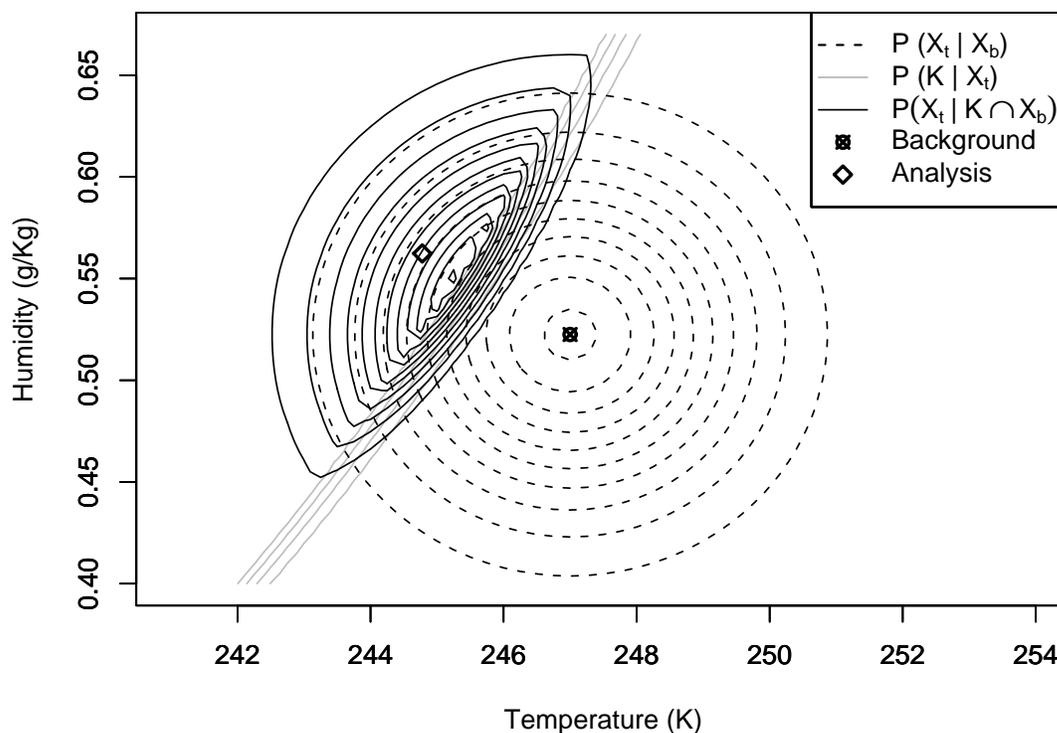




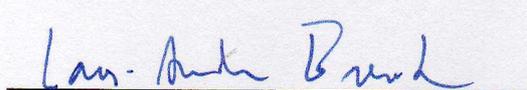
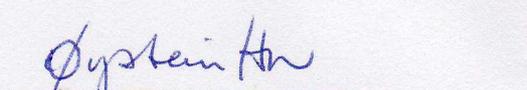
Assimilating Binary Cloud Cover Observations in Variational Data Assimilation Systems

Andrea Storto and Frank Thomas Tveter





Title Assimilating Binary Cloud Cover Observations in Variational Data Assimilation Systems	Date Oslo, January 30 2009
Section Remote Sensing	Report no. no. 1/2009
Author(s) Andrea Storto and Frank Thomas Tvetter	Classification <input checked="" type="radio"/> Free <input type="radio"/> Restricted
	ISSN 0332-9879
Client(s) EUMETSAT	Client's reference EUMETSAT fellow, Final report, ABC project
Abstract This is the final report for the EUMETSAT fellowship ABC project (Assimilation of Binary Cloud Cover Observations) at met.no. The ABC project focussed on developing a theory for assimilating binary cloud cover data in Variational data assimilation systems. The report shows that the developed theory applied to Cloud Profiling Radar data from Cloudsat in Aladin 3D-Var had a positive impact on the forecast skills.	
Keywords Data assimilation, binary observations, Cloud Profiling Radar	

Disiplinary signature	Responsible signature
	

Acknowledgements

The scientific work described in this report has been conducted within EUMETSAT Research Fellowship. The authors want to thank the EUMETSAT staff for the encouragement and the help, and Rolf Stulmann in particular for his constant support. The encouragement, the support and the help from Lars-Anders Breivik, Roger Randriamampianina and Harald Schyberg throughout the ABC-project are also greatly appreciated. Support from the HIRLAM community, the ALADIN community and the ECMWF staff has also been of great importance during the project. The CloudSat Data Processing Center is very acknowledged for the support.

Summary

The quality of the humidity analysis in Numerical Weather Prediction (NWP) models is crucial for the correct prediction of local weather and severe weather events, especially in short-range forecasts. The main goal of the “Assimilation of binary cloud cover” (ABC) project is to improve the humidity analysis by developing and validating a fundamentally new method for exploiting cloud cover satellite observations in current variational data assimilation systems. Cloud data have a binary nature (“cloud” versus “no-cloud”) and their assimilation in variational data assimilation systems is not consistent with the requirements of variational data assimilation theory, i.e. continuous derivatives with respect to state parameters (no switching points). The strategy investigated in this report for the use of cloud observations in existing variational data assimilation systems consists of considering the cloud cover observations as binary occurrences, and applying Bayesian decision theory to compute optimal humidity *pseudo-observations* suitable for direct use in variational data assimilation systems. Such an approach can also in principle be extended to other binary products measuring on/off processes as well (e.g. “rain” versus “no-rain”, or even “conceptual” observations like ‘cumulus clouds’ vs. ‘no-cumulus-clouds’). The pseudo-observation strategy is tested on the CloudSat Cloud Profiling Radar (CPR) which is a 94-GHz radar devoted to microphysical studies. A “threshold algorithm” able to retrieve cloud profiles from the radar return power has been implemented and tuned, producing cross-section of cloud fractions available to on-duty forecasters at met.no. Using a large-scale condensation scheme as observation operator and assimilating pseudo-observations only in case of non-zero observation minus guess difference in cloud-fraction space has permitted to produce humidity pseudo-observations, which have been subsequently assimilated in ALADIN/HARMONIE 3D-Var. The results of the assimilation experiments are positive, especially in terms of improved verification skill scores for dynamical parameters (wind, geopotential). The impact of the pseudo-observations is significant in data-sparse area, like over sea where the amount of in-situ humidity observations is very small. It is shown that the sensitivity of the analysis with respect to CloudSat data is comparable with radiosonde measurements of humidity.

Contents

Acknowledgements	i
Summary	ii
List of Tables	v
List of Figures	vi
1 Theory for Assimilating Binary Observations	1
1.1 Introduction	1
1.2 Introductory study	2
1.3 Handling binary data	19
1.4 One dimensional assimilation of cloud cover	32
1.5 Conclusions	45
2 Assimilation of Binary Cloud Cover: Towards the Use of of Cloud-Sat Radar Reflectivities	47
2.1 Rationale	47
2.2 Exploitation of CloudSat data	48
2.3 Theory for assimilation of binary cloud cover applied to CloudSat cloud fractions data	64
3 Use of satellite data in ALADIN-HARMONIE Norway	72
3.1 Introduction	72
3.2 Observations in the reference assimilation system	73
3.3 Satellite radiances assimilation	74
3.4 Variational Bias correction	76
3.5 Importance of the background error covariances assessment	76
3.6 Impact on analysis	78
3.7 Experimental observations	78
3.8 Summary	80
4 Assimilating Humidity Pseudo-observations Derived from the Cloud Profiling Radar Aboard CloudSat in ALADIN 3D-Var	81
4.1 Introduction	81
4.2 Assimilation strategy	84
4.3 The Bayesian analysis	86
4.4 Use of CloudSat observations	92
4.5 Impact of the humidity pseudo-observations on analysis and forecasts	97
4.6 Summary and perspectives	105

References	111
A Attendance of conferences, meetings and schools	117
B Documentation and publications	118
C Software produced	120
D Main international cooperation	121

List of Tables

3.1	Observations and horizontal thinning distance used in HARMONIE-3DVar.	74
3.2	Use of MSG/SEVIRI thermal channels in HARMONIE-3DVar.	75
4.1	Single observation experiments: pressure, CloudSat-observed cloud fraction, background values, pseudo-observation from the Bayesian analysis and observation minus guess and observation minus analysis within the 3D-Var.	97

List of Figures

1.1	Probability distribution of X_t . The histogram refers to the simulated data set, while the grey line to Eq. (1.1).	3
1.2	Probability distribution of X_b . The histogram refers to the simulated data set, while the grey line refers to Eq. (1.5).	5
1.3	Probability distribution of Y . The histogram refers to the simulated data set, while the grey line refers to Eq. (1.5).	5
1.4	Probability distribution of Y_r . The histogram refers to the simulated data set, while the grey line refers to Eq. (1.5).	6
1.5	The thick line reproduces $P(X_t Y \cap X_b)$ as computed from the data set, while the thin line is the analytical solution of Eq. (1.11). The given values for X_b and Y are set to $270K$ and $264K$ respectively. . .	7
1.6	$P(X_t Y \cap X_b)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively.	9
1.7	$P(X_t \cap Y \cap X_b) / P(Y \cap X_b)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively. . .	9
1.8	$P(Y \cap X_b X_t) P(X_t) / P(Y \cap X_b)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively.	10
1.9	$P(Y X_t) P(X_b X_t) P(X_t) / P(Y \cap X_b)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively.	10
1.10	$P(Y \cap X_t) P(X_b X_t) / P(Y \cap X_b)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively.	11
1.11	$P(X_t Y) P(X_b X_t) / P(X_b Y)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively.	11
1.12	Comparison between $P(X_t Y)$ (histogram) as obtained from the data set, and analytical curve of $P(Y X_t)$ (grey line). The given value for Y is set to $264K$	12
1.13	Histogram of the PDF of $(X_t Y \cap X_b)$ from the data set and analytical solution (Eq. 1.25) with $\delta/2 = 0.25$. The given values for $X_b = x_0$ and $Y = y_0$ are set to $270K$ and $264K$ respectively.	13
1.14	As in Figure (1.13), with $\delta/2 = 1.00$	14

1.15	As in Figure (1.13), with $\delta/2 = 4.00$	14
1.16	As in Figure (1.13), with $\delta/2 = 10.00$	15
1.17	As in Figure (1.13), with $\delta/2 = 15.00$	15
1.18	As in Figure (1.13), with $\delta/2 = 20.00$	16
1.19	Standard Deviation on the growing of the interval δ . As reference value (dashed line), the standard deviation of X_t without conditional constraints is equal to 9.24.	16
1.20	Standard Deviation on the growing of the interval δ for both the observation and the first guess.	17
1.21	J' . The minimum is reached when $X_a = \widetilde{X}_a$	18
1.22	$P(X_T (Y > y_0))$ as computed from the data set (histogram) and its analytical solution (grey line), in the case of binary observation. The dashed line is in correspondence of the y_0 value, set to $264K$	21
1.23	$P(X_T X_b)$ as computed from the data set (histogram) and its analytical solution (grey line). The dashed line is in correspondence of the x_0 value, set to $270K$	21
1.24	$P(X_T (Y > y_0) \cap X_b)$ in the case of binary observation as computed from the data set (histogram) and its analytical solution (grey line), when $y_0 = 264K$. The strong dashed line shows x_0 ($270K$), while the slight dashed line y_0	22
1.25	$P(X_T (Y > y_0) \cap X_b)$ in the case of binary observation as computed from the data set (histogram) and its analytical solution (grey line), when both x_0 and y_0K are equal to $270K$ (dashed line).	22
1.26	$P(X_T (Y > y_0) \cap X_b)$ in the case of binary observation as computed from the data set (histogram) and its analytical solution (grey line), when $y_0 = 275K$. The strong dashed line shows x_0 ($270K$), while the slight dashed line y_0	23
1.27	Comparison between $P(X_T (Y > y_0) \cap X_b)$ (solid line), and the Normal PDF with parameters given by Eq. (1.38) (dashed line). Values for y_0 and x_0 are $264K$ and $270K$ respectively.	23
1.28	Comparison between $P(X_T (Y > y_0) \cap X_b)$ (solid line), and the Normal PDF with parameters given by Eq. (1.38) (dashed line). Value for both y_0 and x_0 is $270K$	24
1.29	Comparison between $P(X_T (Y > y_0) \cap X_b)$ (solid line), and the Normal PDF with parameters given by Eq. (1.38) (dashed line). Values for y_0 and x_0 are $275K$ and $270K$ respectively.	24
1.30	Quantile-quantile plot. Comparison between $P(X_T (Y > y_0) \cap X_b)$ and the Normal PDF with parameters given by Eq. (1.38). Value for both y_0 and x_0 is $270K$	25

1.31	Comparison between $P(X_T (Y = y') \cap (X_b = x_0))$ (solid line), $P(X_T (X_b = x_0))$ (dashed line) and $P(X_T (Y = y'))$ (dotted line), when x_0 and y_0 are set to $270K$ and $264K$ respectively.	26
1.32	Comparison between $P(X_T (Y = y') \cap (X_b = x_0))$ (solid line), $P(X_T (X_b = x_0))$ (dashed line) and $P(X_T (Y = y'))$ (dotted line), when both x_0 and y_0 are set to $270K$	27
1.33	Comparison between $P(X_T (Y = y') \cap (X_b = x_0))$ (solid line), $P(X_T (X_b = x_0))$ (dashed line) and $P(X_T (Y = y'))$ (dotted line), when x_0 and y_0 are set to $270K$ and $275K$ respectively.	28
1.34	Expected value and standard deviation range for $(X_T (Y > y_0) \cap (X_b = x_0))$ on the growing of y_0 , with $x_0 = 270K$ (in correspondence of the dashed line).	29
1.35	Expected value and standard deviation range for $(X_T Y = y')$ on the growing of y_0 , with $x_0 = 270K$ (in correspondence of the dashed line).	30
1.36	Histogram and Normal-like PDF (red line) for $P(X_t Y > y_0 \cap X_b)$, $P(X_t Y > y_0)$ (green line), $P(X_t Y = y')$ (dashed line), $P(X_t $ $X_b = x_0)$ (blue line) for $x_0 = 276K$ and $y_0 = 264K$	30
1.37	Histogram and Normal-like PDF (red line) for $P(X_t Y > y_0 \cap X_b)$, $P(X_t Y > y_0)$ (green line), $P(X_t Y = y')$ (dashed line), $P(X_t $ $X_b = x_0)$ (blue line) for $x_0 = 264K$ and $y_0 = 276K$	31
1.38	$P(X_t Y > y_0)$	31
1.39	$g(x)$	32
1.40	$g'(x)$	32
1.41	Contour of theoretical PDF of $(X_t X_b)$, $(K X_t)$ and $(X_t K \cap X_b)$ when the background temperature is $238K$ and $K = 1$	35
1.42	As in Fig. 1.41 when the background temperature is $245K$ and $K = 1$.	36
1.43	As in Fig. 1.41 when the background temperature is $252K$ and $K = 1$.	36
1.44	Analysis and increments as obtained theoretically for different values of the background both for case $K = 1$ (left) and $K = 0$ (right).	37
1.45	Distribution of $K = 1$ occurrences.	39
1.46	The forward model given in Eq. (1.63).	40
1.47	Contour of cloud cover as computed from Eq. (1.63) on the varying of temperature and specific humidity.	40
1.48	Analysis increments on the varying of the background temperature: (a) case $K = 1$ (Cloud observed), (b) case $K = 0$ (No-Cloud observed).	41
1.49	Cost function J for the temperature.	42
1.50	Cost function J for the specific humidity.	42

1.51	Computation of Bayes risk for different values of the variance of RH_{cr} : (a) PDF, (b) Bayes risk for the analysis \widehat{X}_a minus Bayes risk for the background X_b , case $K = 1$	44
1.52	Computation of Bayes risk for different values of the variance of RH_{cr} when a gross error is introduced: (a) PDF, (b) Bayes risk for the analysis \widehat{X}_a minus Bayes risk for the background X_b , case $K = 1$. . .	45
2.1	CloudSat data dissemination: (a) Over 24 hours (20.10.2007); (b) Over a 6-hours assimilation window (19.10.2007 12:00 UTC \pm 3h). . .	49
2.2	CloudSat cloudmask algorithm, multilayer clouds example, valid UTC time 19.11.2007 00:19: (a) CloudSat orbit track; (b) Cross- section of received power and reflectivity factor; (c) Cross-section of retrieved cloud fractions interpolated to constant geopotential levels. .	52
2.3	CloudSat cloudmask algorithm, surface clutter filter example, valid UTC time 30.11.2007 05:16: (a) CloudSat orbit track; (b) Cross- section of received power and reflectivity factor; (c) Cross-section of retrieved cloud fractions interpolated to constant geopotential levels. .	54
2.4	Example of use of CloudSat cloudmask for nowcasting purposes, valid UTC time 05.12.2007 03:00: (a) CloudSat cloudmask and forecast- ers analysis; (b) METEOSAT MSG1 Infrared and Visible channels (3+9+10).	55
2.5	Values of cloud cover computed by the use of the Slingo and Sundqvist expression, (a) for $RH_{cr} = 0.6$; (b) for $RH_{cr} = 0.8$	58
2.6	RH_{cr} as function of height for different cloud-radar sites using lin- ear regression from NWP models fields: (a) Chilbolton cloud-radar using ECMWF/IFS fields; (b) Chilbolton cloud-radar using Météo- France/Arpège fields; (c) Cabauw cloud-radar using ECMWF/IFS fields; (d) Cabauw cloud-radar using Météo-France/Arpège fields; (e) Palaiseau cloud-radar using ECMWF/IFS fields; (f) Palaiseau cloud- radar using Météo-France/Arpège fields.	59
2.7	RH_{cr} computed from subgrid variances of temperature and specific humidity as function of horizontal resolution using aircraft measure- ments from FASTEX campaign. Data from aircraft NOAA-P3 refer to about 800 hPa, while data from aircraft UK-C130 to about 400 hPa. From Lopez (2002).	60
2.8	RH_{cr} as used operationally at ECMWF (IFS global model) and at Météo-France (Arpège global model and ALADIN limited area model).	61

2.9	RMSE and Mean error at some pressure levels. REF is the pressure-dependent expression of RH_{cr} from Tompkins; CTRL uses as predictors η , η^2 , $CAPE$, T and sea, land and soot aerosol optical thickness. EXP01 uses vertical gradient of potential temperature instead of $CAPE$ in CTRL. EXP02 uses the ice-water separation function instead of T in CTRL. EXP03 adds an additional predictor, η^3 , to CTRL. Cloud fraction is multiplied by 10^3	63
2.10	HR, FAR, POD and ETS scores for the cloud fraction threshold 0.01. Same experiments as in Fig. 2.9.	64
2.11	HR, FAR, POD and ETS scores for the cloud fraction threshold 0.50. Same experiments as in Fig. 2.9.	65
2.12	Comparison between CloudSat cloud cover (top), ALADIN forecast, diagnostic scheme (middle) and ALADIN forecast, prognosed scheme (bottom) for CloudSat 04.01.2007,00:59 and ALADIN forecast 2007010318+006.	65
2.13	Comparison between CloudSat cloud cover (top), ALADIN forecast, diagnostic scheme (middle) and ALADIN forecast, prognosed scheme (bottom) for CloudSat 06.01.2007,12:07 and ALADIN forecast 2007010606+006.	66
2.14	Comparison between CloudSat cloud cover (top), ALADIN forecast, diagnostic scheme (middle) and ALADIN forecast, prognosed scheme (bottom) for CloudSat 07.01.2007,12:48 and ALADIN forecast 2007010706+006.	66
2.15	RMSE and Mean error at some height levels. REF uses the prognostic scheme, EXP1 and EXP2 uses Sundqvist formulation for diagnosed cloud cover with RH_{cr} as used at ECMWF and at Météo-France respectively, EXP2 and EXP3 uses Tompkins formulation for diagnosed cloud cover with RH_{cr} as used at ECMWF and at Météo-France respectively.	67
2.16	Distribution of $(C X_t)$	68
2.17	Probability of having a cloud fraction given X_t (left) and probability of X_t given the observation and the background (right) when the number of sub-observations is 10 and background relative humidity is 80%. In black the distribution as counted in the simulated dataset; in red the theoretical solution of Eq. 2.14; in green the same equation when only one sub-observation is assumed.	70
2.18	Analysis increments for a different number of sub-observations; background relative humidity is 80%.	71

3.1	Difference of root mean square errors between a reference experiment and an experiment with SEVIRI data assimilation against radiosonde observations. Where positive (red), SEVIRI observations have a positive impact.	76
3.2	Vertical error correlation of temperature between model level 48 and all the other model levels as function of horizontal scale. Left and right panel show the statistics derived via the NMC method and the ensemble method respectively.	77
3.3	Cross-sections of temperature analysis increments for brightness temperature single-observation experiments (2 K innovation for AMSU-A channel 9 aboard NOAA-18). Left and right panel show the increments using the NMC method and the ensemble method statistics respectively.	78
3.4	Absolute and relative DFS. Red bars refer to Harris and Kelly bias correction scheme experiment for AMSU and SEVIRI; green bars for variational bias correction experiment.	79
3.5	Difference of root mean square errors between a reference experiment and an experiment with GPS-ZTD data assimilation against radiosonde observations. Where positive (red), GPS-ZTD observations have a positive impact.	80
4.1	CloudSat assimilation flowchart	85
4.2	Cloud fraction - relative humidity relationship	86
4.3	Example of cloud fraction intercomparison	90
4.4	Profile of background-error standard deviation of specific humidity	91
4.5	Example of humidity pseudo-observations	93
4.6	Humidity pseudo-observations errors	95
4.7	Single observation experiments	96
4.8	Daily amount of observations	98
4.9	Innovations and residuals statistics	99
4.10	Bias and standard deviation of analysis increments of CloudSat experiment minus NoCloudSat experiment	100
4.11	Degrees of freedom for signal	101
4.12	Synoptic situation and CloudSat data coverage for the mid-summer case-study	102
4.13	Analysis increments for the illustrative case	103
4.14	Cross-section of forecasts with humidity pseudo-observations minus forecasts without	107
4.15	Temporal evolution of forecast biases between the experiments with and without CloudSat-derived pseudo-observations	108

4.16 Verification against SYNOP observations	109
4.17 Radiosonde verification skill scores	110

Chapter 1

Theory for Assimilating Binary Observations

Summary

The purpose of this Chapter is to study the statistical formalism used in data assimilation, and expand this assimilation theory to include binary observations. One dimensional simulated data sets are used for the verification of the theory. In these data sets we *know* the true state of the atmosphere, X_t , the first guess X_b and the scalar observation Y . The expression for $P(X_t | X_b \cap Y)$ is for instance developed and the derivation is verified against the simulated data set. The theory is applied to a simple one-dimensional binary observation condition, and subsequently extended to more realistic binary cloud observations. An objective quality control procedure based on Bayesian risk function is also investigated.

This Chapter is based on the Met.no Research Note 2006/11 (see Appendix B: Storto A. and Tvetter F.T., 2006).

1.1 Introduction

This Chapter reports the status of the “Assimilation of Binary Cloud Cover” (ABC) project after almost one year from the launching. This project is carried out at the Norwegian Meteorological Institute and funded by EUMETSAT through the fellowships program. A pilot study that investigates the benefits of assimilating binary cloud cover data can be found in Tvetter (2004).

In data assimilation theory, a first guess of the atmospheric state is combined with observations to estimate a new and improved atmospheric state, referred to as the “analysis”. The data assimilation problem can be divided into two parts: first to find the probability density distribution for the true state (X_t), given our observation (Y) and first guess (X_b), $P(X_t | X_b \cap Y)$, and then to estimate the state of the atmosphere X_a that has the lowest expected “Bayesian loss”, $l(X_t, X_a)$, according to this distribution. One commonly used loss function is $l(X_t, X_a) = (X_t - X_a)^2$, which will give the analysis that scores best in RMSE verification methods. The

optimal analysis is then the expected state of the true atmosphere.

In this Chapter we focus on the first part of the data assimilation problem, namely the formulation of $P(X_t | X_b \cap Y)$. In order to better understand the formalism that will be used to solve the binary observations assimilation problem, the formalism is first applied to the classical problem where observations with Normal error distributions are assimilated. A simple binary problem is studied afterwards. A simulated data set is used to verify the analytical probabilities expressions. Attempts are also made to derive theory suitable for use in a 3D-Var environment where the software is assuming that the data has Normal error distributions.

In real life, complex variables as cloud cover depend on many control variables, and the choice of the best “forward operator” to use is a pivotal point. The theory we developed is therefore applied to a more realistic case, again one-dimensional, where the binary condition (cloud or no-cloud) depends on two state variables, namely the temperature and the specific humidity. This bi-variate problem is studied using the same assumptions for the dataset, and introducing a simple relation (Slingo, 1987) between cloud cover and relative humidity, with the aim to provide useful information for the implementation of the theory in an operational framework. In this sense, Bayes decision theory supplies the mathematical instruments to perform an objective quality control, in order to remove from the assimilation the observations which are difficult to use, due to the uncertainty in the forward model.

In section 2 the formalism of the data assimilation is introduced and applied to a simulated dataset; in section 3 a simple binary condition is imposed, and the effects on the analysis are evaluated. In section 4 we take into account a more realistic forward model and study quality control issues.

1.2 Introductory study

The starting point of this study will be the construction of a one dimensional data set suitable for verifying the theoretical achievements. We will focus afterwards on the derivation of the state of the atmosphere with the biggest probability, applying the Mean Squared Error criterion for the verification.

1.2.1 Simulating a data set

The simulated data set consists of a series of known “true” states, observations and previous forecasts (first guess) generated by a random generator¹. The true state

¹To be precise, in the following simulations X_t is assumed to be the temperature, varying from $250K$ to $282K$, and Y_r , X_b , Y have variance set to $1K^2$, $10K^2$ and $5K^2$ respectively; Y still refers to the temperature, so the observation operator is the identity operator. The given values for the observation and for the first guess are set to $264K$ and $270K$. These values are also used in plotting the theoretical Normal distributions.

is assumed to be uniform along a prefixed range, while the series of observations and background have Normal, unbiased and independent random errors with respect to the true state.

Let us suppose that the true state X_t of the variable has a uniform distribution (Fig. 1.1) within a fixed interval (a, b) modelled by

$$X_t \sim U(a, b). \quad (1.1)$$

The errors of the observation and the background are assumed unbiased and Normal,

$$Y_r - X_t \sim N(0, \sigma_r^2)$$

$$Y - X_t \sim N(0, \sigma_o^2)$$

$$X_b - X_t \sim N(0, \sigma_b^2)$$

where σ_r is the standard deviation of the reference observations, σ_o is the standard deviation of the scalar observations and σ_b is the standard deviation of the model first guess. The distribution of the observation and first guess *given the true state of the atmosphere* follows therefore a Normal distribution according to

$$Y_r | X_t \sim N(X_t, \sigma_r^2)$$

$$Y | X_t \sim N(X_t, \sigma_o^2)$$

$$X_b | X_t \sim N(X_t, \sigma_b^2). \quad (1.2)$$

The observation value is found by adding a random number generated by a Normal

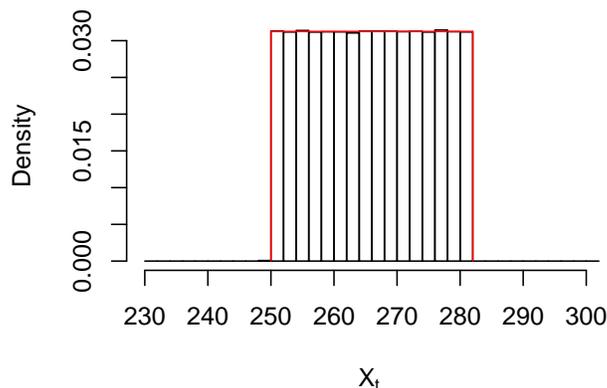


Figure 1.1: Probability distribution of X_t . The histogram refers to the simulated data set, while the grey line to Eq. (1.1).

distribution (with the standard deviation σ_o and variance σ_o^2), to the true state

value X_t . The same procedure has been used to generate the background term X_b with the standard deviation σ_b , ($\sigma_b > \sigma_y$). A “very accurate” observation Y_r , (i.e. a reference observation), was also generated in the same manner using the standard deviation σ_r ($\sigma_r < \sigma_y$).

1.2.2 Total distributions

It is relatively easy to find the total probability density distributions of Y , X_b and Y_r from the simulated data set. When we want to *model* these probability distributions based on the theoretical distributions functions, we start with the probability density functions (PDFs) in Eq. (1.2), giving the total probability distributions of Y , X_b and Y_r as the convolution integrals, along the interval (a, b) , of the probability densities in Eq. (1.2), e.g. the probability of X_b is

$$P(X_b = x) = \int_{-\infty}^{+\infty} P(X_b = x | X_t = u) P(X_t = u) du \quad (1.3)$$

$$= \frac{1}{(b-a)\sqrt{2\pi\sigma_b^2}} \int_a^b e^{-\frac{(u-x)^2}{2\sigma_b^2}} du \quad (1.4)$$

where u is the auxiliary integration variable. Eq. (1.4), after the subdivision of the interval (a, b) in n intervals δ long, becomes in discrete form

$$P(X_b = x) = \lim_{\delta \rightarrow 0} \left[\frac{1}{\delta(b-a)} \sum_{i=1}^{(b-a)/\delta} \phi \left(\frac{x - [a + \delta(i - \frac{1}{2})]}{\sigma_b} \right) \right]. \quad (1.5)$$

Figures (1.2), (1.3) and (1.4) show the probability distributions estimated from the simulated data set along with the corresponding distributions calculated from the model. We observe that the data set distribution and the analytical models distributions match closely. We also see that the smaller the variance the closer to the uniform distribution the PDFs are for Y_r , Y and X_b . Thus (see (Fig. 1.2)) it is possible to consider $P(X_b = x)$ uniform between $(a + k\sigma_b)$ and $(b - k\sigma_b)$ if $\Phi(x < -k)$ is negligible, for instance with $k = 3$, where $\Phi(z)$ is the cumulative density function (CDF) of the Standard Normal distribution, namely the $P(Z < z)$ if $Z \sim N(0, 1)$. So, the Y_r PDF (accurate reference observation) is very similar to the X_t uniform distribution, where the X_b PDF moves more away from this distribution. This property will be of basic importance later on, when we will use the Normal assumption for describing the distribution of X_t given the background information.

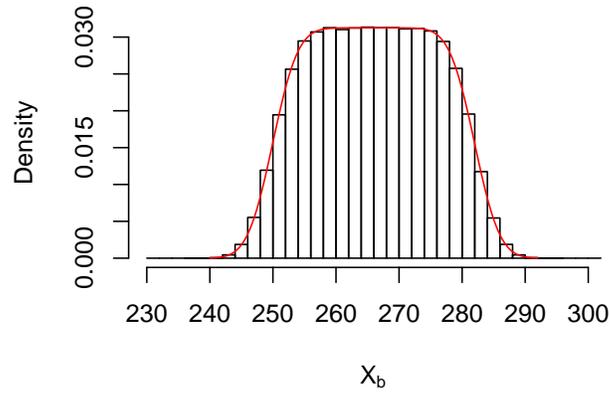


Figure 1.2: Probability distribution of X_b . The histogram refers to the simulated data set, while the grey line refers to Eq. (1.5).

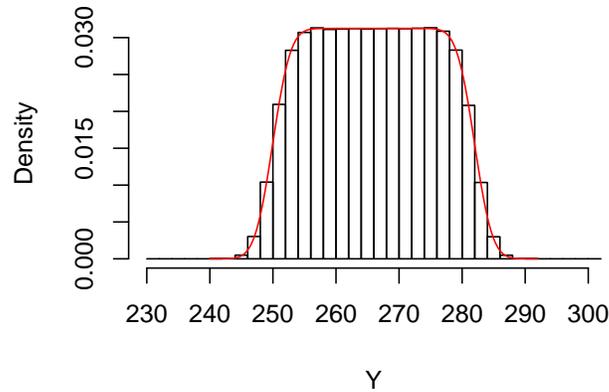


Figure 1.3: Probability distribution of Y . The histogram refers to the simulated data set, while the grey line refers to Eq. (1.5).

1.2.3 Derivation of $P(X_t | Y \cap X_b)$

Now we study the distribution of $P(X_t | Y \cap X_b)$, which is the goal of the first part of the data assimilation problem. If we assume independence between the observation error and the background error, i.e.

$$P(Y \cap X_b | X_t) = P(Y | X_t) P(X_b | X_t). \quad (1.6)$$

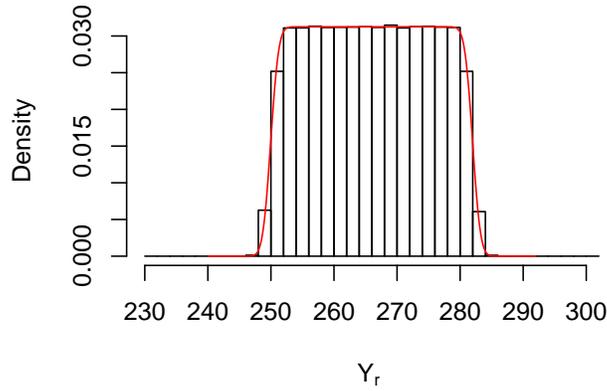


Figure 1.4: Probability distribution of Y_r . The histogram refers to the simulated data set, while the grey line refers to Eq. (1.5).

we may write

$$\begin{aligned}
 P(Y | X_b \cap X_t) &= \frac{P(Y \cap X_b \cap X_t)}{P(X_b \cap X_t)} \\
 &= \frac{P(Y \cap X_b | X_t) P(X_t)}{P(X_b | X_t) P(X_t)} \\
 &= \frac{P(Y | X_t) P(X_b | X_t)}{P(X_b | X_t)} \\
 &= P(Y | X_t).
 \end{aligned}$$

Following a Bayesian analysis, and using the equation above we have

$$P(X_t | Y \cap X_b) = \frac{P(X_t \cap Y \cap X_b)}{P(Y \cap X_b)} \quad (1.7)$$

$$= \frac{P(Y | X_b \cap X_t) P(X_t \cap X_b)}{P(Y \cap X_b)} \quad (1.8)$$

$$= \frac{P(Y | X_t) P(X_t \cap X_b)}{P(Y \cap X_b)} \quad (1.9)$$

$$= \frac{P(Y | X_t) P(X_t | X_b)}{P(Y | X_b)} \quad (1.10)$$

where we may write

$$\begin{aligned} P(Y | X_b) &= \int_{-\infty}^{+\infty} P(Y | X_t) P(X_t | X_b) dX_t \\ &= \int_{-\infty}^{+\infty} P(Y | X_t) \frac{P(X_b | X_t) P(X_t)}{P(X_b)} dX_t \end{aligned}$$

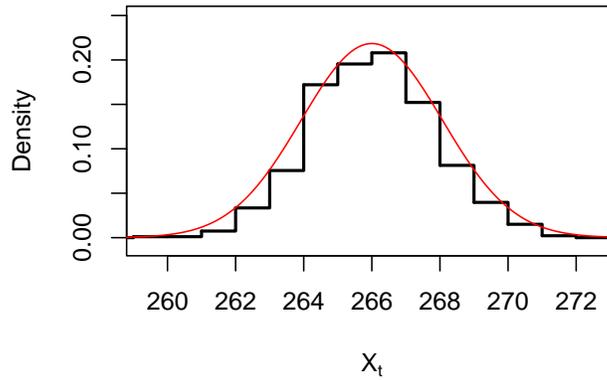


Figure 1.5: The thick line reproduces $P(X_t | Y \cap X_b)$ as computed from the data set, while the thin line is the analytical solution of Eq. (1.11). The given values for X_b and Y are set to $270K$ and $264K$ respectively.

Figure (1.5) shows the probability distribution of the left hand side (LHS) of Eq. (1.10) estimated from the simulated data set and the right hand side (RHS) of Eq. (1.10) calculated analytically. To calculate the analytical distribution we must express Eq. (1.10) in terms of the theoretical distributions in Eq. (1.2). Bayes' rule gives,

$$\frac{P(Y | X_t) P(X_t | X_b)}{P(Y | X_b)} = \frac{P(Y | X_t) P(X_b | X_t) P(X_t)}{P(Y | X_b) P(X_b)}. \quad (1.11)$$

The classical approach to the solution of $P(X_t | Y \cap X_b)$ is given by

$$P(X_t | Y \cap X_b) = \frac{P(X_t \cap Y \cap X_b)}{P(Y \cap X_b)} \quad (1.12)$$

$$= \frac{P(Y \cap X_b | X_t) P(X_t)}{P(Y \cap X_b)} \quad (1.13)$$

$$= \frac{P(Y | X_t) P(X_b | X_t) P(X_t)}{P(Y \cap X_b)} \quad (1.14)$$

$$= \frac{P(Y \cap X_t) P(X_b | X_t)}{P(Y \cap X_b)} \quad (1.15)$$

$$= \frac{P(X_t | Y) P(Y) P(X_b | X_t)}{P(X_b | Y) P(Y)} \quad (1.16)$$

$$= \frac{P(X_t | Y) P(X_b | X_t)}{P(X_b | Y)}. \quad (1.17)$$

Note how X_b and Y have switched places in Eq. (1.17) compared to Eq. (1.10).

Equation (1.10) is more suitable when observations have strong non-linear relations with the model variables. In the case of binary cloud-cover observations (“cloud” or “no-cloud”), it is not possible to compute X_t from the observation Y (model space) since X_t can not be unequivocally found from Y , because many combinations of the true state can cause the same value for the observation. Note that the binary cloud cover observation operator is usually defined as $P(Y | X_t)$ (observations space).

Most of the methods currently used for estimating the background errors (consider for instance the *Lönnberg-Hollingsworth* and derived methods, the *NMC* method and the *Ensemble-based* methods) model the error evolution assuming to know the true state of atmosphere. Therefore, they retrieve first guess (forecast) error statistics suitable for use in the classical Eq. (1.17), and refer to $P(X_b | X_t)$ instead of $P(X_t | X_b)$. However, the expressions can in principle be related by Bayes’ rule,

$$P(X_b | X_t) = P(X_t | X_b) \frac{P(X_b)}{P(X_t)} \quad (1.18)$$

and it is usually valid to assume that the first guess error is relatively small so that $P(X_b) \simeq P(X_t)$, which gives $P(X_b | X_t) \simeq P(X_t | X_b)$ although we will use this approximation only in the next section.

In Figures (1.6) to (1.11), $P(X_t | Y \cap X_b)$ and the RHS of the equations (1.12), (1.13), (1.14), (1.15) and (1.17) are shown as PDFs in the simulated data set. The grey line in all the figures is the analytical model for $P(X_t | Y \cap X_b)$ based on Eq. (1.11). We observe that all the PDFs correspond to each other.

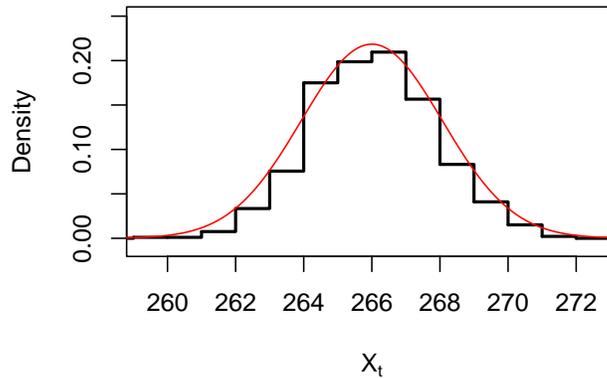


Figure 1.6: $P(X_t | Y \cap X_b)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively.

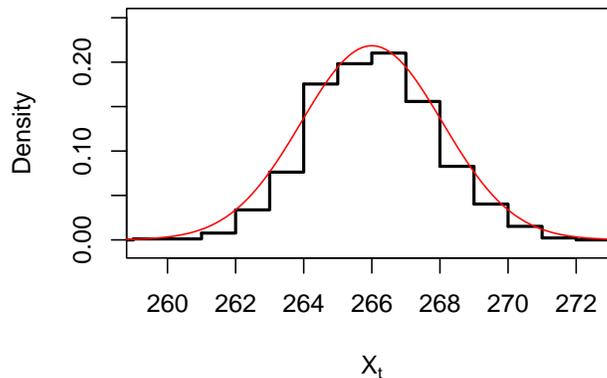


Figure 1.7: $P(X_t \cap Y \cap X_b) / P(Y \cap X_b)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively.

1.2.4 Normal approximations

In classical data assimilation theory, it is assumed that all the involved distributions are Normal. This makes the problem much easier to solve. As an exercise here, we will find the appropriate Normal approximations to our simulated data set, and study where these approximations are valid.

Focusing on the distributions $P(X_t | Y)$ and $P(X_t | X_b)$, under the hypothesis that Y and X_b lie far enough from the borders of the interval (a, b) we can consider $P(Y)$ and $P(X_b)$ constants and about equal to $P(X_t)$. Applying the Bayes' rule

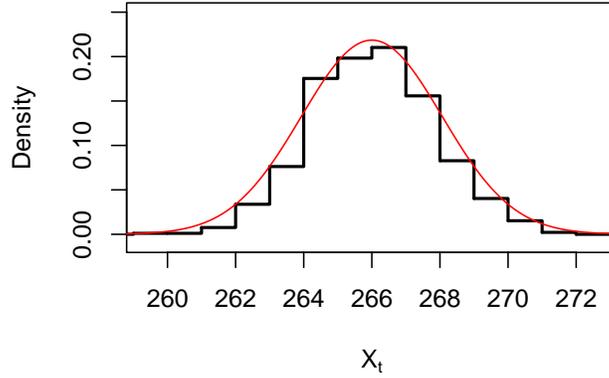


Figure 1.8: $P(Y \cap X_b | X_t) P(X_t) / P(Y \cap X_b)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively.

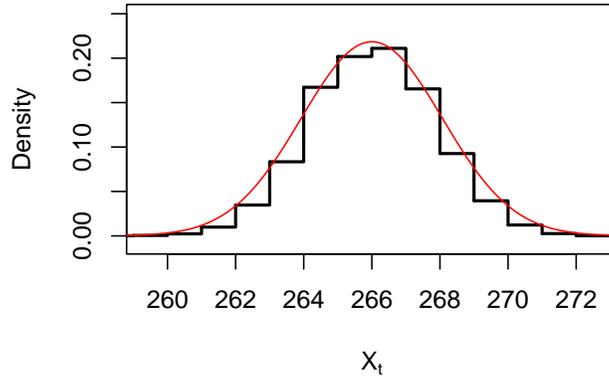


Figure 1.9: $P(Y | X_t) P(X_b | X_t) P(X_t) / P(Y \cap X_b)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively.

and referring to Eq. (1.2), they become

$$\begin{aligned}
 P(X_t | Y) &= \frac{P(Y | X_t) P(X_t)}{P(Y)} \approx P(Y | X_t) = N(0, \sigma_o^2) \\
 P(X_t | X_b) &= \frac{P(X_b | X_t) P(X_t)}{P(X_b)} \approx P(X_b | X_t) = N(0, \sigma_b^2). \quad (1.19)
 \end{aligned}$$

The curve of $P(X_t | Y)$ is identical to $P(Y | X_t)$ (see Fig. 1.12). Substituting this expression into Eq. (1.10) and considering that for Eq. (1.2) and Eq. (1.19) we have

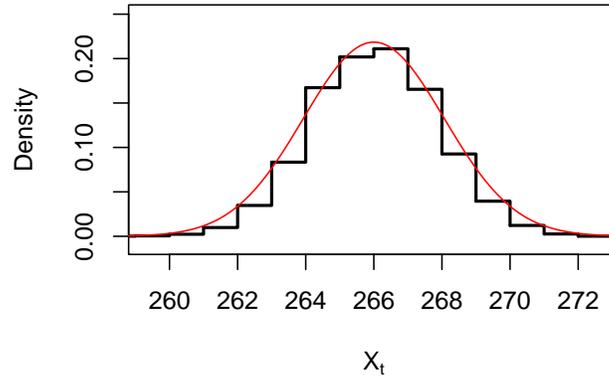


Figure 1.10: $P(Y \cap X_t) P(X_b | X_t) / P(Y \cap X_b)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively.

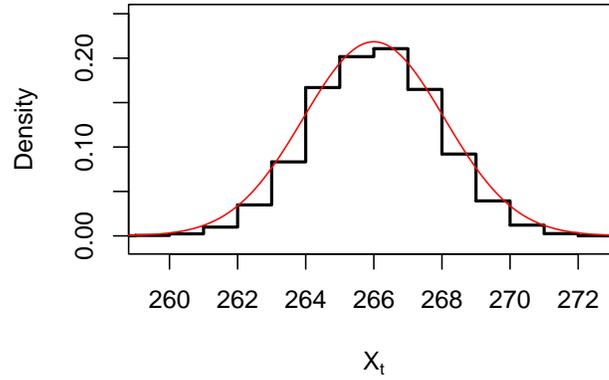


Figure 1.11: $P(X_t | Y) P(X_b | X_t) / P(X_b | Y)$ as computed from the data set (thick line) and analytical solution of based on Eq. (1.11) (thin line). The given values for X_b and Y are set to $270K$ and $264K$ respectively.

$$\begin{aligned}
 P(Y | X_b) &= \int_{-\infty}^{+\infty} [P(Y | X_b \cap X_t) P(X_b | X_t)] dx \\
 &= \int_{-\infty}^{+\infty} [P(Y | X_t) P(X_b | X_t)] dx
 \end{aligned} \tag{1.20}$$

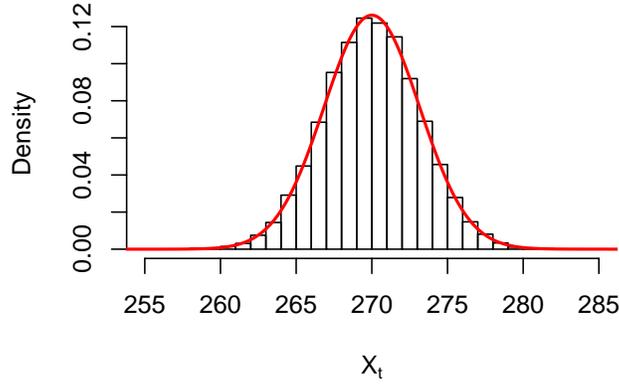


Figure 1.12: Comparison between $P(X_t | Y)$ (histogram) as obtained from the data set, and analytical curve of $P(Y | X_t)$ (grey line). The given value for Y is set to $264K$.

follows that

$$P(X_t | Y \cap X_b) = \frac{(2\pi\sigma_o\sigma_b)^{-1} e^{-\frac{(X_t-Y)^2}{\sigma_o^2}} \cdot e^{-\frac{(X_t-X_b)^2}{\sigma_b^2}}}{\int_{-\infty}^{+\infty} (2\pi\sigma_o\sigma_b)^{-1} e^{-\frac{(x-Y)^2}{\sigma_o^2}} \cdot e^{-\frac{(x-X_b)^2}{\sigma_b^2}} dx} \quad (1.21)$$

and finally we obtain that

$$(X_t | Y \cap X_b) \sim N\left(\frac{\sigma_b^2 Y + \sigma_o^2 X_b}{\sigma_b^2 + \sigma_o^2}, \frac{\sigma_b^2 \sigma_o^2}{\sigma_b^2 + \sigma_o^2}\right) \quad (1.22)$$

which is the approximate analytical solution for the problem.

The interested reader can verify that the two approaches represented by Eq. (1.17) and Eq. (1.10) also agree when the approximations of Eq. (1.19) are applied.

1.2.5 Sampling Interval

For the case of the simulated data set, the PDFs are computed by counting the occurrences in intervals. So, the constraints that appear in conditional probabilities are, in practise, considered in term of intervals δ , for instance $P(X_t | y - \delta/2 < Y < y + \delta/2 \cap x_0 - \delta/2 < X_b < x_0 + \delta/2)$ where y_0 and x_0 are specific values for the observation and first guess. The choice of the interval width influences the searched PDF: evidently, the bigger the interval, the more the resulting

PDF converges towards the unconditioned probability, for instance

$$\lim_{\delta \rightarrow \infty} P(X_t | y_0 - \delta/2 < Y < y_0 + \delta/2 \cap x_0 - \delta/2 < X_b < x_0 + \delta/2) \rightarrow P(X_t).$$

Without loss of generality, we now consider the observation $Y = y_0$. The rigorous formulation of the PDF of $(X_t | Y)$ is $P(X_t | y_0 - \delta/2 \leq Y \leq y_0 + \delta/2)$ rather than $P(X_t | Y = y_0)$, where δ is exactly the global width of the interval.

To take into account the influence of the width of δ , we can divide δ in $n = \delta/\Delta l$ smaller intervals, each a Δl long, and the probability function becomes

$$P(X_t | y_0 - \delta/2 \leq Y \leq y_0 + \delta/2) = \frac{P(y_0 - \delta/2 \leq Y \leq y_0 + \delta/2 | X_t)P(X_t)}{P(y_0 - \delta/2 \leq Y \leq y_0 + \delta/2)} \quad (1.23)$$

$$= \int_{y_0 - \delta/2}^{y_0 + \delta/2} P(Y = v | X_t) \frac{P(X_t)}{P(Y = v)} dv \quad (1.24)$$

$$\approx \sum_{i=0}^n \frac{P(Y = y_0 - \delta(i\Delta l - \frac{1}{2}) | X_t)P(X_t)}{P(Y = y_0 - \delta(i\Delta l - \frac{1}{2}))} \quad (1.25)$$

where v is the auxiliary integration variable.

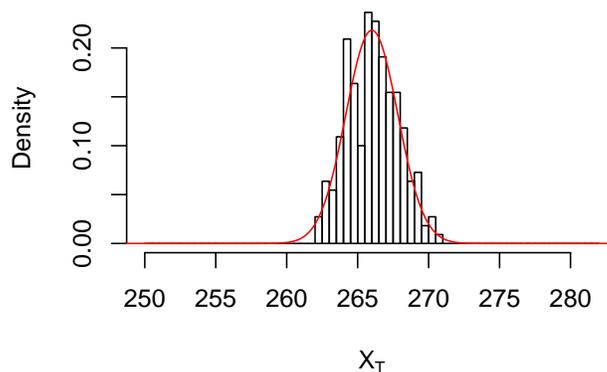


Figure 1.13: Histogram of the PDF of $(X_t | Y \cap X_b)$ from the data set and analytical solution (Eq. 1.25) with $\delta/2 = 0.25$. The given values for $X_b = x_0$ and $Y = y_0$ are set to 270K and 264K respectively.

The previous expression represents the convolution integral of the distribution with mean $y_0 - \delta(i\Delta l - \frac{1}{2})$. We can notice that if δ increases $E(X_t | Y \cap X_b)$ gets closer to – or included in – $(v_0 - \delta/2, y_0 + \delta/2)$, since $P(X_t | Y)$ can be considered uniform in a larger area of the interval (a, b) . We also observe that the standard deviation of $P(X_t | Y)$ increases.

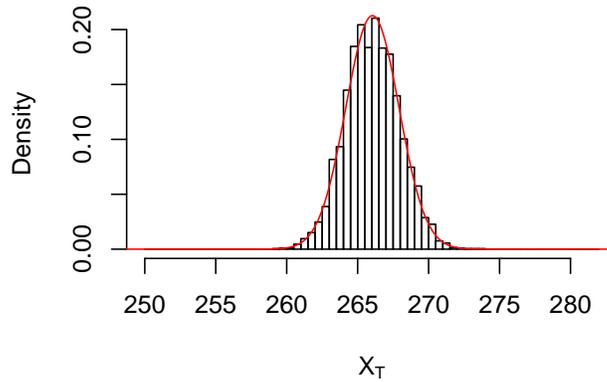


Figure 1.14: As in Figure (1.13), with $\delta/2 = 1.00$.

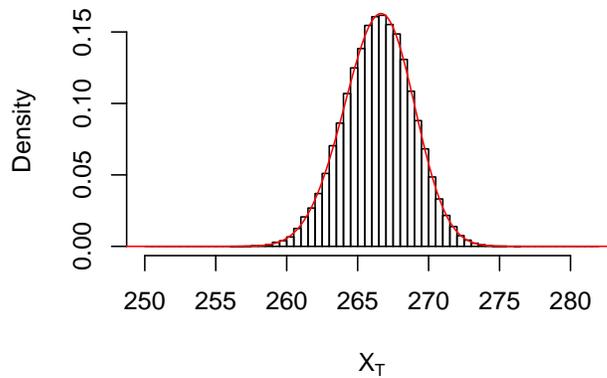


Figure 1.15: As in Figure (1.13), with $\delta/2 = 4.00$.

Figures (1.13) to (1.18) just show $P(X_t | Y \cap X_b)$ when the interval δ increases. Figure (1.19) shows how the standard deviation increases on the growing of δ , until it gets constant to the value of the standard deviation of X_t . Moreover, if the conditional probability of X depends on two variables Y and X_b Eq. (1.25) becomes

$$P(X_t | y_0 - \delta/2 \leq Y \leq y_0 + \delta/2 \cap x_0 - \delta/2 \leq X_b \leq x_0 + \delta/2) =$$

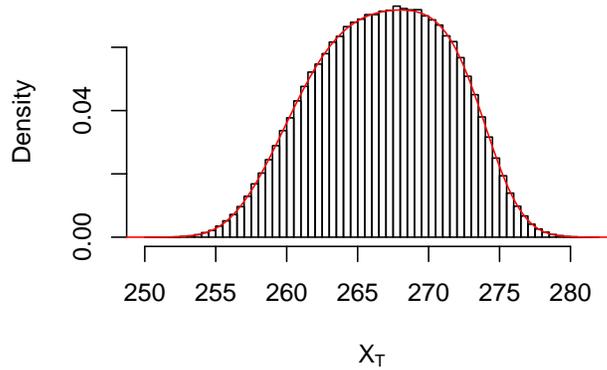


Figure 1.16: As in Figure (1.13), with $\delta/2 = 10.00$.

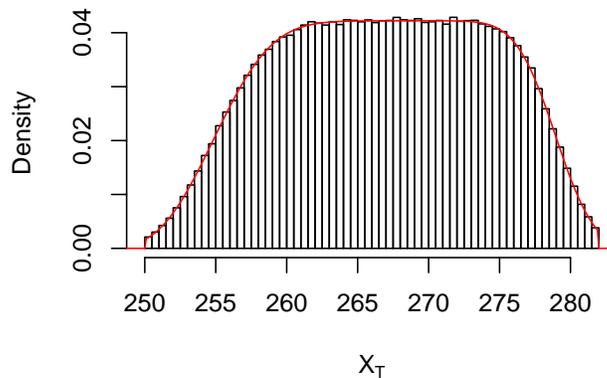


Figure 1.17: As in Figure (1.13), with $\delta/2 = 15.00$

$$\begin{aligned}
 &= \int_{y_0 - \delta/2}^{y_0 + \delta/2} \int_{x_0 - \delta/2}^{x_0 + \delta/2} \frac{P(Y = v \cap X_b = w \mid X_t) P(X_t)}{P(Y = v \cap X_b = w)} dv dw \\
 &\approx \sum_{i=0}^n \frac{P(Y = y_0 - \delta(i\Delta l - \frac{1}{2}) \cap X_b = x_0 - \delta(i\Delta l - \frac{1}{2}) \mid X_t) P(X_t)}{P(Y = x_0 - \delta(i\Delta l - \frac{1}{2}) \cap X_b = x_0 - \delta(i\Delta l - \frac{1}{2}))} \quad (1.26)
 \end{aligned}$$

where v and w are the auxiliary integration variables, while y_0 and x_0 are the given values for Y and X_b respectively.

Figure (1.20) shows the contour plot of the standard deviation for the two-variable case, where the y-axis represents the δ used for the first guess and the x-axis the corresponding δ for the observation. Again, the value to which the standard

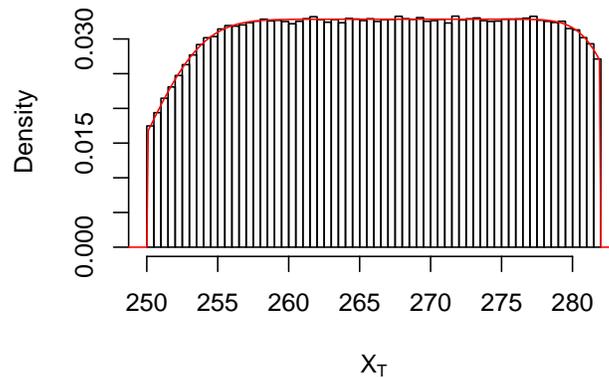


Figure 1.18: As in Figure (1.13), with $\delta/2 = 20.00$.

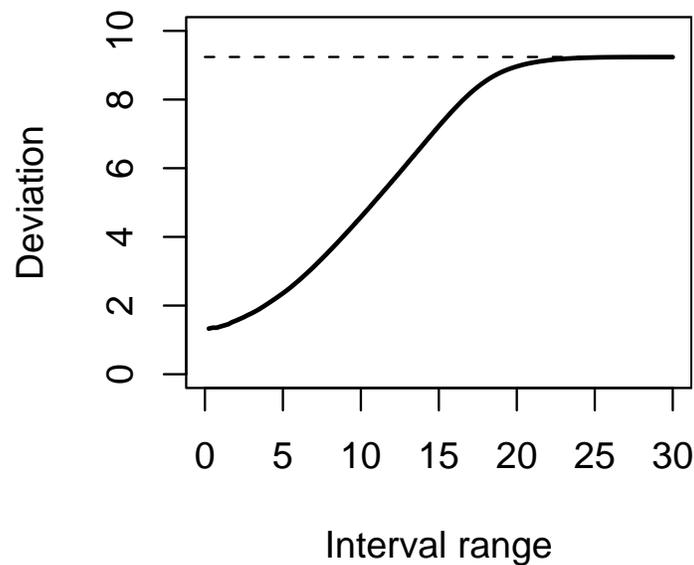


Figure 1.19: Standard Deviation on the growing of the interval δ . As reference value (dashed line), the standard deviation of X_t without conditional constraints is equal to 9.24.

deviation approaches is the unconditioned $P(X_t)$. The standard deviation is slightly changing if the interval for the observation increases with respect to the background. The variable with the biggest variance (i.e. the first guess in this case) has the main influence on the resulting probability.

1.2.6 Verification

To study the error of the analysis realistically, we consider the reference observations Y_r , that are assumed unbiased, independent of the observation Y and the first guess X_b , and with a small standard deviation with respect to the true state,

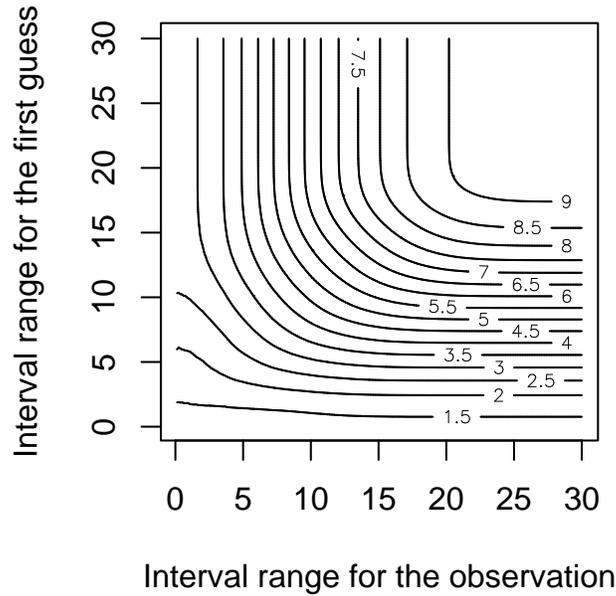


Figure 1.20: Standard Deviation on the growing of the interval δ for both the observation and the first guess.

X_t .

With the Mean Squared Error (MSE) verification penalty function in mind, we define the penalty function J' for the analysis X_a according to

$$J' = E [(Y_r - X_a)^2 | Y \cap X_b]. \quad (1.27)$$

The MSE verification method will favour the analysis which gives the lowest penalty function J' . Let us also define a similar penalty function with respect to the true state instead of the reference observations,

$$J = E [(X_t - X_a)^2 | Y \cap X_b]. \quad (1.28)$$

In their analytical form we get

$$J = \int_{-\infty}^{+\infty} (Y_r - X_a)^2 P(Y_r | Y \cap X_b) dY_r \quad (1.29)$$

$$J' = \int_{-\infty}^{+\infty} (X_t - X_a)^2 P(X_t | Y \cap X_b) dX_t. \quad (1.30)$$

We notice that $Y_r - X_t \mid Y \cap X_b$ has a symmetric distribution around zero, so that $E[(Y_r - X_t)(X_t - X_a) \mid Y \cap X_b] = 0$. We may write

$$\begin{aligned}
 J' &= E [((Y_r - X_t) + (X_t - X_a))^2 \mid Y \cap X_b] \\
 &= E [(Y_r - X_t)^2 \mid Y \cap X_b] \\
 &\quad + 2E [(Y_r - X_t)(X_t - X_a) \mid Y \cap X_b] \\
 &\quad + E [(X_t - X_a)^2] \\
 &= \sigma_r^2 + E [(X_t - X_a)^2 \mid Y \cap X_b] \\
 &= \sigma_r^2 + J,
 \end{aligned} \tag{1.31}$$

where we remember that σ_r is the reference observation error standard deviation. We define the “optimal analysis”, the value of $X_a = \widetilde{X}_a$, for which J has minimum value. We find from Eq. (1.29)

$$\begin{aligned}
 \min_{X_a} J &\Rightarrow \frac{\partial J}{\partial X_a} \Big|_{\widetilde{X}_a} = 0 \\
 &\quad - 2 \int X_t P(Y_r \mid Y \cap X_b) dX_t \\
 &\quad + 2\widetilde{X}_a \int P(Y_r \mid Y \cap X_b) dX_t = 0
 \end{aligned} \tag{1.32}$$

$$\widetilde{X}_a = E[X_t \mid Y \cap X_b]. \tag{1.33}$$

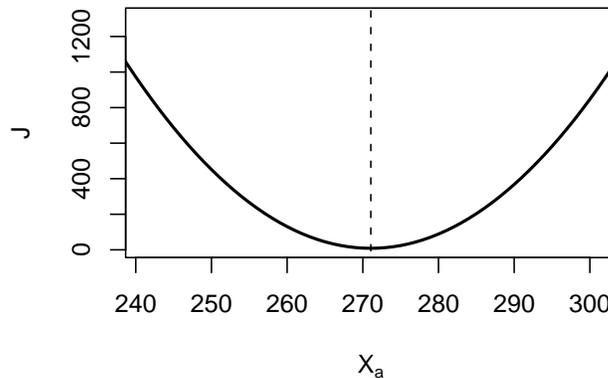


Figure 1.21: J' . The minimum is reached when $X_a = \widetilde{X}_a$

As we see, J and J' differ each other only for σ_r^2 , that is however independent from the analysis itself. This means that an analysis which is optimal with respect to the reference observation will also be optimal with respect to the true state. Figure

(1.21) shows the behaviour of the J' function.

In the case of Normal distributions (Eq. 1.21) we see that an analysis that is defined to give the maximum probability density, will also have the smallest MSE. In fact, the denominator in Eq. (1.10) is then independent from X_t , so the maximum for $P(X_t | Y \cap X_b)$ is obtained when the numerator is maximum

$$\begin{aligned} \max_{X_t} P(X_t | Y \cap X_b) &= \max_{X_t} \left[(2\pi\sigma_o\sigma_b) \exp\left(-\frac{(X_t - Y)^2}{\sigma_o^2}\right) \cdot \exp\left(-\frac{(X_t - X_b)^2}{\sigma_b^2}\right) \right] \\ &= \min_{X_t} \left[\frac{(X_t - Y)^2}{\sigma_o^2} + \frac{(X_t - X_b)^2}{\sigma_b^2} \right] = \frac{\sigma_b^2 Y + \sigma_o^2 X_b}{\sigma_b^2 + \sigma_o^2} \end{aligned}$$

which is equivalent to $E(X_t | Y \cap X_b)$ since it is the mean of the Normal distribution of Eq. (1.22). In the common 3D-Var implementations, all the probability density functions are assumed to be Normal and the cost function is defined so that it gives a minimum for the maximum probability density. As we have seen above, the 3D-Var approach will then yield the optimal MSE analysis.

1.3 Handling binary data

The concept of binary data can be introduced starting from the previous results. An easy way to think of binary observations is to consider the expression given in Eq. (1.24) infinitely increasing the upper limit of the integral. In other words, the observation condition is true if it lies over a threshold and false if it lies below. Thus, Eq. (1.24) becomes

$$\begin{aligned} P(X_t | Y \geq y_0) &= \frac{P(Y \geq y_0 | X_t) P(X_t)}{P(Y \geq y_0 | X_t)} \\ &= \int_{y_0}^{+\infty} P(Y = v | X_t) \frac{P(X_t)}{P(Y = v)} dv. \end{aligned} \quad (1.34)$$

Using the definitions given in Eq. (1.1) and Eq. (1.2) and the result of Eq. (1.19), we obtain

$$P(X_t | Y \geq y_0) = \frac{1}{\sqrt{2\pi\sigma_o^2}} \int_{y_0}^{+\infty} e^{-\frac{(X_t - v)^2}{2\sigma_o^2}} dv. \quad (1.35)$$

Such a PDF is the cumulative probability function for the Normal distribution with mean in y_0 and variance σ_o^2 . Figure 1.22 represents the simulated and theoretical PDF for Eq. (1.35), when y_0 is set to $264K$. We observe that the probability grows from 0 up to his maximum value from about 260 to $268K$ and then it holds steady. This range becomes more narrow when the standard error of the observation

decreases.

Let's consider the equation for $P(X_T | (Y > y_0) \cap X_b)$. Following the same approach, we set X_b to a value x_0 . We obtain, from the last equation and the results of Eq. (1.24) and Eq. (1.25),

$$\begin{aligned}
 P(X_t | (Y \geq y_0) \cap X_b) &= \\
 &\int_{y_0}^{+\infty} \int_{x_0-\delta/2}^{x_0+\delta/2} P(Y = v | X_t) P(X_b = w | X_t) \frac{P(X_t)}{P(Y = v \cap X_b = w)} dv dw \quad (1.36) \\
 &= \frac{1}{\sqrt{2\pi\sigma_o^2\sigma_b^2}} \int_{y_0}^{+\infty} \int_{x_0-\delta/2}^{x_0+\delta/2} e^{-\frac{(X_t-v)^2}{2\sigma_o^2}} e^{-\frac{(X_t-w)^2}{2\sigma_b^2}} \frac{\sqrt{2\pi(\sigma_o^2 + \sigma_b^2)}}{(b-a) \exp\left(-\frac{(w-v)^2}{2\sigma_o^2 + \sigma_b^2}\right)} dv dw. \quad (1.37)
 \end{aligned}$$

Figures (1.24) to (1.26) show $P(X_T | (Y > y_0) \cap X_b)$ for three different y_0 . If y_0 is much smaller than x_0 , the resulting PDF is very similar to $P(X_T | X_b)$ (Fig. 1.23), since $P(X_T | (Y > y_0))$ is already at the maximum value for most $(X_t | X_b)$, and the binary observation condition doesn't affect the resulting PDF. On the contrary, if y_0 is equal to, or greater than, x_0 , (see the figures 1.25 and 1.26 for the other two cases with $y_0 = 270K$ and $y_0 = 275K$ respectively), the binary observation condition controls noticeably the resulting PDF. Finally, we observe that such a PDF presents a Normal-like behaviour. The comparison with a Normal distribution is carried out imposing the coincidence between the expected values of the actual PDF of $(X_T | (Y > y_0) \cap X_b)$ and its Normal approximation. We also define the estimate of variance for X_t using the Maximum Likelihood Estimate theory:

$$\begin{aligned}
 \hat{X}_t &= E(X_T | (Y > y_0) \cap X_b) \\
 \hat{s}^2 &= \frac{1}{n} \sum_{i=1}^n (X_{t_i} - \hat{X}_t)^2 \quad (1.38)
 \end{aligned}$$

We now compare $P(X_t | Y)$ with the distribution that follows the law $N(\hat{X}_t, \hat{s}^2)$. Figure (1.27), (1.28) and (1.29) show the two PDFs for different values of y_0 , while (1.30) the Quantile-quantile plot. All of them demonstrate the goodness of the Normal approximation. This is a very useful property, which will be exploited in the next section, with regards to the variational assimilation. In fact, having a Gaussian-like probability density function for the analysis even with a non-Gaussian distribution for the observation error, allows us to retrieve a *pseudo-observation* with Gaussian error, suitable for variational assimilation.

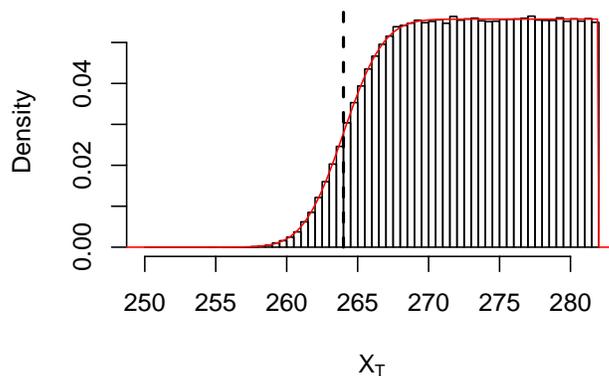


Figure 1.22: $P(X_T | (Y > y_0))$ as computed from the data set (histogram) and its analytical solution (grey line), in the case of binary observation. The dashed line is in correspondence of the y_0 value, set to $264K$.

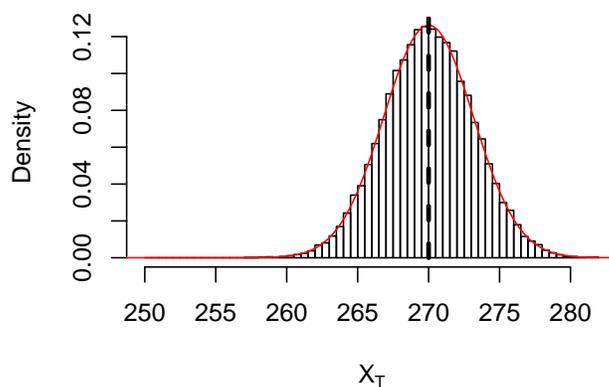


Figure 1.23: $P(X_T | X_b)$ as computed from the data set (histogram) and its analytical solution (grey line). The dashed line is in correspondence of the x_0 value, set to $270K$.

1.3.1 Pseudo-observations

At this moment, most of the Limited Area Models (LAMs) implement the variational theory for the assimilation of the observations and we have to make the assimilation theory for binary observations consistent with the variational assimilation theory. The aim of the variational assimilation consists of finding X_t in order

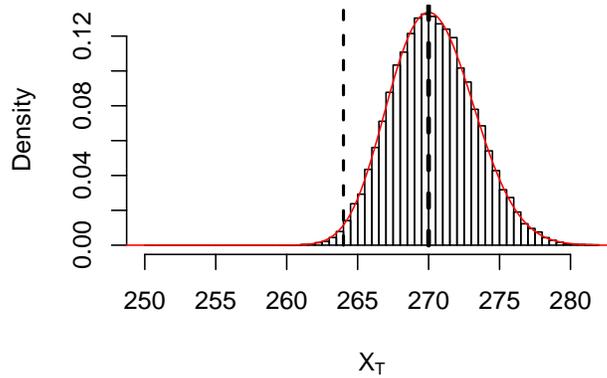


Figure 1.24: $P(X_T | (Y > y_0) \cap X_b)$ in the case of binary observation as computed from the data set (histogram) and its analytical solution (grey line), when $y_0 = 264K$. The strong dashed line shows x_0 ($270K$), while the slight dashed line y_0 .

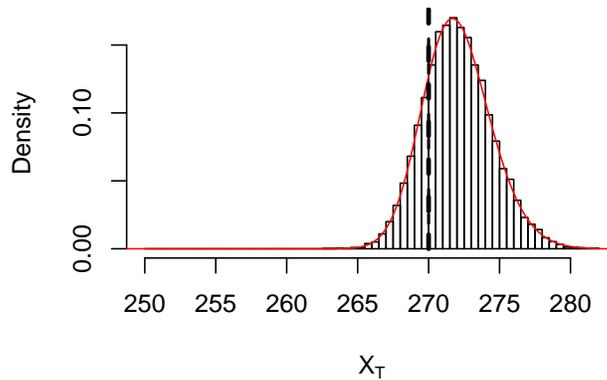


Figure 1.25: $P(X_T | (Y > y_0) \cap X_b)$ in the case of binary observation as computed from the data set (histogram) and its analytical solution (grey line), when both x_0 and y_0K are equal to $270K$ (dashed line).

to have the maximum value for $P(X_T | Y \cap X_b)$, namely we find

$$\begin{aligned} \max_{X_t} P(X_T | Y \cap X_b) &= \frac{P(Y | X_t)P(X_T | X_b)}{P(Y | X_b)} = \\ \min_{X_t} J &= -\ln P(X_T | Y \cap X_b) = -\ln [P(Y | X_t)] - \ln [P(X_t | X_t)] + k \end{aligned} \quad (1.39)$$

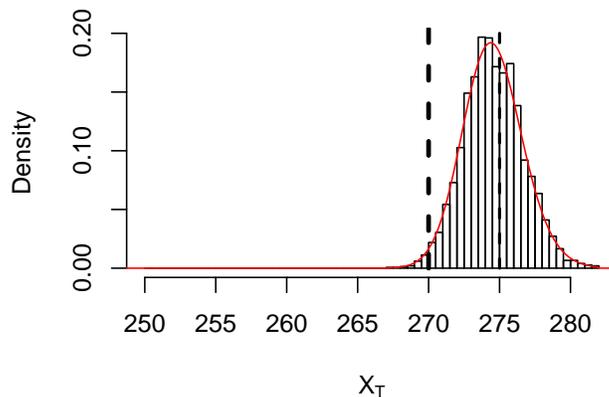


Figure 1.26: $P(X_T | (Y > y_0) \cap X_b)$ in the case of binary observation as computed from the data set (histogram) and its analytical solution (grey line), when $y_0 = 275K$. The strong dashed line shows x_0 ($270K$), while the slight dashed line y_0 .

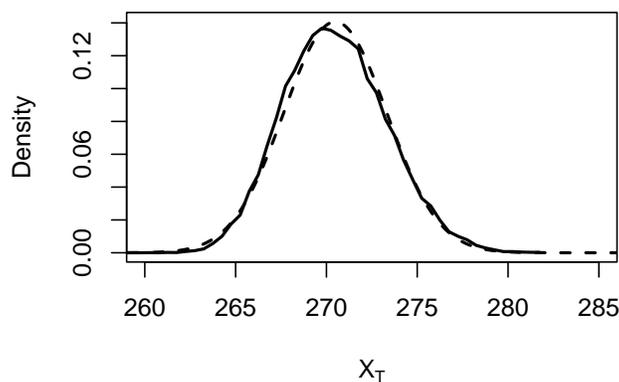


Figure 1.27: Comparison between $P(X_T | (Y > y_0) \cap X_b)$ (solid line), and the Normal PDF with parameters given by Eq. (1.38) (dashed line). Values for y_0 and x_0 are $264K$ and $270K$ respectively.

where k is a constant. The cost function J , introduced in the last equation, with regards to Eq. (1.2) and Eq. (1.19), is defined

$$J = \frac{1}{2} \frac{(Y - X_T)^2}{\sigma_o^2} + \frac{1}{2} \frac{(X_b - X_T)^2}{\sigma_b^2}. \quad (1.40)$$

The variational assimilation² assumes that all the conditional PDF of the true state are Normal. So, in order to assimilate binary observations, the PDF described by

²Now we consider the problem of finding X_t still one-dimensional.

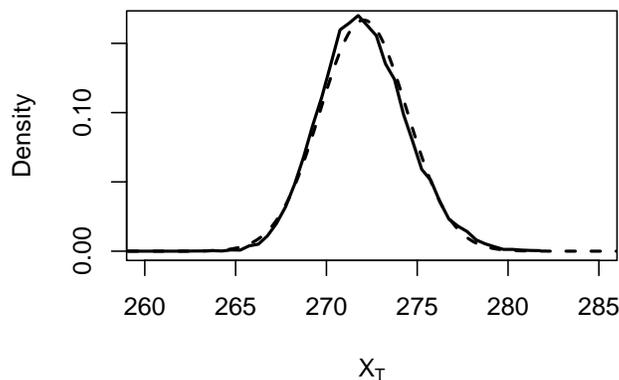


Figure 1.28: Comparison between $P(X_T | (Y > y_0) \cap X_b)$ (solid line), and the Normal PDF with parameters given by Eq. (1.38) (dashed line). Value for both y_0 and x_0 is $270K$.

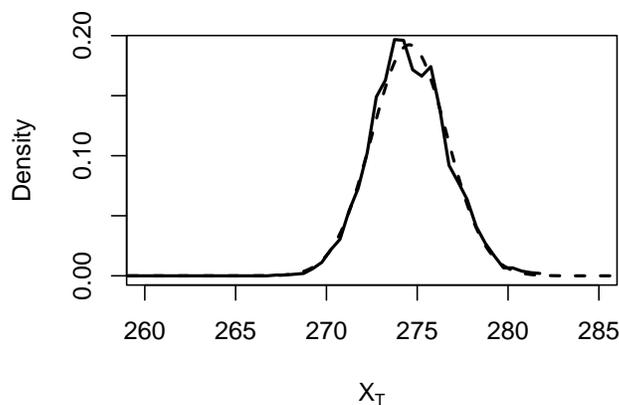


Figure 1.29: Comparison between $P(X_T | (Y > y_0) \cap X_b)$ (solid line), and the Normal PDF with parameters given by Eq. (1.38) (dashed line). Values for y_0 and x_0 are $275K$ and $270K$ respectively.

Eq. (1.37) is not suitable since it is not a Normal distribution. A possible approach is to use a *pseudo-observation* y' , so that

$$E[X_T | (Y > y_0) \cap (X_b = x_0)] = E[X_T | (Y = y') \cap (X_b = x_0)]$$

where the PDF of $(X_T | (Y = y') \cap (X_b = x_0))$ is Normal. This condition means that the analysis is still optimal, referring to the derivation of (1.33). Considering the Normal distribution of the true state given the observation and the first guess,

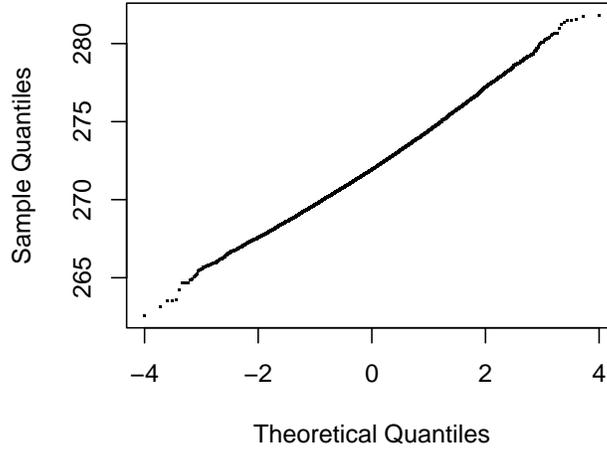


Figure 1.30: Quantile-quantile plot. Comparison between $P(X_T | (Y > y_0) \cap X_b)$ and the Normal PDF with parameters given by Eq. (1.38). Value for both y_0 and x_0 is $270K$.

Eq. (1.22), the definitions in Eq. (1.38) and $y' = E(X_t | Y = y')$, we find y' :

$$\begin{cases} \sigma_{o'}^2 = \frac{\sigma_o^2 s^2}{\sigma_o^2 - s^2} \\ y' = \frac{(\sigma_b^2 + \sigma_{o'}^2) \hat{X}_t - \sigma_{o'}^2 x_0}{\sigma_b^2} \end{cases}$$

Note how the background strongly affects the definition of y' . This can be seen how the need of additional information in the use of binary observations, that are poor with respect to a continuous observation. Figures (1.31), (1.32) and (1.33) show the $N(y', \sigma_{o'}^2)$, namely $P(X_T | (Y = y'))$, against $P(X_T | (Y = y') \cap (X_b = x_0))$ and $P(X_T | (X_b = x_0))$, while figures (1.34) and (1.35) show $E(X_T | (Y > y_0) \cap X_b)$ and y' on the growing of y_0 . The range of standard deviation is also represented for these two figures. We observe that $E(X_T | (Y > y_0) \cap X_b)$ is constant and equal to $E(X_T | X_b)$ until y_0 is small enough, depending on σ_o^2 and σ_b^2 , and then it increases while its standard deviation decreases.

1.3.2 Sensitivity to small shifts

To find a probability density function for $P(Y | X_t)$ such as $P(X_t | Y \cap X_b)$ was Normal, we study the functions family given by

$$P(Y | X_t) = f(x_t) = e^{\alpha x_t^2 + \beta x_t + \gamma}. \quad (1.41)$$

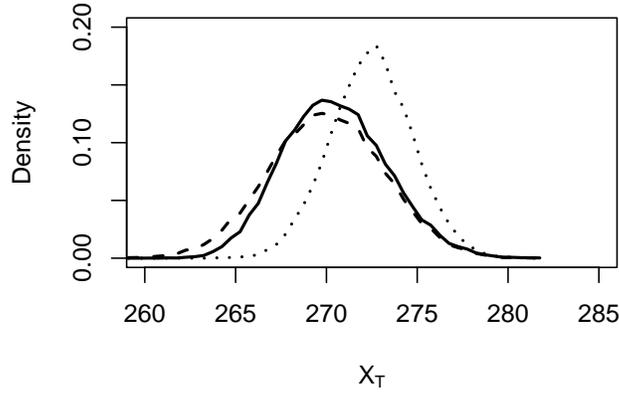


Figure 1.31: Comparison between $P(X_T | (Y = y') \cap (X_b = x_0))$ (solid line), $P(X_T | (X_b = x_0))$ (dashed line) and $P(X_T | (Y = y'))$ (dotted line), when x_0 and y_0 are set to 270K and 264K respectively.

Considering the equation

$$\begin{aligned} P(X_t | Y \cap X_b) &= \frac{P(Y | X_t)P(X_t | X_b)}{P(Y | X_b)} \\ &= aP(Y | X_t)P(X_t | X_b), \end{aligned} \quad (1.42)$$

where

$$a = [P(Y | X_b)]^{-1} \sim const \quad (1.43)$$

we can impose

$$\frac{1}{\sigma_a} e^{-\frac{1}{2} \left(\frac{x-m_a}{\sigma_a} \right)^2} = a \frac{1}{\sigma_b} e^{-\frac{1}{2} \left(\frac{x-m_b}{\sigma_b} \right)^2} e^{\alpha x_t^2 + \beta x_t + \gamma} \quad (1.44)$$

m_a , and σ_a are the expected value and the standard deviation of $P(X_t | Y \cap X_b)$ while m_b , and σ_b are the expected value and the standard deviation of $P(X_b | X_t)$

Thus we have for α , β and γ :

$$\begin{aligned} \alpha &= \frac{1}{2\sigma_b^2} - \frac{1}{2\sigma_a^2} \\ \beta &= \frac{m_a}{\sigma_a^2} - \frac{m_b}{\sigma_b^2} \\ \gamma &= \frac{m_b^2}{2\sigma_b^2} - \frac{m_a^2}{2\sigma_a^2} + \frac{1}{2} \ln \frac{\sigma_b^2}{\sigma_a^2} - \ln a \end{aligned}$$

In order to study the shift of $P(Y | X_t)$ on the x axis as function of m_a and m_b , we impose $\sigma_a^2 \sim \sigma_b^2$, neglecting the quadratic terms. We obtain for the parameters

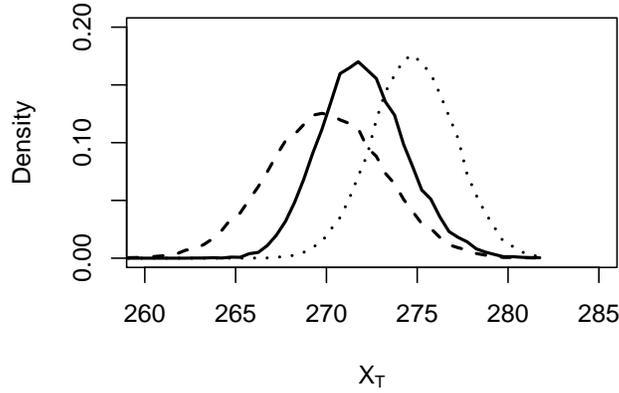


Figure 1.32: Comparison between $P(X_T | (Y = y') \cap (X_b = x_0))$ (solid line), $P(X_T | (X_b = x_0))$ (dashed line) and $P(X_T | (Y = y'))$ (dotted line), when both x_0 and y_0 are set to $270K$.

in $f(x_t)$:

$$\begin{aligned}\alpha &= 0 \\ \beta &= \frac{m_a - m_b}{\sigma_b^2} \\ \gamma &= \frac{m_b^2 - m_a^2}{2\sigma_b^2} - \ln a\end{aligned}$$

Then, the slope of $P(Y | X_t)$ is

$$\frac{dP}{dX_t} = \beta e^\gamma e^{\beta X_t}, \quad (1.45)$$

and the increment of the probability $P(Y | X_t)$ is

$$\frac{\frac{dP(Y|X_t)}{dX_t}}{P} = \frac{d \ln P(Y | X_t)}{dX_t} = \beta \quad (1.46)$$

To study the relationship between the function previously founded, the actual binary observations PDF and the conditional probability for X_t , we exclude two cases,

- when $y_0 \ll x_0$ (e.g. if $x_0 - y_0 > 2.5(\sigma_o + \sigma_b)$, see Fig. 1.36);
- when $y_0 \gg x_0$ (see Fig. 1.37).

In the first case, $P(X_t | Y > y_0 \cap X_b)$ will be in practise equal to $P(X_t | X_b)$ since the binary observation does not affect the resulting probability density function. On the other side, the second case represents a critical case since $P(X_t | X_b)$ and

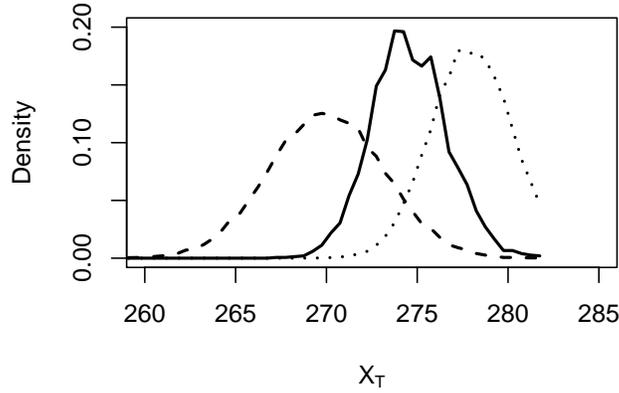


Figure 1.33: Comparison between $P(X_T | (Y = y') \cap (X_b = x_0))$ (solid line), $P(X_T | (X_b = x_0))$ (dashed line) and $P(X_T | (Y = y'))$ (dotted line), when x_0 and y_0 are set to $270K$ and $275K$ respectively.

$P(Y | X_t)$ are very small in the range of interest, and $P(X_t | Y > y_0 \cap X_b)$, as product of these two PDFs, is dramatically affected by their precision (see in Fig. 1.37 how vary the histogram with respect to the Normal curve). Under these hypothesis, (1.46) provides a simple mathematical interpretation of the shift that occur for different values of the observation threshold y_0 . In fact, if $X_a = m_a$ and $X_b = m_b$, and assuming again that $\sigma_a^2 \sim \sigma_b^2$, we have that

$$X_a - X_b = \sigma_b^2 \frac{d \ln P(Y | X_t)}{dX_t}. \quad (1.47)$$

When in fact the probability of Y approaches to the maximum value, the slope of $\ln P(Y | X_t)$ and the $(X_a - X_b)$ is infinitesimal.

1.3.3 Some numerical approximations

For numerical optimisations and future developments, it can be useful to approximate the PDF of $P(X_t | Y > y_0)$ to a function $g(x)$. We suppose that:

- such a family be exponential;
- the exponent be linear.

So, we find the family function: e^{ax+k} , that in a more convenient way can be rewritten $e^{a(x-y_0)+b}$. We also require that $g(x = y_0) = D/2$, where D is the probability $P(X_t \gg y_0 | Y > y_0)$, i.e. the probability of X_t values that have maximum probability to occur, given an observation value y_0 and its standard deviation σ_o . It follows

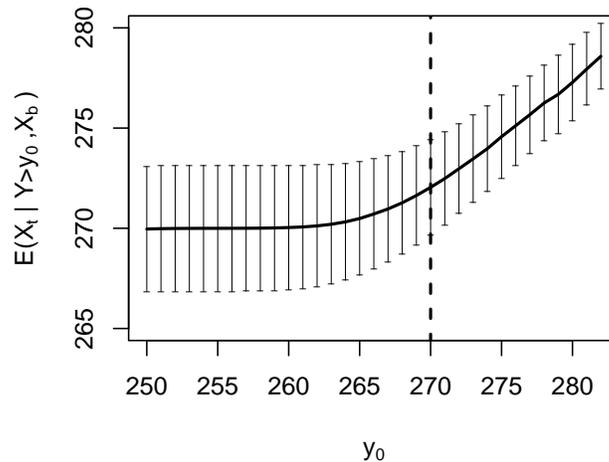


Figure 1.34: Expected value and standard deviation range for $(X_T | (Y > y_0) \cap (X_b = x_0))$ on the growing of y_0 , with $x_0 = 270K$ (in correspondence of the dashed line).

that $b = \ln(D/2)$, and we write once again $g(x)$,

$$g(x) = \frac{D}{2} e^{a(x-y_0)}. \quad (1.48)$$

To be more precis, now we split the function we are looking for in two parts, imposing continuity in y_0 and symmetry with respect to the axis $x = y_0$ and $y = D/2$, that is:

$$\frac{D}{2} - g(y_0 - x) = g(x - y_0) - \frac{D}{2}$$

and reminding that for values greater than c (where c is the upper limit for the uniform $P(X_t)$), we finally obtain:

$$g(x) = \begin{cases} \frac{D}{2} e^{a(x-y_0)} & \text{if } x < y_0 \\ D \left(1 - \frac{1}{2} e^{a(y_0-x)}\right) & \text{if } y_0 < x < c \\ 0 & \text{if } x > c \end{cases}$$

To find the value for D , we impose a normalisation, since the integral of $g(x)$ over x has to be equal to 1. Simplifying the problem, for symmetry we have that (see Fig 1.38)

$$\int_{-\infty}^{+\infty} g(x) du = \int_{-\infty}^{y_0} g(x) du + \left[D(c - y_0) - \int_{-\infty}^{y_0} g(x) du \right] = D(c - y_0) \quad (1.49)$$

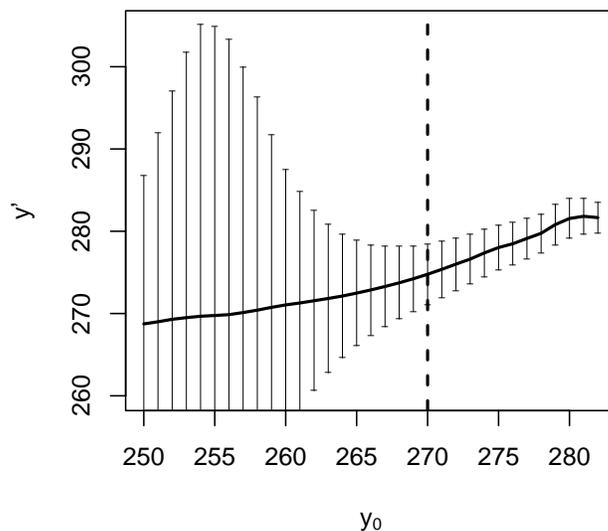


Figure 1.35: Expected value and standard deviation range for $(X_T | Y = y')$ on the growing of y_0 , with $x_0 = 270K$ (in correspondence of the dashed line).

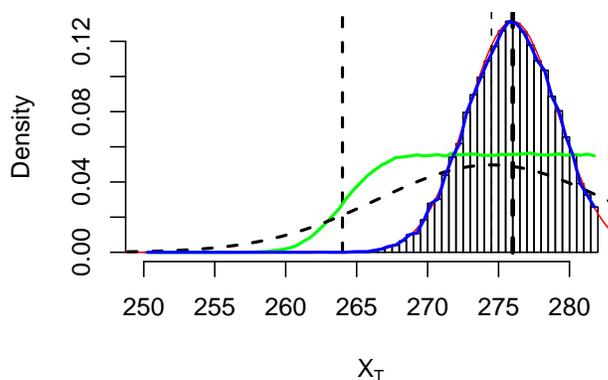


Figure 1.36: Histogram and Normal-like PDF (red line) for $P(X_t | Y > y_0 \cap X_b)$, $P(X_t | Y > y_0)$ (green line), $P(X_t | Y = y')$ (dashed line), $P(X_t | X_b = x_0)$ (blue line) for $x_0 = 276K$ and $y_0 = 264K$.

and so

$$D = \frac{1}{c - y_0} \tag{1.50}$$

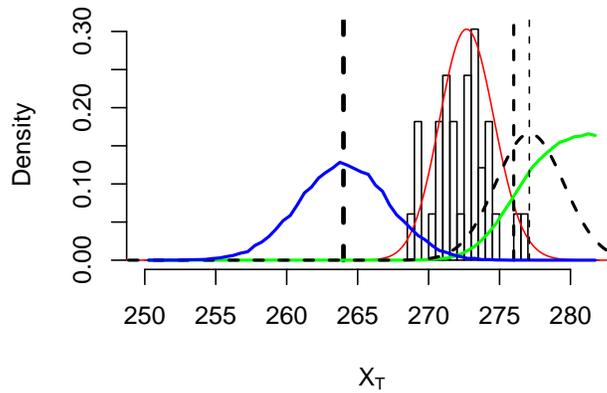


Figure 1.37: Histogram and Normal-like PDF (red line) for $P(X_t | Y > y_0 \cap X_b)$, $P(X_t | Y > y_0)$ (green line), $P(X_t | Y = y')$ (dashed line), $P(X_t | X_b = x_0)$ (blue line) for $x_0 = 264K$ and $y_0 = 276K$.

Note that the exact solution differs very slightly:

$$\int_{-\infty}^{+\infty} g(x) = D(c - y_0) + \frac{D}{2a} e^{a(y_0 - c)}$$

and

$$D = \frac{1}{c - y_0 + \frac{1}{2a} e^{a(y_0 - c)}}$$

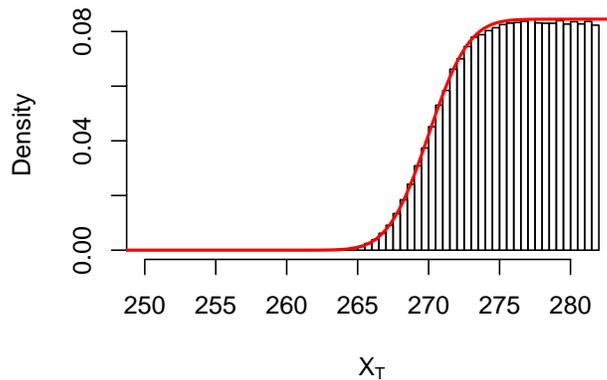


Figure 1.38: $P(X_t | Y > y_0)$.

Finally, Fig. 1.40 shows the approximation $g'(x)$ of $g(x)$, obtained without splitting the function, but considering that $g(x)$ affects $P(X_t)$ mainly when it gets

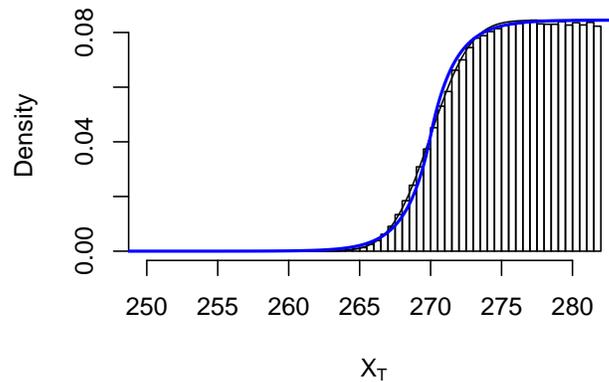


Figure 1.39: $g(x)$.

low values; thus the same function of Eq. (1.49) – for $x < y_0$ – is used unless $g'(x)$ is greater than D .

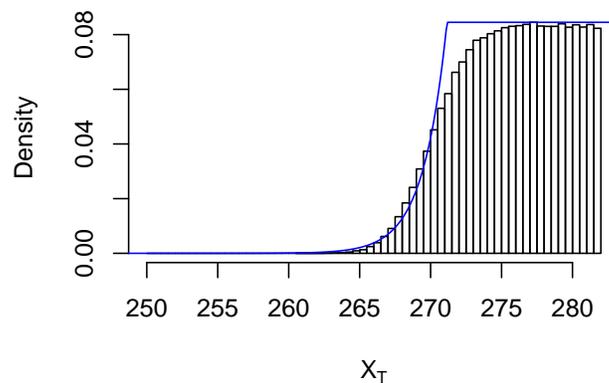


Figure 1.40: $g'(x)$.

1.4 One dimensional assimilation of cloud cover

Up to now we have considered to deal with identity observation operator. We also have assumed that the observation could affect only one physical quantity. To approach the problem of assimilating cloud cover data more realistically, we extend the precedent formalism to the case that two variables are influenced by the observation. In terms of variational algorithms, these variables are called “control state variables” since only these will enjoy the analysis increment at the end of

the assimilation. It is reasonable that the two “control state variables” will be the temperature T and the specific humidity³ q . In addition we will take into account a realistic forward model for computing the cloud cover from T and q . First of all, in order to extend the previous formalism to a multi-variate case, let’s consider the PDF⁴ for the background and the observation, given the true state of the atmosphere

$$P(Y | X_t) = \frac{1}{2\pi\sqrt{|\mathbf{R}|}} \exp\left(-\frac{1}{2}[(Y - \mathbf{H}(\mathbf{X}_t))^T \mathbf{R}^{-1} (Y - \mathbf{H}(\mathbf{X}_t))]\right) \quad (1.51)$$

$$P(X_b | X_t) = \frac{1}{2\pi\sqrt{|\mathbf{B}|}} \exp\left(-\frac{1}{2}[(\mathbf{X}_b - \mathbf{X}_t)^T \mathbf{B}^{-1} (\mathbf{X}_b - \mathbf{X}_t)]\right) \quad (1.52)$$

where:

- \mathbf{X}_t (and \mathbf{X}_b) is the true-state (and background) vector, $\mathbf{X}_t = (T, q)^T$;
- \mathbf{H} is the “forward operator”, moving the control variables into the observation space. This operator is given by the product of the interpolation operator by the observation operator. We will assume to handle with observation measured in the same point of X ; therefore the interpolation operator will be the identity matrix, and \mathbf{H} becomes the cloud cover model;
- \mathbf{R} is the observation errors covariances matrix, that include both the representativeness (non-perfect cloud cover operator) and the instrumental error; since we have only one observation, following the notation in the previous section $\mathbf{R} = \sigma_o^2$;
- \mathbf{B} is the background error covariances matrix. We will assume that the temperature and humidity errors be uncorrelated, so that our \mathbf{B} matrix will be a diagonal matrix consisting of the temperature and humidity error variances.

The formalism for finding $P(X_t | Y \cap X_b)$ is identical to the one-dimension case, except that the observation error is now affected by the additional contribute of the forward model (that however is thought of as unbiased and Gaussian).

1.4.1 Proposed approach

The basic idea is to find an analysis for the temperature and humidity given the cloud cover binary condition and the information from the former forecast. Once we have this analysis, whose errors are assumed Gaussian distributed, we retrieve a “pseudo-observation”, to use in the variational assimilation together with the

³We choose the specific and not the relative humidity for consistency with assimilation systems and forecast models.

⁴In Eq. (1.51) and Eq. (1.52) the multivariate normal distribution is for the bi-variate case, since the power of 2π at the denominator is $n/2$, being n the dimension of Y or X_b .

conventional observations. The formulation of the analysis represents the first step for the detection of the “pseudo-observation”. We consider to directly deal with the binary observation K , i.e. we already have the 0 (no-cloud) or 1 (cloud) information about cloudiness. Therefore we focus on the probability density function for the true state of atmosphere given the background and the “status of cloud/no cloud”. We also define our forward model as the probability of having cloud given the true state of the atmosphere. Thus, we have:

$$H(X_t) \equiv P(K = 1 | X_t) \quad (1.53)$$

$$P(K = 0 | X_t) = 1 - H(X_t) \quad (1.54)$$

$$\begin{aligned} P(K = 1 | X_b) &= \int_{-\infty}^{+\infty} P(K = 1 | X_t) P(X_t | X_b) dX_t \\ &= \int_{-\infty}^{+\infty} H(X_t) P(X_t | X_b) dX_t \end{aligned} \quad (1.55)$$

and for the PDF of X_t

$$P(X_t | K = 1 \cap X_b) = \frac{P(K = 1 | X_t) P(X_t | X_b)}{P(K = 1 | X_b)} \quad (1.56)$$

$$= \frac{H(X_t) P(X_t | X_b)}{\int_{-\infty}^{+\infty} H(X_t) P(X_t | X_b) dX_t} \quad (1.57)$$

Note that this procedure is immediately suitable also for the real satellite observation Y , since from a Bayesian point of view we have

$$P(Y | X_t) = P(K = 1 | X_t) P(Y | K = 1) + (1 - P(K = 1 | X_t)) P(Y | K = 0) \quad (1.58)$$

We can define our analysis \widetilde{X}_a as the expected value of the PDF of the true state

$$\widetilde{X}_a = E[X_t | K \cap X_b]. \quad (1.59)$$

Repeating exactly the same formalism in the previous section, it can be shown that our analysis \widetilde{X}_a is optimal with respect to the RMSE verification. Using Eq. (1.53) and Eq. (1.55), the analytical solution of Eq. (1.59) is given by

$$\begin{aligned} P(X_t | K \cap X_b) &= \frac{P(K | X_t) P(X_t | X_b)}{P(K | X_b)} \quad (1.60) \\ &= \frac{\mathbf{H}(\mathbf{X}_t) \frac{1}{2\pi|\mathbf{B}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{X}_t - \mathbf{X}_b)^T \mathbf{B}^{-1}(\mathbf{X}_t - \mathbf{X}_b)\right]}{\int_{-\infty}^{+\infty} \mathbf{H}(\mathbf{X}_t) \frac{1}{2\pi|\mathbf{B}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{X}_t - \mathbf{X}_b)^T \mathbf{B}^{-1}(\mathbf{X}_t - \mathbf{X}_b)\right] d\mathbf{X}_t}. \end{aligned}$$

This theoretical analysis is shown in Figures 1.41 to 1.43 against the background PDF and the distribution⁵ of $(K|X_t)$ for different background temperatures. In these plots the three PDF (for the background, the observation and the analysis) are drawn. Since is only shown the case $K = 1$, the points on the left of the grey lines (contour of $H(X_t)$) give $P(K | X_t) = 1$, while on the right the $P(K | X_t) = 0$. Figure 1.44 also presents the analysis \widetilde{X}_a and the analysis increments $(\widetilde{X}_a - X_b)$ on the growing of the background temperature. When the probability to have cloud (no-cloud) given the background is high and our binary observation leads to cloud (no-cloud) condition, the increments are close to zero and the binary observation does not add any useful information to the background (Fig. 1.41). When instead the probability to have cloud (no-cloud) given the background is very low and we have $K = 1$ ($K = 0$), the analysis is forced to assure the cloud status, so that the increments is high and decreases (increases) linearly with the background (Fig. 1.43). In the intermediate cases (Fig. 1.42), the increment is less dramatic and grows following a quadratic function. Note also that for a continuous observation, the curves in Fig. 1.44 are linear.

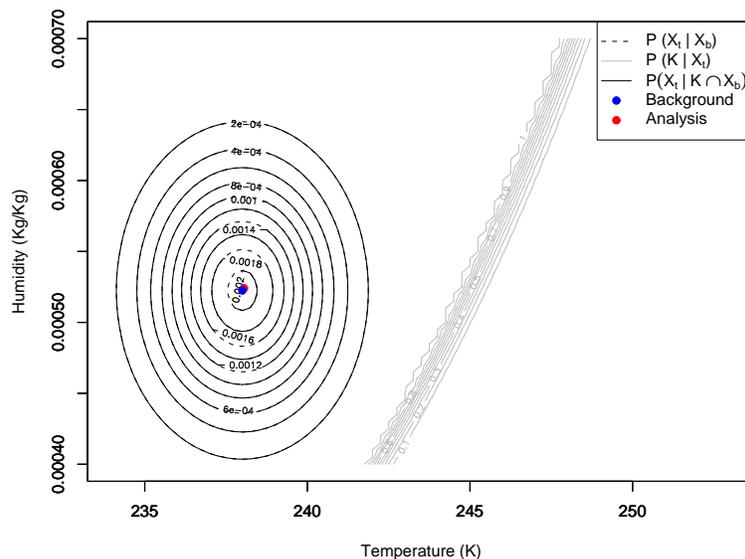


Figure 1.41: Contour of theoretical PDF of $(X_t | X_b)$, $(K | X_t)$ and $(X_t | K \cap X_b)$ when the background temperature is $238K$ and $K = 1$.

Finally, in order to find the *pseudo-observation* Y' , the classical formulation of

⁵In all this examples the binary observation is assumed to be set to “cloud” ($K = 1$). The forward model $H(X_t)$ that we have used will be presented later on.

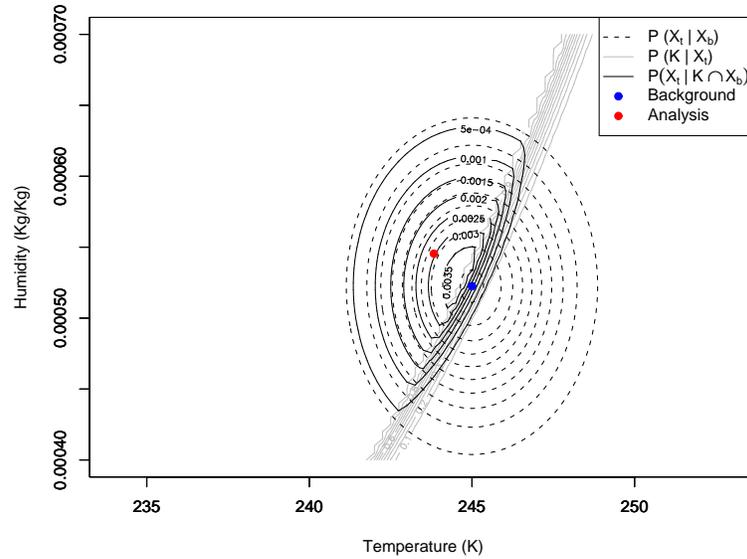


Figure 1.42: As in Fig. 1.41 when the background temperature is $245K$ and $K = 1$.

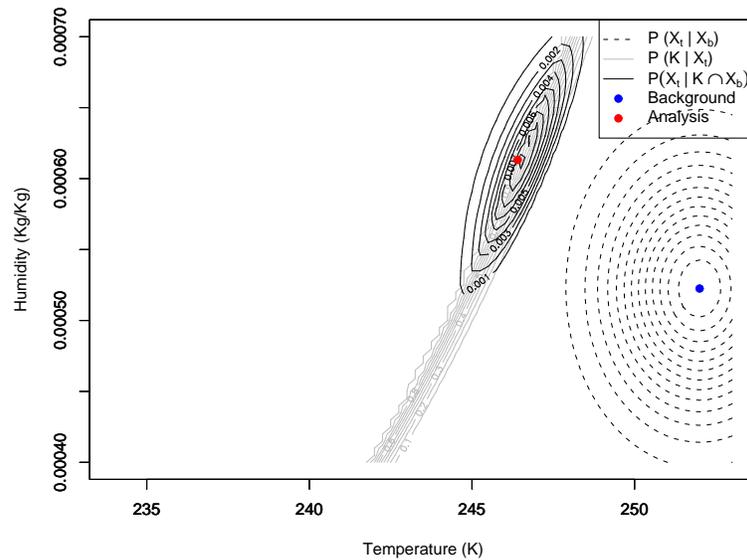


Figure 1.43: As in Fig. 1.41 when the background temperature is $252K$ and $K = 1$.

the analysis from the variational theory has to be inverted⁶

$$\mathbf{R} = \mathbf{H}\mathbf{B}(\mathbf{B} - \mathbf{A})^{-1}\mathbf{B}\mathbf{H}^T - \mathbf{H}\mathbf{B}\mathbf{H}^T \quad (1.61)$$

$$\widetilde{\mathbf{X}}_{\mathbf{a}} - \mathbf{X}_{\mathbf{b}} = \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}) (\mathbf{Y}' - \mathbf{H}\mathbf{X}_{\mathbf{b}}), \quad (1.62)$$

⁶In practise, this formula has to be replaced with more efficient minimization algorithms.

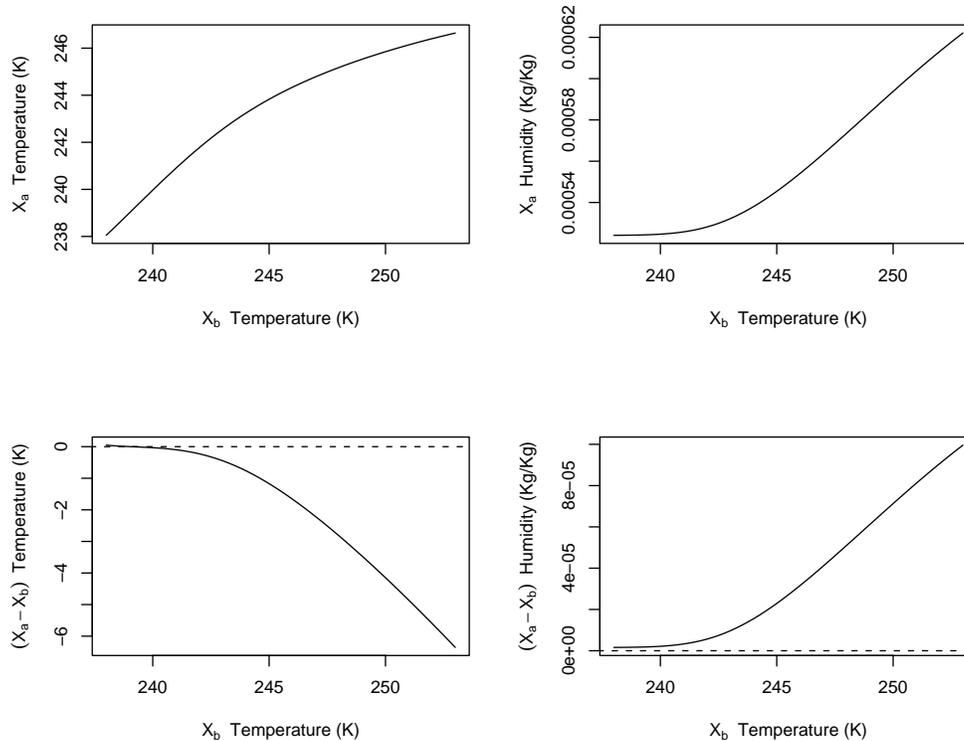


Figure 1.44: Analysis and increments as obtained theoretically for different values of the background both for case $K = 1$ (left) and $K = 0$ (right).

where $\tilde{\mathbf{S}}$ is the estimated variance for $P(X_t | K = 1 \cap X_b)$.

1.4.2 On the forward model

Many models have been developed during last years to compute cloud cover. A few of these formulations are prognostic, integrating the cloud cover over the time either in a four dimensional variational assimilation framework (Jakob, 1999) or in a forecast model (Tompkins, 2002). In terms of a three dimensional variational (3DVAR) assimilation, that is our reference assimilating system, such an approach can not be undertaken, so we will concentrate on diagnostic formulation for the cloud cover. Climate modellers have given great importance to the cloud cover computations in an effort to better reproduce the radiative balance. However, every approach aims to find some reasonable predictors for cloud cover, proposing a relation between these predictors and the cloud cover itself. Unfortunately, most of these formulations choose as cloud cover predictors quantities related to microphysical processes, like the cloud water content or the liquid water content. These variables are not control variables in the three dimensional data assimilation (at least as developed for several LAMs). Unless to explicitly compute these variables,

they are not available during the analysis.

It seems reasonable for now to choose a forward operator as simple as possible and eventually in the future going back to this problem. Further, the vertical integration of the forward operator represents a delicate problem even more than the forward operator itself.

Many studies (e.g. Walcek, 1994) suggest that simple relations between the cloud cover and the relative humidity can be established. The basic idea is that the cloud is formed when the the relative humidity is equal to 100%; anyway, since the humidity that we deal with in NWP is a measure averaged in each grid cell, we assume that clouds can exist also for values of relative humidity smaller than 100%. A simple relation for the cloud cover c , known the threshold RH_{cr} below which no cover can form, is given by:

$$c = \left[\frac{RH - RH_{cr}}{1 - RH_{cr}} \right]^2 \quad (1.63)$$

In the former equation, relative humidity RH is obtained from the control variables T and q as follows

$$RH = 100 \frac{q}{q_s} = 100 \frac{q [p - (1 - 0.622) e_s(T)]}{0.622 e_s(T)}$$

where p is the pressure and $e_s(T)$ the saturation water vapour pressure, computed for instance through an empirical expression (Lowe, 1977). Equation (1.63) has been used operationally in ECMWF (Slingo, 1987) and replaced only in 1993 by a prognostic scheme (Tiedtke, 1993). Anyway the definition of the parameter RH_{cr} is problematic since it depends on many other factors, first of all the pressure level and the related nature of the cloud and also on local physiographic characteristics. As example, Teixeira (2001) found that for the mid-tropospheric atmosphere (500 – 400mb) cloud cover can not be neglected even for relative humidities less than 60%. We will return to this problem later on in this section.

1.4.3 Configuration and results

The simulation performed for this one-dimensional case, as in the previous sections, considers the true state of atmosphere uniformly distributed,

$$X_t \sim U(\mathbf{a}, \mathbf{b}), \quad (1.64)$$

while the background follows the Normal distribution

$$X_b - X_t \sim N(0, \mathbf{B}), \quad (1.65)$$

where temperature and humidity are considered uncorrelated for simplicity. For K we consider the expression

$$K = \begin{cases} 0 & \text{if } H(X_t) < \varepsilon, \\ 1 & \text{if } H(X_t) \geq \varepsilon. \end{cases} \quad (1.66)$$

where $\varepsilon \sim U(0, 1)$. This last condition is imposed to ensure that a few cases of tremendous inconsistency between the theoretical forward model and the K condition can occur.

The forward model is based on Eq. (1.63). Figure 1.46 shows the cloud cover on the growing of the relative humidity, while Fig. 1.47 shows the contour for different values of the two control variables. Note that this model is designed to reproduce the cloud cover in only one layer. The total cloud cover is normally obtained through the maximum-random overlap algorithm, that assumes to have maximum horizontal overlap for the cloud-cover between vertically adjacent cloudy layers, while random horizontal overlap is taken into account between non-adjacent cloudy layers. Therefore the abrupt slope in the forward model curve is normally damped by the multi-layers computation. However, in the case we're studying, it leads to a quasi-binary set of values for $H(X_t)$, even if for the computation of the background (Eq. 1.65) we had to reduce the variance of the specific humidity as computed through a 20 days NMC statistics in order to operate in a more significant interval of humidity. Figure 1.45 shows the distribution of $K = 1$ on the varying of humidity and temperature.

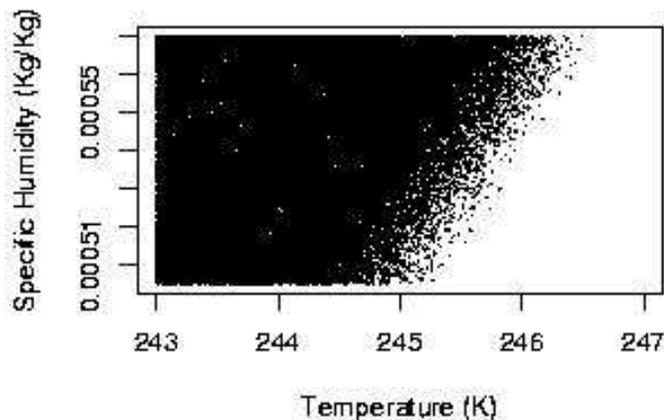


Figure 1.45: Distribution of $K = 1$ occurrences.

Results for analysis increments (see (Fig. 1.48)) match with the analytical solution of Eq. (1.60), leading to the same conclusions. The right wing of $(\widetilde{X}_a - X_b)$ in Fig. 1.48(a) and the left wing in Fig. 1.48(b) does not have a continuous trend because these cases occur rarely even in a large dataset.

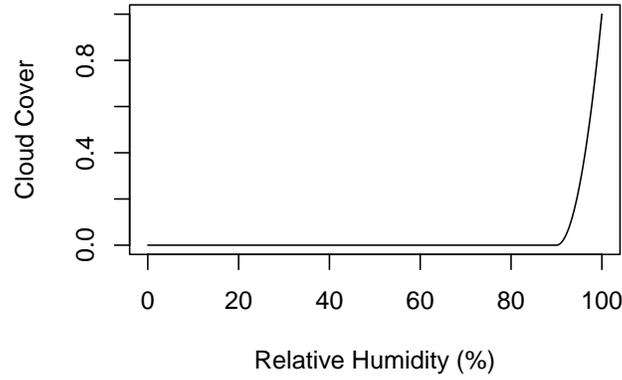


Figure 1.46: The forward model given in Eq. (1.63).

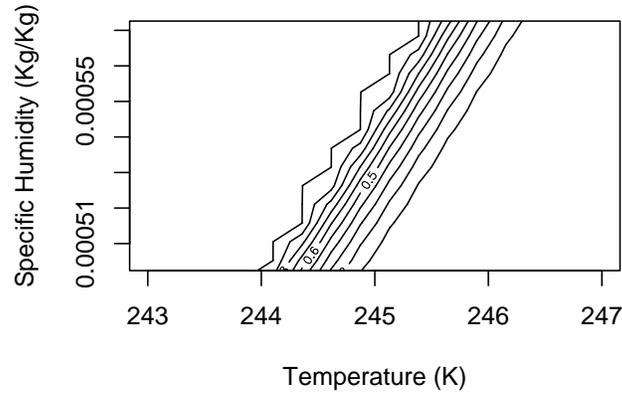


Figure 1.47: Contour of cloud cover as computed from Eq. (1.63) on the varying of temperature and specific humidity.

Now we consider the verification penalty function J

$$J = E \left[(X - \widetilde{X}_a)^2 \right] = E \left[E \left[(X - \widetilde{X}_a)^2 \mid K \cap X_b \right] \right] = E [J \mid K \cap X_b]. \quad (1.67)$$

Note that this J function corresponds to the Mean Squared Error. To be optimal our analysis should satisfy the requirement

$$\widetilde{X}_a = \min J \mid K \cap X_b, \quad (1.68)$$

where we now implicitly have that \widetilde{X}_a is a function of K and X_b . To compute (1.67)

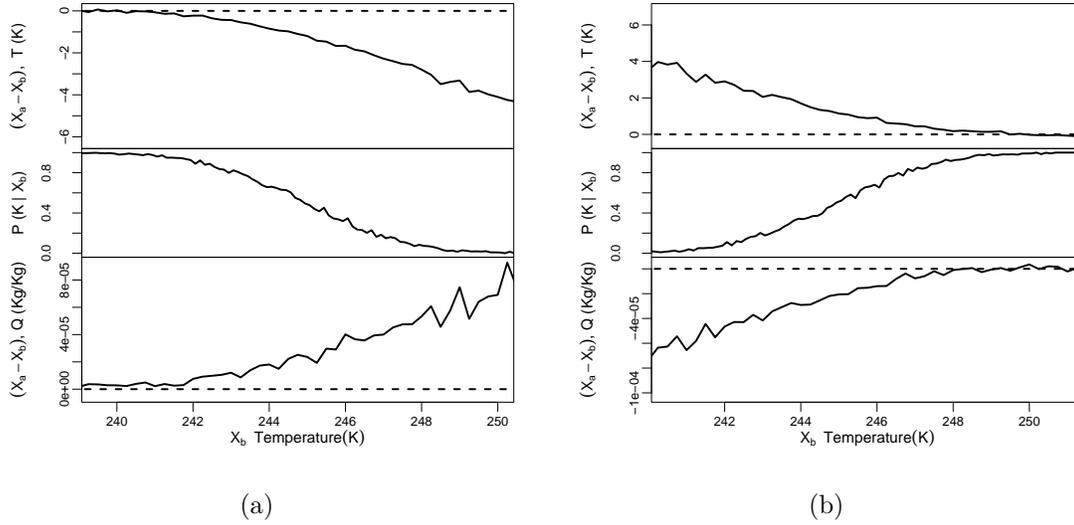


Figure 1.48: Analysis increments on the varying of the background temperature: (a) case $K = 1$ (Cloud observed), (b) case $K = 0$ (No-Cloud observed).

we use the following formula:

$$J | K \cap X_b = E \left[(X - \widetilde{X}_a)^2 | K \cap X_b \right] \quad (1.69)$$

$$= E \left[X^2 | K \cap X_b \right] - 2E \left[X | K \cap X_b \right] \widetilde{X}_a + \widetilde{X}_a^2 \quad (1.70)$$

$$= Var \left[X | K \cap X_b \right] + E \left[\widetilde{X}_a - X | K \cap X_b \right]^2. \quad (1.71)$$

We see that minimum $(J | K \cap X_b)$ is achieved for $\widetilde{X}_a = E[X | K \cap X_b]$.

Figures 1.49 and 1.50 show the cost function J for temperature and humidity respectively⁷. The dashed line corresponds to the analysis that matches closely with the minimum of J , while the dotted line corresponds to the first guess.

1.4.4 How RH_{cr} affects the analysis?

With reference to the cloud cover forward model, we now have the problem to choose the threshold below which no cloud formation is admitted. It can be useful to follow a Bayesian approach and evaluate the influence of the parameter RH_{cr} on the analysis. In terms of the PDF for RH_{cr} , studying the case which has tremendous impact on the analysis can provide useful information for an objective quality control of cloud cover data as well. Our assumption will be basically that the cloud cover can be correctly diagnosed by Eq. (1.63), but the parameter RH_{cr} is uncertain, and thus we introduce a probability density function. Since RH_{cr} is

⁷For a test case where the background is set to $243.7K$ and $5.225 \cdot 10^{-4}KgKg^{-1}$ for temperature and humidity respectively.

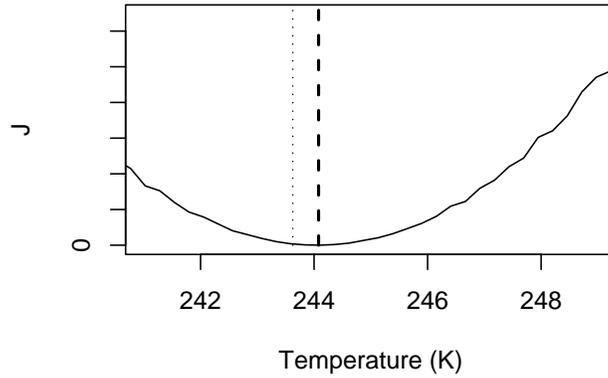


Figure 1.49: Cost function J for the temperature.

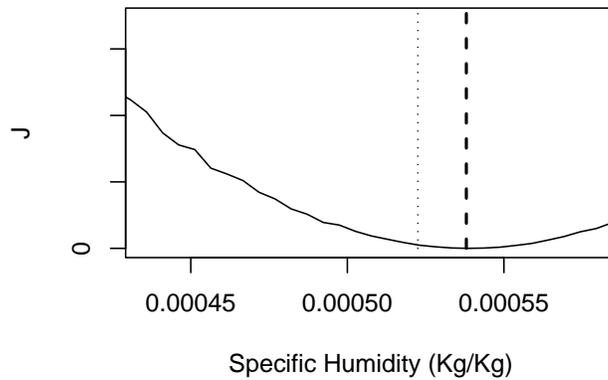


Figure 1.50: Cost function J for the specific humidity.

bounded between two limits, 0 and 1, the natural candidate for the RH_{cr} PDF is the Beta distribution

$$P(RH_{cr}) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} RH_{cr}^{a-1} (1 - RH_{cr})^{b-1}. \quad (1.72)$$

For commodity, we will refer hereafter to the reparameterized Beta distribution with parameters α and β ⁸

$$P(RH_{cr}) = \frac{\Gamma(\beta)}{\Gamma(\alpha\beta)\Gamma((1-\alpha)\beta)} RH_{cr}^{\alpha\beta-1} (1 - RH_{cr})^{(1-\alpha)\beta-1}. \quad (1.73)$$

⁸ α is equal to $a/(a+b)$; β is equal to $(a+b)$.

Now the expected value and the variance of such a distribution will be respectively α and $\alpha(1 - \alpha)/(1 + \beta)$. The probability to have a cloud if we know the relative humidity is (after the assumption (1.53))

$$\begin{aligned} P(K|RH) &= \int_0^1 P(K | RH \cap RH_{cr}) P(RH_{cr}) dRH_{cr} \\ &= \int_0^1 \left[\frac{RH - RH_{cr}}{1 - RH_{cr}} \right]^2 P(RH_{cr}) dRH_{cr}. \end{aligned} \quad (1.74)$$

However, the parameters α and β are unknown, and in real life we fix the value for the critical relative humidity, \widehat{RH}_{cr} . In particular, depending on the vertical level, we choose a \widehat{RH}_{cr} , that will be the expected value for the distribution of RH_{cr} (also equal to α). Then, it's possible to define the Bayes risk associated to two possible decisions: choose \widetilde{X}_a (from Eq. (1.59)) or X_b as analysis. This last condition is equivalent to reject the observation. With regards to the MSE verification, for the analysis \widetilde{X}_a

$$\begin{aligned} B(\widetilde{X}_a) &= E \left[E \left[(X_t - \widetilde{X}_a)^2 | X_b \cap K \cap RH_{cr} \right] \right] \\ &= \int_{-\infty}^{+\infty} \int_0^1 E \left[(X_t - \widetilde{X}_a)^2 \right] P(X_t | X_b \cap K \cap RH_{cr}) dX_t dRH_{cr} \\ &= \int_{-\infty}^{+\infty} \int_0^1 E \left[(X_t - \widetilde{X}_a)^2 \right] \frac{P(K | X_t \cap RH_{cr}) P(X_t | X_b)}{P(K | X_b \cap RH_{cr})} dX_t dRH_{cr} \end{aligned} \quad (1.75)$$

where

$$P(K | X_b \cap RH_{cr}) = \int_{-\infty}^{+\infty} \int_0^1 P(K | X_t \cap RH_{cr}) P(X_t | X_b) dX_t dRH_{cr} \quad (1.76)$$

to compare with the Bayes risk deriving from the choice of using X_b as analysis

$$\begin{aligned} B(X_b) &= E \left[E \left[(X_t - X_b)^2 | X_b \right] \right] \\ &= \int_{-\infty}^{+\infty} E \left[(X_t - X_b)^2 \right] P(X_t | X_b) dX_t \end{aligned} \quad (1.77)$$

We have compared the difference between the Bayes risk associated to the analysis and that one to the background for different RH values of the variance of RH_{cr} (Fig. 1.51a and Fig. 1.51b).

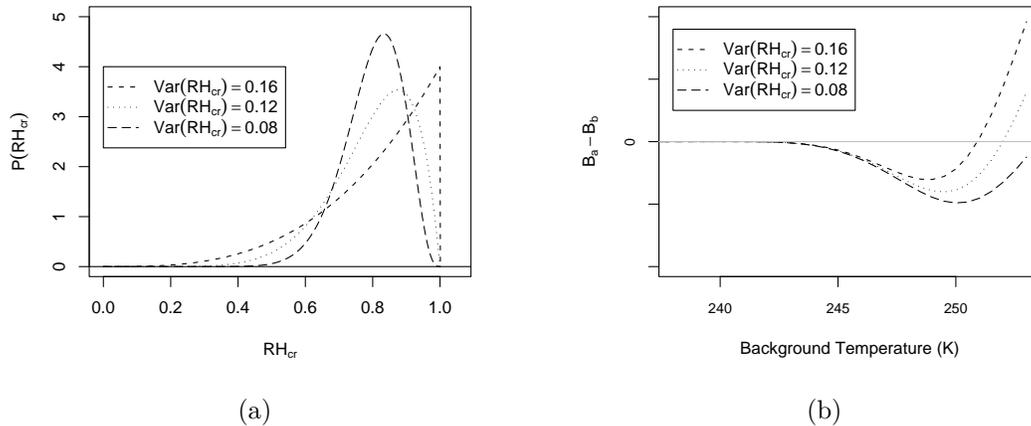


Figure 1.51: Computation of Bayes risk for different values of the variance of RH_{cr} : (a) PDF, (b) Bayes risk for the analysis \widehat{X}_a minus Bayes risk for the background X_b , case $K = 1$.

1.4.5 An objective background quality control

It seems useful to introduce a simple quality control to discard the suspicious binary observations and remove those ones resulting in redundancy with the background. Following the results of the former subsection, the quality control criterion is to remove from the assimilation all the observations for which the Bayes risk associated to the analysis \widetilde{X}_a is even bigger than the risk associated to keeping the background as analysis. This can happen since our forward model roughly model the cloud cover, neglecting many factors that affect the fraction of cloud. Let's introduce a gross error that takes into account the weakness of the forward model. We let this error in the Beta distribution of RH_{cr} as constant additional probability κ for every RH_{cr} :

$$P(RH_{cr}) = \frac{\frac{\Gamma(\beta)}{\Gamma(\alpha\beta)\Gamma((1-\alpha)\beta)} RH_{cr}^{\alpha\beta-1} (1 - RH_{cr})^{(1-\alpha)\beta-1} + \kappa}{1 + \kappa} \quad (1.78)$$

The effect of such an error can be visualised in Fig. 1.52(a). The three PDF are flattened with respect to the same ones in Fig. 1.51(a). The Bayes risk functions in (1.75) and (1.77) are thus recomputed, and they show (Fig. 1.52b) that now the impact of the variance is very slight. For higher values of temperature (that correspond to low values of RH_{cr} , all the PDFs tend to κ , and the difference in the risk function tend therefore to be the same for every value of variance. In the same figure we have cut the right wing. In fact, we can impose as objective quality control criterion, to take \widetilde{X}_a as analysis only if $(B_a - B_b)$ is smaller than a value ϵ . Otherwise, we will discard the observation. Note that this last condition

applies when $X_a - X_b$ is dramatically large, therefore this quality control aims to avoid unrealistic analysis (see Fig. 1.43). To also avoid the assimilation of binary observation that do not contain any information with respect to the background, we can operate similarly, imposing the condition that $(B_a - B_b)$ should be greater than a small threshold below which the analysis is in practise equal to the background. This condition can also prevent the analysis from overestimating the importance of the background itself (that otherwise would be twice taken into account).

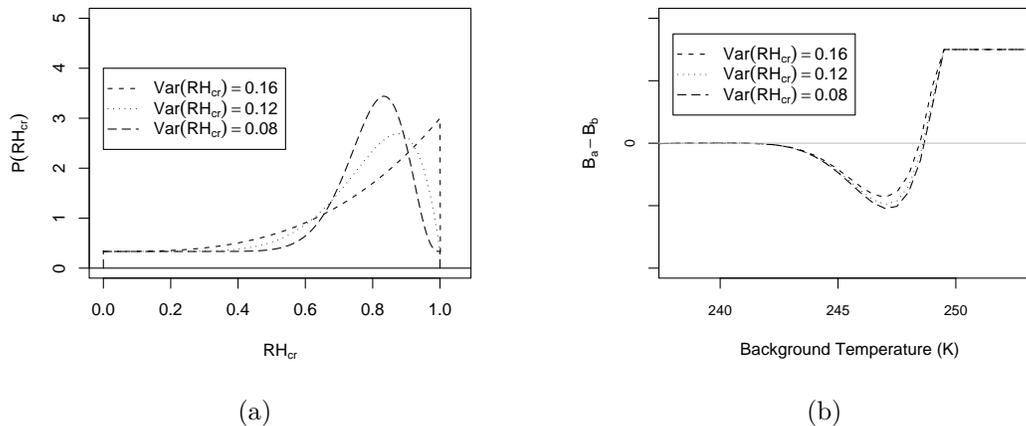


Figure 1.52: Computation of Bayes risk for different values of the variance of RH_{cr} when a gross error is introduced: (a) PDF, (b) Bayes risk for the analysis \hat{X}_a minus Bayes risk for the background X_b , case $K = 1$.

1.5 Conclusions

The formulation of the probability density function of the true state of the atmosphere is the first part of the data assimilation theory. To achieve this, the PDF has been derived following two approaches. The main difference is that the first one computes the probability of the observation given the true state, operating therefore in the “observations space”, while the second includes on the contrary the probability of the true state given the observation, answering the Bayesian solution of the data assimilation problem. Again, the first one is the only possible choice when the observation operator is not invertible, so it is largely used in modern variational assimilation techniques.

Although the two formulations are very similar, they coincide if we use the Normal approximation for the true state PDF. This approximation is verified against a simulated data set, which match closely both the theoretical and the approximated distributions. As consequence, the Normal approximations are accurate enough to use instead of the theoretical formulations of the PDFs, unless the standard

deviations of the background or the first guess are very large. Using the Mean Square Error verification, the analysis has been verified to be optimal, with regards to a reference observation too.

A simple binary observation problem has also been studied, where the observed value here represents the threshold point between the zero and the maximum probability. How the PDF changes depends on the standard deviation of the observation. The resulting PDF for the true state has a Normal behaviour, that is a very important property. With regards to the variational assimilation theory, the conditional probabilities of the true state, given the background and the observation, are in fact assumed Normal. Thus, it is possible to assimilate the binary observations without further simplifications. However, the probability density for the true state known the binary observation is not Normal, so a *pseudo-observation* has been retrieved by imposing the coincidence between the two expected value. Such a condition allows to keep optimal the analysis. The definition of the *pseudo-observation* is noticeably affected by the background; this approach allows in fact to assimilate the binary observation, poorer in terms of information with respect to continuous observations, through the intense use of the background. Unfortunately, the error in the pseudo observation is consequently correlated with the error in the first guess, but the resulting analysis is still optimal in the single observation case.

Same results can be found when we deal with a more realistic situation and when we suppose to assimilate binary cloud cover data. In fact, although in the context of the 3D-Var framework the choice of forward model has to obey several constraints, first of all that it has to depend only on “control state variables”, simple relations between cloud cover and relative humidity are well-known and have been intensively used also in recent years.

The analysis we obtain are forced by this forward model to reproduce the observation. Three main cases occur, depending on the value of the background: i) if the probability to have the binary condition given the background is maximum, the observation itself does not add useful information, and the analysis increment is infinitesimal; ii) if the probability to have the binary condition given the background tends to zero, the analysis is forced to follow the binary condition and the analysis increment is big and very sensitive to the forward model; iii) in the other cases, the analysis increment has a reasonable value.

The Bayesian decision theory gives useful insights and suggests that too big analysis increments should be avoided, and an objective quality control based on the computation of the Bayes risk function can therefore be constructed. Note that we can also observe redundancy between the background and the observation, a problem peculiar of binary data.

Chapter 2

Assimilation of Binary Cloud Cover: Towards the Use of of CloudSat Radar Reflectivities

Summary

CloudSat radar provides high resolution and high quality profiles of reflectivities, that can be used to detect vertical structure of clouds. The theory for the assimilation of binary cloud cover can be therefore applied to the radar received power. A strategy to process the radar data and to use them in assimilation is described in this Chapter.

2.1 Rationale

Assimilation of binary cloud cover offers an attractive perspective for the improvement of moisture analysis. It is theoretically possible to use even simple cloud cover relationship to retrieve humidity profiles, once that the cloud cover is known. However, dealing with spaceborne observations from AVHRR or other high resolution sounders leads to vertical integration problems, since the synchronous use of several channels can only partially overtake the vertical mislocation of the clouds; furthermore, the typical cloud classification from these instruments can provide a cloud top height and a tentative thickness, taking anyway into account only the first cloud seen from the satellite.

Since the cloud cover errors do not follow a Gaussian distribution, and, furthermore, the binary function that describes it cannot be computed the gradient, these observations are not suitable for use in three(four)-dimensional variational assimilation systems, where errors are assumed Gaussian and observations operators have to be linearised. We can however define our analysis as

$$X_a = \int_{-\infty}^{+\infty} X_t P(X_t | K \cap X_b) dX_t \quad (2.1)$$

where K is the binary observation, X_t is the true state of the atmosphere and X_b is the a priori knowledge (background) of the state of the atmosphere. In case of cloud types or cloudmask observation (e.g. from AVHRR) the problem is not well-posed because the solution of Eq. 2.1, given one observation K , involves the vertical integration of many variables, and to ride this problem over one should make some assumptions for the vertical integration, read overlap, of the clouds. Such assumptions can compromise the assimilation of the observations itself.

As intermediate test for the feasibility of assimilating cloud cover into assimilation systems and for testing different strategies that can be followed, the use of CloudSat reflectivities appears of benefit. In facts, the radar aboard CloudSat provides vertical profiles of cloud fraction with extreme precision, if compared to other satellite cloud products, and at very high horizontal and vertical resolution; as active instrument, the 94 GHz Cloud Profiling Radar (CPR) can also correctly detect multi-layer cloud systems. This instrument is therefore an optimal candidate for the assimilation of cloud fraction data, and a strategy is subsequently presented. In section 2.2 the pre-processing algorithm for CloudSat data and a review of the use of the observation operator is described; in section 2.3 we extend the theory for the assimilation of binary observations to CloudSat cloud fraction; in section 4.4 the implementation in an experimental assimilation system is presented.

2.2 Exploitation of CloudSat data

2.2.1 CloudSat and the on-board 94 GHz radar

In April 2006 NASA launched the CloudSat satellite, whose mission aims to provide vertical distribution of cloud layers for model validation and better understanding of microphysical properties of clouds (Stephens et al., 2002). CloudSat joins the so-called “A-train”, a constellation of 6 satellites flying in close proximity (among them, Aqua and Calipso). Reflectivities measured by the 94 GHz Cloud Profiling Radar (CPR), aboard CloudSat, are calibrated to provide profiles of cloud fraction at high resolution (about 2500 meters along the track and 250 meters on the vertical). The radar frequency of 94 GHz has been shown to be perfectly suitable for the the study of clouds microphysics (Clothiaux et al., 1995).

These clouds profiles represent therefore an optimal framework for testing the assimilation of cloud cover data, because of the high resolution, both along the track and on the vertical, and the high quality of the radar data. Successful assimilation of CloudSat radar data can also be of interest for designing next-generation of polar orbiting satellites. However, it should be noted that CloudSat CPR data are disseminated with an average delay of 8 hours, making their use far from being operational.

CloudSat has a nearly circular orbit at 705 Km of altitude with a 99 minutes period that repeats every 233 orbital revolutions (16 days). The 94 GHz radar is a nadir-pointing active instrument that profiles the atmosphere along the orbit track. In order to have an idea about the dissemination of CloudSat data, Fig. 1(a) reports the dissemination over 1 day, and Fig. 1(b) the dissemination for a 6 hours time window, both the figures accounting observations only inside the computational NWP model domain, hereafter ALADIN11. The latter one reproduces a typical data availability in a three-dimensional variational assimilation time-slot.

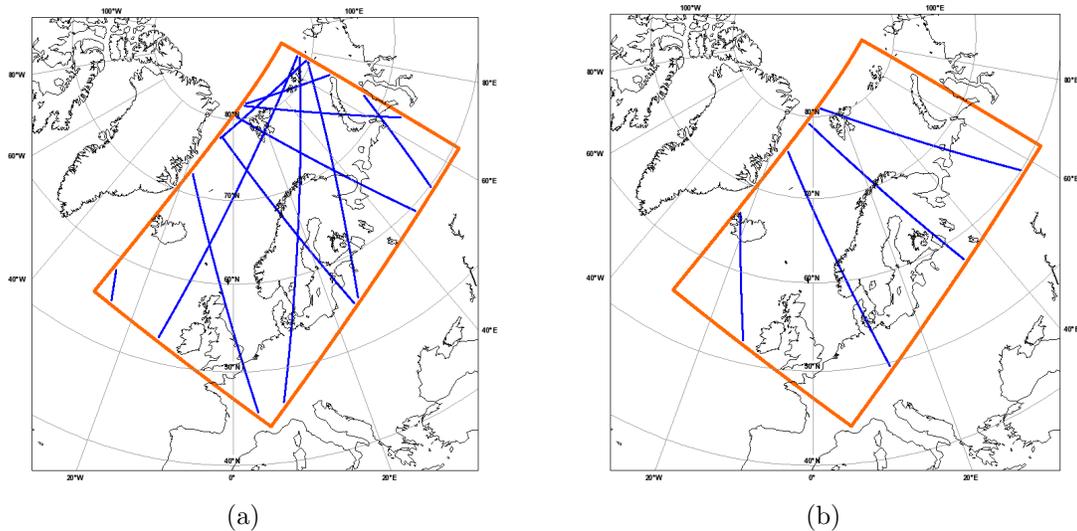


Figure 2.1: CloudSat data dissemination: (a) Over 24 hours (20.10.2007); (b) Over a 6-hours assimilation window (19.10.2007 12:00 UTC \pm 3h).

2.2.2 Retrieval of cloud fraction profiles

The algorithm for the retrieval of cloud profiles from the radar received power consists of several modules that will be briefly reviewed consecutively. First, the State of Health (SOH) aboard CloudSat is checked to correctly reports the status of the satellite. The pre-processing software performs the following checks:

- Validity of GPS data;
- Validity of temperature values;
- Normality of radar telemetry quality;
- Normality of peak power;
- CPR not in calibration manoeuvre;
- State Of Health (SOH) correctly received;

- Star trackers switched on;
- Absence of Non-Incrementing Second Anomaly (NISC).

Then, all the data supplied in Level1-FL (near real-time First-Look data) are extracted and each profile is checked to correctly report the expected number of radar bins. Two major computations are then performed to calculate profiles of reflectivity factors and profiles of cloud fractions.

The reflectivity factor is computed through the radar equation (Meneghini and Kozi, 1990):

$$dbZ = 10 \log_{10} \left(\frac{\alpha^4 \eta}{\pi^5 K^2} 10^{18} \right) \quad (2.2)$$

where α is the radar wavelength, K^2 is function of the dielectric constant (around 0.75 at 94 GHz), and η is the radar backscatter, given by

$$\eta = \frac{P_R - P_N}{P_T} r^2 C \quad (2.3)$$

where P_R , P_N , and P_T are respectively the received power, the noise power and the transmitted power, r is the distance of the radar to the cloud or the surface; C is the radar constant, computed by the CloudSat Data Processing Center from pre-launch measurements, see Li et al. (2007), equal to

$$C = \frac{(4\pi)^3}{\alpha^2 G^2 \Delta \Omega}$$

where G is the antenna gain, Δ is the pulse integral and Ω is the integral of the normalised antenna pattern. In Eq. 2.3, the transmitted power P_T is the one averaged for an entire CloudSat granule (the track between two consecutive passages over the Equator), rather than the instantaneous transmitted power. The noise power is the mean of the received power in several stratospheric levels (from about 17 to 23 Km of height) that are assumed to be hydrometeors-free; the standard deviation of the noise power is computed as well, for use in the cloudmask retrieval.

The cloudmask is computed for every profile using a threshold algorithm, that reads:

$$\begin{cases} CF = 1 & \text{if } P_r - P_n > k_1 \sigma_n \\ CF = \frac{P_r - P_n - k_0 \sigma_n}{(k_1 - k_0) \sigma_n} & \text{if } k_0 \sigma_n < P_r - P_n < k_1 \sigma_n \\ CF = 0 & \text{if } P_r - P_n < k_0 \sigma_n \end{cases} \quad (2.4)$$

where σ_n is the standard deviation of noise power, and k_0 and k_1 are parameters empirically tuned. This simple scheme has been validated against the CloudSat Data Processing Center Level2 products that take advantage of a more advanced

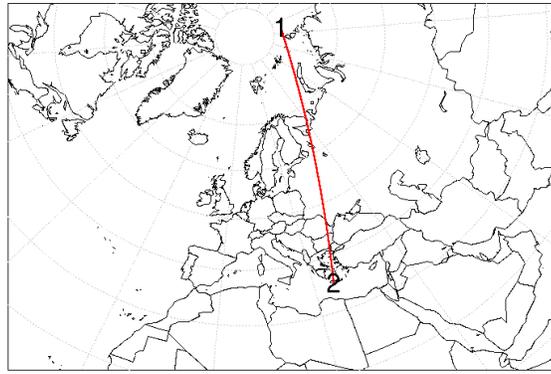
two-step algorithm with an along-track integration scheme to compute power probability within subsets of received power distribution (Clothiaux et al., 1998), and, additionally, has been compared with the combined CloudSat-Calipso-MODIS cloud products, that exploits the lidar reflectivities (from Calipso) to detect thin clouds and refine the cloudmask through the high-resolution two-dimensional cloudmask provided through the MODIS sounder aboard Aqua.

Since the surface can be even 5 orders of magnitude more reflective than hydrometeors (Marchand et al., 2008), the first 3 to 5 bins above surface (up to 0.7 to 1.2 Km) are dominated by the surface return. A filter has been therefore implemented to take out the surface return. The filter subtracts the “surface clutter profile”, estimated on a flat surface for the first 5 bins above the surface, to the received power. Successively, for profiles where such an estimate cannot be considered reliable (the criterion is an orography gradient greater than 2% or a Digital Elevation Map, DEM, reporting an altitude greater than 2000 m) the filter discards cloudy bins if the adjacent profiles along the track, for the same height above the surface, do not show the presence of hydrometeors. This procedure is applied to the first 5 bins above the surface.

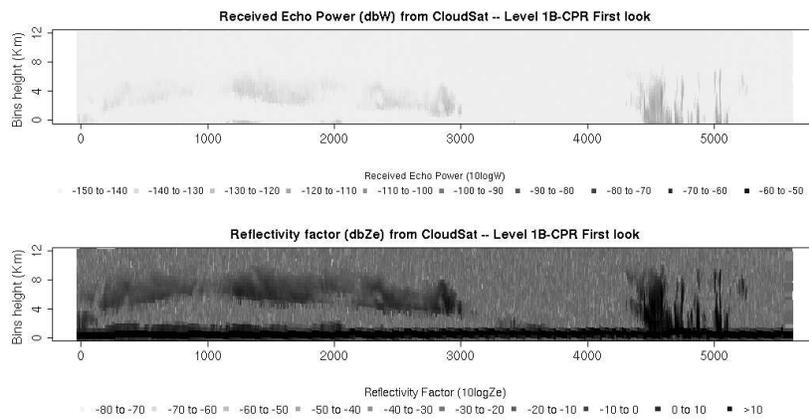
2.2.3 Examples of cloudmask production

To show how the cloudmask algorithm works, we report a couple of interesting examples. Following figures show the CloudSat orbit, and cross-sections of received power, reflectivity factor and cloud fraction, for an area covering extending from 35N to 90N of latitude and from 45W to 100E of longitude. Cloud fractions are interpolated to constant geopotential levels, keeping the same vertical resolution of 240 m as in the original CPR data.

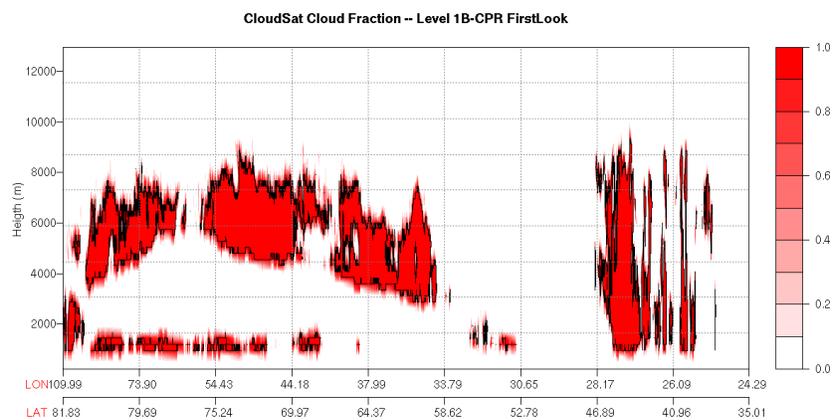
In Fig. 2.2 is shown how the CloudSat cloudmask is able to identify multi-layers cloudy profiles. This feature represents one of the most relevant benefits in the use of spaceborne radar for detecting clouds. Profiles between 65N and 80N latitudes show the presence of mid-level clouds associated with low-level clouds. This can be easily noted for the received power cross-section and the reflectivity factor cross-section, as well as in the resulting cloudmask, and represents an useful advance if compared with usual satellite cloud products (e.g. the Nowcasting-SAF products). Another interesting aspect is the strongly binary nature of the cloud fraction as given by the cloudmask. This feature is consistent with the more advanced cloud products from CloudSat-Calipso-MODIS, and it should be noted how these data mostly report the punctual presence or absence of hydrometeors. This aspect has to be taken into account when assimilating CloudSat data since NWP models report a cloud fraction averaged into a model gridbox, leading to a statistical nature of the cloud cover parameter.



(a)



(b)



(c)

Figure 2.2: CloudSat cloudmask algorithm, multilayer clouds example, valid UTC time 19.11.2007 00:19: (a) CloudSat orbit track; (b) Cross-section of received power and reflectivity factor; (c) Cross-section of retrieved cloud fractions interpolated to constant geopotential levels.

In Fig. 2.3 is shown the capability of the algorithm to filter out the surface clutter from the received power. In particular, over Mongolia (first part of the cross-section, 35N to 50N latitude) and over Greenland the surface return is notably stronger than the hydrometeors layers, and it's correctly filter out by the cloudmask algorithm.

As another example, Fig. 2.4 show the potential use of CloudSat cloudmask for nowcasting purposes. The CloudSat product is now used operationally by forecasters at the Norwegian Meteorological Institute, and can be visualised in real-time by the DIgital ANALysis (Diana) visualisation software. We compare the cloud fraction cross-sections (where the height is projected on the right side of the satellite track) with the subjective analysis by forecasters and with MSG-1 combined visible and infrared channels (3+9+10). The cloud system derived from the warm front developing from UK, remarkable in both METEOSAT images and forecasters analysis, shows the typical advancing structure in the profiles close to Oslo as reported by the CloudSat cloudmask.

2.2.4 Treatment of cloud fractions data for assimilation

The cloud fraction data obtained by the previous algorithm can be used for real-time visualisation and thus be useful for nowcasting purposes. In order to assimilate these data in variational assimilation models, a few more steps are needed to ensure the quality of the observations and to minimise the spatial correlation of the cloud cover observations.

For all the profiles correctly received, a two-step domain check is performed to discard data outside the ALADIN11 domain. The procedure is divided in two step for computer efficiency. The first step roughly check data to be inside a latitude/longitude grid. The second step projects each data point to the model plane, following the Lambert Conformal projection. A “red-zone”, 50 Km large, is taken into account to discard data too close to the domain lateral boundaries.

Since the CloudSat resolution along the orbit track is much higher than the NWP model resolution (about 2.5 Km against 11 Km for the ALADIN11 domain), the observations are averaged and interpolated to the model points. The weighing function is Gaussian; to avoid that one CloudSat profile is used more than once during the interpolation, for each model gridpoint only observations inside a gridbox as large as the model resolution and centred on the model point are used. For each cloud fraction CF_M in the model space:

$$CF_M = \frac{\sum_{i=0}^N w_i f_i}{\sum_{i=0}^N w_i}$$

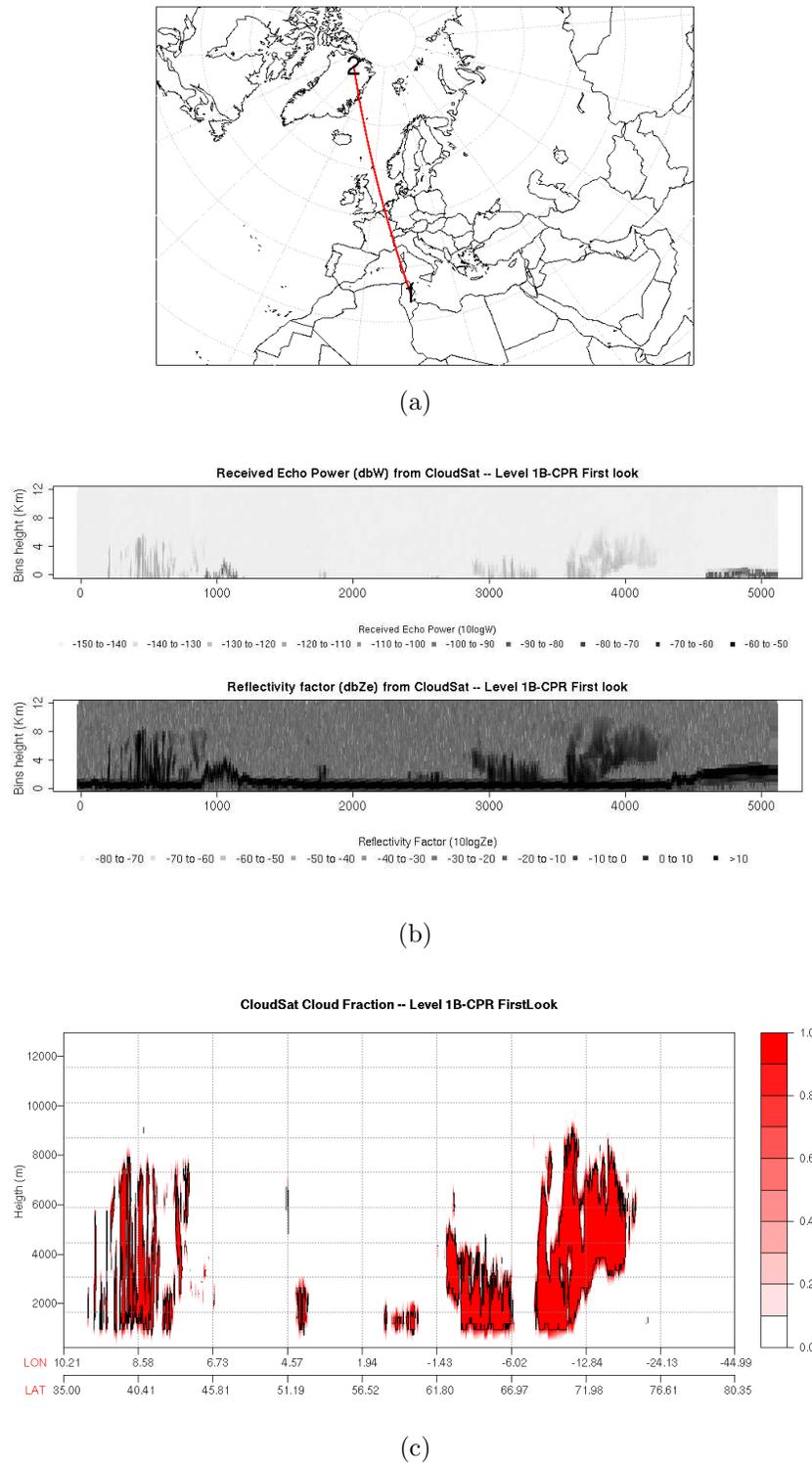
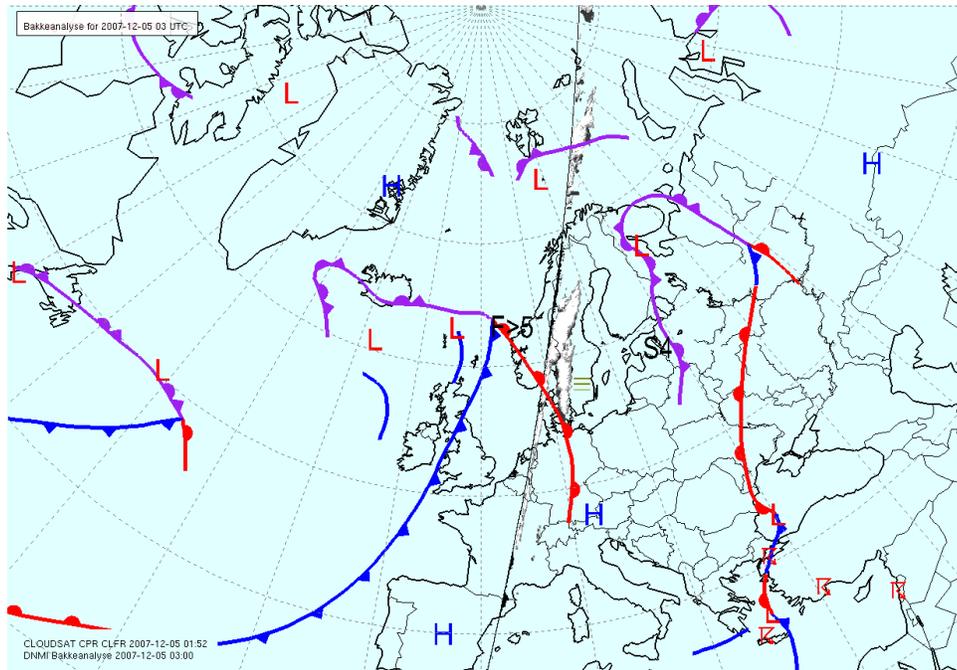
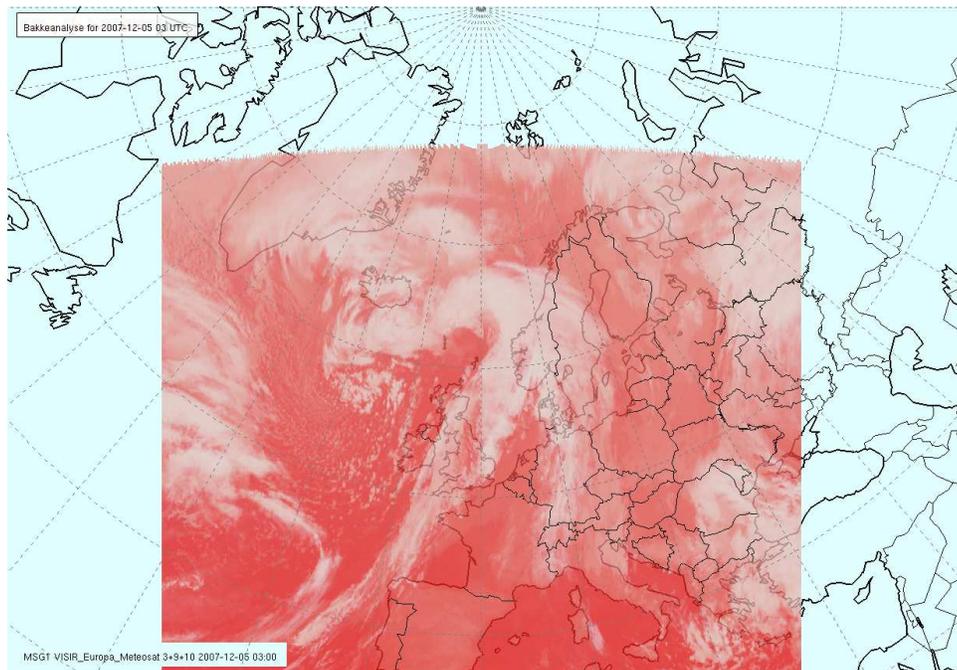


Figure 2.3: CloudSat cloudmask algorithm, surface clutter filter example, valid UTC time 30.11.2007 05:16: (a) CloudSat orbit track; (b) Cross-section of received power and reflectivity factor; (c) Cross-section of retrieved cloud fractions interpolated to constant geopotential levels.



(a)



(b)

Figure 2.4: Example of use of CloudSat cloudmask for nowcasting purposes, valid UTC time 05.12.2007 03:00: (a) CloudSat cloudmask and forecasters analysis; (b) METEOSAT MSG1 Infrared and Visible channels (3+9+10).

where, if d is the geometrical distance between the model point and the i -th observation,

$$w_i = \begin{cases} e^{-\frac{d^2}{2}} & \text{if } f_i \text{ is inside the box} \\ 0 & \text{otherwise} \end{cases}$$

This procedure leads to super-observations generated by averaging different amounts of CloudSat observations; to discard poor super-observation generated by too few origin data, a thinning routine is adopted. The domain is divided in a number of square grid-boxes, whose side is 50 Km, inside which only one super-observation is retained. The choice of 50 Km large grid-boxes issues from a compromise between having dense observations and reducing the spatial correlation of the observations, and it's a value used operationally for thinning of humidity sounders (AMSU-B and MHS). The algorithm keeps the superobservation generated with the largest number of data, and, if eventually two or more of them have the same density, impose the closeness to the box centre as additional criterion.

Processing the cloud fraction data requires the knowledge of a few parameters from the NWP model, hereafter first guess. The following fields are read from a forecast initialised 6 ± 1.5 hours before the CloudSat profiles:

- Orography (spectral 2D field);
- Surface pressure (spectral 2D field);
- Surface temperature (gridpoints 2D field);
- Specific humidity (spectral 3D field);
- Temperature (spectral 3D field);
- Cloud fraction, prognosed (gridpoints 3D field);

The large scale microphysics scheme adopted in the forecast model for the computation of cloud fraction is known as the ‘‘Lopez microphysics’’ scheme (Lopez, 2002). In this scheme there are two prognosed variables, large-scale cloud condensate (suspended liquid water and ice) and precipitation content (rain and snow). The scheme uses additional water species (cloud liquid water, cloud ice, rain, snow and water vapour) to parametrise large-scale condensation, auto-conversion, collection, evaporation of precipitation and precipitation downfall. Some tunings made in the operational implementation regard the ice-water separation function, the relative humidity minimum value for the formation of cloud condensate (Bouteloup et al., 2005).

Spectral fields are moved to grid-point space according to the ALADIN definition of the space of bi-Fourier coefficients (e.g. Berre, 2000). For the generic

parameter s :

$$s(i_x, i_y, z) = \sum_{m=-M}^M \sum_{n=-N}^N s(m, n, Z) \exp \left[2\pi i \left(\frac{mi_x}{J} + \frac{ni_y}{Q} \right) \right]$$

where (i_x, i_y) is the horizontal position vector in physical space; z is the vertical index; J and Q are the number of points of the extended domain in each direction; m and n are the zonal and meridional wavenumbers; $s(m, n, Z)$ is a bi-Fourier spectral coefficient; M and N are the maximum zonal and meridional wavenumbers. The elliptic truncation adopted in ALADIN reads as:

$$\left(\frac{m}{M} \right)^2 + \left(\frac{n}{N} \right)^2 \leq 1$$

After the inverse spectral transform, geopotential is computed through a discrete version of the hypsometric equation, starting from the lowest level (surface):

$$z_i = z_{i+1} + \frac{R_d T_v}{g} \ln \frac{p_{i+1}}{p_i}; \quad T_v = T (1 + 0.622q) \quad (2.5)$$

where z is the geopotential, R_d is the gas constant for dry air ($= 287.04 \text{ J K}^{-1} \cdot \text{kg}^{-1}$), g is the standard acceleration of gravity ($= 9.80665 \text{ m} \cdot \text{s}^{-2}$); p is the pressure; T_v and T are respectively the virtual temperature and temperature; q is specific humidity. Relative humidity RH is obtained from specific humidity;

$$RH = \frac{pq}{0.622e_s(T)}$$

where $e_s(T)$ is the saturation vapour pressure for a temperature T . Values of $e_s(T)$ are tabulated.

For each model point where thinned superobservations are present and valid, first guess upper-air fields are interpolated to original CloudSat levels. The interpolation is linear in the height. Pressure can be either interpolated (default) or computed from CloudSat geopotential (inverting Eq. 2.5).

2.2.5 Cloud fraction forward model

For data assimilation purpose, the cloud fraction computation from the first guess fields should depend on the control vector. Unfortunately, the prognostic scheme of Lopez (2002) contains time evolution of cloud condensate and precipitation, and therefore can be used only in four-dimensional assimilation systems. However, simple relationship between cloud fraction and relative humidity have been used intensively in NWP models. The pre-processing software contains three

scheme for the computation of cloud fraction CF :

$$CF = \left(\frac{RH - RH_{cr}}{1 - RH_{cr}} \right)^2 \quad (2.6)$$

proposed by Slingo (1987);

$$CF = 1 - \sqrt{\frac{1 - RH}{1 - RH_{cr}}} \quad (2.7)$$

proposed by Sundqvist (1978);

$$CF = 1 - \sqrt{\frac{1 - RH}{1 - RH_{cr} - \kappa(RH - RH_{cr})}} \quad (2.8)$$

proposed by Tompkins and Janisková (2004). The latter differs from the Sundqvist relationship because of the introduction of a pressure-dependent variance for the subgrid distribution of cloud cover. The three formulae leads to very similar cloudiness profile given the same parameter RH_{cr} ; Fig. 2.5 compares the Slingo and the Sundqvist formulation when $RH_{cr} = 0.7$. Indeed, the schemes are very sensitive to

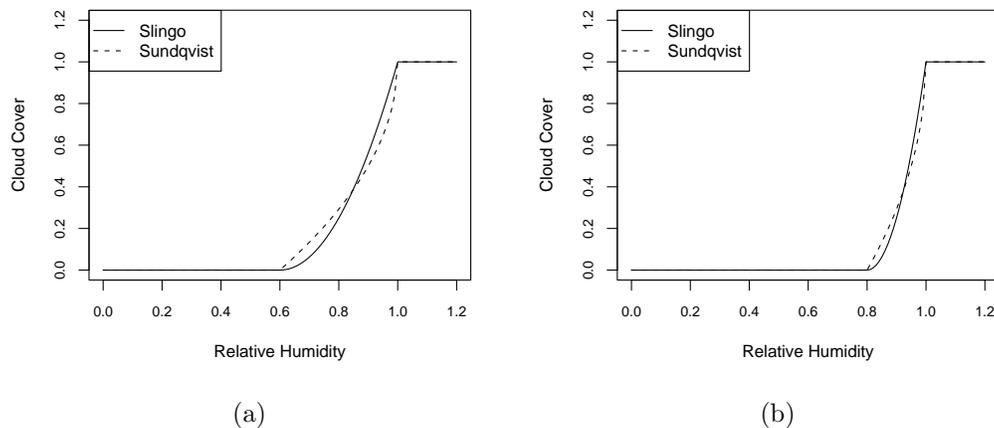


Figure 2.5: Values of cloud cover computed by the use of the Slingo and Sundqvist expression, (a) for $RH_{cr} = 0.6$; (b) for $RH_{cr} = 0.8$.

the tuning of RH_{cr} . To show this, we report the values of RH_{cr} computed through linear regression for observed cloud cover at different heights. It's basically corresponds to the inverse problem, that reads “which value of RH_{cr} should we set to obtain true cloud cover from the model fields?” The cloud cover is measured in three different sites by 94 and 35 GHz zenith-pointing radars and data are kindly provided by the Cloudnet project (Illingworth et al., 2007), that also supplies interpolated values for several parameters from a few NWP models. The relation 2.7

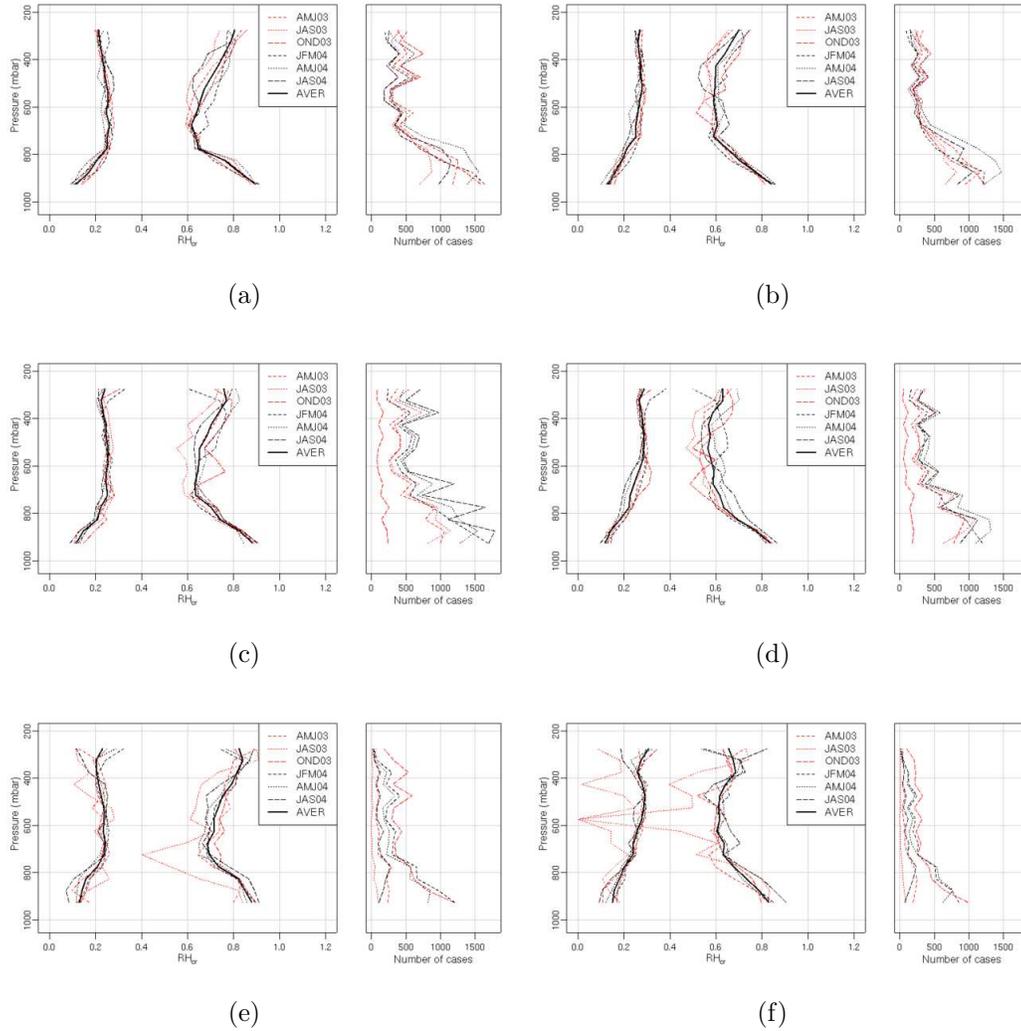


Figure 2.6: RH_{cr} as function of height for different cloud-radar sites using linear regression from NWP models fields: (a) Chilbolton cloud-radar using ECMWF/IFS fields; (b) Chilbolton cloud-radar using Météo-France/Arpège fields; (c) Cabauw cloud-radar using ECMWF/IFS fields; (d) Cabauw cloud-radar using Météo-France/Arpège fields; (e) Palaiseau cloud-radar using ECMWF/IFS fields; (f) Palaiseau cloud-radar using Météo-France/Arpège fields.

has been adopted for the computation. Fig. 2.6 shows the mean and the standard deviation of RH_{cr} for the ECMWF/IFS global model (39 Km of spatial resolution) and the Météo-France/Arpège global model (24 Km of spatial resolution), for different seasons and for the overall period. It is of interest to note the dependency of RH_{cr} from several parameters. First, the model resolution plays an important role given the statistical gridbox-averaged nature of the cloud-cover. Generally, the higher the resolution, the larger the threshold RH_{cr} for the formation of clouds. In principle, at very high cloud-resolving resolutions, RH_{cr} is 1 by definition. Smith (1990) showed that RH_{cr} can be expressed as a function of subgrid covariances of temperature and specific humidity, and Lopez (2002), using FASTEX aircraft in-situ measurements of these quantities, showed that RH_{cr} can be 1 to 0.4 ranging from 1 to several hundreds of kilometres of resolution (Fig. 2.7 shows these results). Second, a seasonal dependency is found. As expected, in summer months when

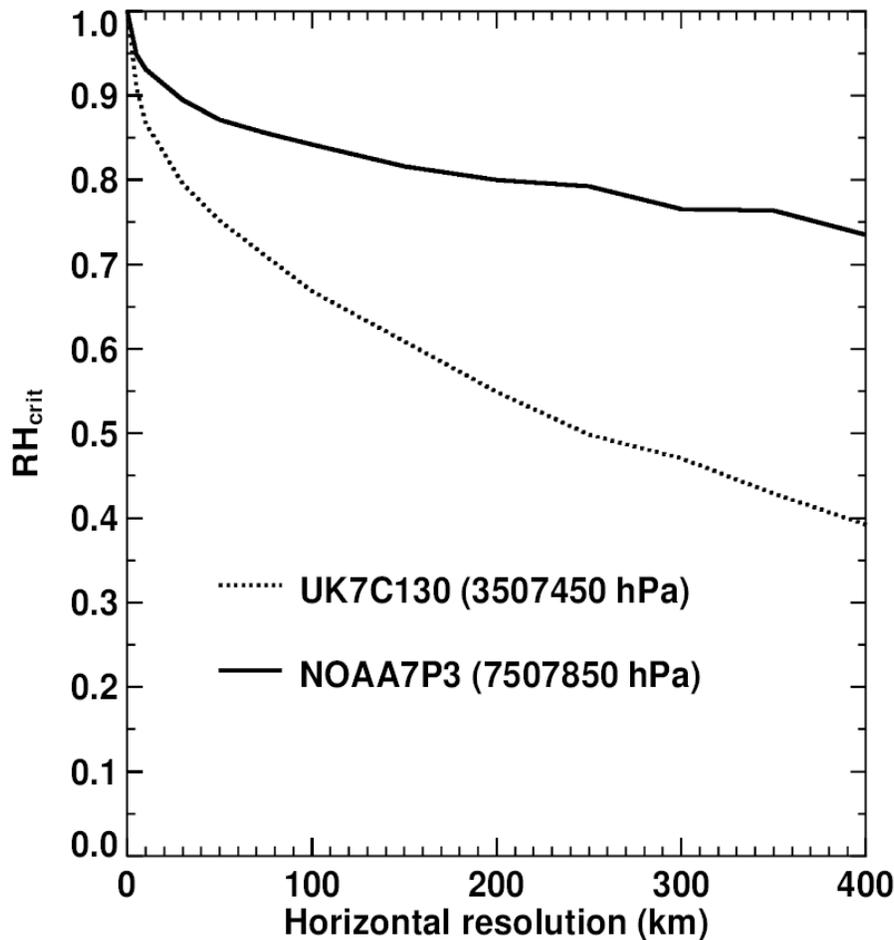


Figure 2.7: RH_{cr} computed from subgrid variances of temperature and specific humidity as function of horizontal resolution using aircraft measurements from FASTEX campaign. Data from aircraft NOAA-P3 refer to about 800 hPa, while data from aircraft UK-C130 to about 400 hPa. From Lopez (2002).

convective activity is more remarkable, formation of clouds is facilitated and RH_{cr}

is smaller. Third, vertical discretization and physics can lead to profiles of RH_{cr} that differ each other. Fig. 2.6 shows how the RH_{cr} profiles differ from a global model to the other for high-level clouds, where Météo-France/Arpège keeps lower values while ECMWF/IFS increases again the critical threshold. This is taken into account in the definition of RH_{cr} for the two NWP models, see Fig. 2.8. However, a well-known behaviour of the RH_{cr} inherited by both the parameterizations and proved by a plenty of experimental studies (e.g. Teixeira, 2001), is that cloud cover cannot be neglected in mid-atmosphere of relative humidities even less than 60%; therefore, RH_{cr} presents a minimum for mid-level clouds.

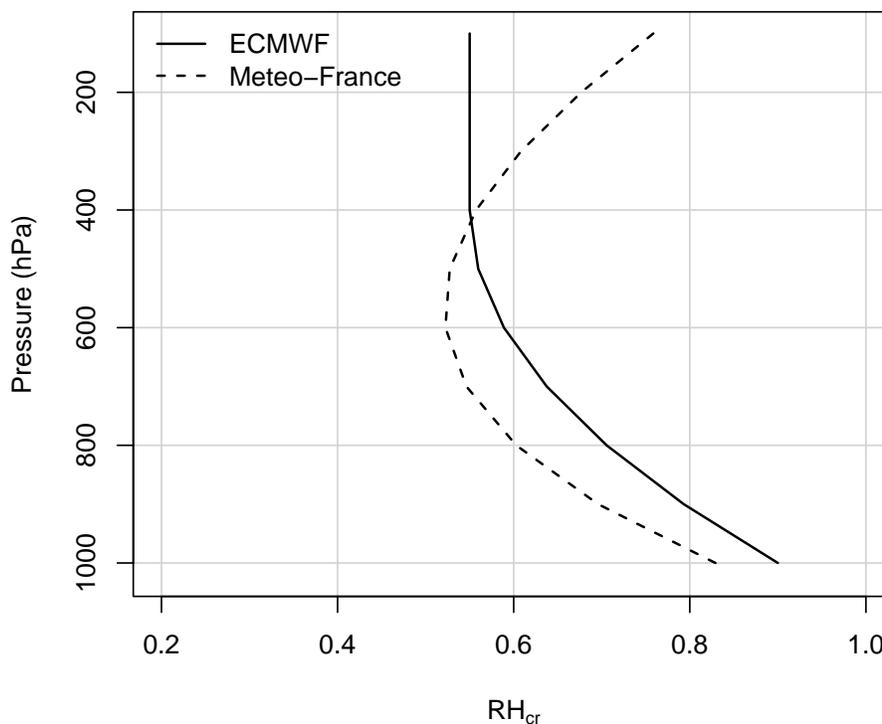


Figure 2.8: RH_{cr} as used operationally at ECMWF (IFS global model) and at Météo-France (Arpège global model and ALADIN limited area model).

Several experiments have been conducted to attempt to specify a non-constant profile for RH_{cr} . The aim is to find an expression for RH_{cr} to have cloud cover values as close as possible to those ones given by a more advanced prognostic scheme, in particular the Lopez (2002) scheme is used as reference prognostic scheme. The proposed method consists of 1) choosing reliable predictors for RH_{cr} ; 2) performing a linear regression for computing the multiplicative coefficients to the predictors that satisfy the Eq. 2.7, assuming that the “true cloud cover” is the one from the prognostic scheme, and assuming that the errors in the 6 hours forecasts for the

vertical profiles of humidity and temperature is negligible; 3) recompute the cloud cover using the predictors and the coefficients for comparison with the prognostic scheme, considering as reference the expression used by Tompkins and Janisková (2004), where $RH_{cr} = f(\eta, \eta^2, \eta^3)$ is a polynomial of third order of the ratio $\eta = p/p_s$ pressure divided by surface pressure. Several predictors have been tested:

- For taking into account the atmospheric instability, either the Convective Available Potential Energy (CAPE) or the vertical gradient of the potential temperature;
- For taking into account the seasonal variability and the ice cloud/liquid water clouds separation, either the temperature or the an ice-liquid water separation function κ , given by

$$\kappa = 1 - \exp \left[- \left((T - 273.16) / \Delta T \right)^2 \right], \text{ for } T < 273.16$$

with ΔT set to 15 according to Bouteloup et al. (2005); κ takes 0 for ice or 1 for water;

- For taking into account the vertical profile of RH_{cr} , either a second order or third order polynomial of the vertical η coordinate;
- For taking into account the role of aerosol concentration in the clouds formation in different areas, three aerosol species (sea aerosol, land aerosol, soot aerosol) optical thickness given by the Tegen et al. (1997) dataset.

The predictors are computed using a two-weeks (01 to 15 January 2007) dataset of ALADIN 6 hours forecast initialised at 00 and 12 UTC using ECMWF initial and lateral boundary conditions. Different predictors are computed for forecasts initialised at 00 UTC (valid time 06 UTC) and at 12 UTC (valid time 18 UTC). This strategy, however, doesn't lead to significant improvements: Fig 2.9, 2.10 and 2.11 report the RMSE and Mean error and the HR, FAR, POD and ETS scores respectively for some experiments and two different cloud fraction thresholds. These last scores are computed building a contingency table for different cloud fraction threshold, (0.01, 0.25, 0.50, 0.75, 0.99) and comparing the experiment cloud fraction value with the one given by the prognostic scheme. The scores are calculated according to WWRP/WGNE (2007). Except for middle-high clouds where the performance of a constant RH_{cr} is not accurate enough, according to the comments formerly pointed for the Cloudnet dataset, the expensive corrective procedure for taking into account geographical and physical differences in the RH_{cr} definition doesn't seem worth of implementation, unless the procedure is refined. More in details, Fig. 2.10 uses a very low threshold (0.01) in order to ensure the separation between "cloud" and

“no-cloud” events. The corrective procedure underestimates the number of cloud occurrences in the low and middle atmosphere. For a larger threshold (as 0.5 in Fig. 2.11), differences between the experiments are much smoother.

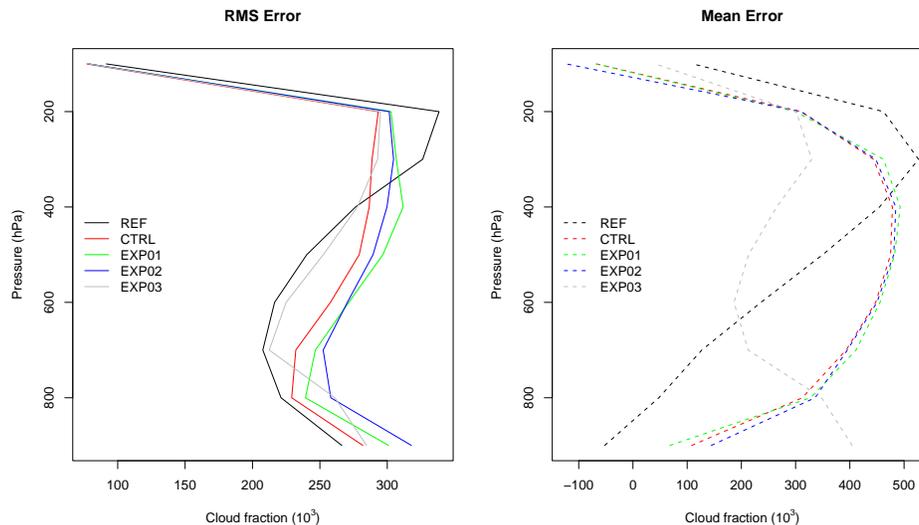


Figure 2.9: RMSE and Mean error at some pressure levels. REF is the pressure-dependent expression of RH_{cr} from Tompkins; CTRL uses as predictors η , η^2 , $CAPE$, T and sea, land and soot aerosol optical thickness. EXP01 uses vertical gradient of potential temperature instead of $CAPE$ in CTRL. EXP02 uses the ice-water separation function instead of T in CTRL. EXP03 adds an additional predictor, η^3 , to CTRL. Cloud fraction is multiplied by 10^3 .

2.2.6 Comparison of CloudSat observed cloud fraction with model prognosed and diagnosed cloud fraction

Superobservations computed in 2.2.4 can be compared with NWP model cloud fractions. This comparison is of interest for assimilating cloud cover data, in order to point out the ability of the forecast model to reproduce the vertical structure of clouds. A one-month comparison has been conducted between the CloudSat cloud fraction super-observations and two cloud cover schemes from the NWP first guess: one prognostic scheme using the “Lopez microphysics”, and one diagnostic using the Sundqvist relation (Eq. 2.7) and the Tompkins definition of RH_{cr} . The model fields, to realistically reproduce an assimilation framework, are 6 hours forecasts Results show that the model is generally able to reproduce the cloud structures as seen by CloudSat. Fig 2.12, 2.12 and 2.14 show a good match between the forecast and the space-borne cloud fractions. Prognosed and diagnosed schemes are able to detect the presence of cloud system. The mislocation and the underestimation, more considerable in the diagnostic scheme, of model fields compared to observations provide an attractive perspective for the assimilation of these data. Once more should be

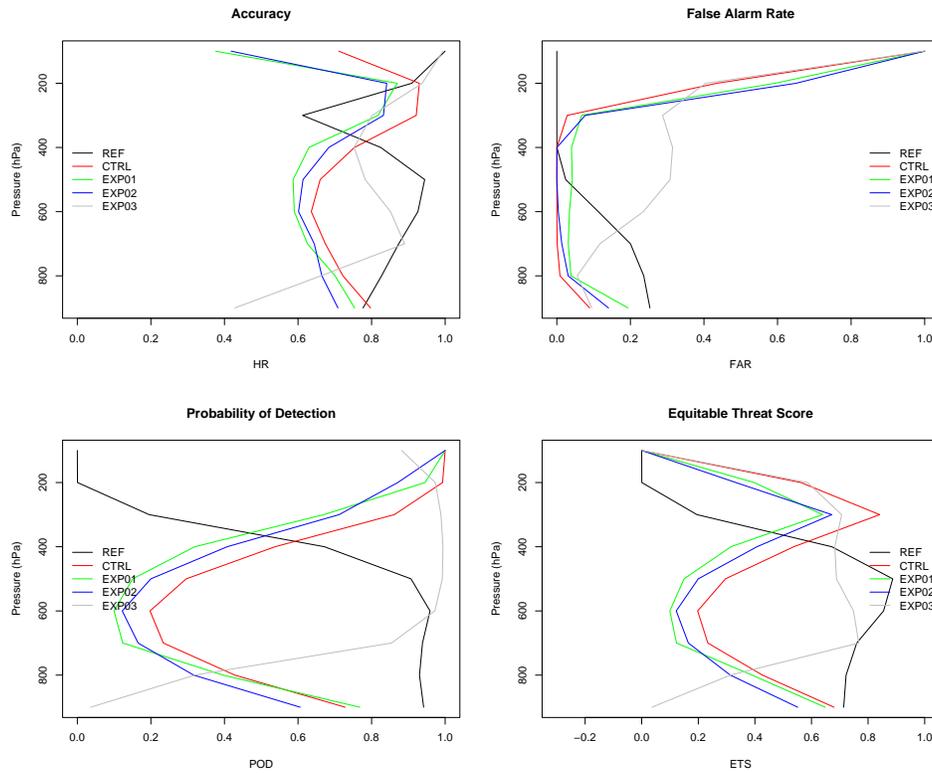


Figure 2.10: HR, FAR, POD and ETS scores for the cloud fraction threshold 0.01. Same experiments as in Fig. 2.9.

noted the strongly binary nature of CloudSat observations against smoother cloud cover fields as seen by the NWP model.

Figure 2.15 also reports the RMSE and the Mean error for different cloud cover relationships and different definition of RH_{cr} . Although the diagnostic schemes suffer of biased error, the Root Mean Square Error suggests that the use of simplified cloud cover schemes is accurate enough for atmospheric cloud characterisation.

2.3 Theory for assimilation of binary cloud cover applied to CloudSat cloud fractions data

2.3.1 The formalism

So far the theory for the assimilation of binary observations has been developed for use when only one observation is assimilated (Storto and Tveter, 2006). The derivation is summarised by the equation

$$P(X_t | K \cap X_b) = \frac{P(K | X_t) P(X_t | X_b)}{P(K | X_b)} \quad (2.9)$$

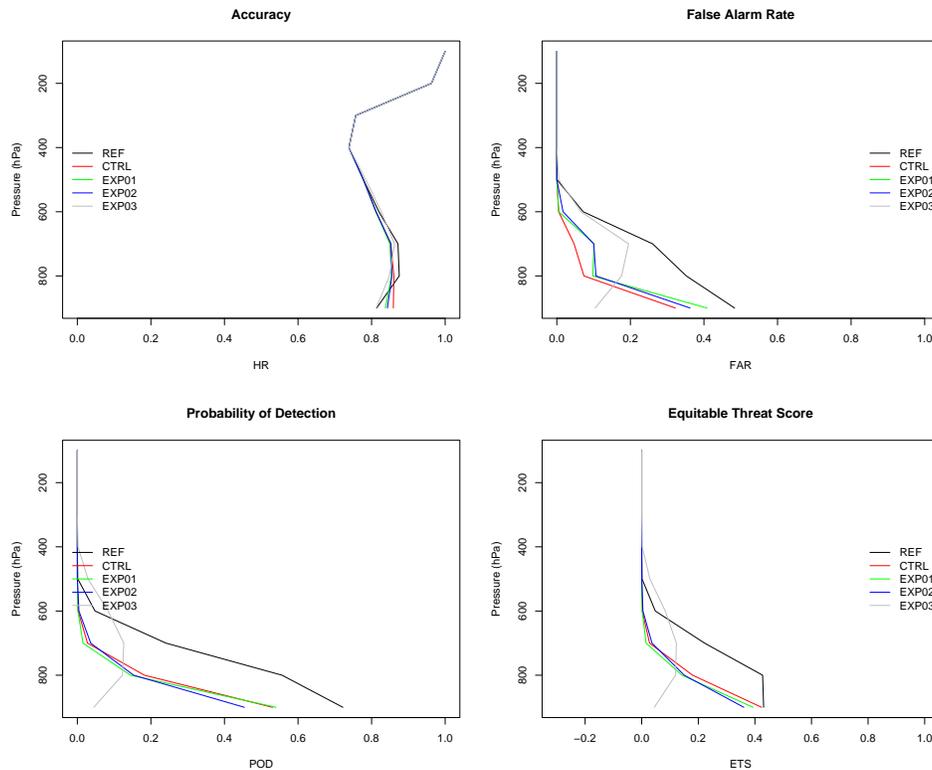


Figure 2.11: HR, FAR, POD and ETS scores for the cloud fraction threshold 0.50. Same experiments as in Fig. 2.9.

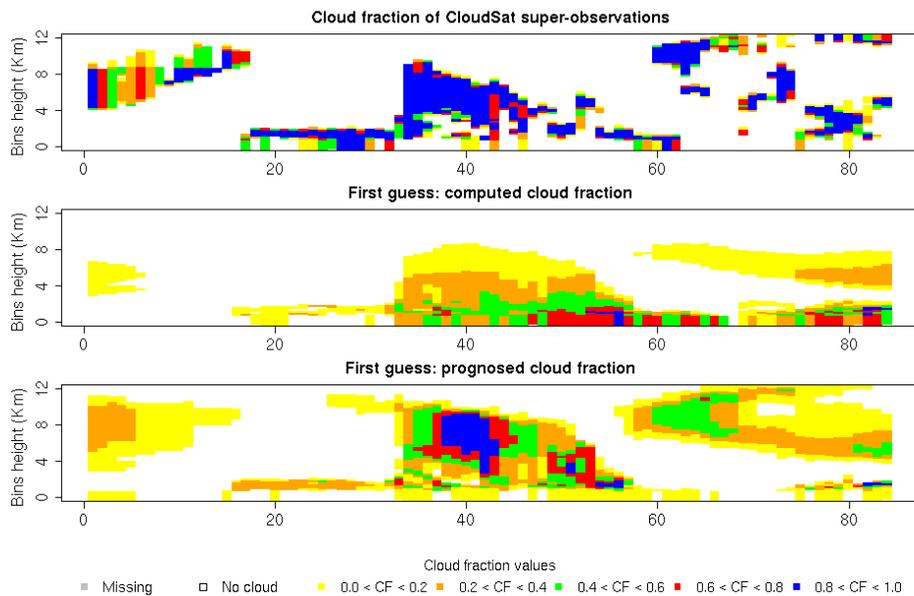


Figure 2.12: Comparison between CloudSat cloud cover (top), ALADIN forecast, diagnostic scheme (middle) and ALADIN forecast, prognosed scheme (bottom) for CloudSat 04.01.2007,00:59 and ALADIN forecast 2007010318+006.

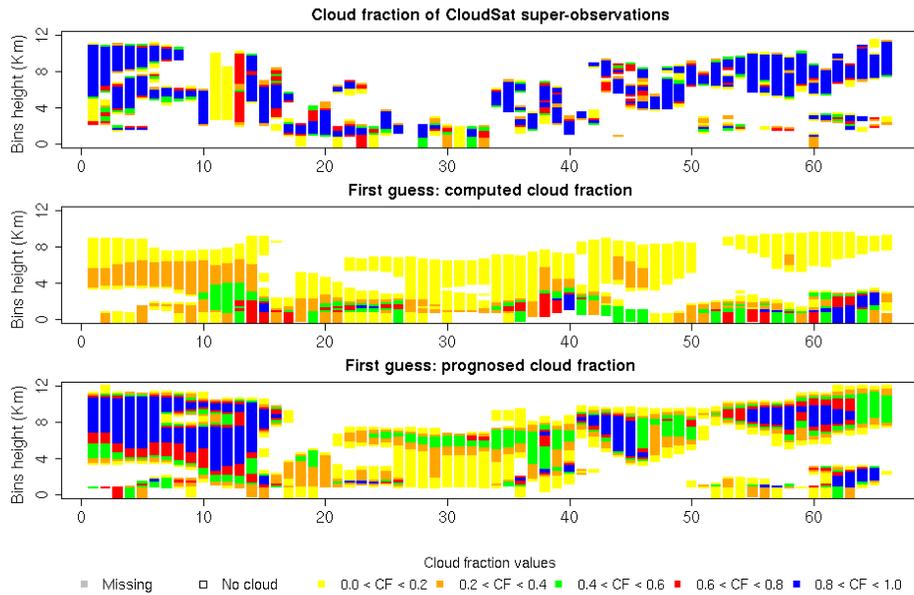


Figure 2.13: Comparison between CloudSat cloud cover (top), ALADIN forecast, diagnostic scheme (middle) and ALADIN forecast, prognosed scheme (bottom) for CloudSat 06.01.2007,12:07 and ALADIN forecast 2007010606+006.

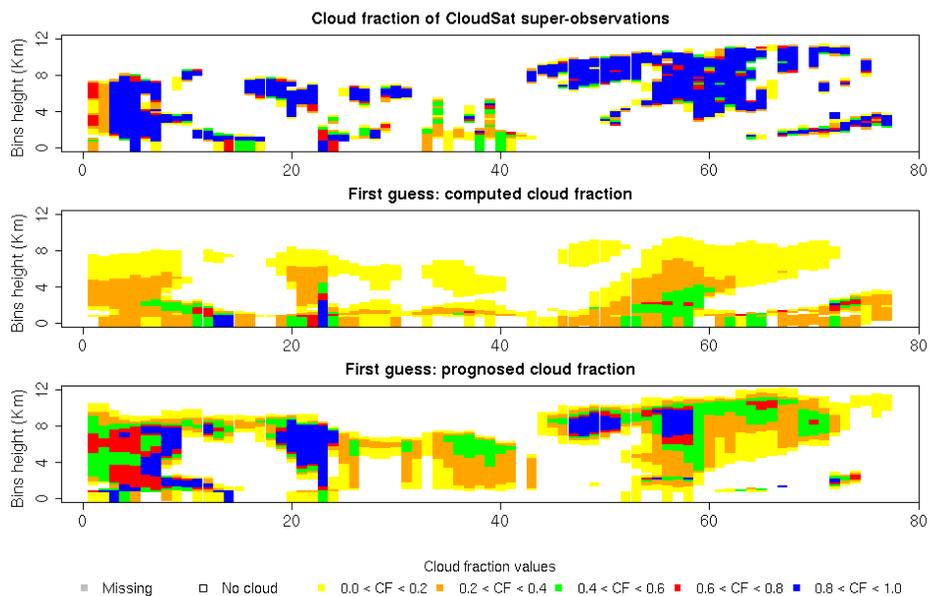


Figure 2.14: Comparison between CloudSat cloud cover (top), ALADIN forecast, diagnostic scheme (middle) and ALADIN forecast, prognosed scheme (bottom) for CloudSat 07.01.2007,12:48 and ALADIN forecast 2007010706+006.

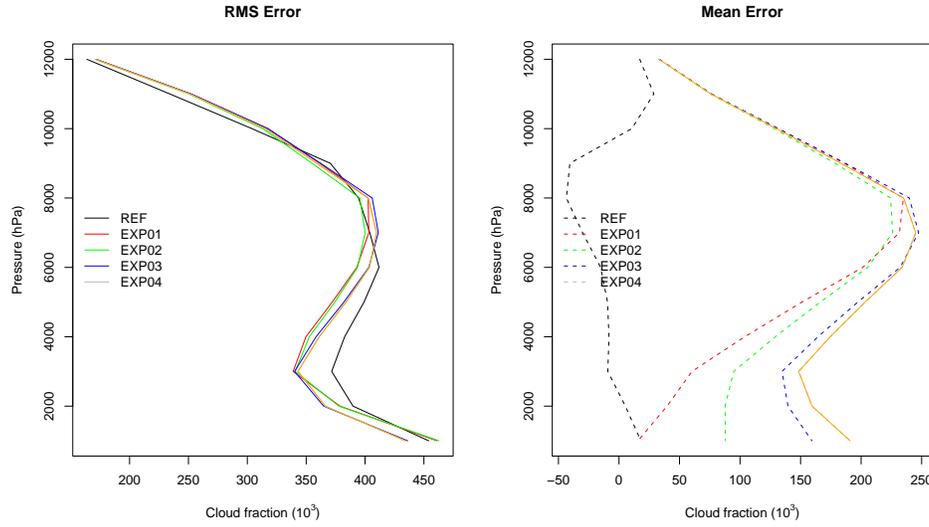


Figure 2.15: RMSE and Mean error at some height levels. REF uses the prognostic scheme, EXP1 and EXP2 uses Sundqvist formulation for diagnosed cloud cover with RH_{cr} as used at ECMWF and at Météo-France respectively, EXP2 and EXP3 uses Tompkins formulation for diagnosed cloud cover with RH_{cr} as used at ECMWF and at Météo-France respectively.

that for K equal to 0 or 1 becomes

$$P(X_t | K \cap X_b) = \begin{cases} \frac{H(X_t)P(X_t|X_b)}{\int_{-\infty}^{+\infty} H(X_t)P(X_t|X_b) dX_t} & \text{if } K = 1, \\ \frac{[1-H(X_t)]P(X_t|X_b)}{\int_{-\infty}^{+\infty} [1-H(X_t)]P(X_t|X_b) dX_t} & \text{if } K = 0. \end{cases} \quad (2.10)$$

In real life the observations can refer to observed or averaged cloud fraction, the latter defined as the percentage of cloud in a given area (e.g. model gridbox). The observed value can be read therefore as the result of several "cloud"/"no-cloud" binary observations, and the previous formulation has to be extended to the case of n observations K_i . If these n observations are assumed independent, we obtain

$$P(X_t | K_1 \cap K_2 \cap \dots \cap K_n \cap X_b) = \frac{\prod_{i=1}^n P(K_i | X_t) P(X_t | X_b)}{\prod_{i=1}^n P(K_i | X_b)} \quad (2.11)$$

and reminding (2.10)

$$P\left(X_t | X_b \bigcap_{i=1}^n K_i\right) = \frac{[H(X_t)]^a [1-H(X_t)]^b P(X_t | X_b)}{\int_{-\infty}^{+\infty} [H(X_t)]^a [1-H(X_t)]^b P(X_t | X_b) dX_t} \quad (2.12)$$

where a and b are the occurrences of $K = 1$ and $K = 0$ respectively, with $a + b = n$. The cloud fraction C can be defined as the number of cloud observations out of the total number of observations n :

$$C = \frac{a}{a+b} = \frac{a}{n}; \quad a = nC; \quad b = n(1-C) \quad (2.13)$$

and we have

$$P\left(X_t \mid X_b \bigcap_{i=1}^n K_i\right) = \frac{[H(X_t)]^{nC} [1 - H(X_t)]^{n(1-C)} P(X_t \mid X_b)}{\int_{-\infty}^{+\infty} [H(X_t)]^{nC} [1 - H(X_t)]^{n(1-C)} P(X_t \mid X_b) dX_t} \quad (2.14)$$

For an infinite number of observations the limit is:

$$P\left(X_t \mid X_b \bigcap_{i=1}^n K_i\right) = \lim_{n \rightarrow \infty} \frac{[H(X_t)]^{nC} [1 - H(X_t)]^{n(1-C)} P(X_t \mid X_b)}{\int_{-\infty}^{+\infty} [H(X_t)]^{nC} [1 - H(X_t)]^{n(1-C)} P(X_t \mid X_b) dX_t} \quad (2.15)$$

Unfortunately the limit is undetermined; for large values of n , $[H(X_t)]^{nC} [1 - H(X_t)]^{n(1-C)}$ gets asymptotic (see Figure 2.16 for $C = 0.8$), and the same happens for the integral at denominator.

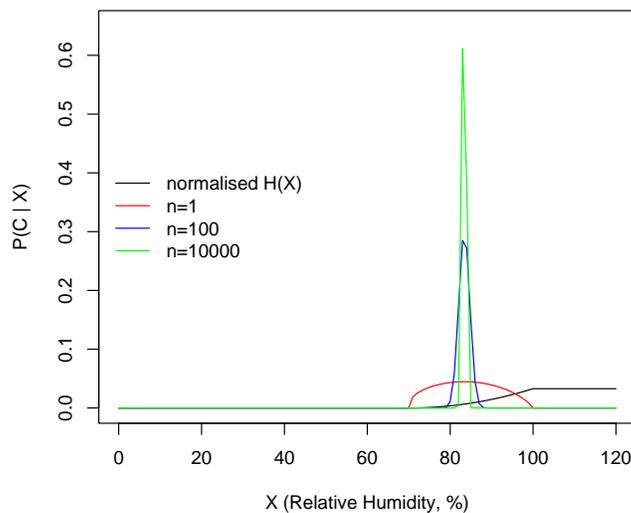


Figure 2.16: Distribution of $(C \mid X_t)$.

The impact of different number of binary “sub-observations” that constitute one averaged observation can be showed by simulating a dataset of observations. Starting from a uniform distribution of “true state of the atmosphere” X_t , it is possible to define the a priori knowledge of the atmosphere (X_b , known as background or

first guess) that follows a normal distribution with mean X_t and standard deviation σ_b . Further, X_t is moved to cloud fraction space through the operator $H(X_t)$, and n binary sub-observations b_i can be defined using n uniform random variables k_i : when $k_i > H(X_t)$, the i -th sub-observation $b_i = 1$ otherwise is 0. The total cloud fraction f is therefore the average of the n b_i sub-observations. If X is the relative humidity, the mathematical formalism reads:

$$\begin{aligned} X_t &\sim U(0, 100) \\ X_b - X_t &\sim N(0, \sigma_b) \\ k_i &\sim U(0, 1) \\ b_i &= 1 \quad \text{if } k_i - H(X_t) > 0 \\ b_i &= 0 \quad \text{if } k_i - H(X_t) \leq 0 \\ CF &= \frac{1}{n} \sum_{i=0}^n b_i \end{aligned}$$

From these n sub-observations, the analysis X_a , defined as $E \left[X_t \mid X_b \bigcap_{i=1}^n K_i \right]$, can be computed through Eq. 2.14, and compared with the actual distribution of occurrences in the simulated dataset.

Fig. 2.17 shows the results when the number of observations is set to 10. The probability distribution of having a cloud fraction given X_t as computed from 2.14 and as counted in the simulated dataset and the probability (theoretical and experimental) of X_t given the observation and the background are plotted for different cloud fraction values. Eq. 2.14 is able to describe the probability distributions interested by the fractional values of clouds; moreover, it's shown (green line in the plots) the error in considering only one sub-observation. This is summarised by Fig. 2.18, that compares the analysis increment (analysis - background) function of observed cloud fraction for different numbers of sub-observations n .

2.3.2 Discussion

At this point an important question to address is “how many binary sub-observations generate a cloud fraction observation”. To answer, it should be noted first of all that from a physical point of view it is sensitive to think of the sub-observations as observations at cloud resolving resolution (1 to 2 Km) for which the parameter cloud fraction is meaningless; at the scale where subgrid fluctuations of water content is negligible, the presence of condensed water is a binary occurrence. This relies also with the definition of RH_{cr} , as pointed in the previous section of the Chapter, and can be reasonably pointed that the number of sub-observations to use

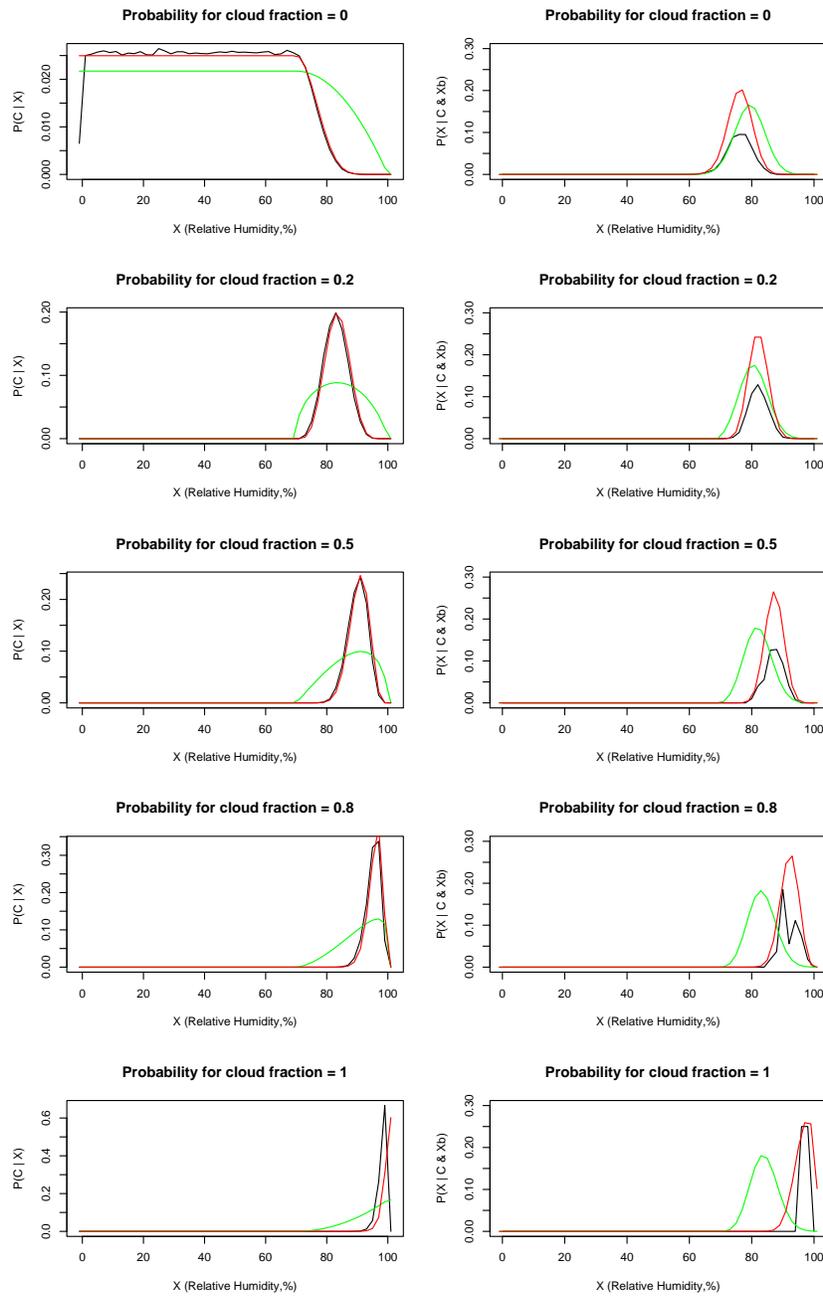


Figure 2.17: Probability of having a cloud fraction given X_t (left) and probability of X_t given the observation and the background (right) when the number of sub-observations is 10 and background relative humidity is 80%. In black the distribution as counted in the simulated dataset; in red the theoretical solution of Eq. 2.14; in green the same equation when only one sub-observation is assumed.

depends on the spatial resolution of the “averaging operator”; in other words, the further is the spatial resolution of the observations from cloud resolving resolution, the more we need a greater number of binary sub-observations.

Another aspect of the use of Eq. 2.14, visible through Fig. 2.18, is that such a decomposition of a cloud fraction observation leads to larger analysis increments.

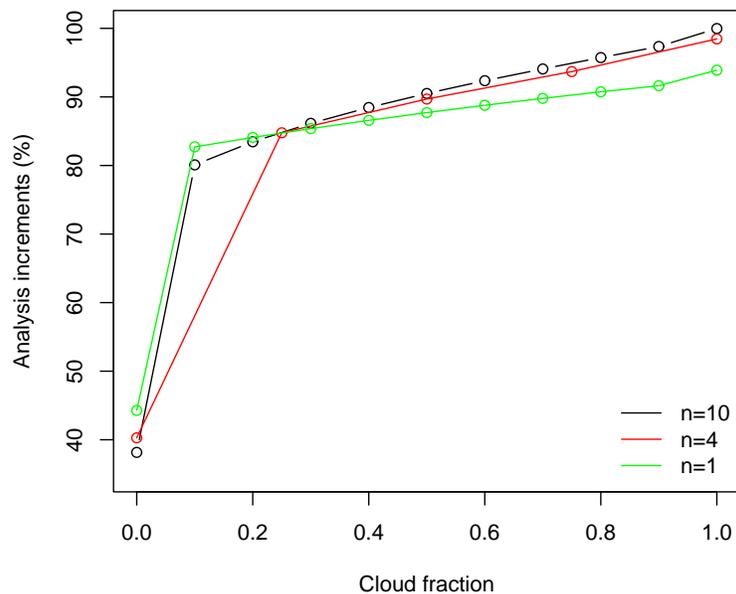


Figure 2.18: Analysis increments for a different number of sub-observations; background relative humidity is 80%.

This depends on the assumption of independence of the sub-observations, each one of them takes the same weight in the computation of the analysis. The larger the number of sub-observations, the more important will be the role of the observations in the analysis; the contribution of the observation probability distribution to the analysis increases because of the increased sharpness of the forward model: with many coherent observations the transition from a “no-cloud” event to a “cloud” event is more sharp.

Chapter 3

Use of satellite data in ALADIN-HARMONIE Norway

Summary

The ALADIN/HARMONIE three-dimensional variational data assimilation system is being implemented at the Norwegian Meteorological Institute. Use of satellite observations is very important to improve numerical weather prediction in high latitude regions. Our poster will present the implementation of most of the available satellite data in the ALADIN/HARMONIE-Norway analysis system (microwave: AMSU-A and AMSU-B/MHS; infrared: Seviri and IASI; GPS: ground-based zenith total delay). This Chapter will also discuss two ways (the so-called NMC and the ensemble-based method) of estimation of background error covariances as well as the use of off-line predictors computation and variational bias correction methods to correct radiance bias.

This Chapter is based on a work presented at the XVI International TOVS Study Conference (see Appendix B: Storto A. and Randriamampianina R., 2008).

3.1 Introduction

Within the ALADIN/HIRLAM cooperation agreement, a new assimilation and forecast system is being developed with the aim of providing a reliable framework for both research and operational purposes, especially for high resolution applications. The system has been named HARMONIE (Hirlam Aladin Regional/Meso-scale Operational NWP In Europe), and his forecast models are now used operationally in many HIRLAM national meteorological services, either using non-hydrostatic physics at cloud-resolving resolution or hydrostatic physics at synoptic scale. The Norwegian Meteorological Institute (Met.no) is putting many efforts in building the assimilation counterpart of the system, whose core is based on the spectral upper-air three-dimensional variational assimilation (3D-Var) of the ALADIN model, op-

erational since 2005 at the Hungarian Meteorological Service and at the French Meteorological Service. The assimilation system (hereafter HARMONIE-3DVar) is currently used at Met.no mainly for two research projects, the Eumetsat funded "Assimilation of binary cloud cover" and the IPY-THORPEX that aims to investigate the importance of remote-sensed observations in forecasting polar lows; there are also plans to operationally run HARMONIE-3DVar in the near future. Basic configuration of the system consists of using a 6 hours forecast from previous cycle as background (first guess), i) updating the sea surface temperature (SST) through the ECMWF SST analysis; ii) extracting and pre-processing all the available and supported observations; iii) performing a surface assimilation based on the ALADIN community Optimal Interpolation software (CANARI) to analyse surface parameters over land (skin temperature, soil water content); iv) performing the spectral upper-air analysis for vorticity, divergence, temperature, specific humidity and surface pressure, v) running the forecast model after proper downscaling of lateral boundary conditions from the ECMWF global model. The upper-air 3D-Var, that we will focus on in the rest of the Chapter, supports at the moment all the conventional observations, Atmospheric Motion Vectors, aircraft in-situ observations, microwave radiances from POESS and Metop platforms, radiances from MSG/SEVIRI Imager, and, in an experimental configuration discussed separately, also infrared radiances from the Infrared Atmospheric Sounding Interferometer (IASI), Zenith Total Delay derived from ground-based GPS stations (GPS-ZTD), humidity retrievals from the CloudSat CPR radar. Assimilation of scatterometer observations from ASCAT aboard Metop is currently under development. In order to optimally exploit the information contained in space-borne instruments, which are very important for limited area assimilation systems, a number of questions should be addressed, such as channel selections for multi-channel instruments, tuning of observational errors, choice and implementation of bias correction strategies, assessment of background error covariances and relative impact of the observing network on analysed fields. In the following part of the Chapter, these issues will be discussed together with an overview of the assimilation system and some remarks about the actual use of some observations; finally, results from sensitivity studies of analysis and forecasts to different observation groups will be presented as diagnostic tool for understanding the relative importance of observations in the system.

3.2 Observations in the reference assimilation system

HARMONIE-3DVar currently supports the assimilation of a number of observations, schematically reported in Table 3.1. The table also presents the average horizontal thinning distance between assimilated observations. All the conventional

observations are assimilated (radiosondes, synoptic land and ship stations reports, buoys and drifting buoys measurements, wind profilers). Additionally, aircraft observations (AMDAR/AIREP), Atmospheric Motion Vectors (AMV) provided by EUMETSAT and derived from Meteosat Second Generation satellites (MSG) are extracted and assimilated, microwave radiances from the Advanced Microwave Sounding Unit (AMSU) and the Microwave Humidity Sounder (MHS) aboard NOAA and Metop polar satellites are exploited. Only MSG/SEVIRI supplies infrared radiances. Further to satellite bias correction, also daytime temperature measurements from radiosondes are bias-corrected through ECMWF flat bias correction values, which depend on instrument characteristics.

Observations	Parameter	Thinning
SYNOP	Z	-
AIREP	U, V, T	25 Km
AMV	U, V	25 Km
DRIBU, BUOY	Z	-
EUROPROFILERS	U, V	-
RADIOSONDES	Z, U, V, T, Q	-
AMSU-A	Tb	80 Km
AMSU-B	Tb	80 Km
MSG/SEVIRI	60 Km	-

Table 3.1: Observations and horizontal thinning distance used in HARMONIE-3DVar.

Quality control and rejection of observations is carried out through a few steps, consisting in duplicated reports check, background quality control, redundancy check and spatial and temporal thinning.

3.3 Satellite radiances assimilation

The observation operator for satellite radiances is the Radiative Transfer for TOV (RTTOV) in his version 8.5, developed by the Numerical Weather Prediction Satellite Application Facility (NWP-SAF). All the AMSU-A channels from 5 to 13 from NOAA-15, -16, -18 and Metop are assimilated, excepted channels 5 and 11 from NOAA-15; for AMSU-B, channels 3, 4 and 5 are assimilated from all the NOAA polar-orbiting satellites and from Metop.

For microwave instruments, bias correction has been performed by applying the Harris and Kelly (2001) scheme, which uses air-mass and scan-angle predictors to compute the bias. The predictors chosen are 1000-200 hPa thickness, 100-50 hPa thickness, skin temperature, integrated water vapour, scan-angle and his square and cubic power. The coefficients have been computed from a two-months period dataset of forecasts initialised by dynamical adaptation from the ECMWF global model. Results from the implementation of variational bias correction scheme will be briefly

discussed later. A detailed study of innovations and residuals statistics for all couples channel/satellite separately for each network (00, 06, 12, 18 UTC) is currently under evaluation, with the aim of blacklisting satellite channels whose observations amount at certain hours is poor in the computational domain and can cause unreliable computation of bias correction coefficients. Assimilation of MSG/SEVIRI infrared radiances takes advantage of the Nowcasting Satellite Application Facility (NWC-SAF) that is used for brightness temperature recalibration, I/O handling. Cloud type and cloud top height products, still from NWC-SAF, are used for eventually black-listing radiances data, according to Table 3.2. Channel 4 is not assimilated because RTTOV does not reproduce accurately radiances for very broad channels (Brunel and Turner, 2003), while the ozone channel is not used at all. Possibility to extend the assimilation of channel 11 also when low-level clouds are detected will be investigated in the future. Only radiances relative to Meteosat-9 scans starting at 05.45, 11.45, 17.45 and 23.45 UTC are considered, and only one pixel over 4 (8 Km resolution) is retained in the observational database. The air-mass scheme for bias correction is the same as the one used for microwave radiances, but bias correction coefficients are assumed latitudinally (along scanline) constant. Radiances that are far away from Meteosat-9 position (latitude $\geq 65^\circ$ N) are rejected.

Channel	Spectral Band	Use
4	IR3.9	Monitored
5	WV6.2	Clear-sky and above mid-level clouds
6	WV7.3	Clear-sky and above mid-level clouds
7	IR8.7	Clear-sky
8	IR9.7	Not Used
9	IR10.8	Clear-sky
10	IR12.0	Clear-sky
11	IR13.4	Monitored

Table 3.2: Use of MSG/SEVIRI thermal channels in HARMONIE-3DVar.

Figure 3.1 shows the impact of MSG-2/SEVIRI assimilation through radiosondes verification scores: the impact is in general slightly positive and, as expected, humidity fields are the most benefited, especially after 24 hours of forecasts.

An ongoing work is dealing with satellite-specific brightness temperature errors using diagnostics from the Desroziers method (Desroziers et al., 2005a); partial results suggest that errors were in general over-estimated and important differences are found between different satellite errors in a limited area model, e.g. Metop/MHS diagnosed errors are significantly smaller than NOAA-16/AMSU-B.

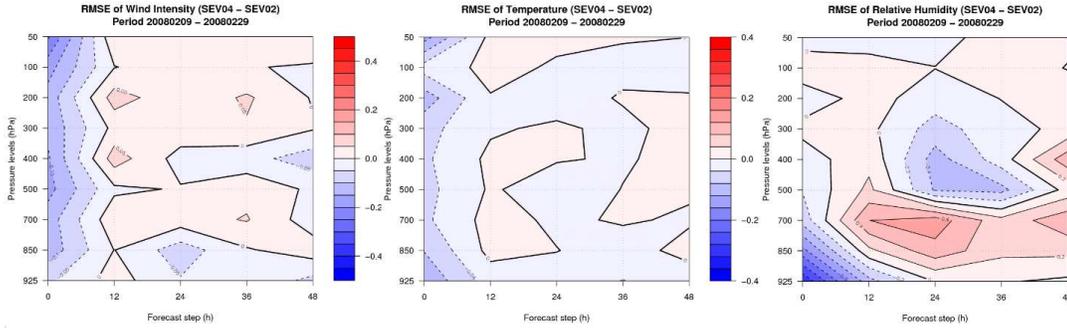


Figure 3.1: Difference of root mean square errors between a reference experiment and an experiment with SEVIRI data assimilation against radiosonde observations. Where positive (red), SEVIRI observations have a positive impact.

3.4 Variational Bias correction

Variational bias correction has been implemented and it's now used for all satellite radiances. Main advantages of such a method are the improved separation between model and observation contribution to the total bias, and the possibility of automatically computing bias correction coefficient, otherwise very expensive for high spectral resolution instruments (i.e. AIRS and IASI). In the Met.no configuration, bias coefficients are initialized from previous assimilation cycle and the departures of from such values are minimised in the 3DVar cost function as additional term. At the end of the minimization, the coefficients are then suitable for initialising next assimilation coefficients. At implementation time, this procedure is iterative, starting from zeroed bias coefficients (cold start), and has been observed to converge to reliable values (i.e. unbiased observation minus analysis differences) in less than a one-month period of 6 hourly assimilation. Results (not shown) show a very positive impact on assimilation statistics and forecasts scores. The use of coefficients from 24 hours old instead of from the previous 6 hours old cycle to respect network characteristics (amount of radiances and scan-angles distribution in the LAM domain) is currently under evaluation.

3.5 Importance of the background error covariances assessment

Specification of background error covariances (B matrix) for use in the 3DVar algorithm may affect the impact of observations, since spatial auto-covariances as well as cross-covariances between different state parameters lead in turn to different weights given to observations in the analysis system. In HARMONIE-3DVar we have obtained background error statistics through the application of two different methods: the "NMC" method, which derives error statistics from a dataset of differences

of couples of forecasts valid at the same time but initialised at different time (we used 48-24 hours forecasts); the ensemble method that in the Met.no configuration uses differences of 6 hours ensemble forecasts minus the ensemble mean, using 10 members derived from downscaling ECMWF/IFS ensemble analysis; these ones had been obtained through observations perturbation (to simulate analysis errors) and spectral backscatter scheme (to simulate forecast errors) from Isaksen (2007) experiment. In both cases, the B matrix formulation follows Berre (2000), that assumes isotropic and homogeneous but vertically varying covariances and cross-covariances computed through multiple linear regression. Main differences between the two methods rely to the broader vertical correlations of background errors for the NMC method, excepted at very small horizontal scales. Figure 3.2 shows for instance the vertical correlations of temperature between level 48 (about 850 hPa) and the other model levels as function of horizontal scale (wavenumber); statistics retrieved via the NMC method present a large-scale vertical correlation reaching downward the surface and upward around 100 hPa. This finding becomes very noticeable in satellite radiances single observation experiments for high-peaking channels that involve many model levels. Figure Eq. (3.3) reports the temperature analysis increments of a single observation experiment for channel 9 of AMSU-A (aboard NOAA-18) with a brightness temperature innovation of 2 K, using an NMC B matrix computed over three months forecasts in winter (DJF 2006/2007) and an Ensemble B matrix from downscaled ensemble analysis with 10 members for the period from 20061025 to 20071125. We want to stress that much broader vertical error correlations generated by applying NMC method cause unrealistic analysis increments at very high levels, even reaching model top, while ensemble errors do not.

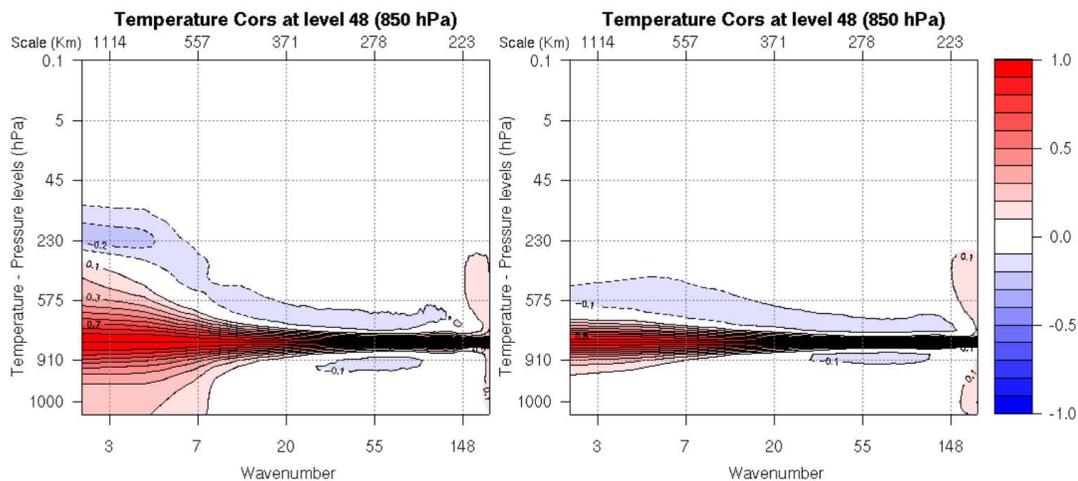


Figure 3.2: Vertical error correlation of temperature between model level 48 and all the other model levels as function of horizontal scale. Left and right panel show the statistics derived via the NMC method and the ensemble method respectively.

3.6 Impact on analysis

As index to study the relative impact of observations in the assimilation system, we use Degrees of Freedom for Signal (DFS, see e.g. Cardinali et al., 2004), that is defined as the derivative of the analysis increments in observation space with respect to the observations. In practice, it is computed through a randomisation technique (Desroziers et al., 2005b).

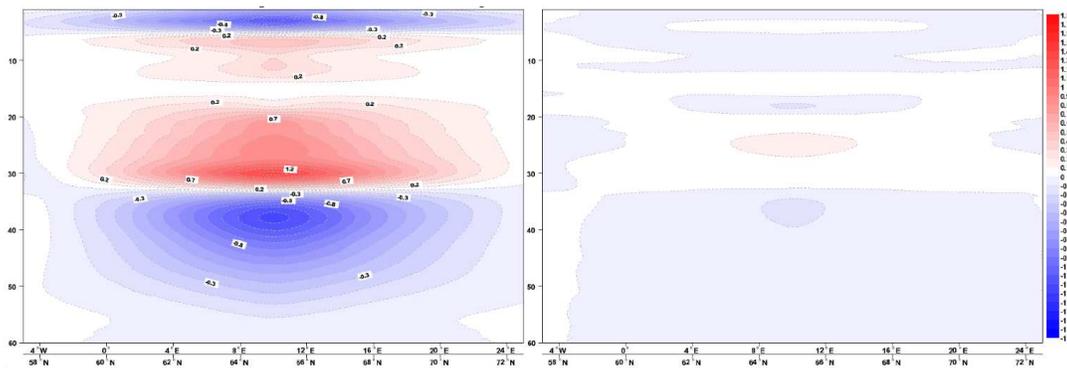


Figure 3.3: Cross-sections of temperature analysis increments for brightness temperature single-observation experiments (2 K innovation for AMSU-A channel 9 aboard NOAA-18). Left and right panel show the increments using the NMC method and the ensemble method statistics respectively.

The perturbation is performed using an unbiased Gaussian random error whose standard deviation equals the observation error; 6 assimilation cycles, 4 days far each other to ensure ergodicity of statistics, have been rerun with perturbed observations. DFS for each observation have been grouped into parameters and types categories to provide information about the weight of observations in the assimilation system. It is also possible to define Relative Degrees of Freedom for Signal as DFS divided by the number of observations in the subset, which indeed represent an index of the theoretical weight given to each single observation. Results (Figure 3.4) show the large importance of wind observations, emphasising the role of aircraft and AMSU-A observations in the HARMONIE-3DVar system. Use of variational bias correction increases the weights given to observations, not only for remote-sensed observations. Humidity measurements and humidity-related observations (SEVIRI Water Vapour channels, AMSU-B) are very important in relative terms, but less crucial in the actual assimilation system because of the small amount, compared to other observations.

3.7 Experimental observations

The assimilation of a number of experimental and new observations is under development within the HARMONIE-3DVar system at Met.no. We summarise in the

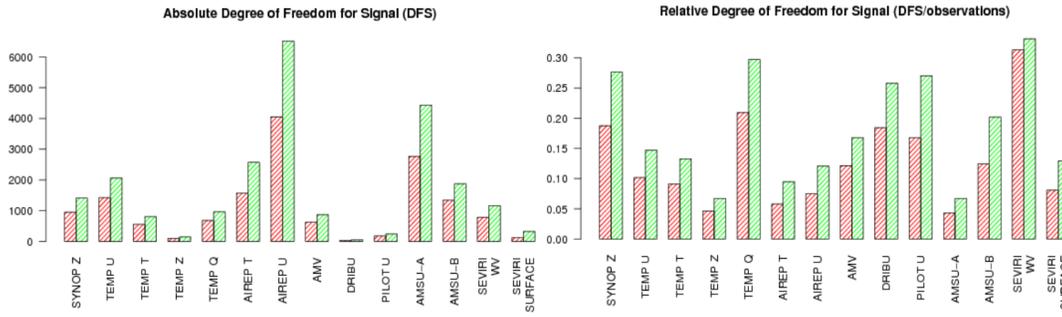


Figure 3.4: Absolute and relative DFS. Red bars refer to Harris and Kelly bias correction scheme experiment for AMSU and SEVIRI; green bars for variational bias correction experiment.

sequel strategies and main results. The reader can refer to Randriamampianina and Storto (2008) in these conference proceedings for issues concerning the assimilation of IASI radiances.

3.7.1 Zenith Total Delay from ground-based GPS stations

Delay of GPS satellite signal measured when ground-based stations point at zenith contains information about the vertical profile of atmospheric refractivity, providing therefore information about temperature profile and integrated vertical moisture content. The observation operator comes from Poli et al. (2007) and links the control state with the delay processed by different centres throughout Europe. Flat bias correction is applied to a number of couples of stations and processing centres whose data supply is regular and whose background departure follows a Gaussian probability density function. Observation errors, specified separately for each station, have been obtained from inflating down empirically the standard deviations of observation minus guess differences and diagnostic statistics (not shown here) suggest that they have been however overestimated. After having screened irregular, unreliable and duplicated stations, 54 stations have been selected inside the HARMONIE-3DVar domain, and the impact of those observations has been studied over a one-month assimilation period. Results (Figure 3.5) show a slightly positive impact of GPS-ZTD, especially for mass fields.

Possibility to improve bias correction procedures by the use of a predictors-based scheme for taking into account observation operator error derived by model orography displacement and other sensitive parameters is currently under evaluation, together with a more robust definition of observation errors.

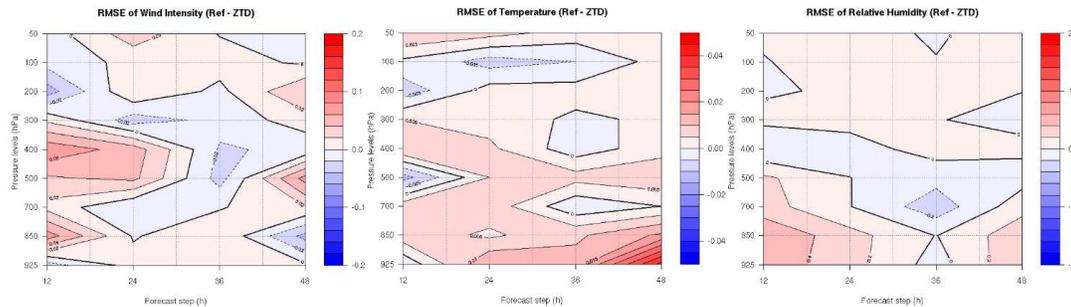


Figure 3.5: Difference of root mean square errors between a reference experiment and an experiment with GPS-ZTD data assimilation against radiosonde observations. Where positive (red), GPS-ZTD observations have a positive impact.

3.8 Summary

As the ALADIN/HARMONIE three-dimensional variational assimilation is being developed at the Norwegian Meteorological Institute, a lot of scientific choices and practical issues have to be coped with. A brief overview of the system, able to assimilate both conventional and remote-sensed data, has been given. Satellite observations are very important for enhancing forecast verification scores: we are now assimilating ATOVS/AMSU-A, AMSU-B and MHS radiances by using the RT-TOV transfer model, and MSG/SEVIRI infrared radiances pre-processed through the NWC-SAF software. Detailed selection of channels for each satellite at each network (0, 06, 12, 18 UTC) is a delicate task in limited area assimilation systems, and will be completed soon. Comparisons between different bias correction strategies has been dealt with: variational procedure leads to easy bias correction procedure for high spectral resolution sounders, and shows positive impact in terms of both unbiased residuals (analysis minus observation) statistics and verification scores. Use of ensemble methods for estimating background error covariances provides less broad vertical correlations than the ones derived via the NMC method. This avoids that high-peaking channels generate unrealistic analysis increments at many vertical levels, reaching the model top. The study of the impact of observation subgroups on the analysis and forecasts has been performed by using randomisation techniques: AMSU-A and wind measurements, especially from airborne instruments, result the most important observations as seen from the analysis, while the impact on forecasts, computed using an RMSE-based cost function, shows the great importance of AMSU-A radiances for all dynamical parameters at all forecast ranges; aircraft and AMV data seem to affect significantly short-range forecasts for temperature fields, while AMSU-B plays an important role for humidity fields. Promising results have been obtained from the assimilation of zenith total delays from ground-based GPS stations.

Chapter 4

Assimilating Humidity Pseudo-observations Derived from the Cloud Profiling Radar Aboard CloudSat in ALADIN 3D-Var

Summary

This Chapter describes an experimental procedure for assimilating CloudSat Cloud Profiling Radar (CPR) observations in ALADIN 3D-Var through the use of humidity pseudo-observations derived from a one-dimensional Bayesian analysis. Cloud data are considered as binary occurrences (“cloud” versus “no-cloud”), which makes the approach feasible to be extended to other cloudiness observations, and to any other binary observation in general. A simple large-scale condensation scheme is used for projecting the prior information from a Numerical Weather Prediction model into cloud fraction space. Verification over a one-month assimilation test period indicates a clear benefit of the pseudo-observation assimilation scheme for the limited CloudSat CPR data set, especially in terms of improved skill scores for dynamical parameters like geopotential and wind.

This Chapter is based on a paper submitted to Meteorological Applications (see Appendix B: Storto A. and Tveter F.T., 2008).

4.1 Introduction

The quality of the humidity analysis in Numerical Weather Prediction (NWP) models is crucial for the correct prediction of local weather and severe weather events, especially in short-range forecasts. Within the global observing system, the only instruments which provide in-situ humidity measurements are the hygrometers that meteorological sondes are equipped with, although other observation types, such as for instance satellite radiances and atmospheric path delay of GPS satellite signals, contain important information about humidity distribution and may have significant impact on medium-range forecasts, also on wind fields (Andersson et al.,

2007). Spaceborne infrared sounders measuring in water vapour absorbing bands contain information about vertical moisture content for clear sky conditions. Unfortunately, the assimilation of radiances in cloudy conditions is still problematic because of the non-linearity of the observation with respect to the state variables and the inaccurate scattering parametrisation in rapid radiative transfer models (RRTMs). Hence, assimilation systems often do not exploit satellite observations in cloudy conditions, that are all the more important for correctly forecasting severe weather events (McNally, 2002). Several attempts are currently being made to address the RRTMs limits in cloudy conditions (see for instance Chevallier et al., 2004).

Cloud cover observations generally contain useful information about the horizontal and vertical structure of the humidity fields, but they are usually not assimilated in NWP models although they are exploited by duty forecasters in subjective forecasting and in nowcasting systems (Golding, 1998). Assimilation of cloud cover observations has been shown to have a positive impact when data source are nowcasting-derived cloud products. At the U.K. Meteorological Office (UKMO), three-dimensional cloud fraction data derived from MOPS, the Moisture Observation Pre-processing system (Wright, 1993), have been assimilated in the Met Office assimilation system (Macpherson et al., 1996) via a nudging scheme. Recent efforts are addressed to move these observations into the variational assimilation scheme (Renshaw, 2007). MOPS data are generated by the combined use of Meteosat infrared sounder, ground-based synoptic stations (SYNOP) reports for cloud base height and radar data for adjusting the cloud ice-liquid water separation. Benefits of assimilating MOPS cloud cover data are evident especially in terms of cloud octas verification, at least up to 18 hours of forecasts. At the European Centre for Medium-Range Weather Forecasts (ECMWF), Janisková et al. (2002a) have shown the possibility of a 1D-Var + 4D-Var assimilation scheme for cloud properties and radiative fluxes observations from the Atmospheric Radiation Measurement (ARM) programme by correcting temperature and specific humidity profiles; this two-step approach has been adopted by Chevallier et al. (2002) for cloud-affected radiances from ATOVS radiometers. Lopez et al. (2006) assessed this strategy also with radar reflectivities, ground-based precipitation measurements and GPS total column water vapour retrievals, adding the temporal dimension as well (2D-Var). At the French Meteorological Service (Météo-France) experimental assimilation of humidity pseudo-profiles (Guidard et al., 2006) derived from Meteosat nowcasting products has been performed during the MAP campaign, showing a positive impact on humidity fields and precipitation forecasts, in terms of both location and intensity. Satellite cloud observations have been also successfully used to adaptively tune the cloud schemes parameters (Norris and da Silva, 2007), and Benedetti et al.

(2003) demonstrated the benefits of assimilating cloud-radar reflectivities in a cirrus model to correct the size distribution of ice-cloud crystals.

In April 2006 the U.S. National Aeronautics and Space Administration (NASA) launched the CloudSat satellite, equipped with an on-board 94-GHz radar, the Cloud Profiling Radar (CPR). Detection of clouds by using millimeter-wavelength radars has already been demonstrated by Pazmany et al. (1994); Clothiaux et al. (1995); Babb et al. (1998); Kollias et al. (2007) but the CPR is the first spaceborne radar devoted to cloud systems studies. The CloudSat mission (Stephens et al., 2002) aims at improving the knowledge of microphysical composition of clouds and providing very accurate observations for the validation and verification of cloud schemes in NWP models. By detecting the vertical structure of clouds, the CPR radar also offers a unique dataset of spaceborne radar echo powers that can be used for detecting cloud distributions at a very high resolution, both along the satellite track and on the vertical. Such cloud data can be therefore used also for assimilation purposes, and this Chapter aims to demonstrate the benefits of assimilating cloud observations from CloudSat. However, dissemination of CloudSat Level 1 data has a delay of about 6 hours, and the exploitation of CPR observations is therefore possible in experimental mode only. The study of the potential of cloud cover observations from the Cloud Profiling Radar is also of interest for both evaluating the assimilation of cloud fraction observations and designing next-generation instruments for low Earth orbit satellites. Another joint NASA-Centre National d'Etudes Spatiales (CNES) mission, Calipso, has started to provide Lidar measurements useful for detecting high-level thin clouds and aerosol concentrations, in synergy with CloudSat. Calipso data and derived products will not be considered in this study.

The Norwegian Meteorological Institute (Met.no) has an ongoing project for the development of assimilation techniques for cloudiness information. The main goal is to establish a theory for assimilating cloud observations which can be applied to several spaceborne platforms. The approach is to use cloud fraction observations in a pre-processing task to retrieve humidity *pseudo-observations* (also called *humidity retrievals*) suitable for direct assimilation in Numerical Weather Prediction systems via variational data assimilation methods.

The assimilation system used in this study is not the Met.no operational HIRLAM 3D-Var (Gustafsson et al., 2001) but the ALADIN 3D-Var (Fischer et al., 2005) assimilation and forecast system in its quasi-operational configuration. This system already supports the assimilation of a number of conventional and remote-sensed observations and the analysis and forecasts performances have already been evaluated over a long period showing a reliable NWP framework. After the introduction, this Chapter consists of 5 sections that describe respectively, i) an overview of the assimilation strategy, ii) the Bayesian theory behind the assimilation scheme

and the cloud fraction observation operator, iii) the practical utilisation of CloudSat data and the definition of observation errors for use in 3D-Var, iv) the verification results from a one-month assimilation period and an illustrative case for evaluating the impact of CloudSat observations.

4.2 Assimilation strategy

The motivation for assimilating cloud cover observations instead of raw radar return power data stems from the desire to establish a theory for assimilating cloud data, which in principle can be applied to several cloud products.

Unless important assumptions in variational data assimilation theory are neglected, cloud observations can not be directly assimilated in existing variational data assimilation algorithms. The cloud observation error distributions are non-Normal, and the relationship between the observation and the atmospheric state (humidity and temperature) is very non-linear. Zou (1997) showed that for “on/off” processes such as cloud formation and precipitation, important errors in the linearised (tangent-linear) observation operator, and in his adjoint model for the gradient computation, arise when the full non-linear model is not continuous at switching points. This is even more critical when assimilating “no-cloud” occurrences, which correspond to a large range of humidity values. To overcome such problems without loss of information and keeping a rigorous derivation, a two-step approach consisting of a one-dimensional Bayesian analysis followed by three-dimensional variational assimilation is used. Note that other approaches have been investigated for assimilating cloud observations in a variational framework, for instance by artificially modifying the cloud cover observation operator to have a continuous derivative for all the humidity range (e.g. Wu and Smith, 1992) or by retaining only cloud observations corresponding to humidity values greater than the condensation switching value, and thus computing the observation contribution to the 3D-Var cost function in relative humidity space instead of cloud fraction observation space (Renshaw, 2007).

The procedure chosen for assimilating CloudSat observations consists of several steps through which the received echo powers measured by the Cloud Profiling Radar are converted into profiles of cloud fraction, then a thinning procedure is applied to those data in order to reduce the number of total observations and minimise the horizontal correlation of observation errors. Cloud fraction data are used by a Bayesian analysis, which also takes advantage of the *a priori* knowledge of the atmosphere via a 6 hours forecast (background) from previous assimilation and forecasts cycle to retrieve pseudo-observations of relative humidity. The pseudo-observations are subsequently used in the three-dimensional variational assimilation

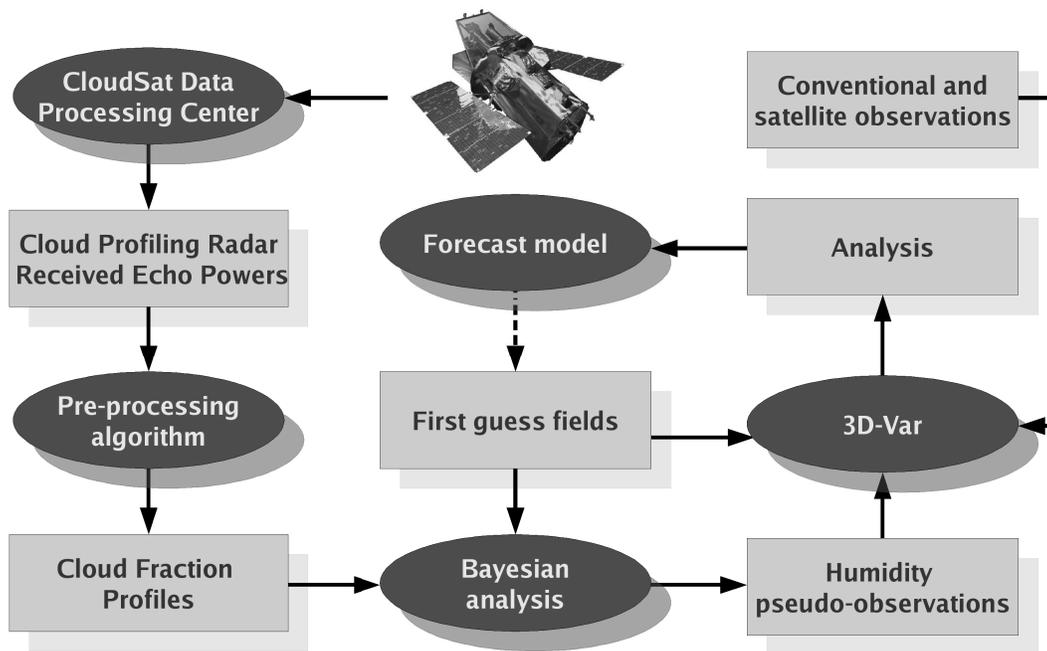


Figure 4.1: Schematic illustration of data sources, tasks and outputs performed during the assimilation algorithm. Square boxes represent data (observations and first guess fields) used in the scheme, while oval-shaped boxes represent processing and computational steps.

system, together with conventional and other remotely sensed observations. The Bayesian analysis that generates pseudo-observations of humidity assumes no spatial correlation among CloudSat observations neither in the vertical, nor along the satellite track, thus performing humidity retrievals layer by layer. The vertical and horizontal thinning does however take the spatial distribution of observations into account. Note that the background is used twice during the assimilation algorithm, which leads to humidity pseudo-profiles with errors theoretically correlated with the background field errors. This is common practise when a 1D-Var + (3)4D-Var scheme is implemented (e.g. Janisková et al., 2002a). Figure 4.1 shows schematically the flow-chart of the assimilation steps as described above. The illustration does not comprise the thinning procedure and the quality control task which take place in between the pre-processing algorithm and 3D-Var.

Though the control vector of the 3D-Var minimisation is formed by vorticity, divergence, temperature, specific humidity and logarithm of surface pressure, all in spectral space, pseudo-observations derived from CloudSat are assimilated as relative humidity data. This choice resides on the observation operator (introduced in section 4.3) which is based on a cloud cover-relative humidity relationship, and also reflects the will of using the full three-dimensional background-error covariances of the 3D-Var system to divide the pseudo-observation information into temperature

and specific humidity contributions.

4.3 The Bayesian analysis

In data assimilation (Bayesian Recursive Estimation), a model of the probability distribution of the state of the atmosphere given the observations and an *a priori* estimate of the atmosphere (i.e. a short-range forecast) is used to identify the analysis that gives the best mean squared error verification scores for the resulting forecasts. Assuming for simplicity that the forecast model error is unbiased, and that the analysis error is linearly propagated forward in time by the forecast model, the problem can be reformulated as a minimum variance analysis problem.

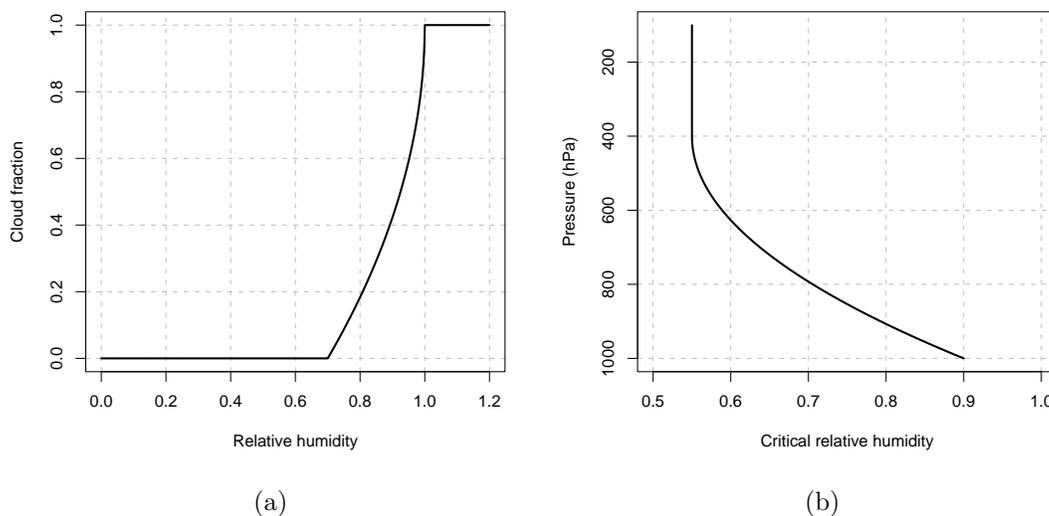


Figure 4.2: Cloud fraction - relative humidity relationship adopted in the Bayesian analysis: (a) Cloud fraction values as function of humidity for case $r_{cr} = 0.7$ (800 hPa); (b) vertical profile of r_{cr} .

According to Bayesian decision theory, the minimum variance analysis, \mathbf{x}^a , is given by the expected state of the atmosphere,

$$\mathbf{x}^a = \mathbb{E} [\mathbf{x}^t | \mathbf{y} \cap \mathbf{x}^b] \quad (4.1)$$

where \mathbf{x}^t is the true state of the atmosphere, \mathbf{y} is the observation vector, \mathbf{x}^b is the *a priori* estimate of the atmospheric state, namely the background⁹, and $\mathbb{E}[\dots]$ is the expectation operator. Conventional observations and the background can be modelled with Normal observation errors. In this case, and according to Lorenc (1986), the expected state of the atmosphere in Eq. (4.1) corresponds to the most

⁹As recommended by Ide et al. (1997), we use the superscript to distinguish between background, analysis and true state of the atmosphere.

probable state of the atmosphere, which also can be found by minimising the cost function

$$J = -\ln(\mathbf{p}(\mathbf{x}^t | \mathbf{y} \cap \mathbf{x}^b)).$$

This cost function approach is commonly referred to as variational data assimilation. Note that when the observation error is very non-Normal, the variational data assimilation approach can not be used directly since the minimisation algorithms rely on a quadratic cost function (i.e. a Normal PDF) to converge.

4.3.1 Theory behind the Bayesian analysis

The cloud fraction observation violates a very basic assumption in variational data assimilation systems, namely that the probability density function (PDF) for the observation given the true state of the atmosphere is a Normal PDF. One way of negotiating the cloud fraction observations into variational data assimilation systems is by reformulating cloud observations into pseudo-observations based on the cloud fraction observations,

$$y_p \equiv E[\mathbf{x}^t | k \cap \mathbf{x}^b] = \int \mathbf{x}^t \mathbf{p}(\mathbf{x}^t | k \cap \mathbf{x}^b) d\mathbf{x}^t \quad (4.2)$$

where k is a single cloud fraction observation that has the value 0 (no-cloud) or 1 (cloud), and then assimilating these pseudo-observations. Next we wish to express $\mathbf{p}(\mathbf{x}^t | k \cap \mathbf{x}^b)$ in terms of the probability distribution for having the observation given the true state of the atmosphere, $\mathbf{p}(k | \mathbf{x}^t)$, and the probability distribution for having the true state of the atmosphere given the background, $\mathbf{p}(\mathbf{x}^t | \mathbf{x}^b)$.

We use a simplified model for the probability of having the true state given the first guess, $\mathbf{p}(\mathbf{x}^t | \mathbf{x}^b)$, that only depends on the state variables in the observation operator, and we assume that this distribution is Normal,

$$\mathbf{p}(\mathbf{x}^t | \mathbf{x}^b) \sim \exp\left[-\frac{1}{2}(\mathbf{x}^b - \mathbf{x}^t)^\top \mathbf{B}^{-1}(\mathbf{x}^b - \mathbf{x}^t)\right],$$

where \mathbf{B} is the background-error covariance matrix. If we assume that the observation- and the background-error are independent, i.e.

$$\mathbf{p}(k \cap \mathbf{x}^b | \mathbf{x}^t) = \mathbf{p}(k | \mathbf{x}^t) \mathbf{p}(\mathbf{x}^b | \mathbf{x}^t),$$

the PDF of k conditioned to \mathbf{x}^b and \mathbf{x}^t does not depend on \mathbf{x}^b ,

$$\begin{aligned}
\mathbf{p}(k | \mathbf{x}^b \cap \mathbf{x}^t) &= \frac{\mathbf{p}(k \cap \mathbf{x}^b \cap \mathbf{x}^t)}{\mathbf{p}(\mathbf{x}^b \cap \mathbf{x}^t)} \\
&= \frac{\mathbf{p}(k \cap \mathbf{x}^b | \mathbf{x}^t) \mathbf{p}(\mathbf{x}^t)}{\mathbf{p}(\mathbf{x}^b | \mathbf{x}^t) \mathbf{p}(\mathbf{x}^t)} \\
&= \frac{\mathbf{p}(k | \mathbf{x}^t) \mathbf{p}(\mathbf{x}^b | \mathbf{x}^t)}{\mathbf{p}(\mathbf{x}^b | \mathbf{x}^t)} \\
&= \mathbf{p}(k | \mathbf{x}^t),
\end{aligned}$$

where the Bayes Theorem has been used. Applying again the Bayes Theorem and using the previous result yields

$$\begin{aligned}
\mathbf{p}(\mathbf{x}^t | k \cap \mathbf{x}^b) &= \frac{\mathbf{p}(\mathbf{x}^t \cap k \cap \mathbf{x}^b)}{\mathbf{p}(k \cap \mathbf{x}^b)} \\
&= \frac{\mathbf{p}(k | \mathbf{x}^b \cap \mathbf{x}^t) \mathbf{p}(\mathbf{x}^t \cap \mathbf{x}^b)}{\mathbf{p}(k \cap \mathbf{x}^b)} \\
&= \frac{\mathbf{p}(k | \mathbf{x}^t) \mathbf{p}(\mathbf{x}^t \cap \mathbf{x}^b)}{\mathbf{p}(k \cap \mathbf{x}^b)} \\
&= \frac{\mathbf{p}(k | \mathbf{x}^t) \mathbf{p}(\mathbf{x}^t | \mathbf{x}^b)}{\mathbf{p}(k | \mathbf{x}^b)}.
\end{aligned}$$

The denominator in the expression above does not depend on the true state of the atmosphere, being merely a normalisation factor. Note that if it is difficult to find any relationship between the observation and the model state, then $\mathbf{p}(k | \mathbf{x}^t)$ becomes almost flat (constant). In this case we see that the pseudo-observation is equal to the background, and the observation provides no additional information to the analysis.

We define the observation operator $H(\mathbf{x}^t) \equiv \mathbf{p}(k = 1 | \mathbf{x}^t)$ and the complement $\mathbf{p}(k = 0 | \mathbf{x}^t) = 1 - H(\mathbf{x}^t)$. In doing this, we implicitly assume that the cloud cover diagnosed by the observation operator represents the probability of having a cloud, and this is commonly adopted in almost all cloudiness formulations (see e.g. Tompkins, 2005 for a review of cloud cover parameterizations). If the probability distribution $\mathbf{p}(\mathbf{x}^t | k \cap \mathbf{x}^b)$ only depends on a limited number of state variables, Eq. (4.2) can be explicitly solved. Note that there is a linear relationship between the pseudo-observation and the state of the atmosphere. The pseudo-observation is therefore quite suitable for use in variational data assimilation if it can be associated with a Normal observation error.

Inspired by the continuous function for $H(\mathbf{x}^t)$, one can construct a non-binary cloud fraction (CF) observation by defining this as a spatial average of n independent

neighbouring binary cloud observations, k_i , i.e.

$$CF = \frac{1}{n} \sum_{i=1}^n k_i. \quad (4.3)$$

Using the observation operator definition above gives

$$\begin{aligned} \mathbf{p} \left(\sum_{i=1}^n k_i \mid \mathbf{x}^t \right) &= \binom{n}{\sum k_i} \mathbf{p} \left(\bigcap_{i=1}^n k_i \mid \mathbf{x}^t \right) \\ &= \binom{n}{\sum k_i} [H(\mathbf{x}^t)]^{\sum k_i} [1 - H(\mathbf{x}^t)]^{\sum(1-k_i)}, \\ \mathbf{p} \left(\mathbf{x}^t \mid \mathbf{x}^b \cap \sum_{i=1}^n k_i \right) &= \frac{[H(\mathbf{x}^t)]^{\sum k_i} [1 - H(\mathbf{x}^t)]^{n - \sum k_i} \mathbf{p}(\mathbf{x}^t \mid \mathbf{x}^b)}{\int [H(\mathbf{x}^t)]^{\sum k_i} [1 - H(\mathbf{x}^t)]^{n - \sum k_i} \mathbf{p}(\mathbf{x}^t \mid \mathbf{x}^b) d\mathbf{x}^t} \end{aligned} \quad (4.4)$$

where the omitted summation limits are from $i = 1$ to n . Note the binomial coefficient $\binom{n}{\sum k_i}$ which briefly appears in the derivation. In practise, considering only one observation (i.e. $n = 1$ so that $\sum_{i=1}^n k_i = CF$), the previous equation provides an explicit formula for the computation of the PDF for the true state of the atmosphere conditioned to the cloud fraction CF and to the background. Note finally that having n *equal* independent cloud observations corresponds to having a single cloud observation that is very accurate. We observe in this case that the “effective” observation operator, $[H(\mathbf{x}^t)]^n$, is more “sharp” than a single observation operator, $H(\mathbf{x}^t)$, as indicated earlier, namely that the uncertainty in the binary observation is accounted for by the observation operator itself.

4.3.2 The cloud fraction observation operator

The observation operator $H(\mathbf{x}^t)$ is used to compute the probability of having a cloud given the “true” value of the control vector. We have chosen a simple diagnostic expression for $H(\mathbf{x}^t)$ which depends on the relative humidity r , as in the large-scale condensation scheme developed by Sundqvist et al. (1988), summarised by

$$H(\mathbf{x}^t) = 1 - \sqrt{\frac{1-r}{1-r_{cr}}} \quad (4.5)$$

where r_{cr} is the relative humidity threshold which represents the switching value for the clouds to exist. The reason behind the choice of Eq. (4.5) resides on the need of using a simple and cheap diagnostic scheme for cloud cover, involving as few model parameters as possible; the use of more sophisticated schemes is however possible and should be investigated in the future. Similar formulations of the cloud fraction relationship have been extensively used as diagnostics expressions in several global models (by Slingo, 1987 for the ECMWF global model, and by Janisková et al.,

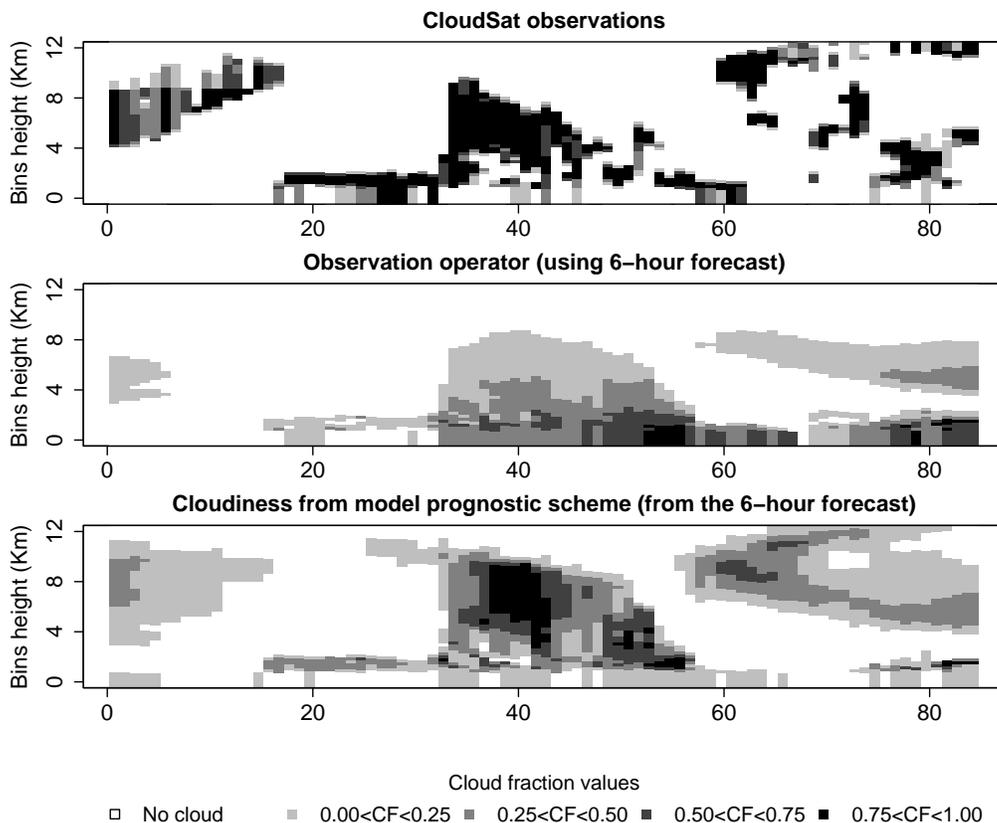


Figure 4.3: Example of cloud fraction intercomparison between the CloudSat observed cloud fraction (top panel), the Sundqvist diagnostic scheme (middle panel) applied to a 6-hour forecast, and the prognostic microphysical scheme (Lopez, 2002) (bottom panel) from a 6-hour forecast. The Sundqvist diagnostic scheme is used as observation operator for computing the humidity pseudo-observations.

2002*b* later for assimilation purposes; by Smith, 1990 for the UKMO Unified Model); Wood and Field (2000) showed that such parameterizations are in good agreements with aircraft data. Recently, Tompkins and Janisková (2004) proposed a large-scale condensation scheme for data assimilation purposes that uses the same equation for the cloud fraction as in Eq. (4.5) with an additional parameter to take into account the vertically-varying standard deviation for the r_{cr} parameter.

Estimation of the r_{cr} threshold is a delicate task; the existence of such a threshold originates from the statistical meaning of cloud fraction in an NWP model, where clouds-related processes are typically sub-grid ones, and the grid-point humidity is the cloud fraction averaged sum of a saturated air moisture content and an unsaturated one. As shown by Lopez (2002) and Tompkins (2003) through the study of temperature and humidity subgrid variances from aircraft in-situ observations, the humidity threshold is therefore strongly dependent on the model resolution, as

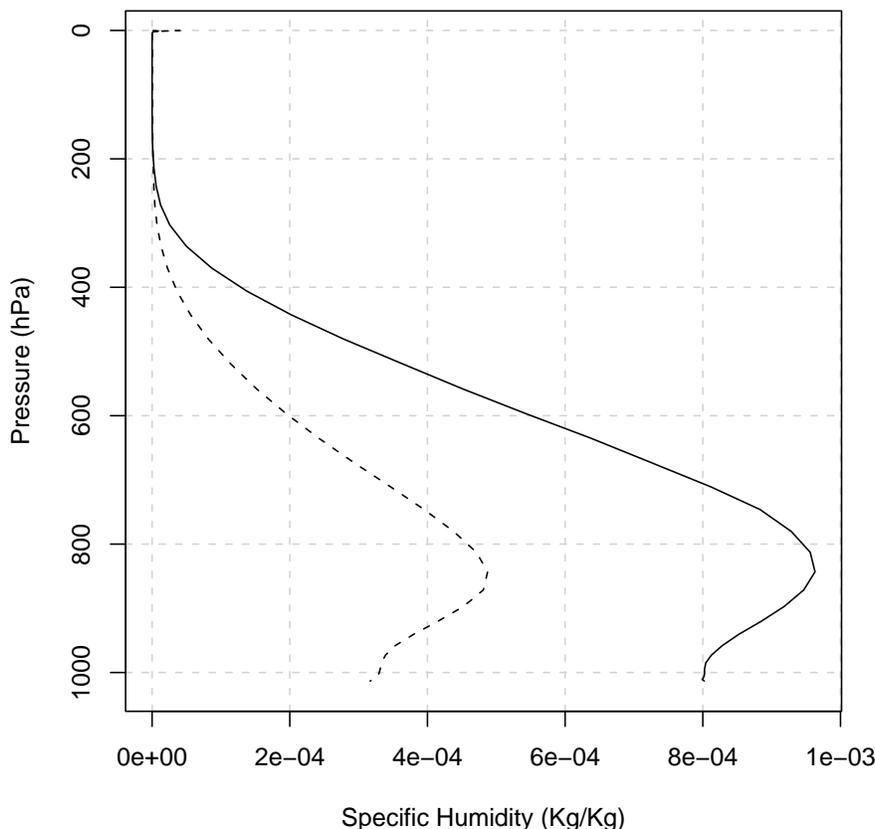


Figure 4.4: Profile of background-error standard deviation of specific humidity for use in the Bayesian analysis. Full line refers to a summer three-month dataset (JJA); dashed line refers to a winter threemonth dataset (DJF).

it should equal 1 at cloud-resolving scale by definition and decrease with the resolution. Seasonal variability does also affect the value for what regards convective activity and ice-liquid separation of cloud water. Finally, several observations campaigns (Walcek, 1994 and Teixeira, 2001) showed the vertical dependence of r_{cr} , that typically reaches a minimum in the middle atmosphere, where clouds can form even for relative humidity values smaller than 60%. Our choice of r_{cr} has been empirically tuned to match as close as possible the model cloud fraction prognostic scheme (Lopez, 2002), following also suggestions in Bouteloup et al. (2005) for model physics implications at high troposphere, leading to the expression:

$$r_{cr} = \begin{cases} 0.55 & \text{if } p < 400hPa \\ 9.72 \cdot 10^{-7}p^2 - 7.78 \cdot 10^{-4}p + 7.06 \cdot 10^{-1} & \text{if } p > 400hPa \end{cases}$$

with p being pressure in hPa. Sundqvist et al. (1988) also took into account oro-

graphic effects, which are not considered here, by decreasing r_{cr} over land by empirical small values. It is interesting to note that Norris and da Silva (2007) assimilated satellite cloud observations not to correct humidity fields but to adaptively tune the critical relative humidity values in the large-scale condensation parametrisation. The observation operator at a fixed level and the r_{cr} definition are plotted in Figure 4.2.

Cloud fraction profiles calculated from the diagnostic model in Eq. (4.5) have been extensively compared with the forecast model prognostic scheme and with CloudSat observations, and show a good agreement especially in terms of cloud location. Nevertheless, the diagnostic model, because of its statistical nature, underestimates the cloud cover, if compared against for instance the radar observations, which measure a punctual quantity and are strongly binary. An example of inter-comparisons between cloud observations, the forecast model prognostic cloud cover and the diagnosed cloud cover is given in Figure 4.3.

4.4 Use of CloudSat observations

The CPR aboard CloudSat is a 94-GHz nadir-pointing radar. It therefore provides cross-sections of return power along the satellite track. Clothiaux et al. (1995) showed that cloud radar operating at 94 GHz are able to detect almost all types of atmospheric hydrometeors excepted very high and thin clouds (cirrus) and has some attenuation problems when precipitating clouds are present.

4.4.1 Pre-processing of CloudSat data

The cloud detection procedure is a simplified algorithm from Clothiaux et al. (1998), for which cloud fraction is calculated from the ratio between the net return power and the estimated standard deviation of noise power, and it reads

$$CF = \begin{cases} 1 & \text{if } P_T - P_R - P_N > k_1\sigma_n \\ \frac{P_T - P_R - P_N - k_0\sigma_n}{(k_1 - k_0)\sigma_n} & \text{if } k_0\sigma_n < P_T - P_R - P_N < k_1\sigma_n \\ 0 & \text{if } P_T - P_R - P_N < k_0\sigma_n \end{cases} \quad (4.6)$$

where P_T is the transmitted radar power, P_R is the received echo power (return power), P_N and σ_n are respectively the mean and the standard deviation of the noise power, estimated over a number of stratospheric bins assumed to be hydrometeors-free (Marchand et al., 2008). k_1 and k_0 are empirically tuned parameters. The previous algorithm has been validated against Level-2 CloudSat products, which also uses a spatial box-filter to improve the hydrometeors detection (Marchand et al., 2008), and against combined CloudSat-Calipso-MODIS products provided by the

CloudSat Data Processing Center (CDPS) of the Cooperative Institute for Research in the Atmosphere (CIARA). The cloud fraction detection algorithm also comprises a surface clutter filter, that takes use of a CDPS estimated flat surface reflectivity profiles (Marchand et al., 2008) with a joint control among neighbouring radar pixels when the satellite profile refers to complicated orography. Despite the received echo powers delivery delay, the cloud product generated by Eq. (4.6) runs operationally at Met.no providing additional validation information for on-duty forecasters.

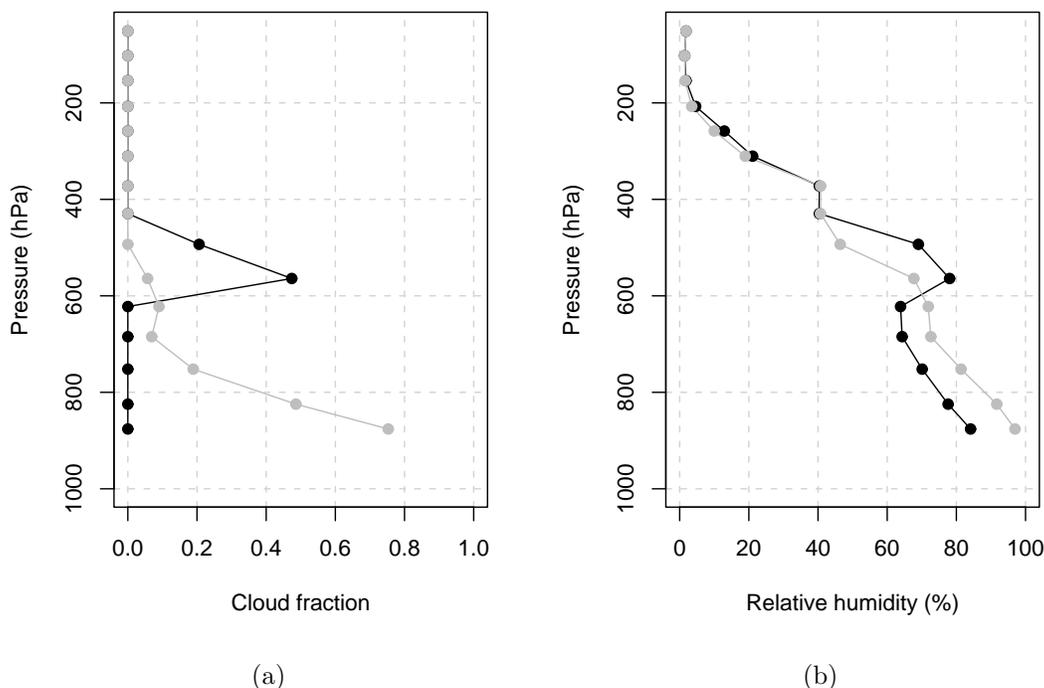


Figure 4.5: Example of humidity pseudo-observations from the Bayesian analysis of CloudSat cloud fraction. Left panel: comparison between first guess (grey line) and CloudSat-derived observations (black line) in cloud fraction space; right panel: relative humidity profiles of first guess (grey line) and pseudo-observations after the Bayesian analysis (black line).

Cloud cover data derived from Eq. (4.6) are spatially averaged in order to make the observation resolution consistent with the one of the model, as CloudSat scans the atmosphere at an approximate resolution of 2.5 Km along the satellite track, while the model resolution is of 11 Km. This procedure also reproduces the statistical average of cloud fraction in a NWP model, where cloud-related processes are typically sub-grid phenomena. Equations (4.2) and (4.4) are applied to CloudSat cloud cover data to retrieve humidity pseudo-observations, assuming only one binary observation ($\sum_{i=1}^n k_i = CF$). In using Eq. (4.4), background errors are assumed vertically and horizontally uncorrelated, and humidity errors standard deviation are

computed through the NMC method (Parrish and Derber, 1992) using 48 - 24 hours forecasts differences for two three-months periods relative to winter and summer seasons (Figure 4.4); error values for intermediate months are interpolated.

Quality check and vertical thinning of humidity pseudo-observations retain only one observations each 50 hPa on the vertical, and which does not deviate more than 0.95 from the first guess in cloud fraction space or 3 times the background-error standard deviation in humidity space. Figure 4.5 shows an example of Bayesian analysis, in particular in the situation of dramatic discrepancy between observed and background cloud profiles; on the left panel a comparison between CloudSat cloud fractions and the background-diagnosed cloud cover is presented, while on the right panel the analysed profile of relative humidity is plotted together with the background profile. Note that in the middle and high atmosphere (in the figure between 200 and 300 hPa), inconsistencies between the different cloud cover schemes in the forecast model and in the Bayesian analysis lead to unjustified analysis increments; however, these situations refers to observations minus guess differences equal to 0 in cloud cover space, and pseudo-observations of humidity are rejected in such a case. This approach also reduces redundancy of the background information in 3D-Var. Further, CloudSat data above 200 hPa are not used in order to avoid problems rising from the possible missed detection of very high and thin clouds, which as mentioned earlier is a well-known weakness in cloud detection based on 94-GHz radars.

4.4.2 Errors specification for use in 3D-Var

Humidity pseudo-observations derived from CloudSat are suitable for direct assimilation in variational assimilation systems. Observation errors for such observations are assumed spatially uncorrelated and are computed following a Monte Carlo approach. The analysis is computed for different values of simulated background and CloudSat cloud fractions perturbed with an unbiased Normal error. The background standard deviation has been set equal to the NMC-derived values. The CloudSat cloud cover observation error is calculated from Eq. (4.6) through propagation of covariances, assuming that cloud fraction errors are only caused by the uncertainty in the power noise estimation:

$$\sigma_{CF}^2 = \sigma_{P_N}^2 \left(\frac{\partial CF}{\partial P_N} \right)^2 + \sigma_{P_R}^2 \left(\frac{\partial CF}{\partial P_R} \right)^2 + \sigma_{P_T}^2 \left(\frac{\partial CF}{\partial P_T} \right)^2 \simeq \sigma_{P_N}^2 \left(\frac{\partial CF}{\partial P_N} \right)^2 \simeq 0.044,$$

σ^2 being the variance. An example of an error profile is given in Figure 4.6 for a clear-sky background profile, as a function of pressure and CloudSat cloud cover observation. It shows that the closer the observed cloud fraction is to 0, the larger

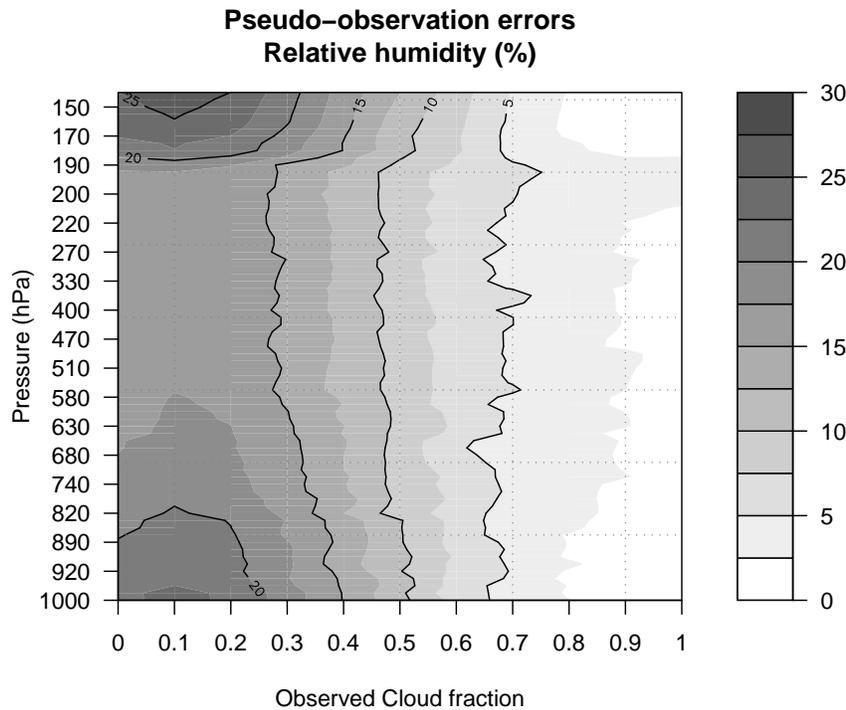


Figure 4.6: Humidity pseudo-observations errors (relative humidity, %) from Monte Carlo simulations as function of pressure and observed cloud fraction. The background profile refers to a mid-summer clear-sky atmosphere.

the error is, because the bigger the uncertainty in the humidity pseudo-observations is in the absence of clouds, which can allow the humidity to range from 0 to the critical threshold. Note that the use of the Monte Carlo method is feasible because humidity pseudo-observations can be explicitly computed. Assuming that cloud fraction errors only depend on the estimation of noise power standard deviation and considering Eq. (4.6) as a “perfect” algorithm leads in general to an underestimation of pseudo-observation errors.

4.4.3 Single observation experiments

The effects of CloudSat pseudo-observation assimilation are initially studied in single observation experiments. Two cases are compared: i) a non-cloudy observation in combination with a cloudy background, leading to a negative humidity analysis increment (EXPSO1); ii) a cloudy observation in combination with a clear-sky background, leading to a positive humidity analysis increment (EXPSO2). Table 4.1 reports observed values of cloud fractions, innovations and residuals¹⁰ at observation location for both the experiments, while Figure 4.7 shows the cross-sections of the analysis increments for the two single observation experiments, for tempera-

¹⁰Analysis increments in observation space.

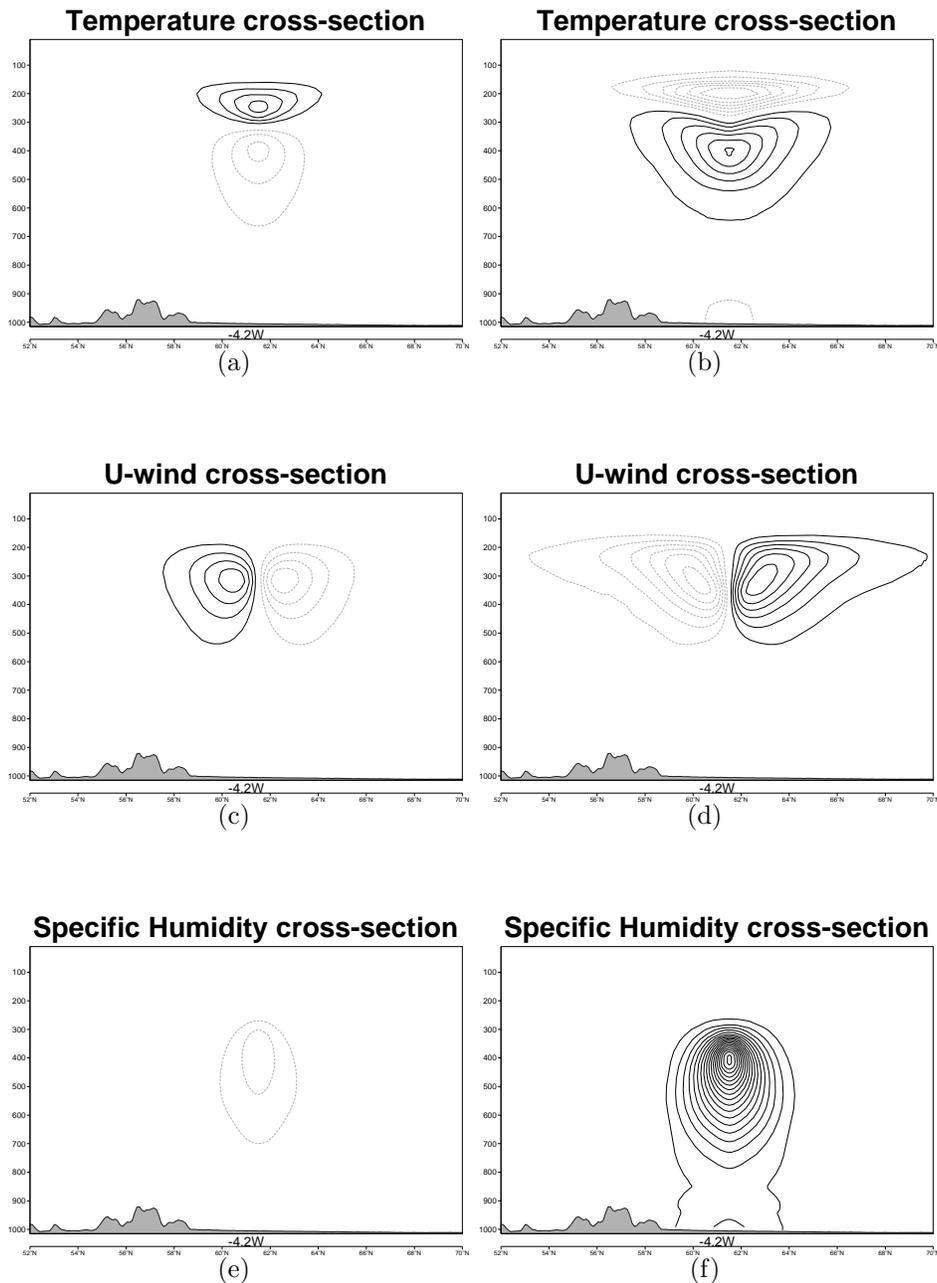


Figure 4.7: Single observation experiments: cross-sections of analysis increments along the meridian correspondent to the observation location (-4.2 W) for 9 degrees latitude in both north and south direction: (a) Temperature for EXP1; (b) Temperature for EXP2; (c) U-wind for EXP1; (d) U-wind for EXP2; (e) Specific Humidity for EXP1; (f) Specific Humidity for EXP2. Contours are plotted every 0.05 K, 0.10 m s $^{-1}$ and 0.025 g Kg $^{-1}$ for temperature, u-wind and specific humidity respectively; full lines correspond to positive increments while dashed lines to negative increments. Vertical axis is the pressure in hPa.

ture, u-component of horizontal wind, and specific humidity. One can notice that according to the humidity retrieval error definition, the closer the observation is to a “no cloud” event, the smaller the impact is on the analysis. This is an important

	EXPSO1	EXPSO2
Pressure (hPa)	324.32	399.20
Observed Cloud Fraction	0.103	0.528
Background Rel. Hum. (%)	88.01	31.95
Background Cloud Fraction	0.484	0.000
Rel. Hum. Pseudo-observation (%)	71.43	70.70
Rel. Hum. Pseudo-observations Error (%)	17.05	8.07
Rel. Hum. Innovations (%)	-16.58	38.75
Rel. Hum. Analysis Increment (%)	-12.83	35.16

Table 4.1: Single observation experiments: pressure, CloudSat-observed cloud fraction, background values, pseudo-observation from the Bayesian analysis and observation minus guess and observation minus analysis within the 3D-Var.

feature of our approach: though all the cloud information are used, a very different and asymmetric weight is given to them, depending primarily on the fact that a non-cloudy observation contains less information than a cloudy observation, or, in other words, an observation which corresponds to moist-unsaturated air is less informative than an observation corresponding to moist-saturated air. The background constraint spreads the observation minus background increments into the fields of the analysed variables (vorticity, divergence, temperature and specific humidity) and it is important to notice that the pseudo-observations directly modify the specific humidity and temperature fields, and indirectly, the wind fields (Berre, 2000). In particular, the background-error covariance matrix (not shown here), computed via the “NMC-method”, has a strong negative vertical correlation of temperature throughout the troposphere, while the temperature-humidity coupling shows a negative cross-correlation in the very low atmosphere and a positive cross-correlation from 900 to 500 hPa and negligible above. As an important consequence, humidity retrievals in the high atmosphere impact temperature fields above, humidity retrievals in the low atmosphere have a large impact near the surface. The wind circulation is always modified.

4.5 Impact of the humidity pseudo-observations on analysis and forecasts

The variational assimilation system used for testing the assimilation of the humidity pseudo-observations is the three-dimensional spectral limited area variational assimilation system ALADIN 3D-Var (Sadiki and Fischer, 2005; Fischer et al., 2005). In the Met.no quasi-operational configuration, the assimilation step is performed every 6 hours, consisting of a) updating the sea surface temperature (SST) through the ECMWF SST analysis, b) performing the surface Optimal Interpolation analysis, c) performing the upper-air spectral three-dimensional variational assimilation.

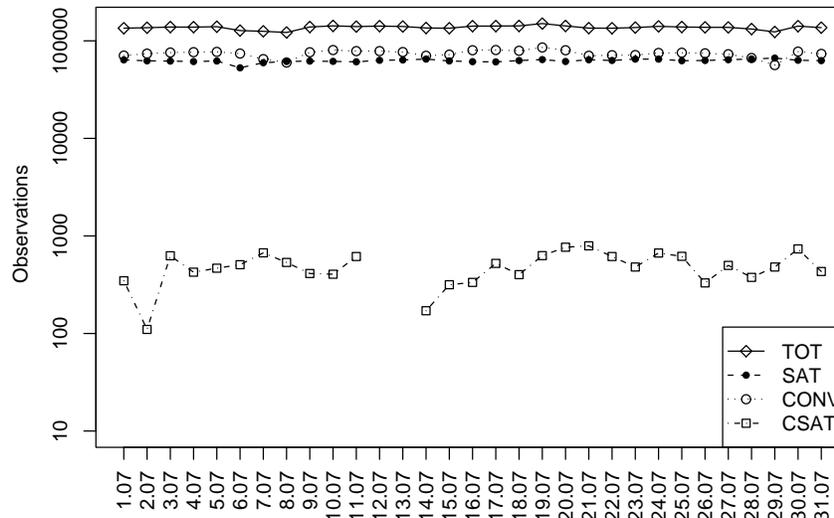


Figure 4.8: Daily amount of observations used in the assimilation for July 2007, in logarithmic scale. TOT is the total number of observations; SAT corresponds to satellite radiances; CONV to conventional observations; CSAT to CloudSat pseudo-observations.

Observations assimilated in the upper-air analysis are all the conventional (land and ship SYNOP reports, wind profilers, radiosondes, buoys), aircraft observations, Atmospheric Motion Vectors (AMV) from MSG satellites, AMSU-A and AMSU-B radiances from the POES satellites and AMSU-A and MHS from Metop. The computational domain has a resolution of 11 Km and extends for 4455x2970 km (405x270 nodes). Vertical discretization consists of 60 hybrid eta levels. A background quality check is performed over all the observations, while bias correction is performed for all the satellite radiances following the air-mass scheme developed by Harris and Kelly (2001), and for radiosonde daytime observations of temperature according to ECMWF coefficients. Background-error statistics have been derived via the NMC method applied to three-months dataset (20070601 to 20070831) of 48 minus 24 hours forecast differences. This configuration represents the Reference experiment (NoCloudSat), against which the CloudSat experiment is compared, which adds the CloudSat-derived humidity pseudo-observations to the reference set of observations. Both the experiments have been carried out for a one-month period, from 20070701 to 20070731.

The daily amount of observations used in 3D-Var for the CloudSat experiment is shown in Figure 4.8. Note that CloudSat-derived humidity observations are very few, 0.34% (about 200 to 300 per day) of all the observations during the July 2007 experiment, while all the satellite radiances (45.62%) are in a number comparable with the conventional, aircraft-borne and AMV observations (54.04%). An inter-

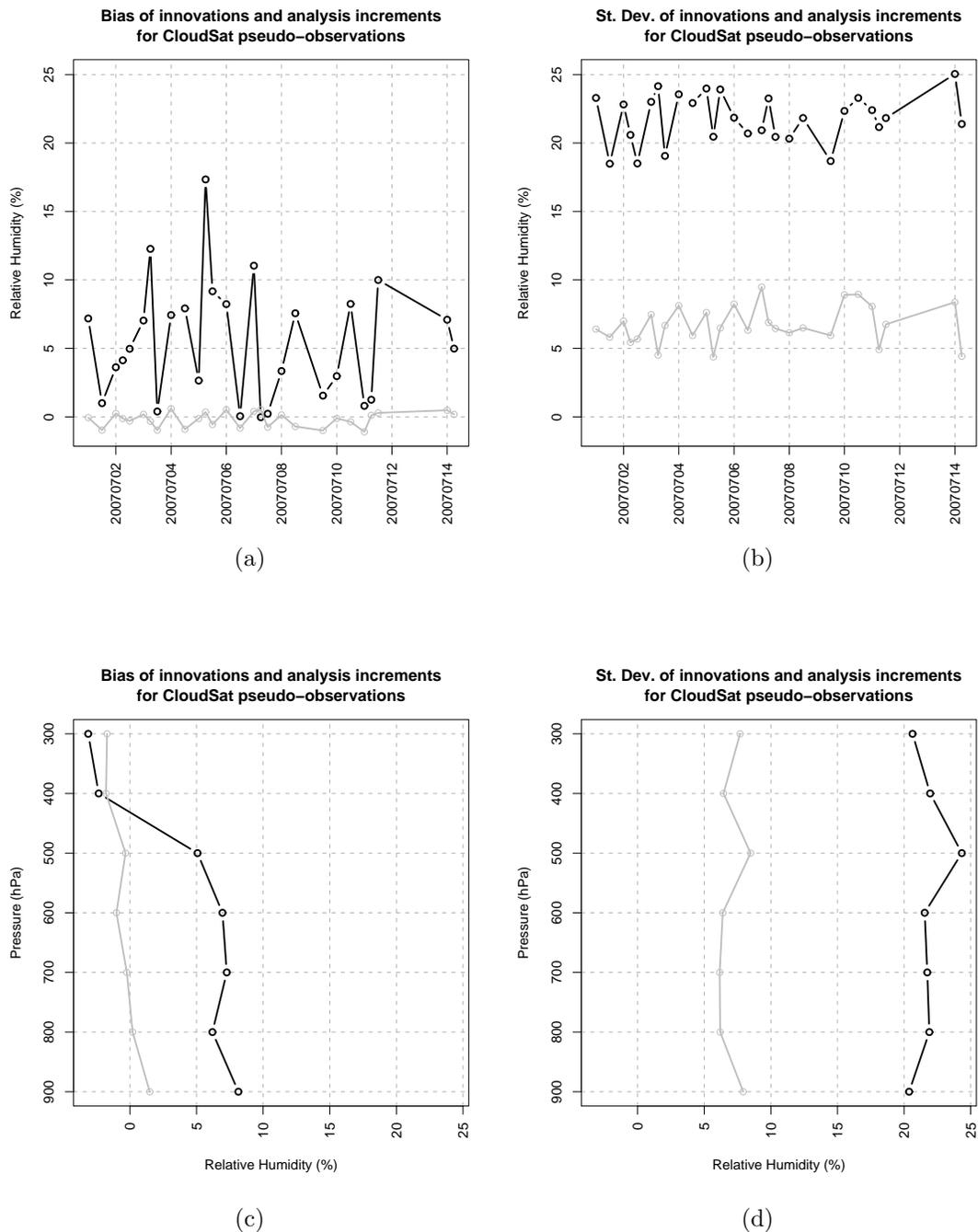


Figure 4.9: Innovations (observation minus first guess, in black) and observation minus analysis (in grey) statistics for CloudSat observations: (a) bias timeseries; (b) standard deviation timeseries; (c) vertical profile of bias; (d) vertical profile of standard deviation.

ruption period in the production of pseudo-observations is present between 12th and 13th July, due to problems in CloudSat data reception.

Observation minus first guess and observation minus analysis statistics (Figure 4.9) show that CloudSat analysis increments are unbiased. The weight of CloudSat observations is very large in the analysis, since the observation minus first guess

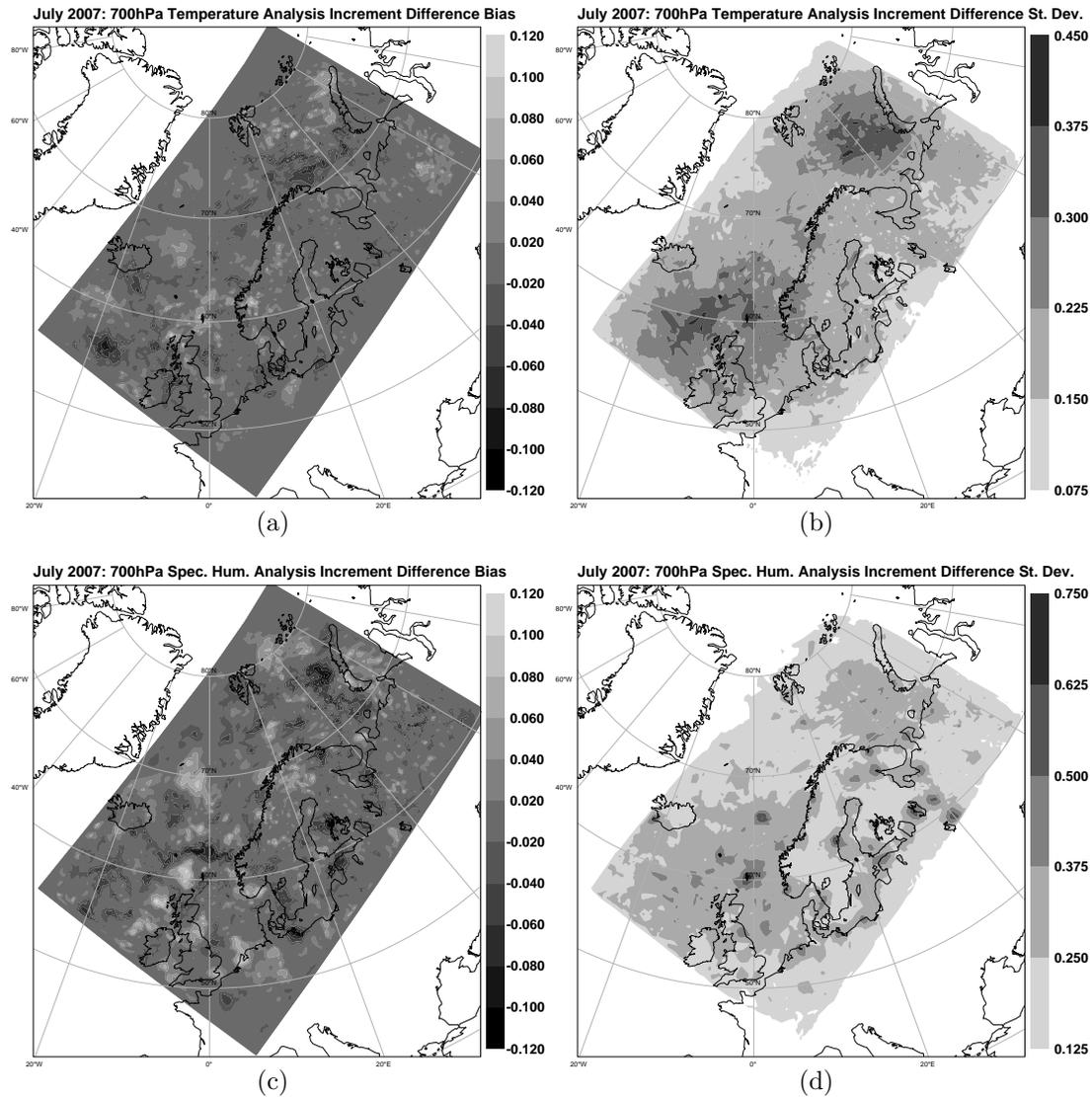


Figure 4.10: Bias and standard deviation of analysis increments (analysis minus first guess) of CloudSat experiment minus NoCloudSat experiment at 700 hPa: (a) bias of differences of temperature increments; (b) standard deviation of differences of analysis increments; (c) bias of differences of specific humidity increments; (d) standard deviation of differences of specific humidity increments.

standard deviation is much larger than the observation minus analysis standard deviation. CloudSat analysis increments for the one-month experimental period are also found to make a major impact on data-sparse areas. Figure 4.10 shows the monthly-averaged bias and standard deviation of the differences between the analysis increments of the CloudSat experiment minus the analysis increments of the NoCloudSat experiment at 700 hPa. They represent two indexes for identifying the most-affected areas by the CloudSat-derived observations. It is clear that areas over sea, especially in areas corresponding to the Norwegian Sea and in the Northwest part of the LAM domain are particularly sensitive to the new observations, with

an overall unbiased behaviour over the domain. It is not possible to say whether the humidity pseudo-observations in general have a drying or wetting effect on the analysis.

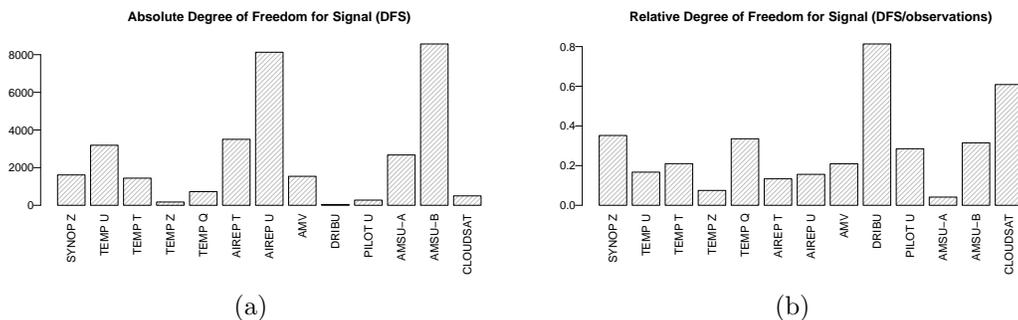


Figure 4.11: Degrees of freedom for signal grouped by observation type and parameter: absolute (a) and relative (b).

Degrees of Freedom for Signal (DFS, see e.g. Cardinali et al., 2004), that are defined as the derivative of the analysis (increments) in observation space with respect to the observations, have been computed through a randomisation technique (Chapnik et al., 2006) that reads:

$$DFS = (\tilde{\mathbf{y}} - \mathbf{y}) \mathbf{R}^{-1} (\mathbf{H}(\tilde{\mathbf{x}}^a - \mathbf{x}^b) - \mathbf{H}(\mathbf{x}^a - \mathbf{x}^b)) \quad (4.7)$$

where $\tilde{\mathbf{y}}$ is the vector of perturbed observations, $\tilde{\mathbf{x}}^a$ is the analysis from perturbed observations, \mathbf{R} is the observational error covariance matrix and \mathbf{H} is the tangent-linear version of the observation operator. The perturbations are performed using an unbiased random error whose standard deviation equals the observation error; 6 assimilation cycles, 4 days distant each other, have been rerun with perturbed observations. Equation (4.7) can be applied to any observation subset. DFS have been calculated with the aim of studying the impact of CloudSat data relatively to the other observation types in the assimilation system; however, DFS shows the self-sensitivity of the observation (sensitivity at observation location), without providing information about the spatial correlation between the analysis and the observations. DFS indicate the importance of the observations in terms of weight, efficacy of the observation operator, and amount of observations.

Results from the DFS computation are shown in Figure 4.11, together with the Relative Degrees of Freedom for Signal, that is, DFS divided by the number of observations. Globally, CloudSat data are not very important in the assimilation system, given the small amount of observation, but their relative weight is comparable with the humidity measurements from the radiosonde network. Relative DFS, which however are meaningless in terms of real impact, show that the humid-

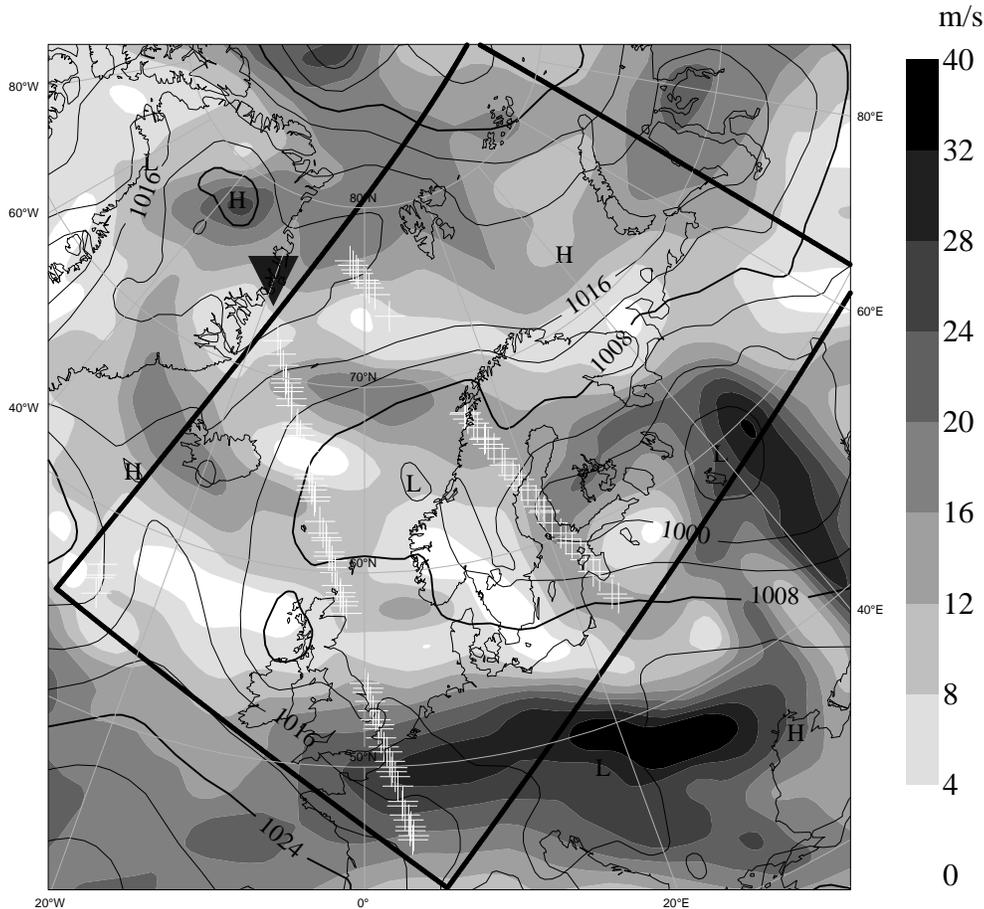
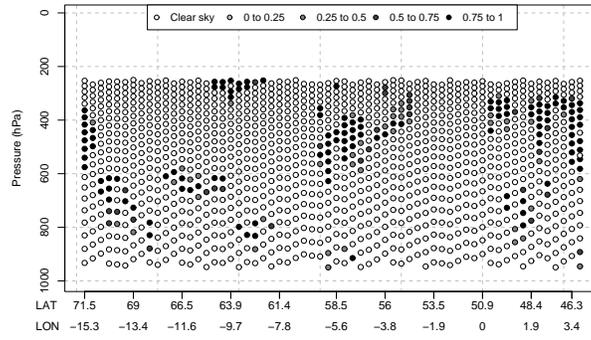


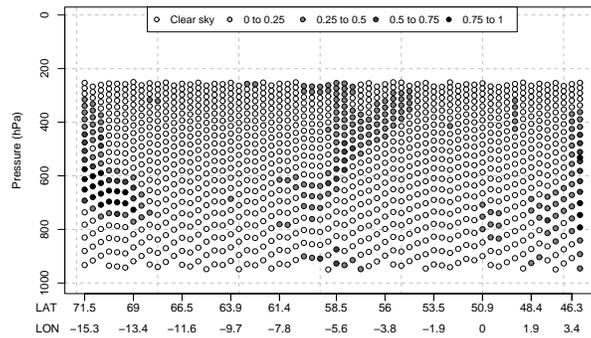
Figure 4.12: Synoptic situation and CloudSat data coverage for the mid-summer case-study (20070709 at 12 UTC). Shaded contours show the wind intensity at 500 hPa (in m s⁻¹), black-line contours show the mean-sea level pressure, white crosses are in correspondence of the humidity pseudo-observations from CloudSat data after background quality check. The black frame indicates the domain extension.

ity pseudo-observations (indicated as CloudSat in the Figure) are the second most important after buoys observations, and this result seems to be primarily related to the availability of the humidity pseudo-observations in data-sparse regions (i.e. over sea).

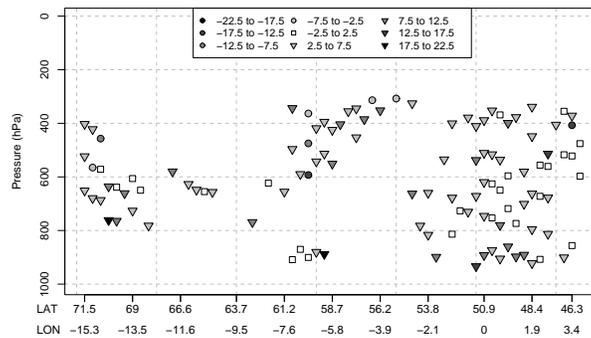
The available humidity pseudo-observations are very few and their impact in terms of modifications of meteorological development cannot easily be identified. The effect of adding CloudSat-derived pseudo-observations is here briefly illustrated for a mid-summer case-study (20070709), which is selected because of a relatively high amount of CloudSat profiles compared to other assimilation windows. In the 6-hours assimilation window centred at 12 UTC, CloudSat passed inside the Norwegian ALADIN domain 3 times. After the background quality check, 412 single-level observations are retained. In Figure 4.12 the profiles containing valid observations are plotted over a map representing the synoptic situation (where contours of 500 hPa wind intensity and mean-sea level pressure are plotted). The pseudo-observations



(a)



(b)



(c)

Figure 4.13: Cross-section of analysis increments due to the CloudSat pseudo-observations for the illustrative case. Analysis increments (bottom panel) are plotted together with cloud fractions detected by CloudSat (top panel) and cloud fractions computed from the first guess (middle panel) using the Sundqvist large-scale condensation scheme. The geographical extension of the cross section coincides with the longest satellite track in Figure 10, indicated there by the solid black triangle, from the eastern coast of Greenland and passing along the eastern cost of the British island, in North-South direction according to the satellite track direction.

analysis increments in observation space (relative humidity) are shown in Figure 4.13 for the central profile of Figure 4.12, together with the CloudSat-derived and the model-diagnosed cloud fractions. It is possible to distinguish a low-level humidity increase in the northern (left) part of the cross-section accompanied by a drying effect of CloudSat observations in the middle part of the atmosphere. The effect of the pseudo-observations on the forecast (Figure 4.14) is evident mainly over sea and after 6 hours of forecasts, they have caused an important gradient of both temperature and moisture visible in the relative humidity analysis increments in the northern part of the cross-section, which in turn lead to a downward heat flux.

It is also interesting to notice that in connection with the vertical profile approximately at (10W, 64N) a few humidity pseudo-observations with positive observation minus first guess relative humidity at about 850 hPa increased the moisture content significantly and caused an important increase of temperature above the layer, because of the negative vertical correlations between specific humidity and temperature, and the negative vertical auto-correlations of temperature errors. Due to the balances induced by the background constraint, a temperature gradient evolved upwards, changing also the wind fields in the high atmosphere (not shown here).

The differences between the control and the experimental forecasts become less important and more noisy after +24 hours, and at +48 hours they are hard to distinguish. This loss-of-memory of the forecast system is a well-known feature of the impact of the humidity observing system (see e.g. Bengtsson and Hodges, 2005) and can be studied through the temporal evolution of the bias between the experiment with CloudSat humidity retrievals and the one without. Figure 4.15 clearly indicates that such a bias is gradually lost after the forecast integration, and becomes very small at +48 hours for temperature, while very rapidly vanishes for cloud fraction (which is anyway a prognostic variable in the forecast model). This is true also for specific humidity, apart from a positive bias in the middle atmosphere after 48 hours of forecasts.

Verification of surface pressure against SYNOP stations has been carried out for all the 00 and 12 UTC 48-hour forecasts from 20070705 to 20070731 (first five days are considered as a warming-up period). Impact of CloudSat is slightly positive on the average (Fig. 4.16) for the surface pressure, while the root mean square error timeseries indicate that there are no significant improvements during extreme weather events. Skin temperature (not shown), which is in general sensitive to the cloud fraction because of the change in the downwelling radiation, is not significantly affected by the CloudSat observations.

Verification against radiosondes (Fig. 4.17) shows an effective improvement of forecasts with respect to geopotential and wind fields, while mass fields have better scores in the middle atmosphere and worse scores in the lower atmosphere.

Such a result can partially be explained by the fact that the simplified cloud cover observation operator does not comprise parameterizations of convective processes that are all the more important in the very low part of the atmosphere. According to our experience, the long background vertical auto- and cross- correlations derived through the “NMC” method may also be cause of unjustified analysis increments near the surface.

4.6 Summary and perspectives

Within a three-dimensional variational data assimilation system, we have developed a method to assimilate cloud observations from CloudSat as humidity pseudo-observations. These pseudo-observations are the product of a Bayesian analysis based on a short-range forecast and the cloud or no-cloud observations. The Bayesian decision theory provides an attractive statistical framework for exploiting binary observations, otherwise difficult to use in variational data assimilation systems. A major strength of this approach is the possibility to assimilate also clear-sky occurrences, which turn out to have a smaller weight in the variational data assimilation. The formulation presented in this Chapter is very general, which implies a potential extension in the future to other instruments that supply cloudiness observations, like for instance, ground-based cloud-radars, cloud products from nowcasting tools, and in principle also any other observation of an on/off meteorological process. Cloud products from satellite infrared sensors are also of potential use in this approach.

The methodology has been applied to received echo powers from the 94-GHz radar aboard CloudSat, the first spaceborne radar devoted to microphysical studies. CloudSat provides very high-quality and high-resolution data for cloud detection, and a simple algorithm is applied to the radar powers to diagnose profiles of cloud fractions. These cloud observations show a very binary behaviour. Unfortunately CloudSat cannot be used operationally because of the delay in data dissemination. One conclusion from this work is that spaceborne cloud radar should be encouraged in designing next-generation of polar-orbiting satellites. Extending the radar viewing geometry to off-nadir profiles to increase data coverage would also be of obvious benefit to Numerical Weather Prediction purposes.

A re-tuned large-scale condensation scheme based on a cloud cover-relative humidity relationship has been used to compute cloud fraction values at observation location from the background within the Bayesian analysis. Despite its simplicity compared to many advanced microphysical schemes that are currently implemented in medium-range forecast models, this condensation scheme is able to sufficiently reproduce the cloud fields in terms of both vertical structure and horizontal exten-

sion.

The results of the assimilation trials are positive, especially for the improvement of verification scores of dynamical parameters (wind, geopotential). For relative humidity and temperature, scores against independent radiosondes show an improvement in the middle atmosphere, a weak impact in the high atmosphere and a negative impact in the lower part of the atmosphere. The large-scale condensation scheme may be inappropriate in the regions of the atmosphere characterised by important convective activity. This weakness can be overcome in the future by implementing a more advanced cloud scheme.

Despite the small amount of CloudSat data, we find that the impact of the pseudo-observations is significant in data-sparse area, like over the sea, and that the self-sensitivity of the analysis with respect to CloudSat data is comparable with radiosonde measurements of humidity. However, we did not find important improvements during severe weather events and it is consequently difficult to show a real case where CloudSat observations played an important role in the analysis and forecast system. The forecast system seems to have a short memory of humidity pseudo-observations, so changes to humidity fields vanish soon.

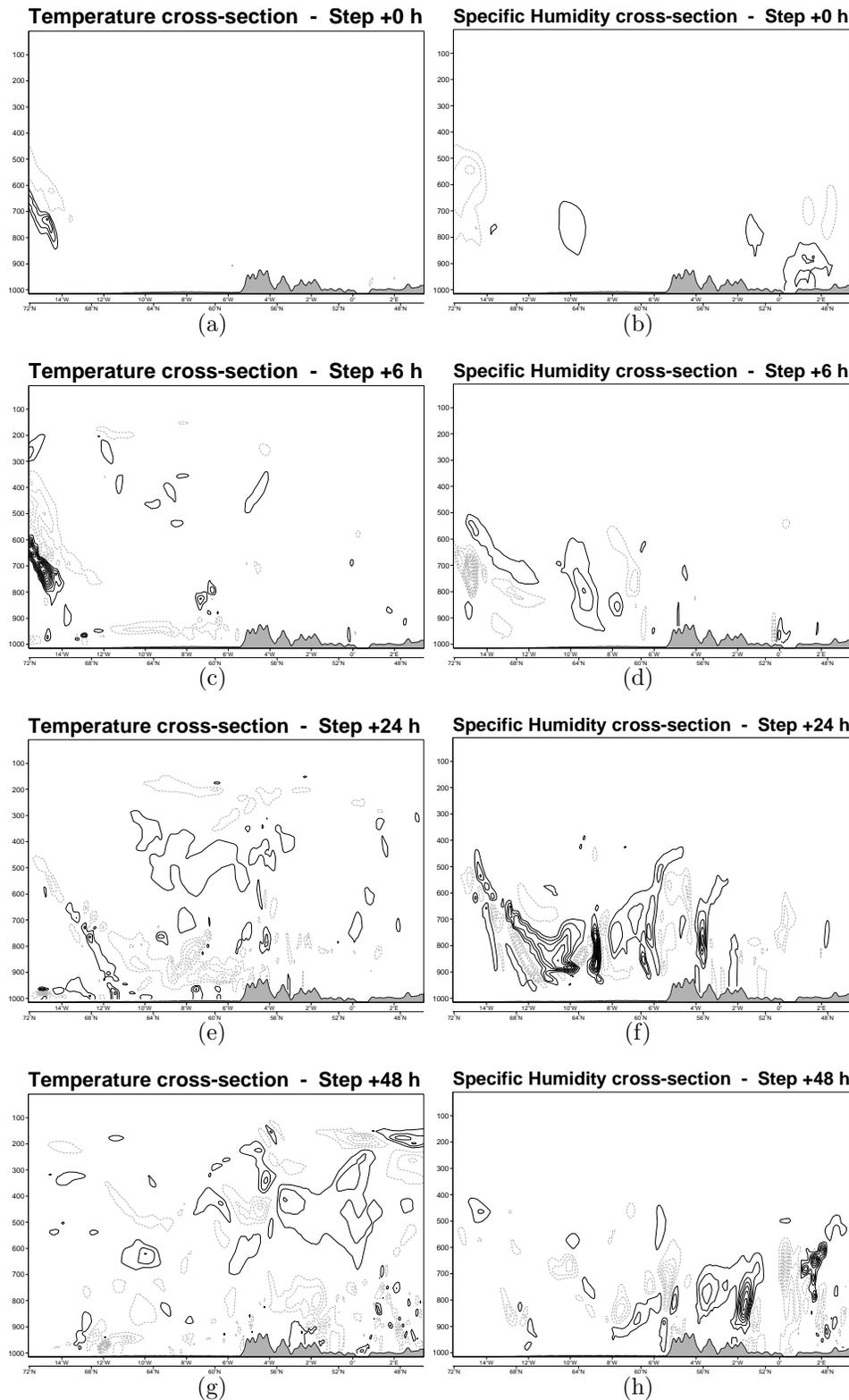
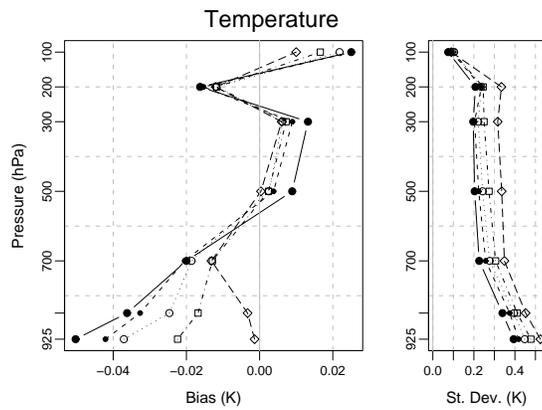
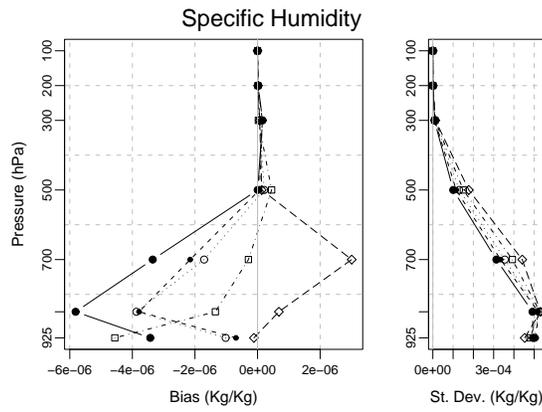


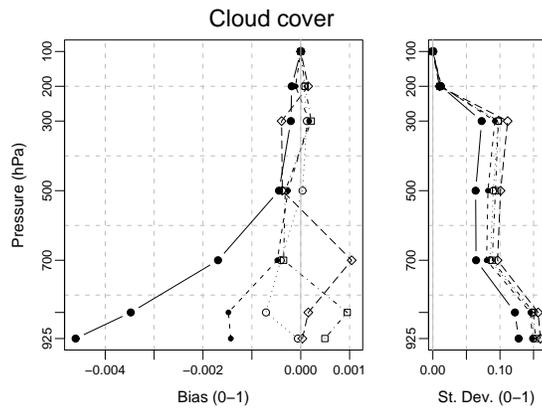
Figure 4.14: Cross-sections of difference of CloudSat experiment forecasts minus No-CloudSat experiment forecasts for the case-study of Figure 11: temperature (a, c, e, g) and specific humidity (b, d, f, h) at analysis time (a, b) and at +06 (c, d) +24 (e, f) and + 48 (g, h) hours since analysis time. Full lines refer to positive differences while dashed lines to negative differences. Vertical axis is the pressure in hPa. Contours are plotted every 0.3 K and 0.3 g Kg⁻¹ for temperature and specific humidity respectively.



(a)



(b)



(c)

Figure 4.15: Temporal evolution of forecast biases and standard deviation of CloudSat experiment forecasts minus NoCloudSat experiment forecasts for the entire simulation period: temperature (a), specific humidity (b) and cloud fraction (c). For each parameter bias and standard deviation are plotted in the left and right panel respectively. Full lines with black solid circles refer to analysis time; dashed lines with solid bullets to 6-hour statistics; dotted lines with circles to 12-hour statistics; dash-dotted lines with squares to 24-hour statistics; long-dashed lines with diamonds to 48-hour statistics.

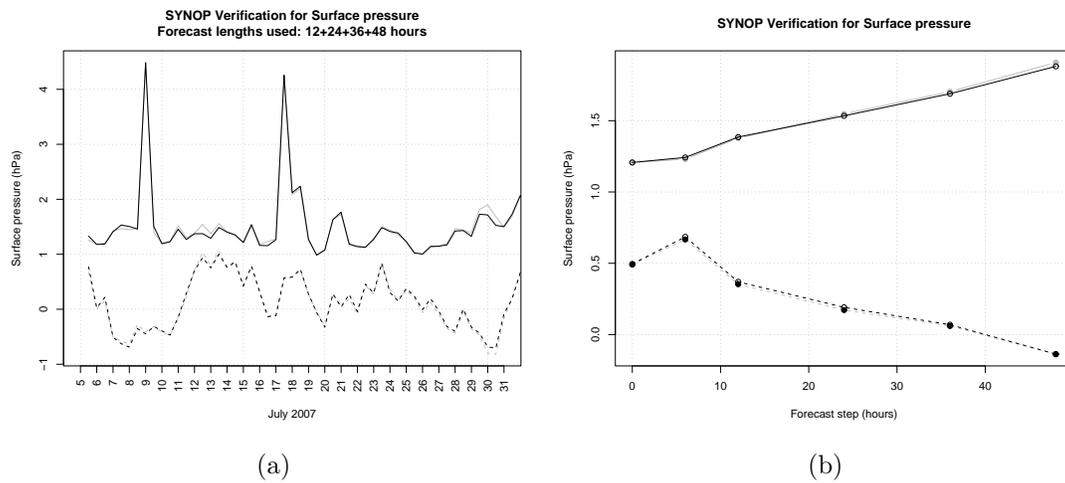


Figure 4.16: Verification against SYNOP stations surface pressure: (a) RMSE and bias timeseries and (b) RMSE and bias against forecast length (right panel). Black lines refer to the CloudSat experiment while grey lines to the NoCloudSat experiment; full lines denote RMSE and dashed lines denote bias.

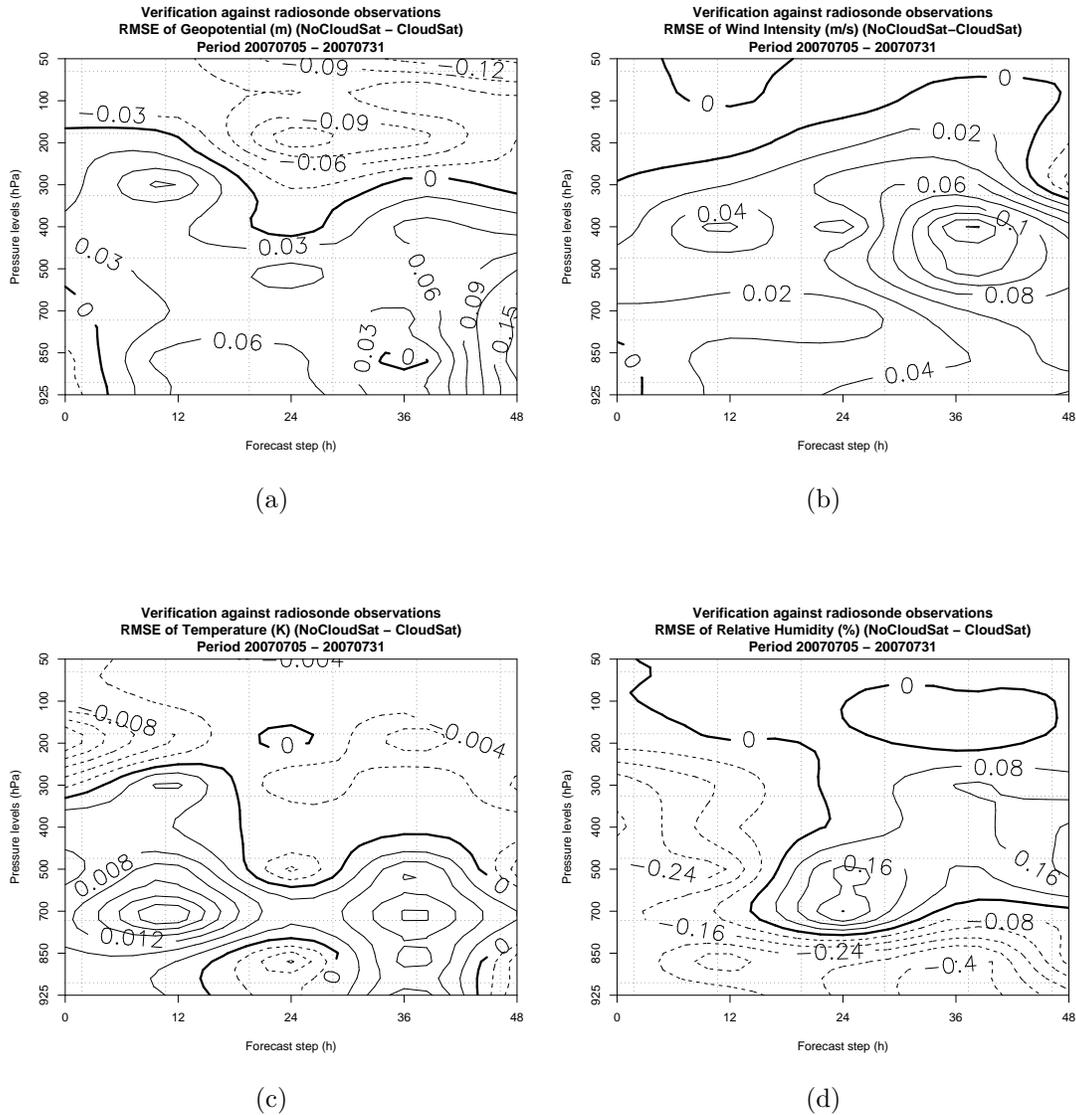


Figure 4.17: Radiosonde verification skill scores: contour of differences between RMSE of reference experiment (NoCloudSat) minus RMSE of CloudSat Experiments for (a) geopotential; (b) wind intensity; (c) temperature; (d) relative humidity. Full lines denote positive values, dashed lines denote negative values and the tick full line corresponds to the zero contour line. Positive values indicate that CloudSat assimilation decreases the root mean square error. Contour intervals are 0.03 m, 0.02 m s⁻¹, 0.004 K and 0.08 % for geopotential, wind intensity, temperature and relative humidity respectively.

References

- ANDERSSON, E., E. HÓLM, P. BAUER, A. BELJAARS, G. KELLY, A. MCNALLY, A. SIMMONS, J.-N. THÉPAUT, AND A. TOMPKINS, 2007. Analysis and forecast impact of the main humidity observing systems. *Q. J. R. Meteorol. Soc.*, **133**, pp. 1473–1485.
- BABB, D., J. VERLINDE, AND B. ALBRECHT, 1998. Retrieval of cloud microphysics parameters from 94-GHz radar doppler power spectra. *J. Atmos. Ocean. Technol.*, **16**, pp. 489–503.
- BENEDETTI, A., G. STEPHENS, AND T. VUKIĆEVIĆ, 2003. Variational assimilation of radar reflectivities in a cirrus model. Part I: Model description and adjoint sensitivity studies. *Q. J. R. Meteorol. Soc.*, **129**, pp. 277–300.
- BENGTSSON, L. AND K. HODGES, 2005. On the impact of humidity observations in numerical weather prediction. *Tellus*, **57A**, pp. 701–708.
- BERRE, L., 2000. Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Mon. Wea. Rev.*, **128**, pp. 644–667.
- BOUTELOUP, F., F. BOUYSSSEL, AND P. MARQUET., 2005. Improvements of Lopez’s prognostic large scale cloud and precipitation scheme. ALADIN Newsletter 28, ALATNET.
- BRUNEL, P. AND S. TURNER, 2003. On the use of Planck-weighted transmittances in RTTOV. In International TOVS Study Conference. ITSC.
- CARDINALI, C., S. PEZZULLI, AND E. ANDERSSON, 2004. Influence-matrix diagnostic of a data assimilation system. *Q. J. R. Meteorol. Soc.*, **130**, pp. 2767–2786.
- CHAPNIK, B., G. DESROZIERS, F. RABIER, AND O. TALAGRAND, 2006. Diagnosis and tuning of observational error in a quasi-operational data assimilation setting. *Q. J. R. Meteorol. Soc.*, **132**, pp. 543–565.
- CHEVALLIER, F., P. BAUER, J.-F. MAHFOUF, AND J.-J. MORCRETTE, 2002. Variational retrieval of cloud profile from ATOVS observations. *Q. J. R. Meteorol. Soc.*, **128**, pp. 2511–2525.
- CHEVALLIER, F., P. LOPEZ, A. TOMPKINS, M. JANISKOVÁ, AND E. MOREAU, 2004. The capability of 4D-Var systems to assimilate cloud-affected satellite infrared radiances. *Q. J. R. Meteorol. Soc.*, **130**, pp. 917–932.

- CLOTHIAUX, E., G. MACE, T. ACKERMAN, T. KANE, J. SPINHIRNE, AND V. SCOTT, 1998. An automated algorithm for detection of hydrometeor returns in micropulse lidar data. *J. Atmos. Ocean. Technol.*, **15**, pp. 1035–1042.
- CLOTHIAUX, E., M. MILLER, B. ALBRECHT, T. ACKERMAN, J. VERLINDE, D. BABB, R. PETERS, AND W. SYRETT, 1995. An evaluation of a 94-GHz radar for remote sensing of cloud properties. *J. Atmos. Ocean. Technol.*, **12**, pp. 201–229.
- DESROZIERS, G., L. BERRE, B. CHAPNIK, AND P. POLI, 2005*a*. Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc.*, **131**, pp. 3385–3396.
- DESROZIERS, G., P. BROUSSEAU, AND B. CHAPNIK, 2005*b*. Use of randomization to diagnose the impact of observations on analyses and forecasts. *Q. J. R. Meteorol. Soc.*, **131**, pp. 2821–2837.
- FISCHER, C., T. MONTMERLE, L. BERRE, L. AUGER, AND S. STEFANESCU, 2005. An overview of the variational assimilation in the ALADIN/France NWP system. *Q. J. R. Meteorol. Soc.*, **131**, pp. 3477–3492.
- GOLDING, B., 1998. Nimrod: A system for generating automated very short range forecasts. *Meteorological Applications*, **5**, pp. 1–16.
- GUIDARD, V., C. FISCHER, M. NURET, AND A. DZIEDZIC, 2006. Evaluation of the ALADIN 3D-Var with observations of the MAP campaign. *Meteorol. Atmos. Phys.*, **92**, pp. 161–173.
- GUSTAFSSON, N., L. BERRE, S. HÖRNQUIST, X.-Y. HUANG, M. LINDSKOG, B. NAVASCUÉS, K. MOGENSEN, AND S. THORSTEINSSON, 2001. Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint. *Tellus*, **53A**, pp. 425–446.
- HARRIS, B. A. AND G. KELLY, 2001. A satellite radiance-bias correction scheme for data assimilation. *Q. J. R. Meteorol. Soc.*, **127**, pp. 1453–1468.
- IDE, K., P. COURTIER, M. GHIL, AND A. LORENC, 1997. Unified notation for data assimilation: operational, sequential and variational. *J. Meteorol. Soc. Japan*, **75**, pp. 181–189.
- ILLINGWORTH, A. J., R. J. HOGAN, E. J. O’CONNOR, D. BOUNIOL, M. E. BROOKS, J. DELANOË, D. P. DONOVAN, J. D. EASTMENT, N. GAUSSIAT, J. W. F. GODDARD, M. HAEFFELIN, H. K. BALTINK, O. A. KRASNOV, J. PELON, J.-M. PIRIOU, A. PROTAT, H. W. J. RUSSCHENBERG, A. SEIFERT,

- A. M. TOMPKINS, G.-J. VAN ZADELHOFF, F. VINIT, U. WILLÉN, D. R. WILSON, AND C. L. WRENCH, 2007. Cloudnet, Continuous evaluation of cloud profiles in seven operational models using ground-based observations. *Bull. Amer. Meteor. Soc.*, **88**, pp. 883–898.
- ISAKSEN, L., 2007. Use of analysis ensembles in estimating flow-dependent background error variances. In Workshop on Flow-dependent aspects of data assimilation. ECMWF Reading (UK).
- JAKOB, C., 1999. Evaluation of cloud cover in the ECMWF reanalysis. *J. Climate*, **12**, pp. 947–959.
- JANISKOVÁ, M., J.-F. MAHFOUF, AND J.-J. MORCRETTE, 2002*a*. Preliminary studies on the variational assimilation of cloud-radiation observations. *Q. J. Roy. Meteor. Soc.*, **128**, pp. 2713–2736.
- JANISKOVÁ, M., J.-F. MAHFOUF, J.-J. MORCRETTE, AND F. CHEVALLIER, 2002*b*. Linearized radiation and cloud schemes in the ECMWF model: Development and evaluation. *Q. J. Roy. Meteor. Soc.*, **128**, pp. 1505–1527.
- KOLLIAS, P., E. CLOTHIAUX, M. MILLER, B. ALBRECHT, G. STEPHENS, AND T. ACKERMAN, 2007. Millimeter-wavelength radars: new frontier in atmospheric cloud and precipitation research. *Bull. Amer. Meteor. Soc.*, **88**, pp. 1608–1624.
- LI, L., S. DURDEN, AND S. TANELLI, 2007. Level 1 B CPR Process Description and Interface Control Document. Tech. Rep. No D-20308, Jet Propulsion Laboratory.
- LOPEZ, P., 2002. Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes. *Q. J. R. Meteorol. Soc.*, **128**, pp. 229–257.
- LOPEZ, P., A. BENEDETTI, P. BAUER, M. JANISKOVÁ, AND M. KÖHLER, 2006. Experimental 2D-Var assimilation of ARM cloud and precipitation observations. *Q. J. R. Meteorol. Soc.*, **132**, pp. 1325–1347.
- LORENC, A., 1986. Analysis methods for numerical weather prediction. *Q. J. Roy. Meteor. Soc.*, **112**, pp. 1177–1194.
- LOWE, P. R., 1977. An approximating polynomial for the computation of saturation vapor pressure. *J. Appl. Meteor.*, **16**, pp. 100–103.
- MACPHERSON, B., B. WRIGHT, W. HAND, AND A. MAYCOCK, 1996. The Impact of MOPS Moisture Data in the U.K. Meteorological Office Mesoscale Data Assimilation Scheme. *Mon. Wea. Rev.*, **124**, pp. 1746–1766.

- MARCHAND, R., G. MACE, T. ACKERMAN, AND G. STEPHENS, 2008. Hydrometeor detection using CloudSat - an Earth orbiting 94 GHz cloud radar. *J. Atmos. Ocean. Technol.*, **25**, pp. 519–533.
- MCNALLY, A., 2002. A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Q. J. R. Meteorol. Soc.*, **128**, pp. 2551–2556.
- MENEGHINI, R. AND T. KOZU, 1990. Spaceborne Weather Radar. *Artechhouse*. 199 pp.
- NORRIS, P. AND A. DA SILVA, 2007. Assimilation of satellite cloud data into the GMAO finite-volume data assimilation system using a parameter estimation method. Part I: Motivation and algorithm description. *J. Atmos. Sci.*, **64**, pp. 3880–3895.
- PARRISH, D. AND J. DERBER, 1992. The National Meteorological Center’s spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, pp. 1747–1763.
- PAZMANY, A., J. MEAD, R. MCINTOSH, M. HERVIG, R. KELLY, AND G. VALI, 1994. 95-GHz polarimetric radar measurements of orographic cap clouds. *J. Atmos. Ocean. Technol.*, **11**, pp. 140–153.
- POLI, P., P. MOLL, F. RABIER, G. DESROZIERS, B. CHAPNIK, L. BERRE, S. HEALY, E. ANDERSSON, AND F.-Z. EL GUELAI, 2007. Forecast impact studies of zenith total delay data from European near real-time GPS stations in Météo France 4DVAR. *J. Geophys. Res.*, **112**, pp. 644–667.
- RENSHAW, R., 2007. Assimilation of Cloud Fraction in Var at the Met Office. In SRNWP Workshop on High resolution data assimilation. 21–23 March 2007, Norrköping, Sweden.
- SADIKI, W. AND C. FISCHER, 2005. A posteriori validation applied to the 3DVar ARPEGE and ALADIN data assimilation systems. *Tellus*, **57A**, pp. 21–34.
- SLINGO, J. M., 1987. The development and verification of a cloud prediction scheme for the ECMWF model. *Q. J. Roy. Meteor. Soc.*, **113**, pp. 899–927.
- SMITH, R., 1990. A scheme for predicting layer clouds and their water content in a general circulation model. *Q. J. R. Meteorol. Soc.*, **116**, pp. 435–460.
- STEPHENS, G., D. VANE, R. BOAIN, G. MACE, K. SASSEN, Z. WANG, A. ILLINGWORTH, E. O’CONNOR, W. ROSSOW, S. DURDEN, S. MILLER, R. AUSTIN, A. BENEDETTI, C. MITRESCU, AND T. C. S. TEAM, 2002. The

- CloudSat mission and the A-Train, A new dimension of space-based observations of clouds and precipitation. *Bull. Amer. Meteor. Soc.*, **83**, pp. 1771–1790.
- STORTO, A. AND F. TVETER, 2006. Theory for assimilating binary observations. Research Note 11/2006, Norwegian Meteorological Institute.
- SUNDQVIST, H., 1978. A parameterization scheme for non-convective condensation including prediction of cloud water content. *Q. J. Roy. Meteor. Soc.*, **104**, pp. 677–690.
- SUNDQVIST, H., E. BERGE, AND J. KRISTJÁNSSON, 1988. Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Mon. Wea. Rev.*, **117**, pp. 1641–1657.
- TEGEN, I., P. HOLLRIG, M. CHIN, I. FUNG, D. JACOB, AND J. PENNER, 1997. Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results. *J. Geophys. Res.*, **102**, pp. 23895–23915.
- TEIXEIRA, J., 2001. Cloud Fraction and Relative Humidity in a Prognostic Cloud Fraction Scheme. *Mon. Wea. Rev.*, **129**, pp. 1750–1753.
- TIEDTKE, M., 1993. Representation of clouds in large-scale models. *Mon. Wea. Rev.*, **121**, pp. 3040–3061.
- TOMPKINS, A., 2005. The parametrization of cloud cover. Tech. memo., ECMWF Research Department.
- TOMPKINS, A. M., 2002. A Prognostic Parameterization for the Subgrid-Scale Variability of Water Vapor and Clouds in Large-Scale Models and Its Use to Diagnose Cloud Cover. *J. Atmos. Sci.*, **59**, pp. 1917–1942.
- , 2003. Impact of temperature and humidity variability on cloud cover assessed using aircraft data. *Q. J. Roy. Meteor. Soc.*, **129**, pp. 2151–2170.
- TOMPKINS, A. M. AND M. JANISKOVÁ, 2004. A cloud scheme for data assimilation. *Q. J. Roy. Meteor. Soc.*, **130**, pp. 2495–2517.
- TVETER, F., 2004. On the assimilation of binary cloud observations in HIRLAM 3D-Var. Research Note 5/2004, Norwegian Meteorological Institute.
- WALCEK, C. J., 1994. Cloud cover and its relationship to relative humidity during a springtime midlatitude cyclone. *Mon. Wea. Rev.*, **122**, pp. 1021–1035.
- WOOD, R. AND P. FIELD, 2000. Relationships between total water, condensed water, and cloud fraction in stratiform clouds examined using aircraft data. *J. Atmos. Sci.*, **57**, pp. 1888–1905.

- WRIGHT, B., 1993. The moisture observation preprocessing system (MOPS). Tech. Rep. No 38, U.K. Meteorological Office Forecasting Research Division.
- WU, X. AND W. SMITH, 1992. Assimilation of ERBE data with a nonlinear programming technique to improve cloud-cover diagnosis. *Mon. Wea. Rev.*, **120**, pp. 2009–2024.
- WWRP/WGNE, 2007. Forecast Verification. Tech. rep., Joint Working Group on Verification, World Meteorological Organization. Available at www.bom.gov.au/bmrc/wefor/staff/eee/verif.
- ZOU, X., 1997. Tangent linear and adjoint of “on-off” processes and their feasibility for use in 4-dimensional variational data assimilation. *Tellus*, **49A**, pp. 3–31.

Appendix A

Attendance of conferences, meetings and schools

- ALADIN 3D-Var Working Week at Hungarian Meteorological Service, Budapest, Hungary (June 2006).
- NetFam Summer School on “Non-hydrostatic dynamics and fine scale data assimilation”, St. Petersburg, Russia (June 2006).
- HIRLAM/ALADIN/AROME Mesoscale Working Week at FMI, Helsinki, Finland (November 2006).
- Eumetsat 2006 Fellow Day, presenting “Status of the ABC project”, Darmstadt, Germany (December 2006).
- HIRLAM 4D-Var Training Week for observation experts, presenting “Assimilating binary cloud cover data”, Norrkoping, Sweden (February 2007).
- Eumetnet Workshop on “High resolution data assimilation with emphasis on the use of moisture-related observations”, presenting “Preliminary study for assimilating binary cloud cover data from AVHRR”, Norrkoping, Sweden (March 2007).
- ECMWF Training Course on “Data assimilation and use of satellite data”, Reading, UK (April 2007).
- Eumetsat 2007 Fellow Day, presenting “Status of the ABC project”, Darmstadt, Germany (December 2007).
- 2008 Eumetsat Meteorological Satellite Conference, presenting “Assimilation of remote-sensed cloudiness observations”, Darmstadt, Germany (September 2008).
- Eumetsat 2008 Fellow Day, presenting “Status of the ABC project”, Darmstadt, Germany (December 2008).

Appendix B

Documentation and publications

- STORTO A. AND TVETER F.T., 2006. Theory for assimilating binary cloud cover, *Met.no Research Note*, **11/2006**.
- STORTO A., 2007. Preliminary study for assimilating binary cloud cover data from AVHRR radiometers. *EUMETNET/SRNWP Workshop on High resolution data assimilation*, 21 - 23 March 2007, Norrkoping, Sweden.
- STORTO A. AND LINDSKOG M., 2007. Experiences with the ALADIN 3D-VAR. *HIRLAM All-Staff Meeting/17th ALADIN Workshop*, 23 - 26 April 2007, Oslo, Norway
- STORTO A. AND RANDRIAMAMPINANINA R., 2008. Sensitivity of the ALADIN-HARMONIE/Norway analysis and forecast systems to different observations. *HIRLAM All-Staff Meeting/18th ALADIN Workshop*, 7 - 10 April 2008, Bruxelles, Belgium.
- RANDRIAMAMPINANINA R. AND STORTO A., 2008. Aladin-Harmonie/Norway and its assimilation system. *HIRLAM All-Staff Meeting/18th ALADIN Workshop*, 7 - 10 April 2008, Bruxelles, Belgium (in *HIRLAM Newsletter* **54**).
- RANDRIAMAMPINANINA R. AND STORTO A., 2008. Investigating the assimilation of IASI data in a limited area model. *XVI International TOVS Study Conference*, 7 - 13 May 2008, Angra dos Reis, Brazil.
- STORTO A. AND RANDRIAMAMPINANINA R., 2008. Use of satellite data in ALADIN/HARMONIE-Norway. *XVI International TOVS Study Conference*, 7 - 13 May 2008, Angra dos Reis, Brazil.
- STORTO A., 2008. Assimilation of cloudiness observations. *Eumetsat Meteorological Satellite Conference*, 8 - 12 September 2008, Darmstadt, Germany.
- RANDRIAMAMPINANINA R. AND STORTO A., 2008. Monitoring the use of IASI data in a limited area assimilation system. *Eumetsat Meteorological Satellite Conference*, 8 - 12 September 2008, Darmstadt, Germany.

- STORTO A. AND TVETER F.T., 2008. Assimilating Humidity Pseudo-observations Derived from the Cloud Profiling Radar Aboard CloudSat in ALADIN 3D-Var, submitted to *Meteorological Applications*.

Appendix C

Software produced

- Preprocessing of AVHRR-PPS software for comparison with model-equivalents cloudiness data.
- Processing of SEVIRI infrared radiances (cloud-masking, brightness temperature re-calibration, geographical thinning) for use in three-dimensional variational assimilation.
- Processing of CloudNet ground-based cloud-radar data for comparison with model-equivalents cloudiness data.
- Preprocessing of GPS-ZTD data for use in three-dimensional variational assimilation.
- Processing of conventional data for use in HARMONIE-3DVar.
- Processing of CloudSat data for computing humidity pseudo-observations to use in HARMONIE-3DVar.

Appendix D

Main international cooperation

- CloudSat Data Processing Center (CDPC) of the Cooperative Institute for Research in Atmosphere (CIRA).
- Météo-France/GMAP for the implementation and optimisation of the ALADIN/HARMONIE model.
- Swedish Meteorological and Hydrological Institute (SMHI), Danish Meteorological Institute (DMI), Finnish Meteorological Institute (FMI) for the HARMONIE model.