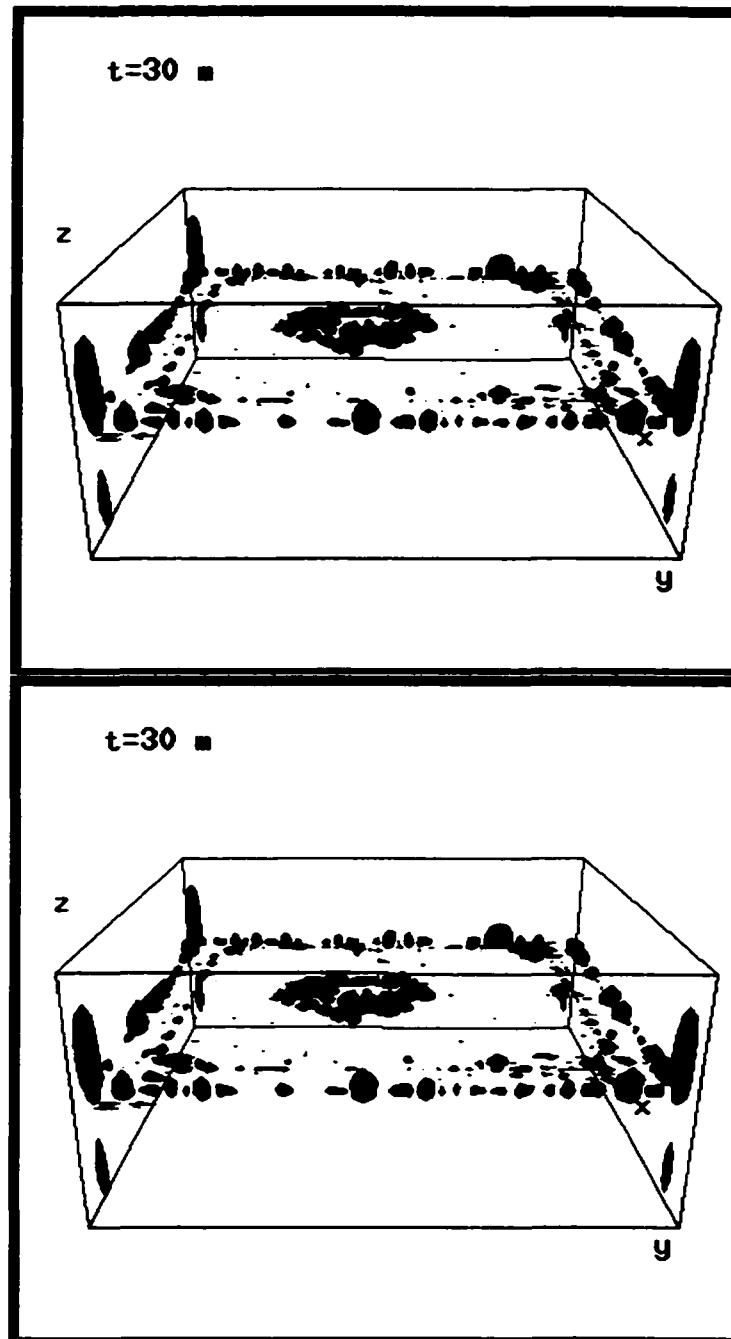


Figure 5.24: Same as figure 5.12 for the vertical wind adjoint solution, but for snow mixing ratio as the forcing variable.

5.7 Conclusion to the 4DVar prelude

Very promising results for the assimilation of cloud radar data into a regional model were obtained with the RAMS adjoint system (RAMAS). Sensitivity studies showed that temperature is the variable with the largest impact on ice amount, especially if the forcing is applied to upper tropospheric smaller crystals (pristine ice). These results confirmed what was found in the sensitivity assessment performed with the CIGMA adjoint. The system response to the inclusion of cloud radar data showed not only a large response in the cloud ice amount and distribution, but also to other model variables responsible for cloud formation and persistence. These results indicate that if any of these variables were to be used as control variables in the optimization procedure, the initial and boundary conditions on these environmental fields would be improved so as to obtain a cloud prediction closer to reality. Since temperature and specific humidity are the variables which are commonly used as the control variables in NWP variational data assimilation systems, these results imply that the assimilation of cloud radar data would be beneficial to advance NWP cloud prediction. Improvement in cloud forecasting is bound to improve the characterization of the other components of the hydrological cycle such as evaporation and precipitation, which are related to clouds, as well as characterization of other environmental conditions, as the results of this chapter show. These conclusions can be extended to all other cloud types detected with active and passive sensors.

In addition to this important, but somewhat obvious outcome of the cloud radar data assimilation, another fundamental implication is the possibility to extract the highest amount of information from the data by using the model as a constraint. As pointed out in chapter 4, cloud radar data only give an indication of the vertical distribution of ice amount. With additional information such as a second radar frequency or other type of measurements, other cloud-related quantities such as particle size and number concentration can be derived as shown in the retrieval presented in appendix A. However, when radar data are optimally used in conjunction with a full mesoscale cloud model, far more information can be extracted from those measurements. Although they do not directly contain any indication of the temper-

ature or specific humidity fields, optimal values of the initial temperature or specific humidity fields can be extracted via the variational assimilation, as shown in the previous sections. Links between cloud variables and dynamical and thermodynamical variables can be explored in the context of the variational system, and specific model aspects can be improved, guided by the assimilation results. In this regard, variational data assimilation is much more than optimal model initialization, and becomes a powerful modeling tool itself.

Chapter 6

Summary and concluding remarks

This research addresses some specific issues related to the optimal combined use of cloud models and cloud radar observations toward improvements in modeling and predicting cirrus clouds, and more generally toward improvements in numerical weather forecasting. While the answer to these questions is not univocal, we believe that the methods suggested here have great esthetic value and the results yield ample scientific relevance.

6.1 Summary of results

A major component of this dissertation was the development of a variational assimilation system. A detailed overview of variational principles was provided in chapter 2 to familiarize the reader with the notation and terminology specific to variational assimilation. Comparisons with other data assimilation techniques were also briefly discussed. All components of a variational assimilation system were described, in particular the *adjoint* model, which is the core of any variational algorithm.

In chapter 3, the RAMS model skill to simulate cirrus cloud formation and evolution was investigated with novel techniques. This investigation was instrumental in understanding key processes in cirrus modeling and identifying major types of deficiencies that could be addressed with a more effective use of observations. The first set of experiments were conducted using an idealized cirrus test case. This simplified scenario, also used in the context of the modeling activity of the GCSS–WG2, allowed a better understanding of the physical processes involved in cirrus formation and maintenance. RAMS results for the idealized cirrus test cases were compared with other cirrus models involved in the GCSS–WG2 activity. All models ran the

same cases, with the same initial conditions and forcing. The comparisons showed a large spread in the results for IWP and cloud geometrical thickness, as well as for dynamical and radiative quantities. Differences in the way the various models treated formation, maintenance and dissipation of the cirrus cloud were marked, and this was manifested in the intercomparisons. The main conclusion from this activity was that the way microphysical processes are represented and linked to thermodynamical, dynamical and radiative processes is a key factor in the accurate representation of ice cloud systems.

Sensitivity of the RAMS model to microphysical parameterizations was also investigated, and it was found that the model response in cloud IWP and geometrical thickness was sensitive to the fall velocity parameterization. The model was also strongly affected by the choice of bulk ice categories included in the simulation. RAMS was then used to simulate an observed cirrus case. Radar observations of subtropical cirrus clouds from the ARM-UAV Spring 1999 experiment (Stephens et al., 2000a) were used in the evaluation of the model performance. For the first set of experiments, the model was run in a 2D configuration with a horizontally homogeneous initial condition and a periodic lateral boundary condition. Results showed that cloud formation did not occur without the implementation of large-scale forcing in the form of a prescribed cooling at cloud levels, corresponding to a gentle large-scale uplift of 3 cm s^{-1} . The model cloud did not reproduce accurately the observed cloud, with the exception of gross features such as horizontally-averaged IWC. Discrepancies were due to the three-dimensional channel configuration and the lack of more realistic large-scale forcing. The model was subsequently run in a 3D configuration, with horizontally non-homogeneous initial conditions; a nudging technique was used to assimilate model data provided by the ECMWF forecast over the experiment area. A tuning of the nudging strength was necessary in order to obtain good agreement with observations, although the fine structure of the cloud was not well represented in the model simulations due to the coarse resolution. Formation and persistence of the cirrus cloud system were directly related via the nudging to large-scale fields. A model error analysis was performed for these 3D simulations using a statistical approach. Model biases for the chosen configuration

were identified and error covariance matrices for the domain average radar reflectivity and IWC fields were computed. Despite the tuning of the nudging coefficients, it was found that the model exhibited a slight wet bias (i.e. over predicted IWC) throughout the cloud, except at cloud base where the radar reflectivity was lower than the observed value. The main results of this chapter were the assessment of the regional model skill through both the investigation of the model response to change in physical parameterizations and the application of a statistical analysis to understand model biases and errors.

Chapter 4 introduced a simplified ice cloud model for radar data assimilation experiments. This model is an original extension of the work of Mitchell (1988) and is based on conservation equations for the first and second moments of the mass size distribution. It predicts the time and vertical evolution of the parameters of a modified gamma size distribution describing a precipitating ice cloud in which vapor deposition, aggregation and sedimentation processes are active. The inputs are ambient variable profiles (temperature, pressure and specific humidity). Comparison of model fields with ground-based radar observations and in situ microphysical observations from various field campaigns (ARM-UAV Spring 1999 in Kauai, and Spring 1998 and Fall 1997 IOPs at the ARM SGP site) were conducted to evaluate model performance. The ice growth model showed a good degree of realism when operating under ambient conditions close to water supersaturation. The adjoint of the cloud model was hence derived and used to investigate the model sensitivity to various inputs and the feasibility of 1–1D variational assimilation of radar reflectivities. Sensitivity studies showed a moderate model response to specific humidity and a large response to temperature variations. The latter was related to the choice of parameterization used for cloud variable initialization. In particular, initial particle number concentration, derived from an exponential function of supersaturation was highly dependent on temperature. This result highlighted the importance of key model parameterizations in determining model response to changes in ambient variables. A low sensitivity was displayed by the model to changes in vertical motion. The direct impact of vertical velocity on model radar reflectivity was unrealistically negligible. We believe that this was due to the lack of a prognostic equation for

supersaturation, hence the lack of an indirect effect of increasing vertical velocity on specific humidity and ice amount, as well as to the lack of turbulence and radiation interactions.

Variations of model output due to perturbations in input parameters such as the coefficients that define the fall speed–diameter and the mass–diameter relationship were also investigated. Results show that certain aspects of the modeled ice cloud depend strongly on the parameters selected, which in turn depend on assumptions regarding the shape of the crystals that form the cloud. For example, cloud base position was strongly controlled by the choice of the fall speed coefficients.

Variational cloud radar data assimilation experiments with the the cloud model and its adjoint were performed. An iterative optimization procedure, based on a gradient method, was implemented to recover optimal initial conditions for the cloud model that minimized the weighted difference of modeled and observed radar reflectivities (cost function). First tests were performed with synthetic data and two different sets of control variables: cloud variables and specific humidity. Initial profiles of the control variables were perturbed by a given percentage, and the optimization algorithm was started with these perturbed initial conditions. It was found that the algorithm was able to converge to the true initial conditions in most cases within a reasonable number of iterations. However, when the cloud variables were used as control variables, only initial condition for the characteristic diameter was recovered, while the number concentration was hardly improved at all by the inclusion of radar reflectivities, due to the low sensitivity of the radar backscattered signal to the number concentration as opposed to the large sensitivity to crystal size. It was concluded that a different choice of cloud control variable, such as IWC could be more adequate. Better results were obtained when specific humidity was used as a control variable since the initial number concentration was related to the available supersaturation. An improvement in specific humidity profiles implied as well an improved initial particle number concentration profile, indicating that the use of ambient variables as control variables is more effective. Another important finding of experiments with synthetic data was the fact that, even with optimal initial conditions, the model was not capable of creating a cloud where there was

no cloud. When the perturbation of the initial condition was accompanied by a vertical misplacement of the initial profiles, the convergence to the true profile was worse. This represents a limitation of the assimilation system that can be ascribed to the lack of prognostic equations for ambient variables in the model. In general, it appeared that the assimilation was more effective in recovering true profiles when the perturbation was positive rather than when perturbation was negative, implying that inclusion of data was more effectual in reducing an excess of cloud rather than increasing a deficit. Experiments with synthetic data also included the investigation of the sensitivity to assumptions on observational and mapping error magnitude, which showed that even for large uncertainties in the “bogus” reflectivity the assimilation was still successful in recovering the true initial profile; the inclusion of additional information provided by the cloud visible optical depth, which did not bring any substantial improvement to the assimilation result; and the inclusion of a background term, which showed how making correct error assumptions is crucial in variational data assimilation in order to obtain a meaningful optimal solution. The assimilation system was also tested for optimal model parameter estimation, focusing on one of the fall velocity parameters. The variational system was successful in recovering the true initial value of the parameter from a perturbation as high as 50% by using measurements—the radar reflectivities—that are not directly sensitive to particle fall speed. This result opens the door to new applications of radar reflectivity measurements, not accessible without the use of a cloud model in conjunction with the radar mapping. Finally, real radar observations were used to perform further assimilation experiments. Assimilation of radar data was successful and allowed the retrieval of optimal initial conditions which improved model prediction. This original result demonstrated both the feasibility and the benefits of cloud radar data assimilation into cloud models.

Implications of radar data assimilation were further explored in chapter 5 where RAMS and its adjoint (RAMAS) were used in preliminary 4DVar experiments. The application of a full-blown mesoscale model allowed for the investigation of the impact of ice cloud radar data on the three-dimensional mesoscale environment. Results from this chapter are highly experimental and novel since it was the

first time that the RAMAS software was ever used for assimilation of cirrus clouds (and the first time as well that ice clouds were the subject of an assimilation exercise). The only previous study on cloud assimilation with RAMAS focused on lower level water clouds observed from satellite-based passive sensors (Vukićević et al., 2001). Experiments included sensitivity studies—which are omnipresent throughout this work—conducted with the RAMS adjoint. Variations in ice amount induced by perturbations in ambient fields were quantified and discussed. Results partially confirmed what was found in chapter 4. However, the magnitude of the sensitivity was smaller due to the action of compensating processes acting in the complex mesoscale model. Preliminary variational experiments with the use of synthetic data reflectivities were also conducted. Results demonstrated that inclusion of ice cloud radar data impacts significantly all ambient variables from potential temperature to specific humidity and wind. The influence was more important the more direct the link between cloud ice amount, on which radar reflectivity depends, and the specific ambient variable under consideration. For this reason the impact on temperature variations was larger than the impact on horizontal wind variations. The influence of data inclusion was felt not only at cloud levels, but also outside of the cloud region, especially at the boundary of the domain. These results are relevant for future assimilation of cloud radar data into weather forecasting models.

6.2 Relevance of this research to weather forecasting

The importance of data assimilation for numerical weather forecasting is well known. Standard meteorological variables are routinely assimilated into the major NWP models such as the ECMWF and NCEP models. Recently, research has focused on improving key aspects of the forecast models that had been previously qualified as second order effects, such as moist processes involving clouds. Effort has been put into both advancing the representation of the links between clouds and precipitation at convective scales (cumulus parameterizations), and representation of large-scale clouds such as low-level stratus and high-level cirrus clouds, that are not directly related to precipitation, but considered important in the atmospheric hydrological

cycle, in the atmospheric general circulation and in the global energy cycle. These advancements in cloud description in these models have prompted a redefinition of data assimilation into NWP. In particular, atmospheric data other than temperature, specific humidity, and winds are now sought to be assimilated. Clouds and rainfall are palatable candidates, by virtue of the availability of global datasets from satellites. Assimilation of these variables may become operational in the next few years. Under this perspective, the research presented in this study is relevant to weather forecasting improvement. Cloud radar data will be available on a global scale in the next future with the deployment of the CloudSat mission. The techniques presented and illustrated in this dissertation, with the necessary refinements to satisfy operational requirements, can be used in NWP assimilation systems to ingest CloudSat data. It will then be possible to fully exploit the potential of cloud radar data to advance weather prediction.

6.3 Future work

Plans for future work include refinements of the CIGMA system. In particular, the derivation of the Hessian matrix is already well on its way. The Hessian will be used for preconditioning in order to improve the convergence of the minimization process. Since the model is computationally inexpensive, the exact derivation of the Hessian is possible. This will allow some comparative studies with preconditioning performed with various approximations to the Hessian that are commonly used. Development of other observational operators for the inclusion of different types of observations, for example Doppler radar reflectivities and radiances, is planned. We also intend to characterize model errors and include of the model error term in the assimilation system. Further assimilation experiments to take advantage of the opportunities afforded by a simplified system before embarking in full four-dimensional variational experiments with the RAMAS model are also in the plan. In particular, the model need to be refined in its basic formulation. For example, the introduction of a parameterization for homogeneous nucleation will improve the degree of realism of the model in representing cirrus clouds which form at low temperatures. Likewise,

the introduction of explicit parameterizations for turbulence/radiation/microphysics interaction will allow us to use the single column model to derive initial Ice Nuclei (IN) concentration from cloud radar observations.

We eagerly hope to apply the tools and the instruments discussed in this dissertation to variational assimilation of real radar data that are already available, for example from the ARM sites, or will be available in the near future from upcoming field experiments. A proposal has been recently submitted to NASA in response to a call for proposals relative to the CRYSTAL-FACE experiment. We intend to actively participate in the pre- and post-experiment modeling activity, on the basis of what was outlined in chapters 3 and 5. 4DVar experiments will not be restricted to ice clouds, but will encompass other cloud types, building on the work by and in collaboration with Dr. T.Vukićević and her co-workers at CIRA. In particular, we would like to address the main issue encountered in the exploration of radar data assimilation which is the limited capability of the assimilation system in improving the initial conditions when the cloud ice amount was underpredicted or no cloud at all was present in the model simulation. A way to address this issue is to enhance the mesoscale model skill in cloud prediction by using it, for example, in a LES configuration with bin microphysics and increased horizontal and vertical resolution. This solution requires a longer computational time and large computer storage capabilities, but it would allow us to better represent the cloud physics by reducing model errors and to extract more information from the measurements. In the study by Vukićević et al. (2001), this expedient is used to improve a RAMS simulation of a stratus cloud over the central US and obtain a better representation of the mesoscale variability (particularly in cloud droplet concentration) which was present in the satellite radiances. The LES bin microphysics version of RAMS was run over the same domain as a bulk microphysics single-moment version, and it was shown that the LES run had a higher degree of horizontal and vertical inhomogeneity. Not only the simulation was more realistic, but the data assimilation was also more effective. The benefits of data assimilation are indeed much greater if the model errors are not dominant or well characterized. Another way to overcome the limitation highlighted chapter 4 and 5 is to include in the assimilation additional

measurements that would complement the cloud radar data, such as radiances or other types of measurements which are sensitive to the humidity and temperature fields, even in the absence of clouds. The success of the assimilation will still be dependent on the model skill and on how far from cloudy conditions the initial fields are: if the profiles are too dry, the model will not be able to produce clouds, even upon assimilation of cloud observations. In principle, however, if the observations contain enough information about the ambient environmental fields, there is no limit on the correction that can be implemented through the assimilation system.

Ultimately, we are interested in performing a combined cloud and precipitation experiment to assess the relative importance of the two components for advancing both cloud and surface rainfall prediction. Data for these experiments, planned in collaboration with T. L'Ecuyer, will be provided by the TRMM satellite. Long-term plans involve the preparation for radar data assimilation on a global scale with CloudSat-like datasets.

These combined efforts will hopefully contribute to completing the colorful mosaic of cloud-environment interactions and a better understanding of Earth's weather and climate.

Bibliography

- Abramovitz, M. and I. A. Stegun: 1974, *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*. Dover Publications, Mineola, NY.
- Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen: 1999, *LAPACK Users' guide*. Philadelphia.
- Anthes, R. A.: 1974, Data assimilation and initialization of hurricane prediction models. *J. Atmos. Sci.*, **31**, 702–719.
- Asselin, R.: 1972, Frequency filter for time integrations. *Mon. Wea. Rev.*, **100**, 487–490.
- Austin, R. T. and G. L. Stephens: 2001, Retrieval of stratus cloud microphysical parameters using millimetric radar and visible optical depth in preparation for CloudSat. Part 1: Algorithm formulation. *J. Geophys. Res.*, submitted.
- Benedetti, A. and G. L. Stephens: 2001, Characterization of model errors in cirrus simulations from a cloud resolving model for application to ice water content retrievals. *Atmos. Res.*, accepted for Special Issue Reno 2000.
- Bennett, A. F. and P. C. McIntosh: 1982, Open ocean modeling as an inverse problem: tidal theory. *J. Phys. Oceanogr.*, **12**, 1004–1018.
- Betts, A. K. and M. J. Miller: 1986, A new convective adjustment scheme. Part I: Observational and theoretical basis. *Quart. J. Roy. Meteor. Soc.*, **112**, 693–709.
- Bischof, C., A. Carle, P. Khademi, and A. Mauer: 1995, The ADIFOR 2.0 System for the Automatic Differentiation of Fortran 77 Programs. Technical Memorandum ANL-MCS-P481-1194, Argonne National Laboratory.
- Boehm, M. T. and J. Verlinde: 2000, Stratospheric influence on upper tropospheric cirrus. *Geophys. Res. Lett.*, **27**, 3209–3212.
- Boehm, M. T., J. Verlinde, and T. P. Ackerman: 1999, On the maintenance of high tropical cirrus. *J. Geophys. Res.*, **104**, 24423–24433.
- Bouttier, F. and P. Courtier: 1999, Data assimilation concepts and methods. Technical report, European Center for Medium-range Weather Forecast.
- Brown, P. R. A. and P. R. Field: 2000, The comparison of cloud-resolving simulations of cirrus cloud with observations. *Proceedings of the 13th International Conference on Cloud and Precipitation*, Reno, NV, 1225–1228.
- Chandrasekar, S.: 1960, *Radiative Transfer*. Dover, New York, NY.
- Cheng, W. Y. Y., T. Wu, and W. R. Cotton: 2001, Large-Eddy simulation of the 26 November 1991 FIRE II cirrus case. *J. Atmos. Sci.*, **58**, 1017–1034.

- Clothiaux, E. E., M. A. Miller, B. A. Albrecht, T. P. Ackerman, J. Verlinde, D. M. Babb, R. M. Peters, and W. J. Syrett: 1995, An evaluation of a 94-GHz radar for remote sensing of cloud properties. *J. Atmos. Ocean. Tech.*, **12**, 201–229.
- Cohn, S. E.: 1997, An introduction to estimation theory. *J. Meteor. Soc. of Japan*, **50**, 147–178.
- Cotton, W. R. and R. A. Anthes: 1989, *Storm and cloud dynamics*. Academic Press, 883 pp.
- Courtier, P.: 1997, Variational methods. *J. Meteor. Soc. Japan*, **75**, 211–218.
- Daley, R.: 1991, *Atmospheric Data Analysis*. Cambridge University Press, New York, NY.
- Danielson, E. F.: 1982, A dehydration mechanism for the stratosphere. *Geophys. Res. Lett.*, **9**, 605–608.
- De Mott, P. J., M. P. Meyers, and W. R. Cotton: 1994, Parameterization and impact of ice initiation processes relevant to numerical model simulations of cirrus clouds. *J. Atmos. Sci.*, **51**, 77–90.
- De Mott, P. J., D. C. Rogers, and S. M. Kreidenweis: 1997, The susceptibility of ice formation in upper tropospheric clouds to insoluble aerosol components. *J. Geophys. Res.*, **102**, 19575–19584.
- Deblonde, G.: 1999, Variational assimilation of SSM/I precipitable water retrievals in the CMC analysis system. *Mon. Wea. Rev.*, **127**, 1458–1476.
- Dee, D. P. and A. M. Da Silva: 1998, Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.*, **124**, 269–295.
- Dowling, D. R. and L. F. Radke: 1990, A summary of the physical properties of cirrus clouds. *J. Appl. Meteor.*, **29**, 970–978.
- Drake, R. L.: 1972, The scalar transport equation of coalescence theory: moments and kernels. *J. Atmos. Sci.*, **29**, 537–547.
- Errico, E. M. and T. Vukićević: 1992, Sensitivity analysis using an adjoint of the PSU/NCAR mesoscale model. *Mon. Wea. Rev.*, **120**, 1644–1660.
- Errico, R. M., T. Vukićević, and K. Raeder: 1993, Comparison of initial and lateral boundary condition sensitivity for a limited-area model. *Tellus*, **45A**, 539–557.
- Eyre, J. R.: 1997, Variational assimilation of remotely-sensed observations of the atmosphere. *J. Meteor. Soc. Japan*, **75**, 331–338.
- Fabry, F. and I. Zawadzki: 1995, Long-term radar observations of the melting layer of precipitation and their interpretation. *J. Atmos. Sci.*, **52**, 838–851.
- Faure, C. and Y. Papegay: 1997, Odyssee Version 1.6. the language reference manual. Technical Memorandum 3211, Institut National de Recherche en Informatique et en Automatique (INRIA).

- Fletcher, N. H.: 1962, *Physics of Rain Clouds*. Cambridge University Press, Cambridge, UK, 386 pp.
- Fouilloux, A. and J. Iaquina: 1998, Mesoscale cirrus cloud modeling and comparison against experimental data collected on 17 april 1994 during the EUCREX campaign. *Mon. Wea. Rev.*, **126**, 2422–2434.
- Fowler, L. D., D. A. Randall, and S. A. Rutledge: 1996, Liquid and ice cloud microphysics in the CSU general circulation model. 1. Model description and results of a baseline simulation. *J. Climate*, **9**, 489–529.
- Gauss, K. F.: 1809, *Theoria motus corporum caelestium*. Published in English by Dover Press, NY, 1963.
- Gerard, E. and R. W. Saunders: 1999, Four-dimensional assimilation of Special Sensor Microwave/Imager total column water vapor in the ECMWF model. *Qyart. J. Roy. Meteor. Soc.*, **125**, 3077–3101.
- Ghemires, M., T. Vukićević, R. Hertenstein, and T. J. Greenwald: 2001, Adjoint strategy for radiance data assimilation using regional atmospheric system. *Proceedings of Fifth Symposium on Integrated Observing Systems*, 122–125, Albuquerque, NM, 14–18 January 2001.
- Giering, R.: 1999, *Tangent linear and Adjoint Model Compiler , Users manual 1.4*.
- Giering, R. and T. Kaminski: 1998, Recipes for adjoint code construction. *ACM Trans. On Math. Software*, **24**, 437–474.
- Gilbert, J. C. and J. Nocedal: 1992, Global convergence properties of conjugate gradient methods of optimization. *SIAM J. Opt.* **2**, 21–42.
- Hall, M. C. G., D. G. Cacuci, and M. E. Schlesinger: 1982, Sensitivity analysis of a radiative convective model by the adjoint method. *J. Atmos. Sci.*, **39**, 2050–2083.
- Hallett, J. and S. C. Mossop: 1974, Production of secondary ice particles during the riming process. *Nature*, **249**, 26–28.
- Harrington, J. K.: 1997, *The effects of radiative and microphysical processes on simulated warm and transition season Arctic stratus*. Ph.D. thesis, Colorado State University.
- Harrington, J. Y., M. P. Meyers, R. L. Walko, and W. R. Cotton: 1995, Parameterization of ice crystal conversion processes due to vapor deposition for mesoscale model using double-moment basis functions. Part I: Basic formulation and parcel model results. *J. Atmos. Sci.*, **52**, 4344–4366.
- Hartmann, D. L., M. E. Ockert-Bell, and M. L. Michelsen: 1992, The effect of cloud type on earth's energy balance: global analysis. *J. Climate*, **5**, 1281–1304.
- Heymsfield, A. J.: 1977, Precipitation development in stratiform ice clouds: a microphysical and dynamical study. *J. Atmos. Sci.*, **34**, 367–381.
- Heymsfield, A. J. and L. J. Donner: 1990, A scheme for parameterizing ice-cloud water content in general circulation models. *J. Atmos. Sci.*, **47**, 1865–1877.

- Heymsfield, A. J. and J. Iaquinta: 2000, Cirrus crystal terminal velocities. *J. Atmos. Sci.*, **57**, 916–938.
- Hogan, R. J., A. J. Illingworth, and H. Sauvageot: 2000, Measuring crystal size in cirrus using 35- and 94-GHz radars. *J. Atmos. Oceanic Technol.*, **17**, 27–37.
- Hollingsworth, A. and P. Lönnberg: 1986, The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: the wind field. *Tellus*, **38A**, 111–136.
- Hou, A. Y., S. Q. Zhang, A. M. D. Silva, and W. Olson: 2000, Improving assimilated global datasets using TMI rainfall and columnar moisture observations. *J. Climate*, **13**, 4180–4195.
- Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc: 1997, Unified notation for data assimilation: operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 181–189.
- Irvine, W. M.: 1975, Multiple scattering in planetary atmospheres. *Icarus*, **25**, 175.
- Jakob, C.: 2001, *Ice clouds in numerical weather prediction: progress, problems and prospects*. Oxford University Press. in *Cirrus*, D.Lynch (Ed.). In press.
- Jakob, C. and J.-J. Morcrette: 1995, Sensitivity of the ECMWF model to the treatment of the ice phase. Number 713, 37–46, workshop on Cloud Microphysics Parameterizations in Global Atmospheric Circulation Models, Kananaskis, Alberta, Canada, 23-25 May 1995.
- Janiskova, M., J.-N. Thépaut, and J.-F. Geleyn: 1999, Simplified and regular physical parameterizations for incremental four-dimensional variational assimilation. *Mon. Wea. Rev.*, **127**, 26–45.
- Jr., J. E. D. and R. B. Schnabel: 1983, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, Englewood Cliffs, NJ.
- Kalkofen, W.: 1985, *Numerical Radiative Transfer*. Cambridge University Press, Cambridge, UK.
- Kaminski, T., R. Giering, and M. Heimann: 1996, Sensitivity of the seasonal cycle of CO₂ at remote monitoring stations with respect to seasonal surface exchange fluxes determined with the adjoint of an atmospheric transport model. *Phys. Chem. Earth*, **21**, 457–463.
- Khvorostyanov, V. I. and K. Sassen: 1998, Cirrus cloud simulation with explicit microphysics and radiation: Part I. Model description. *J. Atmos. Sci.*, **55**, 1808–1821.
- Kirkpatrick, S., C. D. G. Jr, and M. P. Vecchi: 1983, Optimization by simulated annealing. *Science*, **220**, 671–680.
- Klinker, E., F. Rabier, G. Kelly, and J.-F. Mahfouf: 2000, The ECMWF operational implementation of four-dimensional variational assimilation. III: Experimental results and diagnostics with operational configuration. *Quart. J. Roy. Meteor. Soc.*, **126**, 1191–1215.
- Lawson, R. P., R. E. Stewart, and L. J. Angus: 1998, Observations and numerical simulations of the origin and development of very large snowflakes. *J. Atmos. Sci.*

- Le Dimet, F.-X. and O. Talagrand: 1986, Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, **38A**, 97–110.
- L'Ecuyer, T. S. and G. L. Stephens: 2001, An uncertainty model for Monte Carlo retrieval algorithms: application to the TRMM observing system. *Quart. J. Roy. Meteor. Soc.*
- Lenoble, J.: 1985, *Radiative Transfer in Scattering and Absorbing Atmospheres: Standard Computational Procedures*. A. Deepak Publishing, Hampton, VA.
- Lewis, J. M. and J. C. Derber: 1985, The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, **37A**, 309–322.
- Liao, L. and K. Sassen: 1994, Investigation of relationships between Ka-band radar reflectivity and ice and liquid water contents. *Atmos. Res.*, **34**, 231–248.
- Lin, R.-F.: 1997, *A numerical study of the evolution of nocturnal cirrus by a two-dimensional model with explicit microphysics*. Ph.D. thesis, Pennsylvania State University.
- Lions, J. L.: 1971, *Optimal control of systems governed by partial differential equations*. Springer-Verlag, Berlin, Germany.
- Liou, K.-N.: 1980, *An introduction to atmospheric radiation*. Academic Press, New York, NY.
- 1986, Influence of cirrus clouds on weather and climate process: a global perspective. *Mon. Wea. Rev.*, **114**, 1167–1199.
- Liu, C.-L. and A. Illingworth: 2000, Toward more accurate retrievals of ice water content from radar measurements of clouds. *J. Appl. Meteor.*, **39**, 1130–1146.
- Lönnerberg, P. and A. Hollingsworth: 1986, The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: the covariance of height and wind errors. *Tellus*, **38A**, 137–161.
- Lorenc, A. C.: 1981, A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701–721.
- 1984, Analysis methods for the quality control of observations. *Proceedings of ECMWF workshop on The use and quality control of meteorological observations for numerical weather prediction*, Reading, UK, 6–9 November 1984.
- 1986, Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177–1194.
- Luenberger, D. G.: 1984, *Linear and nonlinear programming*. Addison-Wesley, Reading, MA.
- Mace, G. G., T. P. Ackerman, P. Minnis, and D. F. Young: 1998a, Cirrus layer microphysical properties derived from surface-based millimeter radar and infrared interferometer data. *J. Geophys. Res.*, **103(D18)**, 23207–23216.
- Mace, G. G., C. Jakob, and K. P. Monan: 1998b, Validation of hydrometeor prediction from the ECMWF model during the winter season 1997 using millimeter wave radar data. *Geophys. Res. Lett.*, **25**, 1645–1648.

- MacPherson, B.: 1991, Dynamic initialization by repeated insertion of data. *Quart. J. Roy. Meteor. Soc.*, **117**, 965–991.
- Mahfouf, J.-F. and F. Rabier: 2000, The ECMWF operational implementation of four-dimensional variational assimilation. II: Experimental results with improved physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1171–1190.
- Mahrer, Y. and R. A. Pielke: 1977, Effects of topography on sea and land breezes in a two-dimensional numerical model. *Mon. Wea. Rev.*, **105**, 1151–1162.
- Marchuk, G. I.: 1974, Numerical solution of the problems of the dynamics of the atmosphere and the ocean. *Gidrometeoizdat*.
- Marécal, V. and J.-F. Mahfouf: 2000, Variational retrieval of temperature and humidity profiles from TRMM precipitation data. *Mon. Wea. Rev.*, **128**, 3853–3866.
- Marks, C. J. and C. D. Rodgers: 1993, A retrieval method for atmospheric composition from limb emission measurements. *J. Geophys. Res.*, **98**, 14939–14953.
- Matrosov, S. Y.: 1999, Retrievals of vertical profiles of ice cloud microphysics from radar and IR measurements using tuned regressions between reflectivity and cloud parameters. *J. Geophys. Res.*, **104**, 16741–16753.
- Matrosov, S. Y., B. W. Orr, R. A. Kropfli, and J. B. Snider: 1994, Retrieval of vertical profiles of cirrus cloud microphysical parameters from Doppler radar and infrared radiometer measurements. *J. Appl. Meteor.*, **33**, 617–626.
- Matrosov, S. Y., T. Uttal, J. B. Snider, and R. A. Kropfli: 1992, Estimation of ice cloud parameters from ground-based infrared radiometer and radar measurements. *J. Geophys. Res.*, **97**, 11567–11574.
- Ménard, R. and R. Daley: 1995, Application of Kalman smoother theory to the estimation of the error statistics for the 4DVAR problem. *Tellus*, **48A**, 221–237.
- Meyers, M. P., P. J. De Mott, and W. R. Cotton: 1992, New primary ice-nucleation parameterization in an explicit cloud model. *J. Appl. Meteor.*, **31**, 708–721.
- Meyers, M. P., R. L. Walko, J. Y. Harrington, and W. R. Cotton: 1997, New RAMS cloud microphysics parameterization. Part II: The two-moment scheme. *Atmos. Res.*, **45**, 3–39.
- Miller, E. R., J. Wang, and H. L. Cole: 1999a, Correction for dry bias in Vaisala radiosonde RH data. *Proceedings of Ninth ARM Science Meeting Proceedings*, san Antonio, TX, March 22–26, 1999.
- Miller, S. D.: 2000, *A multi-sensor approach to the retrieval and validation of global cloudiness*. Ph.D. thesis, Colorado State University.
- Miller, S. D., G. L. Stephens, and A. C. M. Beljaars: 1999b, A validation survey of the ECMWF prognostic cloud scheme using LITE. *Geophys. Res. Lett.*, **102D**, 16663–16682.
- Miller, S. D., G. L. Stephens, C. K. Drummond, A. K. Heidinger, and P. T. Partain: 2000, A multisensor diagnostic satellite cloud property retrieval scheme. *J. Geophys. Res.*, **105**, 19955–19971.

- Mitchell, D. L.: 1988, Evolution of snow-size spectra in cyclonic storms. Part I: Snow growth by vapor deposition and aggregation. *J. Atmos. Sci.*, **45**, 3431–3451.
- 1991, Evolution of snow-size spectra in cyclonic storms. Part II: Deviations from the exponential form. *J. Atmos. Sci.*, **48**, 1885–1899.
- 1994, A model predicting the evolution of ice particle spectra and radiative properties of cirrus clouds. Part I: Microphysics. *J. Atmos. Sci.*, **51**, 797–816.
- 1996, Use of mass- and area-dimensional power laws for determining precipitation particle terminal velocities. *J. Atmos. Sci.*, **53**, 1710–1723.
- Mitchell, D. L. and S. K. Chai: 1998, The potential dependence of ice nucleation rates on crystal shape. Preprints of Conference on Cloud Physics, 17–21 August, Everett, WA.
- Mobley, C. D.: 1995, *Light and Water*. Academic Press, San Diego, CA.
- Moore, R. W. and T. H. Vonder Haar: 2001, Interannual variability of cloud forcing and meridional energy transport for the northern hemisphere winter from 1984–1990. *J. Climate*, **14** (17), 3643–3654.
- Nakajima, T. and M. D. King: 1990, Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory. *J. Atmos. Sci.*, **47**, 1878–1893.
- Parrish, D. F., J. C. Derber, R. J. Purser, W.-S. Wu, and Z.-X. Pu: 1997, The NCEP global analysis system: recent improvements and future plans. *J. Meteor. Soc. Japan*, **75**, 359–365.
- Passarelli, J., R. E.: 1978a, An approximate analytical model of the vapor deposition and aggregation growth of snowflakes. *J. Atmos. Sci.*, **35**, 118–123.
- 1978b, Theoretical and observational study of snow-size spectra and snowflake aggregation efficiencies. *J. Atmos. Sci.*, **35**, 882–889.
- Pielke, R. A., W. R. Cotton, R. L. Walko, C. J. Tremback, M. E. Nicholls, M. D. Moran, D. A. Wesley, T. J. Lee, and J. H. Copeland: 1992, A comprehensive meteorological modeling system – RAMS. *Meteor. Atmos. Phys.*, **42**, 69–91.
- Platnick, S.: 2000, Vertical photon transport in cloud remote sensing problems. *J. Geophys. Res.*, **105**(D18), 22919–22935.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery: 1992, *Numerical Recipes in C*. Cambridge University Press, 994 pp.
- Pruppacher, H. R. and J. D. Klett: 1997, *Microphysics of cloud and precipitation*. Kluwer Academic Publisher, 954 pp.
- Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons: 2000, The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1143–1170.
- Randall, D. A., J. Curry, P. Duynkerke, S. Krueger, M. Miller, M. Moncrieff, B. Ryan, D. Starr, W. Rossow, G. Tselioudis, and B. Wielicki: 2000, The second GEWEX cloud system study science and implementation plan. Technical report.

- Randall, D. A., Harshvardhan, D. A. Dazlich, and T. G. Corsetti: 1989, Interactions among radiation, convection, and large-scale dynamics in a general circulation model. *J. Atmos. Sci.*, **46**, 1943–1970.
- Rodgers, C. D.: 1976, Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Rev. Geophys. Space Phys.*, **14**, 609–624.
- 1990, Characterization and error analysis of profiles retrieved from remote sounding measurements. *J. Geophys. Res.*, **95**, 5587–5595.
- Rogers, R. F., J. M. Fritsch, and W. C. Lambert: 2000, A simple technique for using radar data in the dynamic initialization of a mesoscale model. *Mon. Wea. Rev.*, **128**, 2560–2574.
- Rossow, W. B., E. Kinsella, A. Wolf, and L. Garder: 1987, International Satellite Cloud Climatology project (ISCCP) Description of Reduced Resolution Radiance Data. Technical report wmo/td-no. 58, World Meteorological Organization.
- Saltelli, A.: 1999, Sensitivity analysis: Could better methods be used? *J. Geophys. Res.*, **104**, 3789–3793.
- Sasaki, Y.: 1970, Some basic formalisms of four-dimensional data assimilation. *Mon. Wea. Rev.*, **98**, 875–883.
- Schubert, W. H.: 1997, Atmospheric dynamics. Technical report, Department of Atmospheric Science, Colorado State University.
- Sherwood, S. C.: 1999, Cirrus moistening in the tropics. *J. Geophys. Res.*, **104**, 11949–11960.
- Stamnes, K. and P. Conklin: 1984, A new multi-layer discrete ordinate approach to radiative transfer in vertically inhomogeneous atmospheres. *J. Quant. Spectrosc. Radiat. Transfer*, **31**, 273–282.
- Stamnes, K., S.-C. Tsay, W. J. Wiscombe, and K. Jayaweera: 1988, Numerically stable algorithm for dom radiative transfer in multiple scattering and emitting layer media. *Appl. Opt.*, **27**, 2502–2509.
- Starr, D. O., A. Benedetti, M. T. Boehm, P. R. A. Brown, K. M. Gierens, E. Girard, V. Giraud, C. Jakob, E. J. V. I. Khvorostyanov, M. Koehler, A. Lare, R.-F. Lin, K. Maruyama, M. Montero, W.-K. T. Y. Wang, and D. Wilson: 2000, Comparison of cirrus cloud models: a project of the GEWEX Cloud System Study (GCSS) working group on cirrus cloud system. *Proceedings of the 13th International Conference on Cloud and Precipitation (ICCP)*, Reno, NV.
- Starr, D. O. and S. K. Cox: 1985, Cirrus clouds. Part I: A cirrus cloud model. *J. Atmos. Sci.*, **42**, 2663–2681.
- Stauffer, D. R. and J.-W. Bao: 1993, Optimal determination of nudging coefficients using adjoint equations. *Tellus*, **45A**, 358–369.
- Stauffer, D. R. and N. L. Seaman: 1990, Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Mon. Wea. Rev.*, **118**, 1250–1277.

- Stephens, G. L.: 1978, Radiation profiles in extended water clouds. Part I: Theory - Part II: Parameterization schemes. *J. Atmos. Sci.*, **35**, 2111–2132.
- 1994, *Remote sensing of the lower atmosphere*. Oxford University Press, Oxford, UK.
- 2001, *Cirrus, climate and global change*. Oxford University Press, in *Cirrus*, D.Lynch (Ed.). In press.
- Stephens, G. L., R. Ellingson, J. Jr. Vitko, W. Bolton, T. P. Tooman, F. P. J. Valero, P. M. P. Pilewskie, G. S. Phipps, S. Sekelsky, J. R. Carswell, S. D. Miller, A. Benedetti, R. B. McCoy, R. F. Jr. McCoy, A. Lederbuhr, and R. Bambha: 2000a, The Department of Energy's Atmospheric Radiation Measurement (ARM) Unmanned Aerospace Vehicle (UAV) Program. *Bull. Amer. Meteor. Soc.*, **81**, 2915–2937.
- Stephens, G. L., C. Jakob, and M. Miller: 1998, Atmospheric ice—a major gap in understanding the effects of clouds on climate. *GEWEX/WCRP News*, **8**.
- Stephens, G. L., S.-C. Tsay, P. W. Stackhouse, and P. J. Flatau: 1990, The relevance of the microphysical and radiative properties of cirrus clouds to climate and climatic feedback. *J. Atmos. Sci.*, **47**, 1742–1753.
- Stephens, G. L., D. G. Vane, and S. J. Walter: 2000b, The CloudSat mission: a new dimension to space-based observations of cloud in the coming millennium. *Bull. Amer. Meteor. Soc.*.
- Stephens, G. L. and P. J. Webster: 1981, Clouds and climate: sensitivity of simple systems. *J. Atmos. Sci.*, **38**, 235–247.
- Sun, J. and N. A. Crook: 1997, Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: model development and simulated data experiments. *J. Atmos. Sci.*, **54**, 1642–1661.
- 1998, Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part II: retrieval experiments of an observed florida convective storm. *J. Atmos. Sci.*, **55**, 835–852.
- Talagrand, O. and P. Courtier: 1987, Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311–1328.
- Tiedke, M.: 1993, Representation of clouds in large-scale models. *Month. Wea. Rev.*, **121**, 3040–3061.
- Trenberth, K. E. and J. M. Caron: 2000, Estimates of meridional atmosphere and ocean heat transports. *J. Climate*.
- Tripoli, G. J. and W. R. Cotton: 1982, Colorado State University three-dimensional cloud/mesoscale model, 1982: Pt.1, General theoretical framework and sensitivity experiments. *J. Rech. Atmos.*, **16**, 185–219.
- van de Hulst, A. C.: 1962, A new look at multiple scattering. Technical report, NASA Institute for Space Studies, New York, NY.

- Verlinde, J. and W. R. Cotton: 1993, Fitting microphysical observations of nonsteady convective clouds to a numerical model: an application of the adjoint technique of data assimilation to a kinematic model. *Mon. Wea. Rev.*, **121**, 2776–2793.
- Vukićević, T.: 1998, Optimal initial perturbations for two cases of extratropical cyclogenesis. *Tellus*, **45A**, 493–510.
- 2000, Class notes on data assimilation. Technical report, Department of Atmospheric Science, Colorado State University.
- Vukićević, T. and R. M. Errico: 1993, Linearization and adjoint of parameterized moist diabatic processes. *Tellus*, **45A**, 493–510.
- Vukićević, T., T. J. Greenwald, R. Hertenstein, and M. Ghemires: 2001, Use of cloudy radiance observations in mesoscale data assimilation. *Proceedings of Fifth Symposium on Integrated Observing Systems*, 113–117, Albuquerque, NM, 14–18 January 2001.
- Vukićević, T. and P. Hess: 2000, Analysis of tropospheric transport during MLOPEX using adjoint technique. *J. Geophys. Res.*, **105**, 7213–7230.
- Vukićević, T. and K. Raeder: 1995, Use of an adjoint model for finding triggers for Alpine lee cyclogenesis. *Mon. Wea. Rev.*, **123**, 800–816.
- Walko, R. L., W. R. Cotton, M. P. Meyers, and J. Y. Harrington: 1995, New RAMS cloud microphysics parameterization. Part I: the single moment scheme. *Atmos. Res.*, **38**, 29–62.
- Walko, R. L. and C. J. Tremback: 2000. *Regional Atmospheric Modeling System Version 4.2 Introduction to RAMS 4.2*.
- Waterman, P. C.: 1981, Matrix–exponential description of radiative transfer. *J. Opt. Soc. Am.*, **71**, 410.
- Whitley, D.: 1994, A genetic algorithm tutorial. *Statistics and Computing*, **4**, 65–85.
- Wielicki, B. A., B. R. Barkstrom, E. F. Harrison, R. B. Lee III, G. L. Smith, and J. E. Cooper: 1996, Clouds and the Earth’s Radiant Energy System (CERES): an Earth Observing System experiment. *Bull. Amer. Meteor. Soc.*, **77**, 853–868.
- Winker, D. M., R. H. Couch, and M. P. McCormick: 1996, An overview of LITE: NASA’s Lidar In-space Technology Experiment. *Proceedings of IEEE*, volume 84, 164–180.
- Winker, D. M. and B. A. Wielicki: 1999, The PICASSO-CENA Mission. *Proc. SPIE*, 2636.
- Wu, T., W. R. Cotton, and W. Y. Y. Cheng: 2000, Radiative effects on the diffusional growth of ice particles in cirrus clouds. *J. Atmos. Sci.*, **57**, 2892–2904.
- Wung, B., J. Verlinde, and J. Sun: 2000, Dynamical and microphysical retrievals from Doppler radar observations of a deep convective cloud. *J. Atmos. Sci.*, **57**, 262–283.
- Young, S. A.: 1995, Analysis of lidar backscatter profiles in optically thin clouds. *Appl. Opt.*, **34**, 7019–7031.

- Zipser, E. J. and R. H. Johnson: 1998, Systematic errors in radiosonde humidities a global problem? 72–73, preprints of the 10th Symposium on Meteorological Observations and Instrumentation, Phoenix, Arizona.
- Zou, X., A. Barcilon, I. M. Navon, J. Whitaker, and D. G. Caccuci: 1993a, An adjoint sensitivity study of blocking in a two-layer isentropic model. *Mon. Wea. Rev.*, **121**, 2833–2857.
- Zou, X. and Y.-H. Kuo: 1996, Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Mon. Wea. Rev.*, 2859–2882.
- Zou, X., I. M. Navon, M. Berger, K. H. Phua, T. Schlick, and F. X. Le Dimet: 1993b, Numerical experience with limited-memory Quasi-Newton and truncated Newton methods. *SIAM J. Optimization*, **3**, 582–608.
- Zou, X., I. M. Navon, and F. X. L. Dimet: 1992, An optimal nudging data assimilation scheme using parameter estimation. *Quart. J. Roy. Meteor. Soc.*, **118**, 1163–1186.
- Zupanski, D.: 1997, A general weak constraint applicable to operational 4DVAR data assimilation system. *Mon. Wea. Rev.*, **125**, 2274–2292.
- Zupanski, D. and F. Mesinger: 1995, Four-dimensional variational assimilation of precipitation data. *Mon. Wea. Rev.*, **123**, 1112–1127.
- Zupanski, M.: 1998, A preconditioning algorithm for four-dimensional variational data assimilation. *Mon. Wea. Rev.*, **126**, 2274–2292.

Appendix A

Radar retrievals of ice cloud properties

A.1 Introduction

In this appendix a brief overview of an optimal-estimation radar retrieval of ice cloud properties is presented. The retrieval was an off-spring of the development of the radar mapping for the assimilation studies, and was motivated by the need for an ice cloud algorithm for CloudSat 94-GHz spaceborne radar. Many elements of the retrieval theory are in common with variational assimilation. The following sections are part of a paper by Benedetti and Stephens (2001) accepted for publication in *Atmospheric Research*.

A.2 Fundamentals of retrieval

The retrieval approach used in this study is adapted from Austin and Stephens (2001) (hereafter, AS01) based on work by Rodgers (1976) and (1990) and Marks and Rodgers (1993). The retrieval is formulated in the context of estimation theory, and allows for the inclusion of *a priori* information. It provides a quantitative estimate of the uncertainty in the retrieved quantities, and the relative influence of the measurements and *a priori* data on the retrievals. In AS01 a retrieval of stratus cloud microphysical parameters is described. The retrieval assumes a log-normal distribution of cloud droplets and uses radar observations and optical depth information. The difference between this work and AS01 is the specific application of the retrieval technique to ice clouds.

The radar measurements and the optical depth values, represented by the vector y , are related to the parameters we wish to retrieve, x (state vector), by a *forward*

model F :

$$\mathbf{y} = F(\mathbf{x}, b) + \epsilon_y \quad (\text{A.1})$$

where b represents the forward model parameters that are not retrieved, and ϵ_y includes the measurement error and the forward model error.

The retrieval algorithm seeks the optimal solution, $\hat{\mathbf{x}}$, by minimizing a quadratic cost function:

$$J = (\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a) + (\mathbf{y} - F(\mathbf{x}, b))^T \mathbf{S}_y^{-1} (\mathbf{y} - F(\mathbf{x}, b)), \quad (\text{A.2})$$

where \mathbf{S}_y is the forward model plus observation error covariance matrix whose diagonal elements represents the uncertainties that define ϵ_y in (A.1). \mathbf{x}_a is the *a priori* profile based on likely or statistical values of the state vector elements, and S_a is the *a priori* covariance matrix, which represents the uncertainty of this profile.

Minimization of the cost function leads to an iterative solution for the optimal state vector:

$$\mathbf{x}^{i+1} = (\mathbf{S}_a^{-1} + \mathbf{K}^i \mathbf{S}_y^{-1} \mathbf{K}^i)^{-1} [\mathbf{S}_a^{-1} \mathbf{x}_a + \mathbf{K}^i \mathbf{S}_y^{-1} (\mathbf{y} - F(\mathbf{x}^i) + \mathbf{K}^i \mathbf{x}^i)] \quad (\text{A.3})$$

where \mathbf{K} represents the Jacobian of the forward model with respect to the state vector ($\mathbf{K} = \frac{\partial F}{\partial \mathbf{x}}$), also known as the kernel matrix. Convergence is achieved when:

$$\Delta \mathbf{x}^T \mathbf{S}_x^{-1} \Delta \mathbf{x} \ll n \quad (\text{A.4})$$

with $\Delta \mathbf{x} = \mathbf{x}^{i+1} - \mathbf{x}^i$; n is the dimension of the state vector. The error covariance matrix \mathbf{S}_x is given by:

$$\mathbf{S}_x = (\mathbf{S}_a^{-1} + \mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K})^{-1} \quad (\text{A.5})$$

The diagonal elements of the matrix S_x provide uncertainty estimates for the retrieved parameters, and off-diagonal elements provide correlations between parameter uncertainties.

After convergence is achieved, a quality test for the goodness of the fit is performed computing the following quantity:

$$\chi^2 = (\mathbf{y} - F(\hat{\mathbf{x}}))^T \mathbf{S}_y^{-1} (\mathbf{y} - F(\hat{\mathbf{x}})) + (\mathbf{x}_a - \hat{\mathbf{x}})^T \mathbf{S}_a^{-1} (\mathbf{x}_a - \hat{\mathbf{x}}) \quad (\text{A.6})$$

Equation (A.6) represents a random variable which follows a χ^2 distribution with n degrees of freedom. A good retrieval should give a χ^2 of order n . For a more rigorous definition, χ^2 tables can be consulted to find the critical values of this variable that are within the desired statistical significance (Siegel and Morgan, 1996). In this work, only IWC retrievals with associate χ^2 significant to the 1% level are shown.

Other valuable information can be derived by computing the matrix quantity:

$$\mathbf{A} = \mathbf{S}_x \mathbf{K}_i^T \mathbf{S}_y^{-1} \mathbf{K}_i. \quad (\text{A.7})$$

Using (A.7), equation (A.3) can be rewritten as

$$\mathbf{x}^{i+1} = \mathbf{x}^i + \mathbf{A} \mathbf{K}_i^{-1} (\mathbf{y} - F(\mathbf{x}^i)) + (\mathbf{I} - \mathbf{A})(\mathbf{x}_a - \mathbf{x}^i), \quad (\text{A.8})$$

where \mathbf{I} is the identity matrix.

The \mathbf{A} matrix provides information on the relative weighting of the measurements and the *a priori* data. Ideally, \mathbf{A} should be equal to the identity matrix, and the retrieval of the state parameters should rely solely on the measurements, unless there is a reason to trust the *a priori* information. The greater the departure from unity, the more heavily the *a priori* profile is used. In reality, however, since measurements and forward model errors might give rise to unphysical retrievals, the introduction of *a priori* information ensures more stability in the inverse solution acting as a constraint on the retrieval (Rodgers, 1976). Usually, *a priori* information and the relative covariance matrix are derived from climatological records. When retrieving those variables for which a global climatology is not yet available, for example IWC, defining a plausible *a priori* with the appropriate error is a challenging task. The following sections try to answer this question by using the CRM results presented in chapter 3 and their ECM as *a priori* information for the retrieval of IWC.

A.3 Forward model equations

The “usual” gamma distribution is assumed to describe the size spectra of ice crystals, which are approximated as spheres:

$$n(D) = N_t \frac{1}{\Gamma(\nu)} \left(\frac{D}{D_n} \right)^{(\nu-1)} \frac{1}{D_n} \exp\left(- \frac{D}{D_n} \right). \quad (\text{A.9})$$

N_t represents the total number concentration, D_n the characteristic diameter and ν is the distribution width. The moments of the distribution, $I(p)$, are:

$$I(p) = \frac{1}{N_t} \int_0^\infty D^p n(D) dD = D_n^p F(p), \quad (\text{A.10})$$

where

$$F(p) = \frac{\Gamma(\nu + p)}{\Gamma(\nu)} \quad (\text{A.11})$$

The IWC is defined in terms of the third moment of the size distribution:

$$IWC = \int_0^\infty \rho_i \frac{\pi}{6} D^3 n(D) dD, \quad (\text{A.12})$$

with ρ_i representing the bulk ice density.

Assuming that ice crystals are approximately Rayleigh scatterers at the W-band radar frequency, the following definition of radar reflectivity Z can be used:

$$Z = \int_0^\infty D^6 n(D) dD \quad (\text{A.13})$$

At visible wavelengths, the extinction efficiency approaches the limit of 2, and the visible extinction coefficient, σ_{ext} , can be written as:

$$\sigma_{ext} = \int_0^\infty \frac{\pi}{4} D^2 n(D) dD \quad (\text{A.14})$$

Using equation (A.10), we can rewrite equations (A.12) through (A.14) as follows:

$$IWC = \frac{\pi}{6} \frac{\Gamma(3 + \nu)}{\Gamma(\nu)} \rho_i N_t D_n^3 \quad (\text{A.15})$$

$$Z = \frac{\Gamma(\nu + 6)}{\Gamma(\nu)} N_t D_n^6 \quad (\text{A.16})$$

$$\sigma_{ext} = \frac{\pi}{2} \frac{\Gamma(\nu + 2)}{\Gamma(\nu)} N_t D_n^2 \quad (\text{A.17})$$

The visible optical depth is computed by integrating the extinction coefficient through the cloud layer:

$$\tau = \int_{z_{base}}^{z_{top}} \sigma_{ext}(z) dz \quad (\text{A.18})$$

Equations (A.15)–(A.18) constitute the basis of the retrieval. In the version of the retrieval presented here, equation (A.15) was used to rewrite the characteristic diameter in terms of the IWC, and the IWC was directly retrieved assuming a bulk ice density of 0.92 gm^{-3} . In principle, all the distribution parameters, N_t , ν and D_n are functions of the vertical coordinate at cloud levels, and $3n$ independent measurements, with n number of cloud levels, are needed to fully retrieve the parameter profiles. In practice, we have a radar reflectivity profile (n measurements) and a column optical depth, for a total of $(n + 1)$ measurements. So it is necessary to make some assumptions to reduce the number of retrievables. In this study, the width of the distribution is assumed constant with height, with a fixed value of 2. The number concentration is also assumed constant with height, and a layer average value is retrieved, along with the IWC profile. More details on the implementation and validation of the retrieval along with a complete error analysis will be part of future work.

A.4 Application of ECM to the retrieval of cirrus cloud properties

A.4.1 Use of RAMS model field as *a priori*

The average IWC profiles and the ECM computed from the CRM simulations can be used as *a priori* information for the retrieval of IWC. The advantage of an optimal estimation retrieval over the use of the IWC–Z relationship of equation (3.9) is that new measurements can be easily added to the retrieval, for example the cloud optical depth, allowing for the retrieval of additional information. Moreover, a full error analysis is readily available by computing the matrix S_x . This is a significant advantage over an empirical approach. Examples of the retrieved IWC from the Kauai radar data are presented in the following section.

The cloud optical depth information was calculated using the retrieval described in Miller et al. (2000), courtesy of Steve Miller. SSP and GOES radiance data were used in the optical depth estimation. Optical depth results for the Kauai cirrus case are shown in Stephens et al. (2000a).

A.4.2 IWC retrievals

The average bias-corrected IWC profile and the ECM computed from the CRM simulations are used as *a priori* information for the retrieval of IWC. A comparison is made between retrievals obtained using RAMS profile versus an arbitrary constant *a priori* profile with large error covariance. The retrieved IWC is also compared to IWC estimated with relation (3.9). The *a priori* layer-average number concentration is chosen arbitrarily and assigned a large error. We hope to relax this assumption in the future by making use of the number concentration value provided by the CRM. In situ measurements are needed to perform an error analysis on number concentration, similar to the one presented in chapter 3 of this study for the IWC.

Figure A.1 shows the retrieved layer average microphysical fields, the IWC (top panel), the number concentration (middle panel) and characteristic diameter (bottom panel) using the RAMS profile as *a priori*. Error bars on layer averages are also shown. The error on the characteristic diameter is propagated from the errors on IWC and number concentration. Figure A.2 shows the same variables, but for an arbitrary constant *a priori* profile. When using a constant profile it is necessary to choose a large variance, since no realistic estimate of the errors can be made. This leads to an **A** matrix nearly equal to the identity matrix. Conversely, when information from the CRM is used, the algorithm makes use of it, and the matrix **A** is not perfectly unity, although its departure from the unit matrix is not excessive, indicating that the *a priori* is not completely driving the retrieval.

The retrieved IWC values in the two cases are comparable, although it appears that the retrieved average number concentration is larger, and correspondingly the average characteristic diameter is lower, when using the RAMS *a priori* profile. The retrieval errors are also similar. It may appear that no substantial improvement is

seen when using the model-derived profile. However, it was found that retrieved values with a significant χ^2 value more frequently occurred when using the RAMS profile rather than the constant profile, resulting in a larger number of “acceptable” retrievals. This indicates that, although a retrieval is in principle possible even in the case of “bad” *a priori* data, a more accurate *a priori* profile is in general preferable.

The IWC retrieval presented in this work was compared with the radar-only IWC retrieval described in LI00 to provide some form of validation, and discuss some of the main differences between these two approaches. The advantage of an optimum estimation retrieval over the use of empirical IWC-Z relationship, is that new measurements can easily be added to the retrieval, for example the cloud optical depth, allowing for the retrieval of additional information. It is found that values of IWC retrieved with the two methods are quite different. Figures A.3 and A.4 show scatterplots of IWC retrieved with the present method versus the estimate of IWC given by inverting equation (3.9). The data are scattered, indicating a poor agreement. When using the RAMS profile, slightly better agreement is obtained. This may be due to the fact that relation (3.9) was used to assign the CRM bias. Discrepancies can be attributed to the use of an invariant IWC-Z relationship that does not account for variations in thermodynamic and microphysics parameters (such as temperature and ice crystal size) across the cloud layer whereas retrievals based on the method introduced in this paper accommodates some of this variation. One possible improvement is to incorporate the temperature or size information, if available, to obtain improved, parameter-dependent coefficients for multiple IWC-Z relationships. It should be further mentioned that the goal of these comparisons is not to determine which retrieval gives the most realistic estimate of IWC, as this requires a validation against independent IWC data; the purpose is rather to place results from current estimation-based retrieval into context with respect to IWC-Z retrieval techniques which are more commonly used.

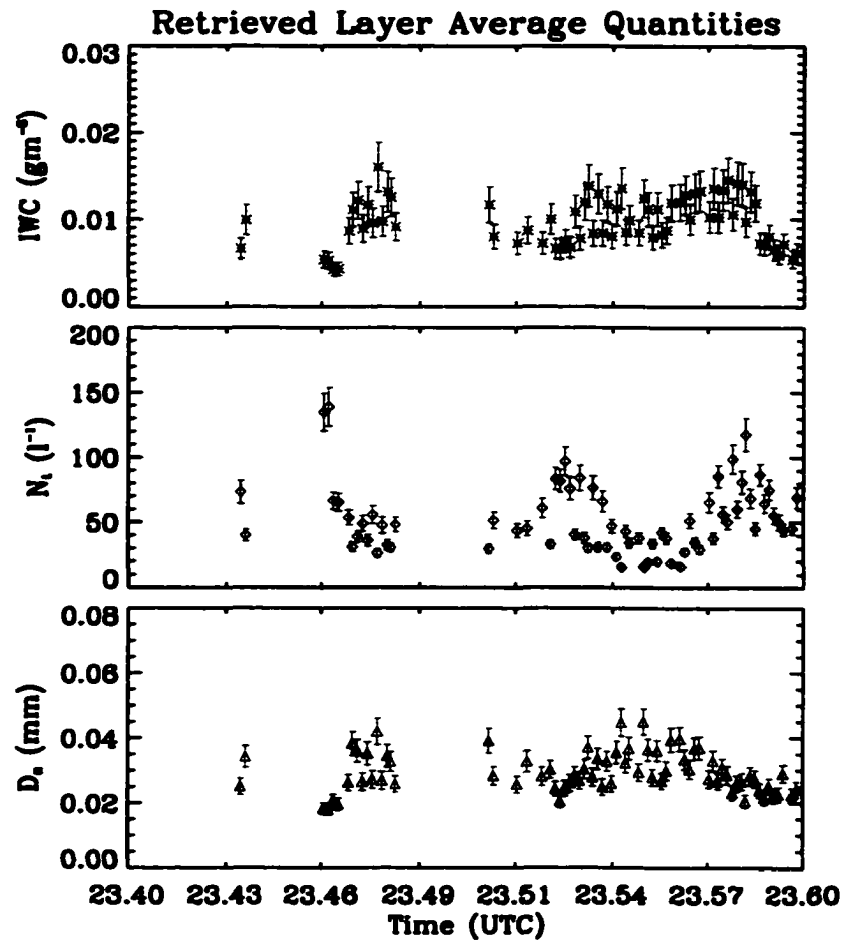


Figure A.1: Microphysical properties for April 30 cirrus, retrieved using CRM *a priori*.

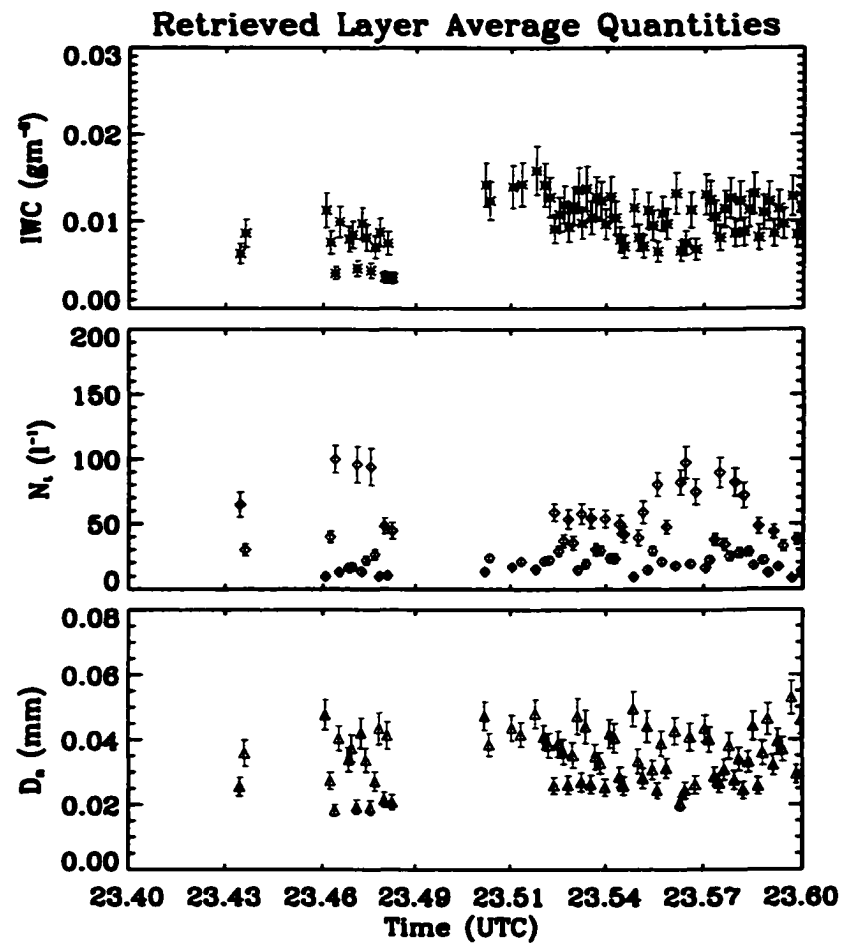


Figure A.2: As in figure A.1, but using constant *a priori*.

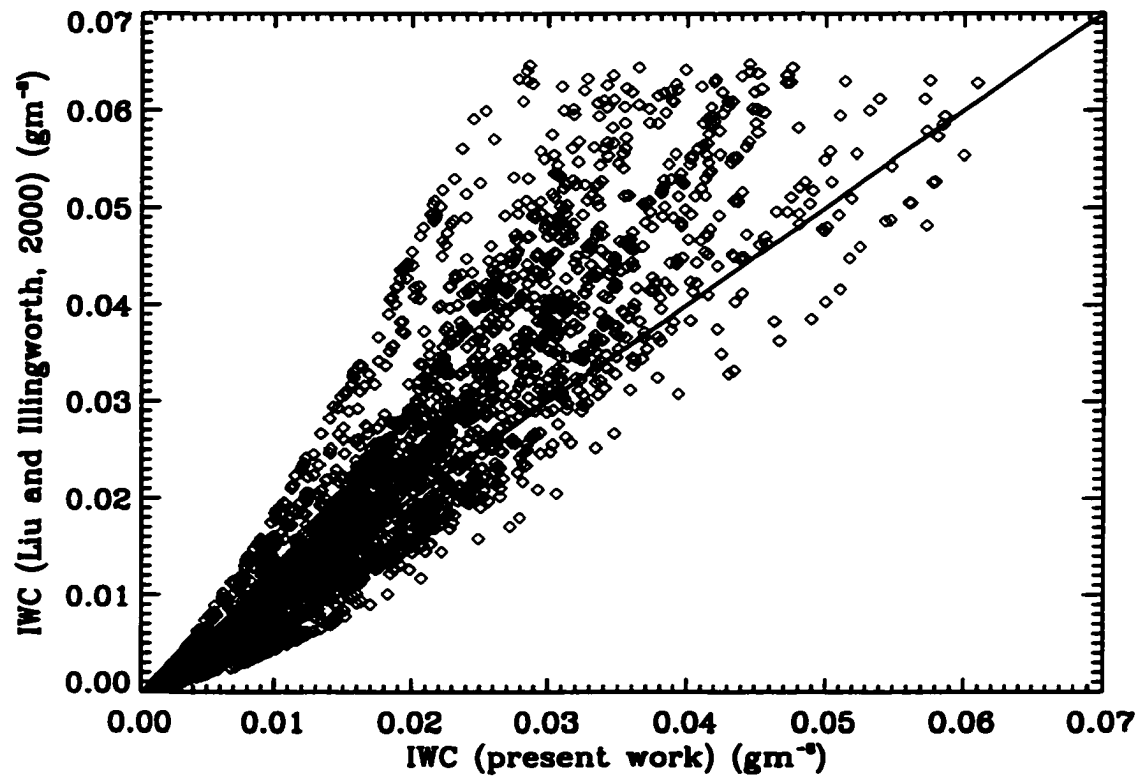


Figure A.3: IWC from optimal estimation retrieval versus IWC derived from IWC-Z relationship. CRM fields were used as *a priori*.

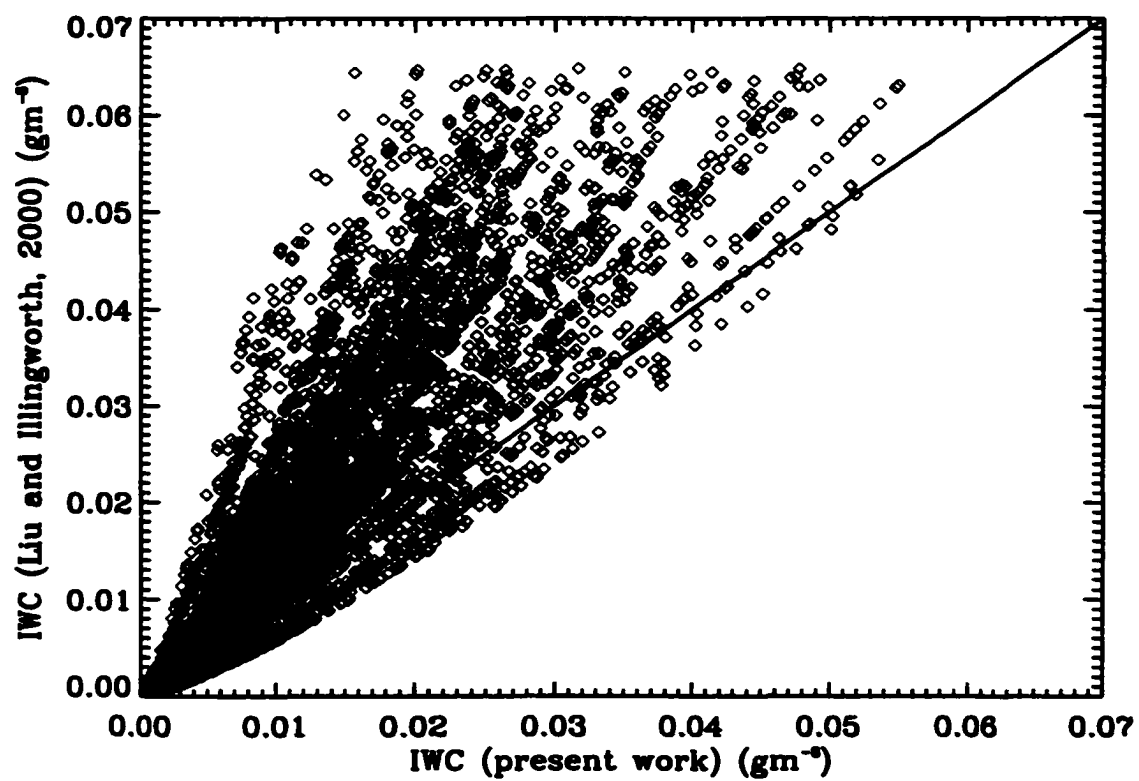


Figure A.4: As in figure A.3, but using the constant *a priori* profile.

Appendix B

More on ice growth model evaluation

B.1 Case study 3: Oklahoma, 26 September 1997

A second cirrus case observed during the Fall 1997 Cloud IOP at the ARM SGP site was considered. The radar image shown in figure B.1 together with the Raman lidar optical depth, shows a cloud that appears quite thin and broken at the initial time, and then grows into a thicker cirrus. In situ observations were available from 18:10UTC to approximately 21:00UTC. No radiosonde was launched during the flight of the UND Citation. The two closest soundings are 17:30UTC, and 23:30UTC. The 17:30UTC sounding was chosen for model initialization. In situ measurements are analyzed over a time interval of half an hour between 18:30 and 19:00UTC (one hour after the sounding). Discrepancies between model and observed fields are going to be enhanced due to this additional factor.

Figure B.2 shows time series of temperature, relative humidity with respect to ice and plane altitude as recorded on board of the UND Citation. Time series of measurements show more fluctuations than the May 8 case, which is a sign of a less horizontally homogeneous cloud.

Figure B.3 presents time series of particle number concentration and mean size measured by the 2D-C probe. The average mean size is around $100 \mu m$ and the average number concentration is around 10 particles per liter with a peak of $70 l^{-1}$.

As previously seen the initialization with unsaturated initial profiles does not provide a good agreement of model predicted cloud concentrations and reflectivities (not shown). When the supersaturation is increased and the profile is brought close to water saturation at cloud levels (see B.4), a closer agreement is obtained for all different characteristic diameter initializations (see figures B.5 and B.7, B.8 and B.6,

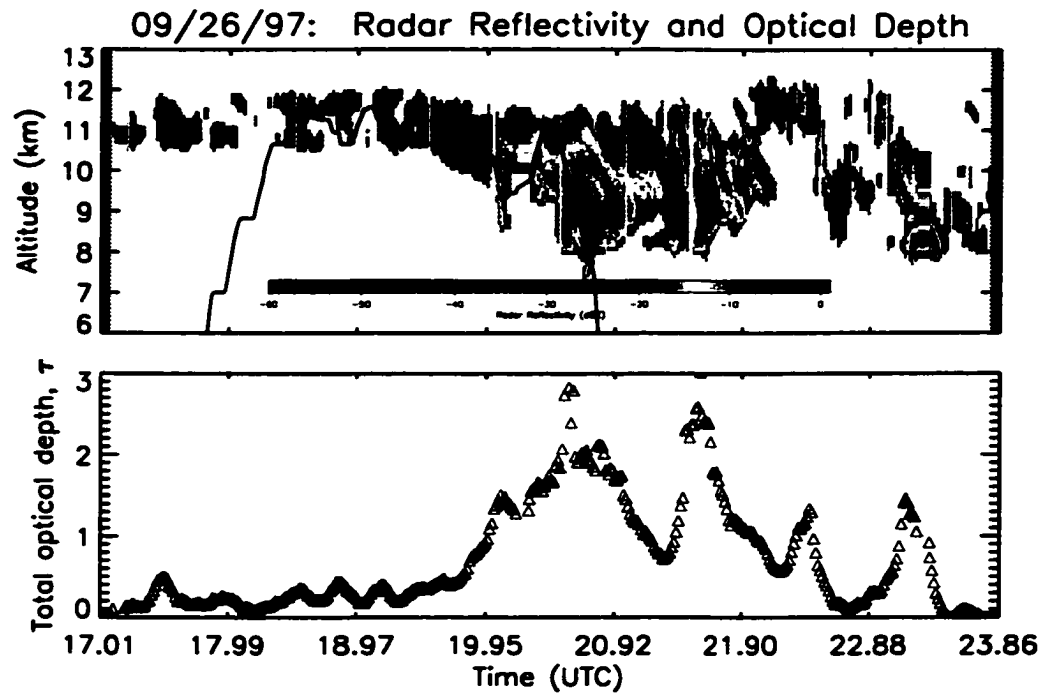


Figure B.1: Radar reflectivity and total optical depth time series for a cirrus cloud observed over the ARM SGP site, Oklahoma, on September, 26 during the Fall 1997 Cloud IOP.

and B.9 and B.10).

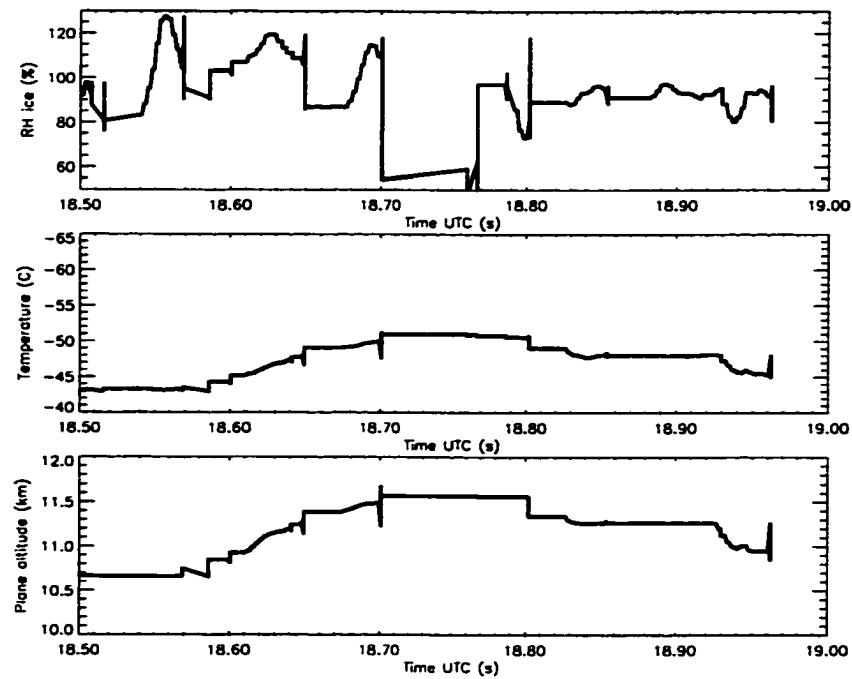


Figure B.2: Relative humidity with respect to ice (upper panel), temperature (middle panel), and plane altitude (lower panel) recorded between 17:30 and 17:48 UTC on board of the UN Citation on May 8, 1998.

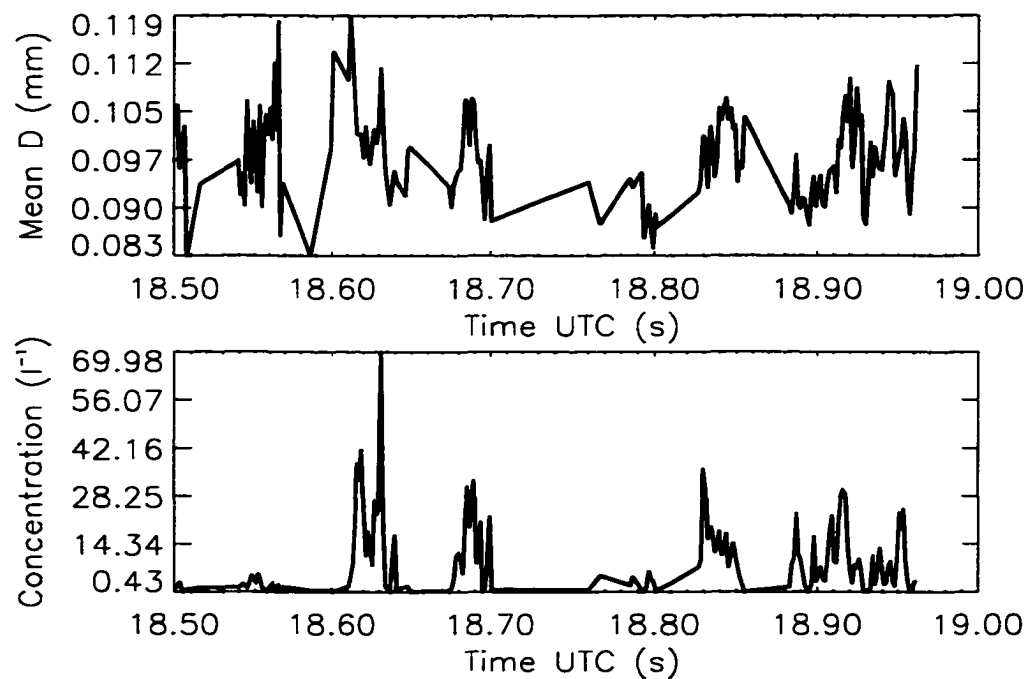


Figure B.3: Ice particle mean size (upper panel), and number concentration (lower panel) measured between 18:30 and 19:00 UTC by the 2D-probe on board of the UN Citation on September. 26 1997.

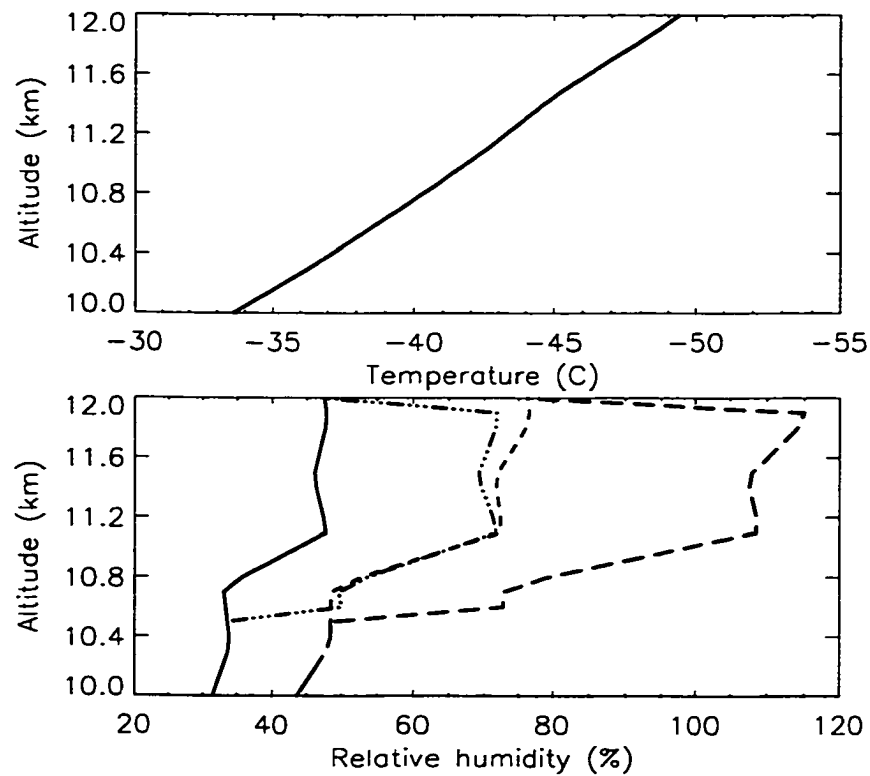


Figure B.4: Atmospheric state variable profiles between 10 and 12 km. Temperature is shown in the upper panel. Relative humidity (RH) is shown in the lower panel. The various profiles represents respectively: original RH wrt water (solid line), original RH wrt ice (short dashed), increased RH wrt water (dot-dashed), and increased RH wrt ice (long dashed).

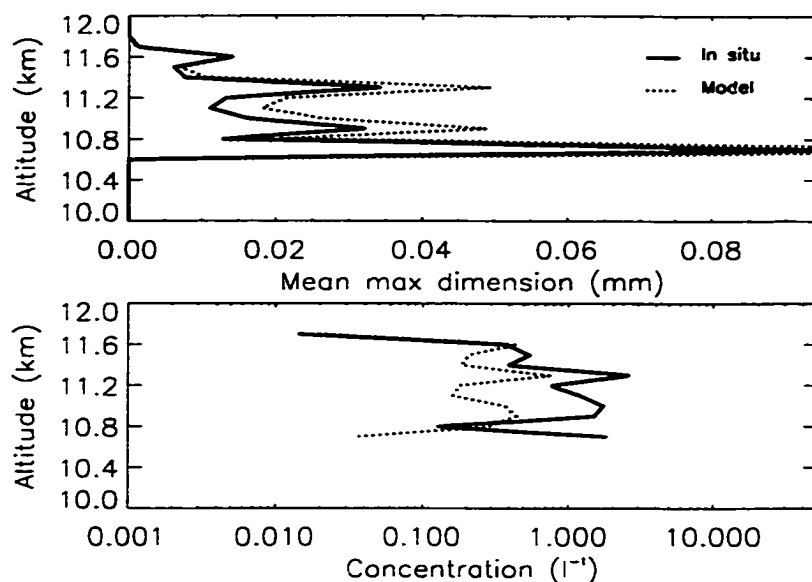


Figure B.5: Comparison between observed (solid line) and model number concentrations and mean sizes (dashed line). Meyers et al. (1992) initialization for N_t and constant D_n ($100\mu m$) initialization for D_n . Increased value of RH wrt water (see text for explanations).

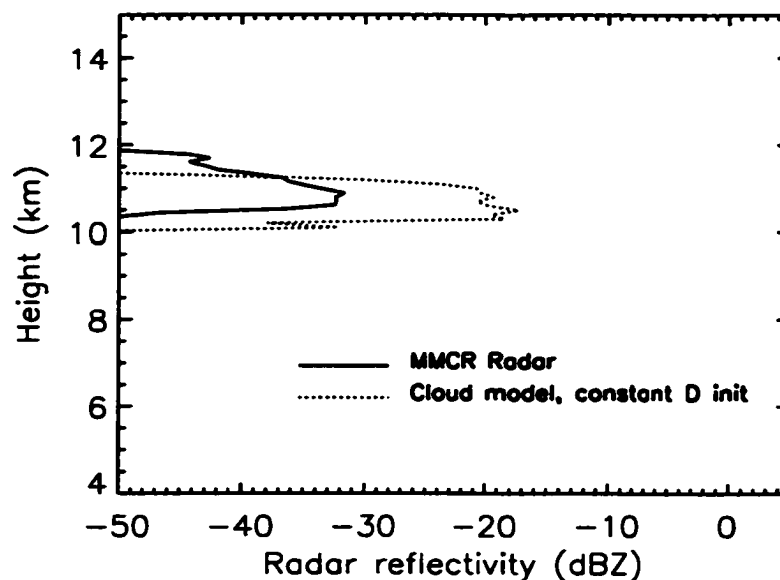


Figure B.6: Same as figure B.5, but for radar reflectivities.

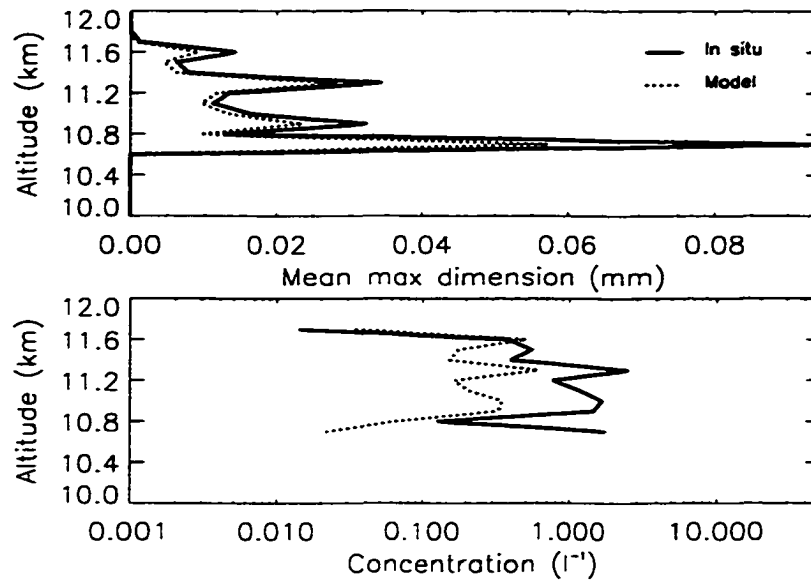


Figure B.7: Same as figure B.5, but for initial value of $D_n = 50\mu m$.

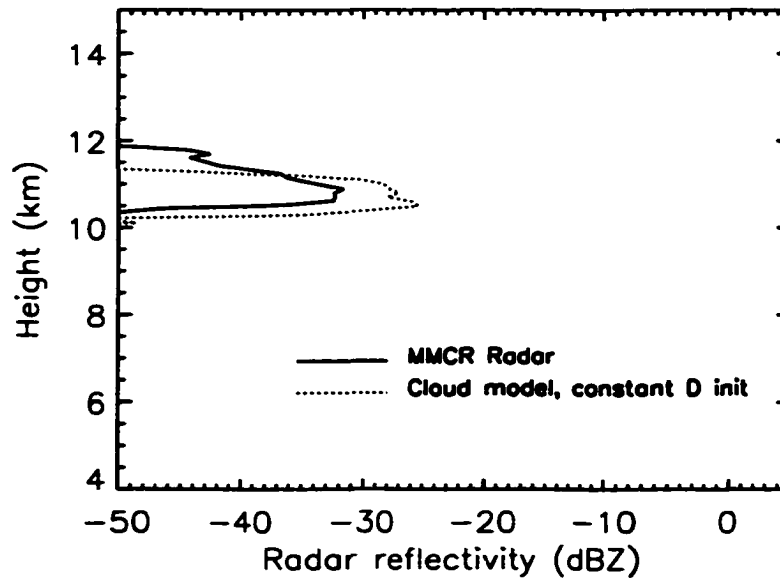


Figure B.8: Same as figure B.6, but for initial value of $D_n = 50\mu m$.

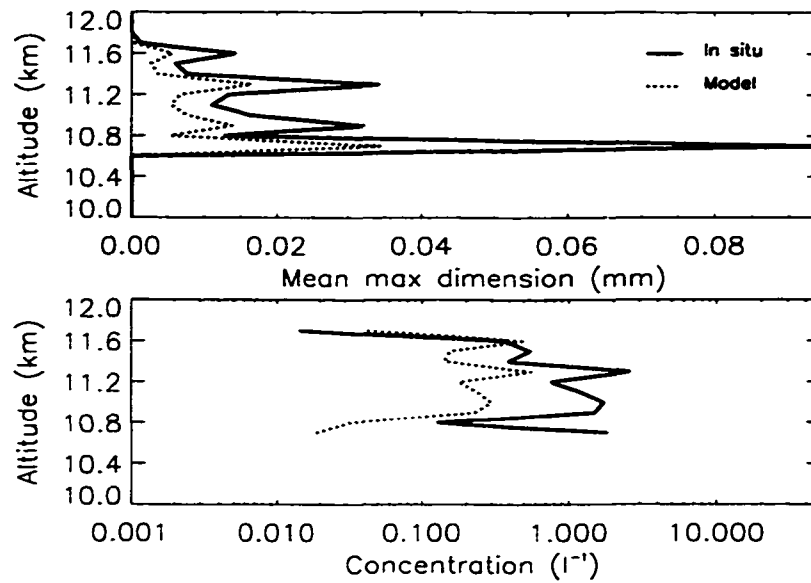


Figure B.9: Same as figure B.5, but for ECMWF initialization for D_n .

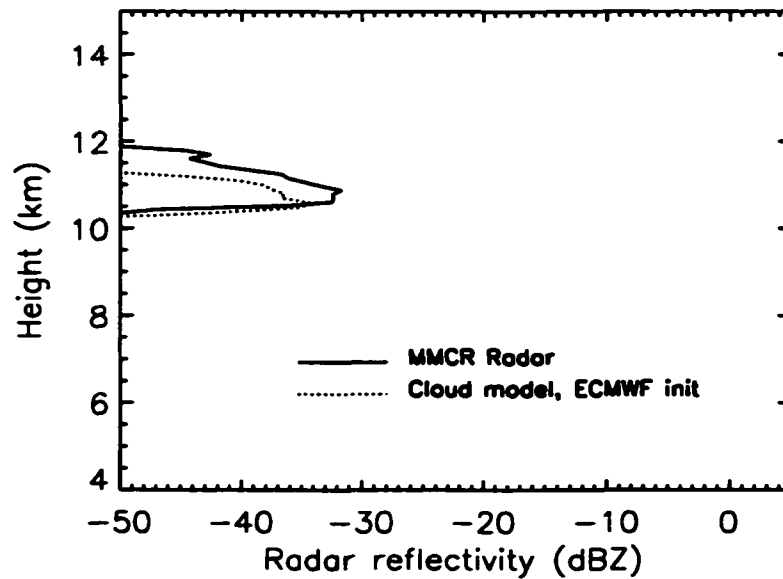


Figure B.10: Same as figure B.6, but for ECMWF initialization for D_n .

Appendix C

Properties of reflected sunlight derived from a Green's function method¹

C.1 Abstract

The inference of optical depth and particle size of clouds and aerosols using remotely sensed reflected radiance at solar wavelengths has received much attention recently. The information these measurements provide is path integrated. However, very little is known about the vertical distribution of this weighting. To characterize it, we first solve the Radiative Transfer Equation (RTE) by a Green's function approach, and then investigate the sensitivity of the weighting to vertical inhomogeneities in the extinction by introducing a function that is closely related to the Green's function, herein called the Contribution Function. This function calculates the contributions to the radiance at the upper boundary of the medium by underlying layers. Three hypothetical clouds of identical optical depth but exhibiting different extinction profiles were used in this study. The Contribution Function was found very sensitive to the extinction profile. The global reflection and transmission matrices used to construct the Green's function, derived using an eigenmatrix method, resulted in an efficient, stable, and accurate method for calculating the emerging radiances that can be extended to multi-layered media.

C.2 Introduction

Measurements of sunlight scattered in the Earth's atmosphere are used to infer information about cloud and aerosol particles as well as information about the abundance

¹To appear in *Journal of Quantitative Spectroscopy and Radiative Transfer*, 2001.

of trace gases in the atmosphere. Retrieval methods developed to infer optical depth and particle size of clouds and aerosol using observations of solar reflectance typically rely on measurements in two or sometimes more spectral regions or bands (e.g. Nakajima and King (1990) and Platnick (2000)). This approach yields path integrated information that is weighted in some way according to the distribution of particles along the path. The weighting accounts for the vertical transport through plane parallel layers in a manner that quantifies the relative information content of each layer to the overall reflected or transmitted signal. A description of the nature of this path weighting of information has been given by Platnick (2000).

The present work extends and generalizes his characterization of the weighting by using a Green's function formulation to calculate contributions to the radiance emerging at the upper boundary of a vertically inhomogeneous medium with a constant single scattering albedo. This solution provides, in a straightforward manner, a way of investigating the sensitivity of these contributions to the prescribed inhomogeneities (such as the vertical variations of the extinction or scattering functions) as a continuous function of the vertical coordinate.

Although many different techniques have been developed to solve the radiative transfer equation (Lenoble, 1985), the Doubling-Adding method (DA) (van de Hulst (1962) and Irvine (1975)) and the eigenmatrix formulation (e.g. Discrete Ordinate Radiative Transfer, DISORT (Stamnes et al., 1988)) are the most commonly used methods for dealing with vertically inhomogeneous media.

The eigenmatrix method solves the RTE by first discretizing the angular variation in the phase function and radiance, and then treating the resulting system of equations as a two-point BVP (Boundary Value Problem). For vertically uniform optical properties, eigenmatrix methods are used to construct its solution. This method dates back to Chandrasekar (1960). The DA method solves the RTE by converting the two-point BVP into an IVP (Initial Value Problem) via a Riccati transformation; this introduces the concept of global reflection (\mathcal{R}) and global transmission (\mathcal{T}) operators, that account for multiple reflections in a layer. For a single layer, homogenous in its optical properties, the DA method is reduced to simple doubling. For multilayered media, the interaction principle is used to compute the

global reflection and transmission matrices for the whole medium via the adding technique, once \mathcal{R} and \mathcal{T} for all layers are known.

The connection between the eigenmatrix and the DA method of solution was described by Waterman (1981). In that work he used the interaction principle to relate the matrix exponential to the global reflection and transmission operators. However, the computation of global \mathcal{R} and \mathcal{T} were approximated for thin ($\tau < 1$) media, due to numerical instabilities encountered in thicker media. In this paper we reformulate Waterman's work by calculating \mathcal{R} and \mathcal{T} without approximation via the eigenmatrix approach in a manner that yields a numerically stable solution for any arbitrary optical depth. Subsequently these stable forms for \mathcal{R} and \mathcal{T} are used to compute the Green's function, from which the goals stated above can be attained. This approach has received some attention in astrophysics (Kalkofen, 1985) and in hydrologic optics (Mobley, 1995).

The outline of the paper is as follows. Section C.3 introduces a stable semi-analytical form for global \mathcal{R} and \mathcal{T} derived from the matrix exponential solution of the sourceless RTE. The solar source function is then added and the Green's function is derived in terms of the global \mathcal{R} and \mathcal{T} . Section C.4 explores the numerical solution of the azimuthally dependent radiances for a single layer medium using the methods described in section C.3. Results are compared against a standard doubling code. The *Contribution Function* and its integral, the *Integrated Contribution Function*, are also introduced using the Green's function solution. We show vertical profiles of these functions and discuss their implications toward inferring extinction profiles from reflected solar radiation. This issue is further addressed by exploring the extent to which different portions of the medium contribute to the observed radiances at the boundary. This is quantified in section C.5 via the simple concept of *penetration optical depth*. A summary and conclusion are provided in section C.6 along with a discussion of the physical significance of these results to the inverse problem of radiative transfer.

C.3 Theoretical Foundations

C.3.1 Basic concepts

This study is posed in terms of the radiative transfer through a single isolated layer of an absorbing and scattering medium that possesses a prescribed vertical variation of optical properties. The layer will be taken to extend vertically from $z = 0$ to $z = H$. The extension to a multi-layered medium is possible by a repeated application of the interaction principle for each layer, provided that the single scattering albedo ω_0 is constant in every layer. The solution for the i^{th} layer is propagated as a boundary condition for the $(i + 1)^{\text{th}}$ layer in exactly the same manner as in the DA method. The resulting form of the solution will be invariant with respect to the number of layers.

Discretization of the Radiative Transfer Equation

Under the assumption that the single layer optical medium is horizontally homogeneous, the transfer of monochromatic diffuse radiation through the layer is governed by the plane-parallel equation of radiative transfer:

$$\begin{aligned} \mu \frac{dI(z, \mu, \phi)}{dz} = & - \sigma_e(z)I(z, \mu, \phi) \\ & + \frac{\sigma_s(z)}{4\pi} \int_0^{2\pi} \int_{-1}^1 P(z, \mu, \phi, \mu', \phi') I(z, \mu', \phi') d\mu' d\phi' + \Sigma(z, \mu, \phi), \end{aligned} \quad (\text{C.1})$$

where $I(z, \mu, \phi)$ is the specific intensity at level z along the direction specified by μ (cosine of the polar angle) and ϕ (azimuth angle), σ_e is the extinction function, σ_s the scattering function, $P(z, \mu, \phi, \mu', \phi')$ is the phase function, and $\Sigma(z, \mu, \phi)$ is the source.

To approximate the solution to (C.1) we discretize the integro-differential equation and then solve the resulting boundary value problem (BVP) with the boundary condition given in section C.4.1 by combining eigenmatrix and interaction principle

methods. With the radiance written as

$$I(z, \mu, \phi) = \sum_{m=0}^M I_m(z, \mu) \cos[m(\phi - \phi_{\odot})], \quad (\text{C.2})$$

where ϕ_{\odot} is the solar azimuth angle, and with similar expansions for the source term and the phase function in conjunction with equation (C.2), we obtain a system of M *uncoupled* equations, one for each term of the azimuthal expansion. Thus the solution of (C.1) reduces to the solution of m BVPs. The radiance field can be divided into N upwelling and N downwelling *streams* for each mode m , producing radiance vector pairs $I_m^+(z)$ and $I_m^-(z)$. The RTE equation thus reduces to its fully discrete matrix form:

$$\frac{d}{dz} \begin{pmatrix} I_m^+(z) \\ I_m^-(z) \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{t}}_m(z) & -\hat{\mathbf{r}}_m(z) \\ \hat{\mathbf{r}}_m(z) & -\hat{\mathbf{t}}_m(z) \end{pmatrix} \begin{pmatrix} I_m^+(z) \\ I_m^-(z) \end{pmatrix} + \begin{pmatrix} \hat{\Sigma}_m^+(z) \\ \hat{\Sigma}_m^-(z) \end{pmatrix}, \quad (\text{C.3})$$

where $\hat{\mathbf{t}}_m(z)$ and $\hat{\mathbf{r}}_m(z)$ represent the local transmission and reflection matrices

$$\hat{\mathbf{t}}_m(z) = -\sigma_e(z)\mathbf{M} + \frac{(\mathbf{I} + \delta_{0,m})}{4}\sigma_s(z)\mathbf{M}\mathbf{P}_m^+\mathbf{W}, \quad (\text{C.4})$$

$$\hat{\mathbf{r}}_m(z) = -\frac{(\mathbf{I} + \delta_{0,m})}{4}\sigma_s(z)\mathbf{M}\mathbf{P}_m^-\mathbf{W}. \quad (\text{C.5})$$

\mathbf{M} is the diagonal matrix whose elements are the inverse of cosines of the quadrature angles, \mathbf{W} is the diagonal matrix whose elements are the quadrature weights and \mathbf{P}_m^+ and \mathbf{P}_m^- are the forward (+) and backward (-) phase function matrices, specified for a given m . \mathbf{I} is the unit matrix.

The source of diffuse radiation at solar wavelength can be written as

$$\hat{\Sigma}_m^{\pm}(z) = \frac{\sigma_s(z)}{4\pi} F_{\odot} \mathbf{M} \mathbf{P}_{\odot m}^{\mp} e^{-\int_z^H \sigma_e(z') dz' / \mu_{\odot}}, \quad (\text{C.6})$$

where $\mathbf{P}_{\odot m}^{\mp}$ are the forward and backward phase function vectors associated with scattering of the direct solar beam. F_{\odot} represents the monochromatic irradiance incident at the top of the layer and the exponential factor is the attenuation of this direct beam to the level z of interest.

Solution of the RTE for a vertically inhomogeneous layer

The usual method of solving (C.3) for a vertically varying optical medium is to divide the atmosphere into a number of distinct but vertically uniform layers. We develop an alternative solution to (C.3) for given specific analytical forms of the vertical variation of the extinction. The solution proceeds under the assumption that the single scattering albedo is constant through the layer. This is reasonable for wavelengths in the visible and near-infrared region of the spectrum.

The profile of the extinction function is defined as $\sigma_e(z) = \frac{d\psi(z)}{dz}$ and by $\sigma_s(z) = \omega_o \frac{d\psi(z)}{dz}$ where ω_o is the single scattering albedo. The condition $\sigma_e(z) > 0, \forall z$ must hold. This specific form of the extinction was chosen to simplify the integration of the exponential solution described below. We now redefine the source term and the local transmission and reflection matrices as

$$\hat{\Sigma}_m(z) = \frac{d\psi(z)}{dz} \Sigma_m(z), \quad (\text{C.7})$$

$$\hat{\mathbf{t}}_m(z) = \frac{d\psi(z)}{dz} \mathbf{t}_m, \quad (\text{C.8})$$

$$\hat{\mathbf{r}}_m(z) = \frac{d\psi(z)}{dz} \mathbf{r}_m, \quad (\text{C.9})$$

where \mathbf{t}_m and \mathbf{r}_m are the local reflection and transmission matrices defined for constant coefficients, and $\Sigma_m^\pm(z)$ is given by

$$\Sigma_m^\pm(z) = \frac{\omega_o}{4\pi} F_\odot \mathbf{M} \mathbf{P}_{\odot m}^\mp e^{-\int_z^H \sigma_e(z') dz' / \mu_\odot}. \quad (\text{C.10})$$

Using equations (C.7)–(C.10) we now rewrite equation (C.3):

$$\frac{d}{dz} \begin{pmatrix} \mathbf{l}_m^+(z) \\ \mathbf{l}_m^-(z) \end{pmatrix} = \frac{d\psi(z)}{dz} \mathbf{A} \begin{pmatrix} \mathbf{l}_m^+(z) \\ \mathbf{l}_m^-(z) \end{pmatrix} + \frac{d\psi(z)}{dz} \begin{pmatrix} \Sigma_m^+(z) \\ \Sigma_m^-(z) \end{pmatrix}, \quad (\text{C.11})$$

where we have introduced the $(2N \times 2N)$ -matrix:

$$\mathbf{A}_m = \begin{pmatrix} \mathbf{t}_m & -\mathbf{r}_m \\ \mathbf{r}_m & -\mathbf{t}_m \end{pmatrix}. \quad (\text{C.12})$$

The integral solution to (C.11) thus follows as

$$\begin{pmatrix} I_m^+(H) \\ I_m^-(H) \end{pmatrix} = e^{\mathbf{A}_m \psi(H)} \begin{pmatrix} I_m^+(0) \\ I_m^-(0) \end{pmatrix} + \int_0^H e^{\mathbf{A}_m [\psi(H) - \psi(z)]} \begin{pmatrix} \Sigma_m^+ \\ \Sigma_m^- \end{pmatrix} \frac{d\psi(z)}{dz} dz, \quad (\text{C.13})$$

where $I_m^\pm(0)$ are the boundary conditions, and $\psi(z)$ represents the integral with respect to z of the extinction coefficient. $\psi(H)$ is the total optical depth, which in remote sensing literature is also indicated with the symbol τ . In the following discussion, the two symbols are interchangeable.

To simplify the notation, hereafter the subscript m is dropped and all dependence on azimuth mode is understood.

Properties of the exponential matrix

A key step in evaluating (C.13) is the evaluation of the exponential matrix $e^{\mathbf{A}\psi(z)}$, and for this purpose it is convenient to *diagonalize* \mathbf{A} . From linear algebra we know that a square matrix \mathbf{A} gives rise to the eigenvalue problem

$$\mathbf{A}\mathbf{X} = \mathbf{X}\Lambda, \quad (\text{C.14})$$

where \mathbf{X} is the matrix whose columns are the *eigenvectors* of \mathbf{A} and Λ is the diagonal matrix whose elements are the *eigenvalues* of \mathbf{A} . It is well known that the eigenvalues of the \mathbf{A} matrix are real and come in pairs $(\pm\lambda_k, k = 1, \dots, N)$ (?).

From equation (C.14), it follows that

$$\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}. \quad (\text{C.15})$$

By expanding the matrix exponential in a Taylor series of the variable $\psi(z)$, we have:

$$e^{\mathbf{A}\psi(z)} = \mathbf{X}e^{\Lambda\psi(z)}\mathbf{X}^{-1}, \quad (\text{C.16})$$

where $e^{\Lambda\psi(z)}$ is a diagonal matrix. The solution of the RTE is thus reduced to computing the eigenvalues and eigenvectors of \mathbf{A} . As is known, this task poses numerical difficulties, but the form of \mathbf{A} can be exploited to reduce the order of the

problem from $2N$ to N by applying the technique of *deflation of polynomial degree* (Stamnes and Conklin, 1984). This not only reduces the computational time, but offers insight into the underlying structure of the exponential matrix itself. Details of this approach applied to the matrix \mathbf{A} are presented in appendix A.

Using the spectral decomposition of \mathbf{A} and the two matrices

$$\mathbf{X} = \begin{pmatrix} \mathbf{u}_+ & \mathbf{u}_- \\ -\mathbf{u}_- & -\mathbf{u}_+ \end{pmatrix},$$

and

$$\mathbf{X}^{-1} = \begin{pmatrix} \mathbf{v}_+ & \mathbf{v}_- \\ -\mathbf{v}_- & -\mathbf{v}_+ \end{pmatrix},$$

where from appendix A we have

$$\begin{aligned} \mathbf{v}_+ &= [\mathbf{I} - (\mathbf{u}_+^{-1}\mathbf{u}_-)^2]^{-1}\mathbf{u}_+^{-1} \\ \mathbf{v}_- &= (\mathbf{u}_+^{-1}\mathbf{u}_-)\mathbf{v}_+, \end{aligned}$$

we can write the exponential matrix as follows:

$$\begin{aligned} \mathbf{X}e^{\Lambda\psi(z)}\mathbf{X}^{-1} &= \begin{pmatrix} \mathbf{u}_+ & \mathbf{u}_- \\ -\mathbf{u}_- & -\mathbf{u}_+ \end{pmatrix} \begin{pmatrix} e^{\Lambda^+\psi(z)} & \mathbf{0} \\ \mathbf{0} & e^{-\Lambda^+\psi(z)} \end{pmatrix} \begin{pmatrix} \mathbf{v}_+ & \mathbf{v}_- \\ -\mathbf{v}_- & -\mathbf{v}_+ \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{e}_{11}(z) & \mathbf{e}_{12}(z) \\ \mathbf{e}_{21}(z) & \mathbf{e}_{22}(z) \end{pmatrix} \end{aligned} \quad (\text{C.17})$$

where

$$\mathbf{e}_{11}(z) = \mathbf{u}_+e^{\Lambda^+\psi(z)}\mathbf{v}_+ - \mathbf{u}_-e^{-\Lambda^+\psi(z)}\mathbf{v}_-, \quad (\text{C.18})$$

$$\mathbf{e}_{12}(z) = \mathbf{u}_+e^{\Lambda^+\psi(z)}\mathbf{v}_- - \mathbf{u}_-e^{-\Lambda^+\psi(z)}\mathbf{v}_+, \quad (\text{C.19})$$

$$\mathbf{e}_{21}(z) = -\mathbf{u}_-e^{\Lambda^+\psi(z)}\mathbf{v}_+ + \mathbf{u}_+e^{-\Lambda^+\psi(z)}\mathbf{v}_-, \quad (\text{C.20})$$

$$\mathbf{e}_{22}(z) = -\mathbf{u}_-e^{\Lambda^+\psi(z)}\mathbf{v}_- + \mathbf{u}_+e^{-\Lambda^+\psi(z)}\mathbf{v}_+. \quad (\text{C.21})$$

Here Λ^+ is the diagonal matrix corresponding to the *positive* eigenvalues of \mathbf{A} .

Thus the elements of the eigenmatrix are linear combinations of decaying and growing exponentials. The growing exponentials are a source of numerical instability when $\psi(z)$ becomes large. This instability can be circumvented by using the interaction form of the solution to (C.11) which requires the specification of the global reflection and transmission matrices.

C.3.2 Global Reflection and Transmission matrices

Recasting the eigenmatrix solution into the form of the interaction principle leads to a direct derivation of the global transmission and reflection operators. Using the interaction principle for a single layer of depth H , and in the absence of sources, we can write:

$$\begin{aligned} \mathbf{l}^+(H) &= \mathcal{T}(0, H)\mathbf{l}^+(0) + \mathcal{R}(H, 0)\mathbf{l}^-(H) \\ \mathbf{l}^-(0) &= \mathcal{R}(0, H)\mathbf{l}^+(0) + \mathcal{T}(H, 0)\mathbf{l}^-(H), \end{aligned} \quad (\text{C.22})$$

where $\mathcal{T}(0, H)$ and $\mathcal{R}(0, H)$ are the global transmission and reflection functions for illumination from below, and $\mathcal{T}(H, 0)$ and $\mathcal{R}(H, 0)$ are for illumination from above.

From the homogeneous solution of the RTE (first term on right hand side of equation (C.13)) we have:

$$\begin{aligned} \begin{pmatrix} \mathbf{l}^+(H) \\ \mathbf{l}^-(H) \end{pmatrix} &= \mathbf{X}e^{\Lambda\psi(H)}\mathbf{X}^{-1} \begin{pmatrix} \mathbf{l}^+(0) \\ \mathbf{l}^-(0) \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{e}_{11}(H) & \mathbf{e}_{12}(H) \\ \mathbf{e}_{21}(H) & \mathbf{e}_{22}(H) \end{pmatrix} \begin{pmatrix} \mathbf{l}^+(0) \\ \mathbf{l}^-(0) \end{pmatrix}, \end{aligned} \quad (\text{C.23})$$

which can be rewritten as

$$\begin{aligned} \mathbf{l}^+(H) &= \mathbf{e}_{11}(H)\mathbf{l}^+(0) + \mathbf{e}_{12}(H)\mathbf{l}^-(0) \\ \mathbf{l}^-(H) &= \mathbf{e}_{21}(H)\mathbf{l}^+(0) + \mathbf{e}_{22}(H)\mathbf{l}^-(0). \end{aligned} \quad (\text{C.24})$$

Solving for $l^-(0)$ and $l^+(H)$, gives:

$$\begin{aligned} l^+(H) &= [\mathbf{e}_{11}(H) - \mathbf{e}_{12}(H)\mathbf{e}_{22}^{-1}(H)\mathbf{e}_{21}(H)]l^+(0) + \\ &\quad \mathbf{e}_{12}(H)\mathbf{e}_{22}^{-1}(H)l^-(H) \\ l^-(0) &= \mathbf{e}_{22}^{-1}(H)l^-(H) - \mathbf{e}_{22}^{-1}(H)\mathbf{e}_{21}(H)l^+(0). \end{aligned} \quad (\text{C.25})$$

Comparing equation (C.22) to (C.25) we establish that:

$$\mathcal{T}(0, H) = \mathbf{e}_{11}(H) - \mathbf{e}_{12}(H)\mathbf{e}_{22}^{-1}(H)\mathbf{e}_{21}(H), \quad (\text{C.26})$$

$$\mathcal{R}(H, 0) = \mathbf{e}_{12}(H)\mathbf{e}_{22}^{-1}(H), \quad (\text{C.27})$$

$$\mathcal{T}(H, 0) = \mathbf{e}_{22}^{-1}(H), \quad (\text{C.28})$$

$$\mathcal{R}(0, H) = -\mathbf{e}_{22}^{-1}(H)\mathbf{e}_{21}(H). \quad (\text{C.29})$$

Thus the global reflection and transmission matrices are completely determined from linear combinations of elements of the eigenmatrix. In principle, these matrices are stable over all ranges of optical depth. However, the form of equations (C.26–C.29) remains unstable. The steps toward achieving a stable form of \mathcal{R} and \mathcal{T} are discussed in appendix B. With the development described there, the stable forms of these matrices are:

$$\begin{aligned} \mathcal{T}(H, 0) &= -\mathbf{u}_+[\mathbf{I} - (\mathbf{u}_+^{-1}\mathbf{u}_-)^2][(\mathbf{u}_+^{-1}\mathbf{u}_-)^{-1}e^{-\Lambda^+\psi(H)}] \\ &\quad \{\mathbf{I} - [(\mathbf{u}_+^{-1}\mathbf{u}_-)^{-1}e^{-\Lambda^+\psi(H)}]^2\}^{-1}\mathbf{u}_-^{-1} \end{aligned} \quad (\text{C.30})$$

and

$$\begin{aligned} \mathcal{R}(H, 0) &= -\mathbf{u}_+[\mathbf{I} - (\mathbf{u}_+^{-1}\mathbf{u}_-)e^{-\Lambda^+\psi(H)}(\mathbf{u}_+^{-1}\mathbf{u}_-)^{-1}e^{-\Lambda^+\psi(H)}] \\ &\quad \{\mathbf{I} - [(\mathbf{u}_+^{-1}\mathbf{u}_-)^{-1}e^{-\Lambda^+\psi(H)}]^2\}^{-1}\mathbf{u}_-^{-1}. \end{aligned} \quad (\text{C.31})$$

For a layer of constant ω_o , even though $\sigma_e(z)$ and $\sigma_s(z)$ are varying as assumed in this work, $\mathcal{R}(H, 0) = \mathcal{R}(0, H)$ and $\mathcal{T}(H, 0) = \mathcal{T}(0, H)$.

A direct numerical comparison with the global reflection and transmission matrices derived using doubling showed that both \mathcal{R} and \mathcal{T} , calculated using single precision, are stable up to a total optical depth of 1000. This technique allows for the evaluation of global \mathcal{T} and \mathcal{R} for the entire layer without any approximation or first order expansion. When the source terms are taken into account, the radiance exiting the top of the layer can be computed efficiently and accurately.

C.3.3 Derivation of the Green's function matrix

In this section we will use the interaction principle and the global reflection and transmission matrices developed above to compute the Green's function. The solution to the RTE equation with sources may be written as

$$\begin{aligned} \begin{pmatrix} I^+(H) \\ I^-(H) \end{pmatrix} &= \mathbf{X}e^{\Lambda\psi(H)}\mathbf{X}^{-1} \begin{pmatrix} I^+(0) \\ I^-(0) \end{pmatrix} + \begin{pmatrix} J^+ \\ J^- \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{e}_{11}(H) & \mathbf{e}_{12}(H) \\ \mathbf{e}_{21}(H) & \mathbf{e}_{22}(H) \end{pmatrix} \begin{pmatrix} I^+(0) \\ I^-(0) \end{pmatrix} + \begin{pmatrix} J^+ \\ J^- \end{pmatrix}, \end{aligned} \quad (\text{C.32})$$

where the source term is given by the following integral:

$$J^\pm = \int_0^H e^{\mathbf{A}[\psi(H)-\psi(z)]} \begin{pmatrix} \Sigma^+ \\ \Sigma^- \end{pmatrix} \frac{d\psi(z)}{dz} dz. \quad (\text{C.33})$$

If we rewrite the solution in the interaction principle form, and substitute the global \mathcal{R} and \mathcal{T} matrices, we have:

$$\begin{aligned} \begin{pmatrix} I^+(H) \\ I^-(0) \end{pmatrix} &= \begin{pmatrix} \mathcal{T}(0, H) & \mathcal{R}(H, 0) \\ \mathcal{R}(0, H) & \mathcal{T}(H, 0) \end{pmatrix} \begin{pmatrix} I^+(0) \\ I^-(H) \end{pmatrix} \\ &+ \begin{pmatrix} \mathbf{I} & -\mathcal{R}(H, 0) \\ \mathbf{0} & -\mathcal{T}(H, 0) \end{pmatrix} \begin{pmatrix} J^+ \\ J^- \end{pmatrix}. \end{aligned} \quad (\text{C.34})$$

Since global \mathcal{R} and \mathcal{T} do not depend on z , we can bring the matrix inside the integral that defines the source term and proceed to multiply the two matrices:

$$\begin{aligned}
& \begin{pmatrix} \mathbf{I} & -\mathcal{R}(H, 0) \\ \mathbf{0} & -\mathcal{T}(H, 0) \end{pmatrix} e^{\mathbf{A}(\psi(H)-\psi(z))} = \\
& \begin{pmatrix} \mathbf{I} & -\mathcal{R}(H, 0) \\ \mathbf{0} & -\mathcal{T}(H, 0) \end{pmatrix} \begin{pmatrix} \mathbf{e}_{11}(H-z) & \mathbf{e}_{12}(H-z) \\ \mathbf{e}_{21}(H-z) & \mathbf{e}_{22}(H-z) \end{pmatrix} = \\
& \begin{pmatrix} \mathbf{e}_{11}(H-z) - \mathcal{R}(H, 0)\mathbf{e}_{21}(H-z) & \mathbf{e}_{12}(H-z) - \mathcal{R}(H, 0)\mathbf{e}_{22}(H-z) \\ -\mathcal{T}(H, 0)\mathbf{e}_{21}(H-z) & -\mathcal{T}(H, 0)\mathbf{e}_{22}(H-z) \end{pmatrix} = \\
& \begin{pmatrix} \mathbf{G}_{++}(H-z) & \mathbf{G}_{+-}(H-z) \\ \mathbf{G}_{-+}(H-z) & \mathbf{G}_{--}(H-z) \end{pmatrix} = \mathbf{G}(H-z). \tag{C.35}
\end{aligned}$$

Expression (C.35) represents the Green's function matrix for the RT problem. Since it contains \mathbf{e}_{11} , \mathbf{e}_{12} , \mathbf{e}_{21} , and \mathbf{e}_{22} , its form is numerically unstable for large optical depths. However if we consider equation (C.35), the Green's function matrix can be subdivided into four submatrices which contain decaying and growing exponentials. As before, we can manipulate each of these submatrices separately in an effort to induce stability.

Beginning with the top left corner submatrix $\mathbf{G}_{++}(H-z)$, if we write $\mathbf{e}_{11}(H-z)$ and $\mathbf{e}_{21}(H-z)$ explicitly we have:

$$\begin{aligned}
& \mathbf{G}_{++}(H-z) = \\
& \mathbf{e}_{11}(H-z) - \mathcal{R}(H, 0)\mathbf{e}_{21}(H-z) = \\
& \mathbf{u}_+ e^{\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_+ - \mathbf{u}_- e^{-\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_- \\
& - \mathcal{R}(H, 0)(-\mathbf{u}_- e^{\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_+ + \mathbf{u}_+ e^{-\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_-) \\
& = \mathbf{u}_+ e^{\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_+ + \mathcal{R}(H, 0)\mathbf{u}_- e^{\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_+ \\
& - \mathbf{u}_- e^{-\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_- - \mathcal{R}(H, 0)\mathbf{u}_+ e^{-\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_- \tag{C.36}
\end{aligned}$$

The reason for the regrouping of the matrices becomes evident by noting that the two terms which contains $e^{-\Lambda^+[\psi(H)-\psi(z)]}$ are inherently stable, since Λ^+ is a diagonal matrix of all positive elements and by definition $\psi(H) > \psi(z)$, $\forall z < H$. However, the two terms containing $e^{\Lambda^+[\psi(H)-\psi(z)]}$ are unstable.

We write these terms as

$$\begin{aligned} \mathbf{u}_+ e^{\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_+ + \mathcal{R}(H, 0) \mathbf{u}_- e^{\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_+ = \\ \mathbf{u}_+ [\mathbf{I} + \mathbf{u}_+^{-1} \mathcal{R}(H, 0) \mathbf{u}_-] e^{-\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_+ \end{aligned} \quad (\text{C.37})$$

If we use global \mathcal{R} derived in section C.3.2 in the above expression, we can proceed to compute a stable form for (C.36):

$$\begin{aligned} \mathbf{G}_{++}(H - z) = \\ \mathbf{u}_+ \{ -[\mathbf{I} - ((\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)})^2]^{-1} (\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)} \\ + (\mathbf{u}_+^{-1} \mathbf{u}_-) e^{-\Lambda^+ \psi(H)} [\mathbf{I} - ((\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)})^2]^{-1} \} (\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(z)} \mathbf{v}_+ \\ - [\mathbf{u}_- + \mathcal{R}(H, 0) \mathbf{u}_+] e^{-\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_- \end{aligned} \quad (\text{C.38})$$

This expression appears complex, but is numerically stable.

Similarly we can divide the submatrix in the top right corner of (C.35) $\mathbf{G}_{+-}(H - z)$, into a stable part and an unstable part and stabilize the latter, making use again of the computed \mathcal{R} :

$$\begin{aligned} \mathbf{G}_{+-}(H - z) = \\ \mathbf{u}_+ \{ -[\mathbf{I} - ((\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)})^2]^{-1} (\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)} \\ + (\mathbf{u}_+^{-1} \mathbf{u}_-) e^{-\Lambda^+ \psi(H)} [\mathbf{I} - ((\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)})^2]^{-1} \} (\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(z)} \mathbf{v}_- \\ - [\mathbf{u}_- + \mathcal{R}(H, 0) \mathbf{u}_+] e^{-\Lambda^+[\psi(H)-\psi(z)]} \mathbf{v}_+ \end{aligned} \quad (\text{C.39})$$

For the submatrix in the bottom left of (C.35), $\mathbf{G}_{-+}(H - z)$, we can apply a similar procedure, this time using \mathcal{J} and obtain:

$$\begin{aligned} \mathbf{G}_{-+}(H - z) = \\ -\mathbf{u}_+ [\mathbf{I} - (\mathbf{u}_+^{-1} \mathbf{u}_-)^2] [\mathbf{I} - ((\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)})^2]^{-1} (\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(z)} \\ [\mathbf{I} - e^{-\Lambda^+[\psi(H)-\psi(z)]} (\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+[\psi(H)-\psi(z)]} (\mathbf{u}_+^{-1} \mathbf{u}_-)] \mathbf{v}_+ \end{aligned} \quad (\text{C.40})$$

For the submatrix in the bottom right corner of (C.35), $\mathbf{G}_{--}(H - z)$:

$$\begin{aligned} \mathbf{G}_{--}(H - z) = & \\ & -\mathbf{u}_+[\mathbf{I} - (\mathbf{u}_+^{-1}\mathbf{u}_-)^2][\mathbf{I} - ((\mathbf{u}_+^{-1}(\mathbf{u}_+^{-1}\mathbf{u}_-)^{-1}\mathbf{u}_-)^{-1}e^{-\Lambda^+\psi(H)})^2]^{-1}(\mathbf{u}_+^{-1}\mathbf{u}_-)^{-1}e^{-\Lambda^+\psi(z)} \\ & [\mathbf{I} - ((\mathbf{u}_+^{-1}\mathbf{u}_-)^{-1}e^{-\Lambda^+[\psi(H)-\psi(z)]})^2]\mathbf{v}_+ \end{aligned} \quad (\text{C.41})$$

One thing to notice is that the above expressions have redundant matrix groups. For example $\mathbf{u}_+^{-1}\mathbf{u}_-$ appears several times, but needs to be calculated only once. Moreover \mathbf{v}_+ and \mathbf{v}_- can be expressed in terms of \mathbf{u}_+ and \mathbf{u}_- , through relations (C.65), derived in appendix A, that we repeat here for convenience:

$$\begin{aligned} \mathbf{v}_+ &= [\mathbf{I} - (\mathbf{u}_+^{-1}\mathbf{u}_-)^2]^{-1}\mathbf{u}_+^{-1} \\ \mathbf{v}_- &= (\mathbf{u}_+^{-1}\mathbf{u}_-)\mathbf{v}_+ \end{aligned}$$

These relations completely determine the Green's function matrix. In section C.4.2 we will apply it to compute the contribution to the upwelling radiance at the upper boundary of a scattering and absorbing medium by the lower layers.

C.4 Numerical Results

We present two applications of the results obtained in the previous sections. The first uses global reflection and transmission matrices to compute the radiance exiting the upper boundary of a medium via the interaction principle. This approach is computationally more efficient than doubling, particularly for optically thick media, yet achieves a similar degree of accuracy. A comparison of radiances computed with a doubling code, and the global \mathcal{R} and \mathcal{T} derived in the present work is summarized in section C.4.1.

The second numerical application pertains to the computation of the Contribution Function and the Integrated Contribution Function, using the Green's function solution, for different profiles of cloud properties. This application is relevant for

remote sensing problems where it is important to know the vertical path weighting of the information contained in the radiance measured at the top boundary. The Green's function that we introduced in section C.3.3 is the generalization of the weighting function commonly used in retrievals based on atmospheric thermal emission. In the case when scattering can be neglected and where internal thermal sources are present, the RTE reduces to its scalar form. The weighting function can then be simply recovered by assuming a given form for the extinction coefficient. This results in an analytical, exponential transmission function. When scattering is dominant, as in the case of reflected solar radiation, defining a weighting function is more complicated. The Green's function formulation for the solution of RTE is useful towards defining this weighting function as will be shown in section C.4.2.

C.4.1 Radiance calculations and comparison with the doubling method

Consider a single layer of total optical depth $\psi(H)$. Assume the source function is only given by the solar contribution and consider a direct solar beam incident from a direction $(\mu_{\odot}, \phi_{\odot})$ impinging at the top of the layer. Using equation (C.10) we can evaluate the integral given in (C.33) and obtain an analytical expression for J^{\pm} :

$$J^{\pm} = \int_0^H \frac{d\psi(z)}{dz} e^{\mathbf{A}[\psi(H)-\psi(z)]} \begin{pmatrix} \Sigma^+ \\ \Sigma^- \end{pmatrix} dz = -e^{\mathbf{A}\psi(H)} \mathbf{S}_1^{\pm} + \mathbf{S}_2^{\pm} \quad (\text{C.42})$$

where

$$\mathbf{S}_1^{\pm} = \left(\frac{\mathbf{I}}{\mu_{\odot}} - \mathbf{A} \right)^{-1} \frac{\omega_{\odot}}{4\pi} F_{\odot} \mathbf{M} \mathbf{P}_{\odot}^{\mp} e^{-\frac{\psi(H)}{\mu_{\odot}}},$$

and

$$\mathbf{S}_2^{\pm} = \left(\frac{\mathbf{I}}{\mu_{\odot}} - \mathbf{A} \right)^{-1} \frac{\omega_{\odot}}{4\pi} F_{\odot} \mathbf{M} \mathbf{P}_{\odot}^{\mp}.$$

According to equation (C.32), the solution for the radiance is given by:

$$\begin{pmatrix} I^+(H) \\ I^-(H) \end{pmatrix} = e^{\mathbf{A}\psi(H)} \begin{pmatrix} I^+(0) \\ I^-(0) \end{pmatrix} - e^{\mathbf{A}\psi(H)} \begin{pmatrix} \mathbf{S}_1^+ \\ \mathbf{S}_1^- \end{pmatrix} + \begin{pmatrix} \mathbf{S}_2^+ \\ \mathbf{S}_2^- \end{pmatrix} \quad (\text{C.43})$$

Rearranging the above expression using the interaction principle as in equation (C.34) we have:

$$\begin{aligned} I^+(H) &= \mathcal{T}(0, H)I^+(0) + \mathcal{R}(H, 0)I^-(H) \\ &\quad - \mathcal{T}(0, H)S_1^+ - \mathcal{R}(H, 0)S_2^- + S_2^+ \end{aligned} \quad (\text{C.44})$$

$$\begin{aligned} I^-(0) &= \mathcal{R}(0, H)I^+(0) + \mathcal{T}(H, 0)I^-(H) \\ &\quad - \mathcal{R}(0, H)S_1^+ + S_1^- - \mathcal{T}(H, 0)S_2^-. \end{aligned} \quad (\text{C.45})$$

The boundary conditions are:

$$I^+(0) = \mathbf{R}_g I^-(0) + A_g \frac{F_\odot}{\pi} e^{-\frac{\psi(H)}{\mu_\odot}} \quad (\text{C.46})$$

$$I^-(H) = 0, \quad (\text{C.47})$$

where we assume that there is no diffuse incoming radiation at the top of the layer and the surface reflection is specified by the matrix \mathbf{R}_g . For a Lambertian surface, \mathbf{R}_g has a very simple form with all equal rows, i.e.

$$\mathbf{R}_g = \frac{A_g}{\pi} \begin{pmatrix} w_1 \mu_1 & w_2 \mu_2 & \dots & w_N \mu_N \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix},$$

where A_g is the ground albedo. For the case of a generic surface, the matrix \mathbf{R}_g will have a more complicated structure, which has to be specified according to the reflecting properties of the surface. For the case of a completely opaque surface, $A_g = 0$. In what follows, we will use this assumption. We are aware of the fact that for optically thin clouds, surface albedo is a major problem in remote sensing using optical channels, and setting $A_g = 0$ is a simplifying assumption that cannot be made in practice. Nonetheless such assumption is applicable in the presence of thicker clouds or over opaque surfaces such as oceans, and results are still relevant for these cases.

The solution for the radiance field is:

$$\begin{aligned}
I^+(H) &= \mathcal{T}(0, H)\mathbf{R}_g I^-(0) + \mathcal{T}(0, H)A_g \frac{F_\odot}{\pi} e^{-\frac{\psi(H)}{\mu_\odot}} \\
&\quad - \mathcal{T}(0, H)S_1^+ - \mathcal{R}(H, 0)S_2^- + S_2^+ \tag{C.48}
\end{aligned}$$

$$\begin{aligned}
I^-(0) &= (\mathbf{I} - \mathcal{R}(0, H)\mathbf{R}_g)^{-1} (\mathcal{R}(0, H)A_g \frac{F_\odot}{\pi} e^{-\frac{\psi(H)}{\mu_\odot}} \\
&\quad - \mathcal{R}(0, H)S_1^+ + S_1^- - \mathcal{T}(H, 0)S_2^-). \tag{C.49}
\end{aligned}$$

Note that for each azimuthal component, m , of the radiance vector (see equation (C.2)) the global reflection and transmission matrices have to be re-evaluated, since they are also azimuthally-dependent. There is nothing new here since all the development in the previous sections can be applied for all terms of the azimuthal expansion, provided that eigenvalues and eigenvectors of the \mathbf{A} matrix are evaluated for all m .

The phase function used in this study is the Henyey–Greenstein, characterized by the asymmetry parameter g (Liou, 1980). Standard linear algebra routines (LAPACK, Linear Algebra PACKage, (Anderson et al., 1999)) were utilized to compute the eigenvalues and eigenvectors of the \mathbf{A} matrix. The radiances were computed for $N = 16$ streams, using Gaussian quadratures. The doubling–adding radiative transfer code used for the comparison of results is documented in Miller et al. (2000).

Radiances obtained with the global reflection and transmission have been compared against those from a doubling code, and agreement between these methods was as expected from numerical uncertainties (see Tables C.4.1.a and C.4.1.b).

There is, however, one exception to the very close agreement obtained. This applies to a particular set of parameters, corresponding to the case of low quadrature angles ($\mu < 0.09$) in conjunction with conservative scattering ($\omega_o \approx 1$), an optically thin layer ($\tau < 1$), and a highly forward peaked phase function ($g \approx 0.85$). All other combinations of optical parameters are not affected by this same problem; convergence of the azimuthal cosine series is obtained in most cases with a few expansion terms.

For optically thick media, the calculation of the global reflection and transmission matrices is more efficient than in the doubling method since the number of matrix multiplications is independent of the optical depth of the layer. The total

number of matrix multiplications, excluding the computation of the eigenvalues and eigenvectors, required \mathcal{R} and \mathcal{T} is 11, regardless of the optical depth of the layer. By comparison, for the doubling method to reach an optical depth of 16 starting from an optical depth $\Delta\tau = 0.001$, 28 matrix multiplication are required (14 for \mathcal{R} and 14 for \mathcal{T}), where this figure is derived from the relation $2^p\Delta\tau = \tau$, and in our example $p=14$. For this case the eigenmatrix approach will be faster than doubling by more than a factor of two.

We would like to also point out that the present method combines the advantages of the DA, via the introduction of the global reflection and transmission matrices, with the advantages of the eigenmatrix approach. In the case where the boundary conditions change, the new radiances exiting the boundaries of the medium can be computed in a straightforward manner using \mathcal{T} and \mathcal{R} . without solving a new BVP as it is required if DISORT is used to perform the calculation. On the other hand, if the optical properties of the layer change, the global reflection and transmission matrices can be recomputed in only 11 matrix multiplications, which will be faster than performing a doubling calculation (once again, this depends on the optical depth of the layer as pointed out in the previous paragraph).

The numerical results above pertain to a single layer atmosphere. Further investigation is necessary to evaluate quantitatively the advantages of the present method of solution over others in the case of a multi-layered medium.

C.4.2 Green's function results: Contribution Function and Integrated Contribution Function

In general, the Green's function solves the integral form of an ordinary differential equation, so we can use it to calculate the solution for the radiance exiting the boundaries of the layer. This application is, however, redundant, since more efficient methods to compute the solution to the RTE have been developed. In this section we apply the Green's function form of the solution in a new manner to establish a way of obtaining information about the vertical distribution of the radiance field.

To this end we return to equation (C.34) and express it in terms of the Green's

function matrix derived in that section:

$$\begin{pmatrix} I^+(H) \\ I^-(0) \end{pmatrix} = \begin{pmatrix} \mathcal{T}(0, H) & \mathcal{R}(H, 0) \\ \mathcal{R}(0, H) & \mathcal{T}(H, 0) \end{pmatrix} \begin{pmatrix} I^+(0) \\ I^-(H) \end{pmatrix} + \int_0^H \begin{pmatrix} \mathbf{G}_{++}(H-z) & \mathbf{G}_{+-}(H-z) \\ \mathbf{G}_{-+}(H-z) & \mathbf{G}_{--}(H-z) \end{pmatrix} \begin{pmatrix} \Sigma^+ \\ \Sigma^- \end{pmatrix} \frac{d\psi(z)}{dz} dz \quad (\text{C.50})$$

For simplicity, as was done for the radiance calculations presented in the previous subsection, we assume a non-reflecting lower boundary ($A_g = 0$), so that the homogeneous part of the solution vanishes and equation (C.50) reduces to:

$$\begin{pmatrix} I^+(H) \\ I^-(0) \end{pmatrix} = \int_0^H \begin{pmatrix} \mathbf{G}_{++}(H-z) & \mathbf{G}_{+-}(H-z) \\ \mathbf{G}_{-+}(H-z) & \mathbf{G}_{--}(H-z) \end{pmatrix} \begin{pmatrix} \Sigma^+ \\ \Sigma^- \end{pmatrix} \frac{d\psi(z)}{dz} dz. \quad (\text{C.51})$$

In satellite applications, we are mostly interested in the radiance exiting the upper boundary, so we write it explicitly:

$$I^+(H) = \int_0^H [\mathbf{G}_{++}(H-z)\Sigma^+ + \mathbf{G}_{+-}(H-z)\Sigma^-] \frac{d\psi(z)}{dz} dz. \quad (\text{C.52})$$

It is clear from (C.52) that the submatrix $\mathbf{G}_{++}(H-z)$ is responsible for the redistribution of the upwelling components of the solar source vector (Σ^+) and $\mathbf{G}_{+-}(H-z)$ is responsible for the redistribution of the downwelling components of the solar source vector (Σ^-) into upwelling radiance. For any particular direction, it is only necessary to examine a single row of the submatrices $\mathbf{G}_{++}(H-z)$ and $\mathbf{G}_{+-}(H-z)$ and sum over all m azimuth components. If however, the satellite were nadir pointing, then only the $m = 0$ component would be sufficient to compute the remotely sensed radiances.

The radiance at the top of the atmosphere is the result of the convolution between the Green's function and the solar source function, so that it is necessary to study this convolution in order to establish what level within the medium mostly influences the observed radiance. We explore this matter with the aid of the *Contribution Function* and its integral.

We rewrite equation (C.52) as

$$I^+(H) = \int_0^H W(H, z) dz, \quad (\text{C.53})$$

where

$$W(H, z) = [\mathbf{G}_{++}(H - z)\Sigma^+ + \mathbf{G}_{+-}(H - z)\Sigma^-] \frac{d\psi(z)}{dz} \quad (\text{C.54})$$

represents the Contribution Function.

Similarly we introduce an Integrated Contribution Function as

$$I_W(H, z) = \int_z^H W(H, z') dz'. \quad (\text{C.55})$$

The integral does not span over the whole layer so that I_W maintains a z -dependence. When integrated over the whole depth of the layer (i.e. $z = 0$), the radiance exiting the upper boundary is recovered. For any other value z , I_W gives the integrated *contribution* to the upwelling radiance at the top of the layer from all levels below H . The radiance computed from the Green's function formulation for the case of a nonreflecting lower boundary (no homogenous solution, but only a particular solution of the RTE) has been compared with previous results from doubling and global \mathcal{R} and \mathcal{T} radiance comparison (see section C.4.1). Good agreement has been found, which indicates that our computation of $W(H, z)$ and its integral is correct.

C.4.3 Examples of reflection from vertically varying clouds

To illustrate an application of the Contribution Function and the Integrated Contribution Function introduced in the previous section, we examine three hypothetical clouds with the following extinction coefficient profiles:

- (a) σ_e constant with height (homogeneous cloud)
- (b) σ_e increasing linearly with height ("stratus-like" cloud)

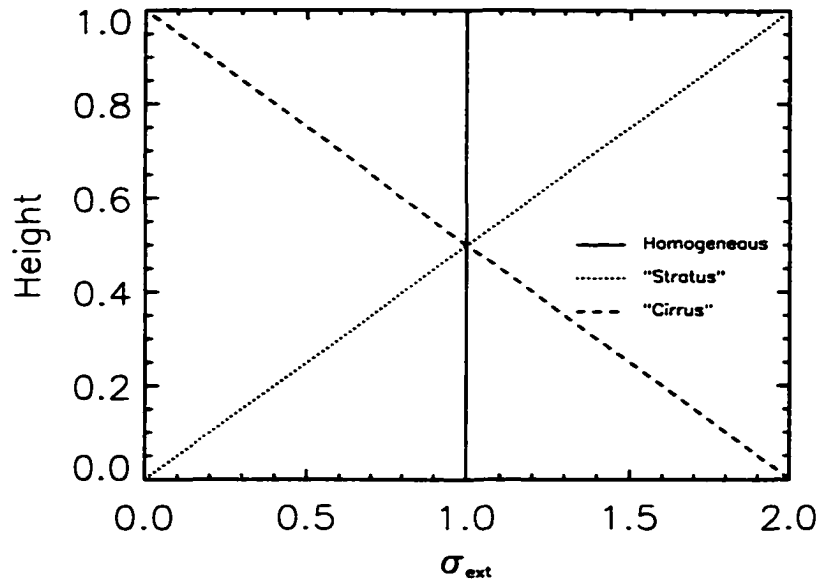


Figure C.1: Vertical profiles of extinction coefficient for three hypothetical clouds (see text for details).

(c) σ_e decreasing linearly with height ("cirrus-like" cloud)

A plot of these extinction profiles is shown in figure C.1. This cloud classification is arbitrary and is suggested by the fact that stratiform clouds tend to have larger droplets near their top and smaller droplets at their base, hence an extinction coefficient increasing with height, while the converse is true for cirrus clouds. It is clear that there are factors other than the σ_e profile that distinguish stratus from cirrus clouds, in particular the scattering phase function, which renders this representation of the optical properties of these clouds unrealistic. We are also aware that the linear form of the σ_e vertical dependence and the assumption of constant single scattering albedo are limiting factors in our analysis. Nevertheless, we believe that useful insight can be gained by examining the different Contribution Function and Integrated Contribution Function profiles obtained for each of the three cases.

We consider two different values of total optical depth, $\tau = 1$ and $\tau = 10$, and single scattering albedo, $\omega_o = 1$ (conservative case) and $\omega_o = 0.2$ (nonconservative case). Again an Henyey–Greenstein phase function is used for the calculations, and the asymmetry parameter is fixed at a value of 0.8. We computed profiles of $W(H, z)$ and $I_W(H - z)$ for a solar zenith angle of 0° and 60° .

Figure C.2 shows the profiles of $W(H, z)$ and $I_W(H - z)$, for the direction $\theta_1 = 8.35^\circ$, for optical depth $\tau = 1$. The upper panel is relative to $\omega_o = 1.0$ and the lower panel to $\omega_o = 0.2$. The cosine of the solar zenith angle is $\mu_\odot = 1.0$. Figure C.3 is the same as figure C.2, except for an optical depth of 10.

If we analyze the case (a) and case (b) profiles, we note that the main difference between the two cases resides mainly in the fact that for the $\tau = 1$ case the Contribution Function and its integral have non negligible values throughout the layer, while the $\tau = 10$ profiles are more concentrated at cloud top, indicating that most of the contribution to the upwelling radiance is coming from the upper portion of the cloud.

A qualitative explanation for this type of behavior lies in the fact that when the medium is optically thin ($\tau < 1$), contributions to the upwelling radiance at the top of the atmosphere involve the entire medium regardless of $\sigma_e(z)$. When the optical depth is large, the radiance tends toward saturation. In this case, an additional layer of small optical depth to the base of the medium would negligibly change the radiance at the upper boundary. The situation is different for the “cirrus” cloud (case (c)): the profiles are quite spread out in both cases indicating that contributions to the total radiance are obtained from a larger portion of the cloud, even in the optically thick case, with respect to the stratus and homogeneous cases.

From these results, we conclude that the Contribution Function and its integral, like the Green’s function, are very sensitive to $\sigma_e(z)$. This has a strong bearing on retrieval of cloud optical depth when remotely sensing different cloud types. It appears that the information that one retrieves from radiance measurements at the top of the layer comes preferentially from different levels, according to the type of cloud that it is being sensed.

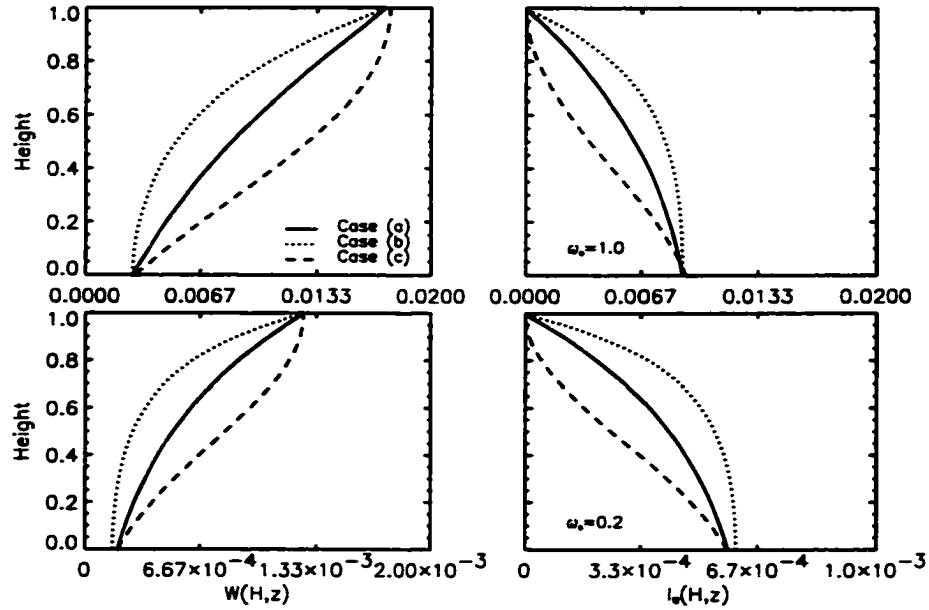


Figure C.2: Profiles of Contribution Function and Integrated Contribution Function in the direction $\theta_1 = 8.35^\circ$ for the three cloud scenarios and two different single scattering albedoes, as labeled (upper panels, $\omega_o = 1$; lower panels, $\omega_o = 0.2$). See text for explanations. Optical parameters: $\tau = 1$, $g = 0.8$. Cosine of solar zenith angle, $\mu_\odot = 1.0$.

An analogous trend is also evident in the nonconservative case as shown in the lower panels of figures C.2 and C.3. The major difference with respect to the conservative case lies in the magnitude of the contributions (note that the horizontal scales change from panel to panel). Although the shape of the profile is similar to the conservative case, in the nonconservative case the amplitude of the two functions is noticeably smaller.

Similar conclusions can be drawn for the case of a solar zenith angle of 60° ($\mu_\odot = 0.5$) (figures C.4 and C.5) as far as the vertical distribution of the contribution

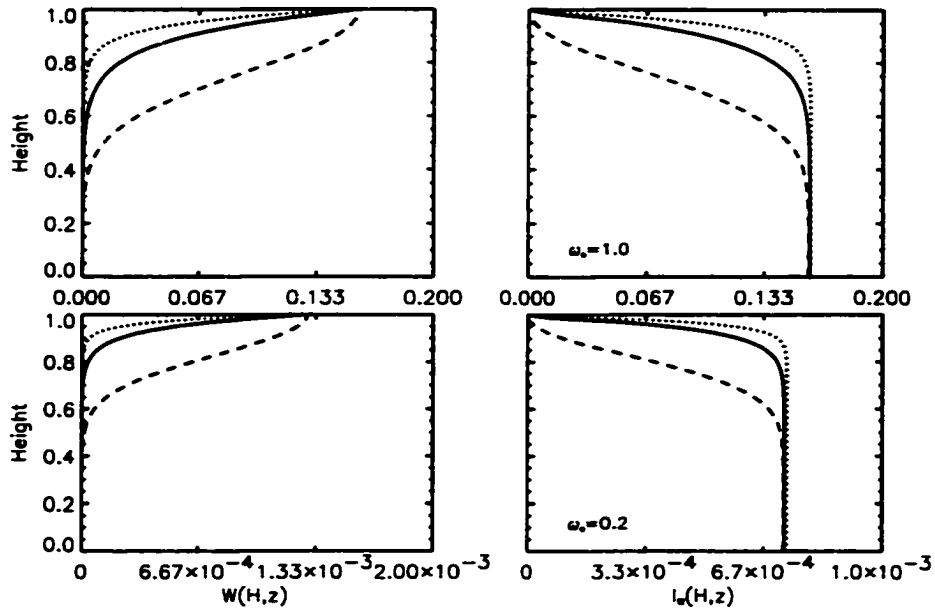


Figure C.3: As in figure C.2, except $\tau = 10$.

goes for the three cloud scenarios. The major difference resides in the magnitude of the contributions in the conservative cases.

In the $\tau = 1$ conservative case (figure C.4, upper panel), the magnitude of the contributions is larger with respect to the corresponding case for $\mu_{\odot} = 1$ (figure C.2, upper panel). In contrast, for the $\tau = 10$ conservative case (figure C.5, upper panel), the magnitude is smaller with respect to the corresponding case for $\mu_{\odot} = 1$ (figure C.3, upper panel). For the nonconservative case, comparing the lower panels of figure C.2 and C.4 and figures C.3 and C.5, no major differences can be noted between the $\mu_{\odot} = 1$ and $\mu_{\odot} = 0.5$ examples. This is due to the fact that for absorbing media, the increase in optical path traveled by the photons due to absorption is more dramatic than the increase in optical path due to a lower solar zenith angle.

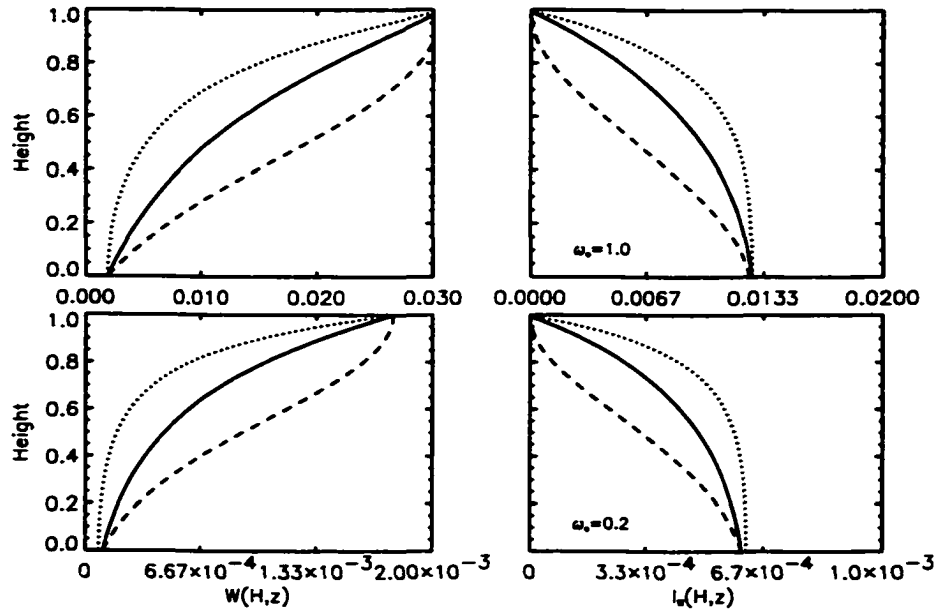


Figure C.4: As in figure C.2, except $\mu_{\odot} = 0.5$.

C.5 Further applications: the penetration optical depth

With the function $I_W(H - z)$ introduced in section C.4.2 we can quantify the percentage of the total upwelling radiance associated with a layer extending downward from the top of the cloud to some reference level z' . Define the optical depth of the layer (z', H) by τ_p . This *penetration* optical depth, τ_p , is a useful way of understanding what portions of the cloud contribute to the measured radiance at cloud top as described by Platnick (2000).

Figure C.6 shows a contour plot of τ_p versus the total cloud optical depth τ of the layer (horizontal axis) and the percentage of total radiance at the upper boundary (vertical axis), for a single scattering albedo of 1 (upper panel) and 0.2 (lower panel) for the homogeneous cloud. The solar zenith angle is 0° . Note that when the percentage is 100%, τ_p equals τ . The conservative and nonconservative

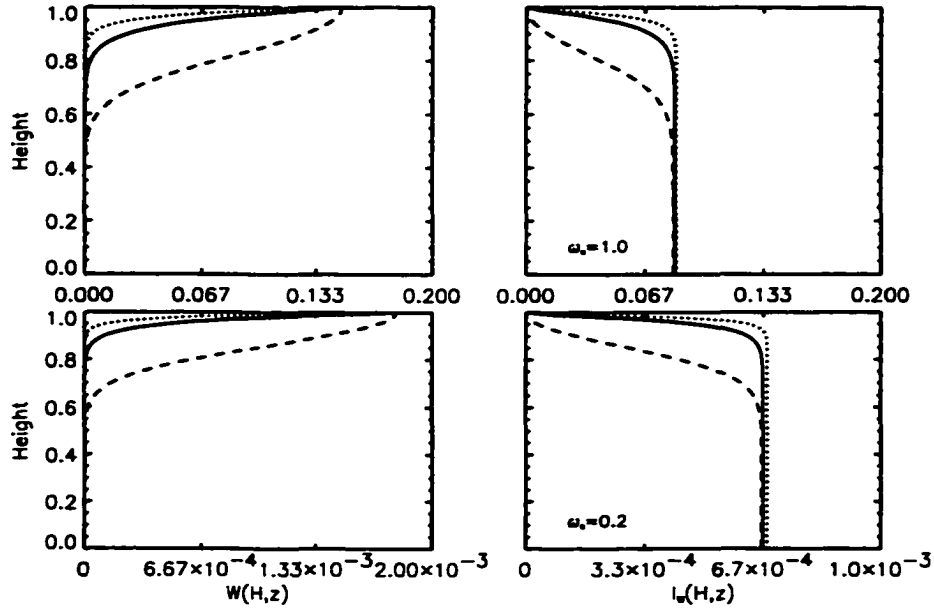


Figure C.5: As in figure C.2, except $\tau = 10$ and $\mu_{\odot} = 0.5$

cases yield different results in terms of τ_p . For example, if the total optical depth is 4, 90% of the total radiance emerges from the layer of optical depth approximately equal to 1.6 when $\omega_o = 1$. For the same percentage and total optical depth, $\tau_p \approx 1.1$ for $\omega_o = 0.2$. This indicates that absorption reduces the level of penetration within the cloud, leading to a more concentrated contribution near cloud top. This result is also shown in the previous section with the profiles of $W(H, z)$ and $I_W(H, z)$. At higher solar zenith angle, 60° , a similar trend is observed (see figure C.7). The increase of solar zenith angle contributes to the reduction of the penetration optical depth because the path length the photons travel is increased by $\frac{1}{\mu_{\odot}}$, where μ_{\odot} is the cosine of the solar zenith angle. This increases the probability of absorption or scattering. For the same example considered before, i.e. total $\tau = 4$ and percentage of total radiance equal to 90%, $\tau_p \approx 1.1$ in the conservative case and $\tau_p \approx 0.7$ in the nonconservative case.

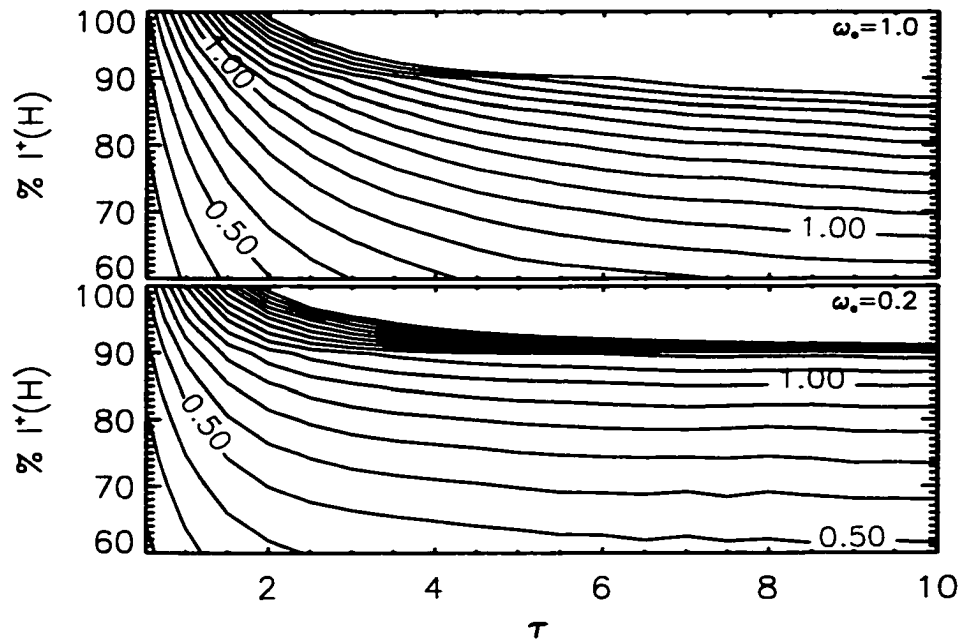


Figure C.6: Contour plot of penetration optical depth (τ_p) as a function of total optical depth and percentage of upwelling radiance at the upper boundary. Two single scattering albedo are shown: $\omega_o = 1$ (upper panel) and $\omega_o = 0.2$ (lower panel). The cosine of the solar zenith angle is 1.0.

The penetration optical depth is the same for the three clouds considered, indicating that it is not possible to discriminate between clouds of different types only from a measurement of the reflected radiance at the top of the atmosphere, unless an assumption is made concerning the form of the extinction coefficient, or multispectral measurements are also performed. For example multi-spectral measurements can be taken over a very narrow spectral interval where the cloud particles scatter conservatively (e.g. $\omega_o = 1$), but where the surrounding atmospheric gas exhibits strong variations in absorption as in the case of the oxygen A-band (750–760 nm). This requires additional information about the cloud height and thickness such as may be provided by an active instrument (e.g. radar). Alternatively, measurements

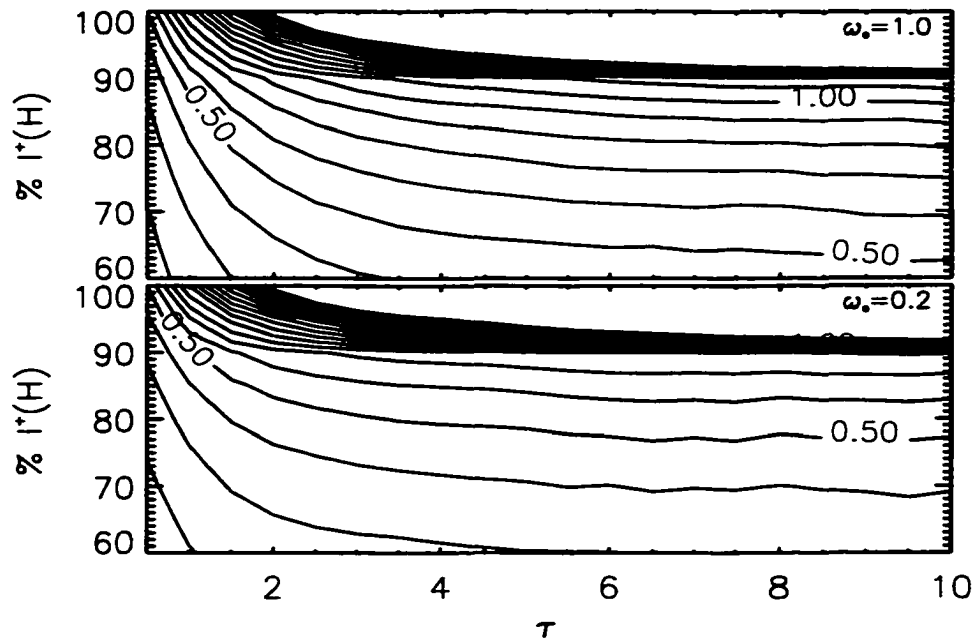


Figure C.7: As in figure C.6, except the cosine of solar zenith angle is 0.5.

can be performed in various region of the spectrum at different wavelengths where there is negligible absorption by atmospheric gases but strong variations of ω_0 of the cloud particles.

C.6 Summary and conclusions

Properties of sunlight scattered from vertically inhomogeneous media have been investigated using a Green's function approach. The eigenmatrix solution of the RTE in concert with the interaction principle have been explored in a new way to obtain the Green's function matrix. To achieve a general and stable form of the Green's function, stabilization of global reflection and transmission matrices, derived from the eigenmatrix, was performed. Radiances computed with the resultant stable global \mathcal{R} and \mathcal{T} , for a range of optical depths, single scattering albedoes and asym-

metry parameters, are in excellent agreement with those calculated from a doubling code.

By definition, the Green's function, when weighted by the source function in the form of a convolution, produces the solution of the RTE. From this form of solution, with the aid of the Green's function we introduced the Contribution Function and its integral to show the distribution of the vertical weighting for different extinction profiles. To explore this aspect, we used a homogenous cloud (σ_e constant with height), a "stratus-like" cloud (σ_e increasing linearly with height) and a "cirrus-like" cloud (σ_e decreasing linearly with height), in all cases maintaining constant ω_o and constant total τ . Analysis of profiles of the Contribution Function and the Integrated Contribution Function has shown that the vertical weighting is sensitive to the extinction profile as well as to the single scattering albedo and to the solar zenith angle.

With the aid of the Integrated Contribution Function, the concept of *penetration* optical depth, that is, the optical depth at which a given percentage of the radiance measured at cloud top is recovered, was introduced. It was found that the penetration optical depth is a strong function of single scattering albedo and solar zenith angle, but is insensitive to the extinction profile. This contrasts with the *effective level* of contribution, which is sensitive to the extinction profile. In particular, for a given penetration optical depth, the corresponding penetration depth is generally greater for cirrus clouds than for the other two clouds considered, since the extinction increases from cloud top down at a lower rate than for the stratus and the homogeneous cloud.

These results suggest that information recovered from path integrated measurements, such as visible upwelling radiances at the top of the atmosphere observed by a satellite instrument, originates from specific regions within the cloud and is dependent on the type of cloud sensed. This study provides a quantitative method for assessing the level from which the information is mainly coming from, provided some educated guess can be made about the cloud scene being observed and the type of extinction profile such a cloud might have. The method introduced in this paper, when combined with profile information provided by active sensors, may lead

to a more definitive way of retrieving extinction profiles in cloud.

C.7 Acknowledgments

This research has been supported under Department of Energy Research Grant DE-FG03-94ER61748 and CloudSat NASA Contract #NAS5-99237. We wish to thank Dr. Steve Miller for providing the radiance results for the comparison from his doubling code, Mick Hopsecger for double-checking the algebra, and Tristan L'Ecuyer for suggesting a compact matrix form to incorporate the surface reflection.

C.8 Appendix 1.

Consider the eigenvalue problem in the vector form

$$\mathbf{A}\vec{x}_r = \lambda_r\vec{x}_r, \quad (\text{C.56})$$

where λ_r is a generic eigenvalue of \mathbf{A} and \vec{x}_r is the eigenvector associated.

If we choose to write \vec{x}_r as constituted of two vectors of dimension N , \vec{u}_r and \vec{v}_r , we have:

$$\begin{pmatrix} \mathbf{t} & -\mathbf{r} \\ \mathbf{r} & -\mathbf{t} \end{pmatrix} \begin{pmatrix} \vec{u}_r \\ \vec{v}_r \end{pmatrix} = \lambda_r \begin{pmatrix} \vec{u}_r \\ \vec{v}_r \end{pmatrix} \quad (\text{C.57})$$

which can be written as

$$\begin{aligned} \mathbf{t}\vec{u}_r - \mathbf{r}\vec{v}_r &= \lambda_r\vec{u}_r \\ \mathbf{r}\vec{u}_r - \mathbf{t}\vec{v}_r &= \lambda_r\vec{v}_r. \end{aligned} \quad (\text{C.58})$$

Summing and subtracting the above equations, we get:

$$\begin{aligned} (\mathbf{t} + \mathbf{r})(\vec{u}_r - \vec{v}_r) &= \lambda_r(\vec{u}_r + \vec{v}_r) \\ (\mathbf{t} - \mathbf{r})(\vec{u}_r + \vec{v}_r) &= \lambda_r(\vec{u}_r - \vec{v}_r). \end{aligned} \quad (\text{C.59})$$

Multiply the first row of (C.59) by $(\mathbf{t} - \mathbf{r})$ and the second row by $(\mathbf{t} + \mathbf{r})$. Upon rearrangement,

$$\begin{aligned} (\mathbf{t} - \mathbf{r})(\mathbf{t} + \mathbf{r})(\vec{u}_r - \vec{v}_r) &= \lambda_r^2(\vec{u}_r - \vec{v}_r) \\ (\mathbf{t} + \mathbf{r})(\mathbf{t} - \mathbf{r})(\vec{u}_r + \vec{v}_r) &= \lambda_r^2(\vec{u}_r + \vec{v}_r). \end{aligned} \quad (\text{C.60})$$

If we introduce the new matrices and vectors

$$\mathbf{B} = (\mathbf{t} - \mathbf{r})(\mathbf{t} + \mathbf{r}), \quad \mathbf{C} = (\mathbf{t} + \mathbf{r})(\mathbf{t} - \mathbf{r}),$$

$$\vec{d}_r = \vec{u}_r - \vec{v}_r, \quad \vec{s}_r = \vec{u}_r + \vec{v}_r,$$

we obtain two new eigenvalue problems:

$$\begin{aligned} \mathbf{B}\vec{d}_r &= \lambda_r^2 \vec{d}_r \\ \mathbf{C}\vec{s}_r &= \lambda_r^2 \vec{s}_r. \end{aligned} \tag{C.61}$$

The eigenvalues of the matrix \mathbf{A} are simply the square root of the eigenvalues of either the matrix \mathbf{B} or \mathbf{C} . Having computed the eigenvalues and eigenvectors of, say, the matrix \mathbf{B} we can recover the eigenvectors of \mathbf{A} . From the first equation of the system in (C.59), the vector \vec{s}_r , eigenvector of \mathbf{C} can be expressed in terms of the vector \vec{d}_r ,

$$\vec{s}_r = \frac{(\mathbf{t} + \mathbf{r})}{\lambda_r} \vec{d}_r,$$

from which follows:

$$\begin{aligned} \vec{u}_r - \vec{v}_r &= \vec{d}_r \\ \vec{u}_r + \vec{v}_r &= \frac{(\mathbf{t} + \mathbf{r})}{\lambda_r} \vec{d}_r, \end{aligned} \tag{C.62}$$

or

$$\begin{aligned} \vec{u}_r &= \frac{1}{2} \left[\mathbf{I} + \frac{(\mathbf{t} + \mathbf{r})}{\lambda_r} \right] \vec{d}_r \\ \vec{v}_r &= \frac{1}{2} \left[-\mathbf{I} + \frac{(\mathbf{t} + \mathbf{r})}{\lambda_r} \right] \vec{d}_r. \end{aligned} \tag{C.63}$$

Since λ_r can be positive or negative, we have another set of equations for $\lambda_r < 0$:

$$\begin{aligned}
\vec{u}_r &= \frac{1}{2} \left[\mathbf{I} - \frac{(\mathbf{t} + \mathbf{r})}{|\lambda_r|} \right] \vec{d}_r \\
\vec{v}_r &= \frac{1}{2} \left[-\mathbf{I} - \frac{(\mathbf{t} + \mathbf{r})}{|\lambda_r|} \right] \vec{d}_r.
\end{aligned} \tag{C.64}$$

Inspection discloses that \vec{v}_r corresponding to $\lambda_r > 0$ is equal to $-\vec{u}_r$ for $\lambda_r < 0$. Similarly \vec{v}_r corresponding to $\lambda_r < 0$ is equal to $-\vec{u}_r$ associated with $\lambda_r > 0$. This implies that we only have to compute \vec{u}_r for both λ_r positive and negative to obtain the full eigenvector matrix.

Let \mathbf{u}_+ be the matrix whose columns are the vectors \vec{u}_r for $\lambda_r > 0$ and \mathbf{u}_- the matrix whose columns are the the vectors \vec{u}_r for $\lambda_r < 0$.

Thus the matrix \mathbf{X} takes the following form:

$$\mathbf{X} = \begin{pmatrix} \mathbf{u}_+ & \mathbf{u}_- \\ -\mathbf{u}_- & -\mathbf{u}_+ \end{pmatrix}.$$

It can be shown that the inverse of \mathbf{X} , \mathbf{X}^{-1} , has a similar form:

$$\mathbf{X}^{-1} = \begin{pmatrix} \mathbf{v}_+ & \mathbf{v}_- \\ -\mathbf{v}_- & -\mathbf{v}_+ \end{pmatrix}.$$

and the elements of \mathbf{X}^{-1} can be computed from the elements of \mathbf{X} using the following matrix relationships

$$\begin{aligned}
\mathbf{v}_+ &= [\mathbf{I} - (\mathbf{u}_+^{-1} \mathbf{u}_-)]^{-1} \mathbf{u}_+^{-1} \\
\mathbf{v}_- &= (\mathbf{u}_+^{-1} \mathbf{u}_-) \mathbf{v}_+
\end{aligned} \tag{C.65}$$

obtained from $\mathbf{X}\mathbf{X}^{-1} = \mathbf{X}^{-1}\mathbf{X} = \mathbf{I}$.

C.9 Appendix 2.

Beginning with the form of the matrix exponential derived in section C.3.1, we derive a stable semi-analytical expression for the global reflection and transmission operators.

Starting from equation (C.28), we rewrite it explicitly as

$$\begin{aligned}\mathcal{T}(H, 0) &= \mathbf{e}_{22}^{-1}(H) \\ &= (-\mathbf{u}_- e^{\Lambda^+ \psi(H)} \mathbf{v}_- + \mathbf{u}_+ e^{-\Lambda^+ \psi(H)} \mathbf{v}_+)^{-1}.\end{aligned}\quad (\text{C.66})$$

Rearranging

$$\begin{aligned}\mathcal{T}(H, 0) &= \{(-\mathbf{u}_- e^{\Lambda^+ \psi(H)} \mathbf{v}_-)[\mathbf{I} - (\mathbf{u}_- e^{\Lambda^+ \psi(H)} \mathbf{v}_-)^{-1}(\mathbf{u}_+ e^{-\Lambda^+ \psi(H)} \mathbf{v}_+)]\}^{-1} \\ &= [\mathbf{I} - (\mathbf{u}_- e^{\Lambda^+ \psi(H)} \mathbf{v}_-)^{-1}(\mathbf{u}_+ e^{-\Lambda^+ \psi(H)} \mathbf{v}_+)]^{-1}(-\mathbf{u}_- e^{\Lambda^+ \psi(H)} \mathbf{v}_-)^{-1} \\ &= -[\mathbf{I} - \mathbf{Q}_1]^{-1} \mathbf{Q}_2^{-1},\end{aligned}\quad (\text{C.67})$$

where

$$\begin{aligned}\mathbf{Q}_1 &= (\mathbf{u}_- e^{\Lambda^+ \psi(H)} \mathbf{v}_-)^{-1}(\mathbf{u}_+ e^{-\Lambda^+ \psi(H)} \mathbf{v}_+) \\ \mathbf{Q}_2 &= (\mathbf{u}_- e^{\Lambda^+ \psi(H)} \mathbf{v}_-).\end{aligned}$$

Expanding the term $[\mathbf{I} - \mathbf{Q}_1]^{-1}$ in series we have:

$$\mathcal{T}(H, 0) = -[\mathbf{I} + \mathbf{Q}_1 + \mathbf{Q}_1^2 + \mathbf{Q}_1^3 + \dots] \mathbf{Q}_2^{-1}.\quad (\text{C.68})$$

By using the expressions for \mathbf{v}_\pm derived in section C.3.1, \mathbf{Q}_1 and \mathbf{Q}_2 can be manipulated to yield:

$$\begin{aligned}\mathbf{Q}_1 &= \mathbf{u}_+ [\mathbf{I} - (\mathbf{u}_+^{-1} \mathbf{u}_-)^2] [(\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)}]^2 \{\mathbf{u}_+ [\mathbf{I} - (\mathbf{u}_+^{-1} \mathbf{u}_-)^2]\}^{-1} \\ \mathbf{Q}_2^{-1} &= \mathbf{u}_+ [\mathbf{I} - (\mathbf{u}_+^{-1} \mathbf{u}_-)^2] [(\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)}] \mathbf{u}_-^{-1}\end{aligned}$$

Substituting in (C.68) and re-summing the geometric series, we obtain the following expression for \mathcal{T} , which involves only *decaying* exponentials:

$$\begin{aligned}\mathcal{T}(H, 0) &= -\mathbf{u}_+ [\mathbf{I} - (\mathbf{u}_+^{-1} \mathbf{u}_-)^2] [(\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)}] \\ &\quad \{ \mathbf{I} - [(\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)}]^2 \}^{-1} \mathbf{u}_-^{-1}.\end{aligned}\tag{C.69}$$

In a similar way we can compute a stable form of the global reflection matrix, \mathcal{R} , starting from equation (C.27):

$$\begin{aligned}\mathcal{R}(H, 0) &= \mathbf{e}_{12}(H) \mathbf{e}_{22}^{-1}(H) \\ &= (\mathbf{u}_+ e^{\Lambda^+ \psi(H)} \mathbf{v}_- - \mathbf{u}_- e^{-\Lambda^+ \psi(H)} \mathbf{v}_+) \\ &\quad (-\mathbf{u}_- e^{\Lambda^+ \psi(H)} \mathbf{v}_- + \mathbf{u}_+ e^{-\Lambda^+ \psi(H)} \mathbf{v}_+)^{-1}.\end{aligned}\tag{C.70}$$

The stable form for \mathcal{T} derived above leads to a stable form for the global reflection matrix \mathcal{R} .

$$\begin{aligned}\mathcal{R}(H, 0) &= -\mathbf{u}_+ [\mathbf{I} - (\mathbf{u}_+^{-1} \mathbf{u}_-) e^{-\Lambda^+ \psi(H)} (\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)}] \\ &\quad \{ \mathbf{I} - [(\mathbf{u}_+^{-1} \mathbf{u}_-)^{-1} e^{-\Lambda^+ \psi(H)}]^2 \}^{-1} \mathbf{u}_-^{-1}.\end{aligned}\tag{C.71}$$

Table C.1: Comparison of radiances computed with a doubling code (D) and with the global \mathcal{R} and \mathcal{T} technique presented in this study (RT). m is the index for the expansion terms of the azimuthal series. μ represents the cosine of the quadrature angles (observing directions). $I^+(H)$ and $I^-(0)$ represent respectively the upwelling and downwelling radiances summed up to the m -th term of the azimuth expansion series.

a. Optical parameters: $\tau = 1$, $\omega_o = 1$, $g = 0$, $\mu_{\odot} = 0.866$

μ	m	D $I^+(H)$	RT $I^+(H)$	% diff.	D $I^-(0)$	RT $I^-(0)$	% diff.
0.9894	0	8.1189E-02	8.1189E-02	0.0	7.3267E-02	7.3315E-02	0.06
0.7554	0	9.4889E-02	9.489E-02	0.0	8.3060E-02	8.3114E-02	0.06
0.0950	0	1.4232E-01	1.4232E-01	0.0	8.3294E-02	8.3382E-02	0.1

b. Optical parameters: $\tau = 1$, $\omega_o = 1$, $g = 0.8$, $\mu_{\odot} = 0.866$

μ	m	D $I^+(H)$	RT $I^+(H)$	% diff.	D $I^-(0)$	RT $I^-(0)$	% diff.
0.9894	0	9.9717E-03	9.9460E-03	0.25	1.6764E-01	1.6752E-01	0.07
0.7554	0	1.6232E-02	1.6209E-02	0.15	1.8942E-01	1.8919E-01	0.12
0.0950	0	4.8565E-02	4.8541E-02	0.25	6.9504E-02	6.8196E-02	1.88
0.9894	3	1.0576E-02	1.0562E-02	0.13	2.8254E-01	2.8223E-01	0.11
0.7554	3	2.1393E-02	2.1344E-02	0.2	6.9865E-01	6.9835E-01	0.04
0.0950	3	8.3972E-02	8.3859E-02	1.34	1.3309E-01	1.3082E-01	1.7
0.9894	7	1.0577E-02	1.0597E-02	0.19	2.8442E-01	2.8408E-01	0.12
0.7554	7	2.1415E-02	2.1382E-02	0.15	8.3781E-01	8.3742E-01	0.05
0.0950	7	8.4466E-02	8.4264E-02	0.24	1.3444E-01	1.3211E-01	1.72
0.9894	11	1.0577E-02	1.0597E-02	0.19	2.8441E-01	2.8408E-01	0.12
0.7554	11	2.1408E-02	2.1402E-02	0.02	8.5923E-01	8.584E-01	0.07
0.0950	11	8.4532E-02	8.4381E-02	0.18	1.3445E-01	1.3218E-01	1.68
0.9894	15	1.0577E-02	1.0597E-02	0.19	2.8441E-01	2.8408E-01	0.12
0.7554	15	2.1406E-02	2.1225E-02	0.84	8.6311E-01	7.7356E-01	10.3
0.0950	15	8.4497E-02	8.4449E-02	0.06	1.3447E-01	1.2792E-01	4.86

LIST OF ABBREVIATIONS

ARM	- Atmospheric Radiation Measurement program
ACR	- Airborne Cloud Radar
CDL	- Cloud Detection Lidar
CERES	- Clouds and the Earth's Radiant Energy System
CIRA	- Cooperative Institute for Research of the Atmosphere
CPI	- Cloud Particle Imager
CVI	- Counter-flow Virtual Impactor
CRM	- Cloud Resolving Model
CRYSTAL	- Cirrus Regional Study of Tropical Anvils and cirrus Layers
CSU	- Colorado State University
DA	- Data Assimilation
DAO	- Data Assimilation Office
ECMWF	- European Center for Medium-Range Weather Forecasts
EOS	- Earth Observing System
ERBE	- Earth Radiation Budget Experiment
FIRE	- First ISCPP Regional Experiment
FACE	- Florida Area Cirrus Experiment (part of CRYSTAL)
IOP	- Intensive Operational Period
ISCCP	- International Satellite Cloud Climatology Project
IWC	- Ice Water Content
IWP	- Ice Water Path
GCM	- General Circulation Model
GEOS	- Goddard Earth Observing System
GOES	- Geostationary Observational Environmental System
LES	- Large Eddy Simulation

- LITE** - Lidar In-space Technology Experiment
- JPL** - Jet Propulsory Laboratory
- MODIS** - Moderate-Resolution Imaging Spectroradiometer
- NASA** - National Aeronautics and Space Administration
- NCEP** - National Centers for Environmental Prediction
- NOAA** - National Oceanic and Atmospheric Administration
- NSA** - North Slope of Alaska (ARM site)
- NWP** - Numerical Weather Prediction
- RAMS** - Regional Atmospheric Modeling System
- SGP** - Southern Great Plains (ARM site)
- SCM** - Single Column Model
- SESAME** - Severe Environmental Storm And Mesoscale Experiment
- SSM/I** - Special Sensor Microwave/Imager
- SSP** - Scanning Spectral Polarimeter
- TMI** - TRMM Microwave Imager
- TOA** - Top Of the Atmosphere
- TRMM** - Tropical Rainfall Measuring Mission
- TWP** - Tropical Western Pacific (ARM site)
- UMass** - University of Massachusetts
- VAR** - Variational (Data Assimilation)