Nowcasting precipitation using variational data assimilation

J. Snoek

Cover Text
A study exploring possibilities to improve short term
numerical weather prediction of precipitation
by assimilating near real time weather observations

Nowcasting precipitation using variational data assimilation

by

J. Snoek

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday December 7, 2017 at 13:00.

Student number:	4172892	
Project duration:	June 1, 2016 – December 7, 2017	
Thesis committee:	Prof. dr. ir. H. W. J. Russchenberg,	TU Delft, supervisor
	Dr. S. R. de Roode,	TU Delft
	M. Severin,	Infoplaza
	Dr. Ir. J. A. E. ten Veldhuis,	TU Delft, external supervisor

This thesis is confidential and cannot be made public until December 31, 2017.

Abstract

Nowcasting, synonymous for short term (0-6h) forecasting of precipitation with high detail in terms of location, timing and intensity, is a dynamic field of research in weather forecasting that is ever so relevant for society in mitigating the impact of severe weather events. Ever since remote sensing platforms like radar and satellites became available to probe the atmosphere, forecasters have attempted to use the high resolution and frequent data acquired by these sensors to estimate the development of precipitation events. Techniques that extrapolate the extent and intensity of radar echoes have been developed to forecast precipitation in the near future. While these techniques show decent skill in forecasting precipitation in the first hour of a forecast, the forecasting skill quickly degrades with time as the lifetime of single convective cells producing precipitation is even shorter and moreover because extrapolation methods are unable to capture the non linear processes that are involved with the formation and dissipation of precipitation. With an increasing amount of resources available for computation, efforts in short term precipitation forecasting have shifted towards numerical weather prediction (NWP) models to more accurately represent physical processes governing precipitation. High resolution (1-4km), limited area models are employed to be able to explicitly describe deep convection, with a decreasing use of convective parametrization as a result. However, these limited area models are usually initialised with data from global numerical weather prediction models that lack the detail, especially in terms of moisture, that is essential for accurate precipitation forecasting. Additionally, these high resolution models first have to adjust to an equilibrium state and initiate vertical motions that are especially relevant for convective precipitation in the so called spinup time, rendering the first hours of a forecast useless from a nowcasting perspective. The necessary level of detail can be added to the model by means of variational data assimilation, where the model state of the atmosphere is combined with recent observations to obtain an optimal estimate of the actual atmosphere, which can then be used to initialize a subsequent high resolution forecast.

In this study it is investigated whether variational data assimilation is able to improve the precipitation forecasting skill of a limited area model covering west-central Europe. To that end, a variety of observations is assimilated into a 4-km resolution Weather Research and Forecasting model (WRF) via the Weather Research and Forecasting model's Community Variational/Ensemble Data Assimilation System (WRFDA) with a 15-minute three dimensional variational assimilation (3DVar) rapid update cycling strategy and a static background error covariance matrix based on summer climatology. Besides conventional observations from e.g. surface weather stations, observations from remote sensing platforms like radar (reflectivity and radial velocity) and Global Positioning System (GPS) zenith total delay (ZTD) are also assimilated. The first assimilation experiment concerns several data denial scenarios, where in each scenario only a single type or select combination of observations is assimilated to isolate their impact on precipitation forecasting skill. In the case of a large squall line it is shown that assimilation of every single type of observation results in an improvement of forecasting skill, except for the radar assimilation of hydrometeors whose impact on the forecast is negligible after having precipitated to the surface. Especially assimilation of humidity based on radar reflectivity yields a significant improvement during the full 6 hours of the forecast. The ability of the other types of observations to positively impact forecasting skill is attributed the large scale forcing by wind convergence along the cold front that is driving the formation of precipitation. Overestimation of precipitation quantity from the scenario without assimilation however is aggravated when radar derived humidity is assimilated, most likely as a consequence of the rapid cycling strategy. Secondly, in a case with localized convection the assimilation of radar based humidity proved crucial to trigger convection and thus improve forecasting skill. In this case other types of observations had a minor impact, as the convection was not initiated from the surface or forced mechanically, but elevated in the mixed layer. Again, precipitation quantity is overestimated when observations, especially from radar, are assimilated. In the final case featuring an extensive warm front producing stratiform precipitation, assimilation of radar humidity and ZTD actually degrades forecasting skill in the first hours of the forecast. The main reason is that the subtle nature of this type of precipitation suffers from the rapid update cycles, creating an unbalanced, spatially inhomogeneous moisture field. The background error covariance horizontal lengthscales that are downscaled to allow for more local adjustments in case of local convection are inappropriate for this large scale and homogeneous type of precipitation.

Acknowledgements

The bulky master thesis report that you are about to throw yourself into is the result of an extensive period of hard work and dedication to produce what is my best attempt to describe what I have investigated during my thesis. While I strive to be as straight to the point as possible, I could not help but sneak in a few extra words where necessary to prove my point.

On a more serious note, I would like to express my gratitude to the people without whom it would not have been possible to write this thesis. First I would like to thank my professor and supervisor Herman Russchenberg for the enjoyable discussions we have had and providing constructive feedback on my work. Additionally I want to thank the other members of the TU Delft in the thesis committee, Stephan de Roode and Marleen ten Veldhuis, for being involved in the numerous feedback meetings and taking the time to critically read my work.

I owe a big thanks to Infoplaza for introducing me to a really interesting current topic in the field of meteorology and providing an insight in how knowledge is actually applied in an operational business. First and foremost I would like to thank Michiel Severin for his commitment to help me throughout my thesis and taking part in my thesis committee representing Infoplaza. Secondly, I owe a lot to Noud Brasjen, a former fellow Geoscience & Remote Sensing student working at Infoplaza, for his valuable support in tackling the problems that I faced during my thesis and sharing his own experiences. I also really appreciated the help of Wilfred Janssen in verifying the meteorological description of the case studies in my report, which is challenging not having a meteorological background. Furthermore I have really enjoyed the pleasant environment to work in, made possible by the friendly colleagues and of course the coffee machine.

In terms of enabling this research project I have to mention some people separately. I thank Dave Offiler and Owen Lewis from the United Kingdom Meteorological Office (UKMO) for granting me access to GPS measurements from the EUMETNET GPS Water Vapour Programme (E-GVAP) and the programme coordinator Henrik Vedel from Danmarks Meteorologiske Institut (DMI) for answering my questions regarding the data. I am also glad that Siebren de Haan from Koninklijk Nederlands Meteorologisch Instituut (KNMI) could share his experiences with data assimilation with me. Besides, I really appreciate the open data policy of the KNMI regarding radar data, which proved to be not so straightforward as I thought it was during my quest for radar data from other European countries. For that reason I would also like to express my gratitude to Maarten Reyniers and Frédéric Jamar at Koninklijk Meteorologisch Instituut (KMI) for supplying raw scan data acquired by the weather radars in Belgium for my case studies.

I also have to give credit to Rick Bekker, stormchaser and photographer who was able to capture the fantastic sight of an approaching shelf cloud at Kinderdijk, the Netherlands, which features on the cover of this report.

Of course I could not have completed this thesis without the incredible support of my family. Although great hardship has come upon us last year concerning my mother's health, I feel the support has never been stronger. I hope that by reaching this milestone I can express my appreciation for their encouragement throughout my study.

Last but certainly not least I am really grateful for the support of my friends and fellow Geoscience &Remote Sensing students in Delft. I think we have grown to be a great and unique group of friends and my time as a student in Delft would definitely not have been the same without you.

For now, I am excited to show you the results of the research I have conducted in my thesis, which I consider to be the culmination of my endeavours to become an engineer. I cared to make it enjoyable to read, however at the same time I hope you learn something in the process as well.

Contents

List of Figures

List of Tables

1

Introduction

In a changing climate more extreme weather events are likely to occur (IPCC, 2013, chapter 12). More intensive precipitation events are predicted primarily because of the increased ability of the air to contain water vapour at higher temperatures. Inherently coupled is the need for accurate precipitation forecasts, which will continue to grow in the coming decades. Forecasting short term precipitation up to six hours ahead is often referred to as **nowcasting**. The objective of nowcasting is to provide a detailed estimate of location, timing and intensity of precipitation events. In the 1960s along with the development of remote sensing techniques the first nowcasting methods were developed by Noel and Fleisher (1960) and Hilst and Russo (1960) based on correlating reflections in successive radar images. Using a steady state assumption, the size and speed of a storm could be used to forecast its trajectory in the near future. In the next decades more complex versions of these extrapolation methods were developed that could recognize individual convective cells each having their own velocity (Rinehart (1981)) and additionally extrapolate trends in extent and intensity. Tsonis and Austin (1981), who developed such a size and intensity trending algorithm nevertheless found negligible improvement in forecasting skill when more sophisticated non linear trending schemes were employed. Wilson et al. (1998) have performed similar experiments using the TITAN (Thunderstorm Identification, Tracking, Analysis and Nowcasting) algorithm developed by Dixon and Wiener (1993) with and without extrapolating trends in size and intensity and similarly did not observe any improvement in forecasting skill, leading them to conclude analogous to Tsonis and Austin (1981) that the processes responsible for change in precipitation through time are not necessarily observable from the past evolution of a radar reflection. In case precipitation is a result of convection, most of the time processes in the planetary boundary layer like convergence are responsible for initiation of precipitation. With this notion in mind so called expert systems have been developed that attempt to predict the initiation and dissipation of convective cells in addition to radar extrapolation, by including a range of additional knowledge based (e.g. from statistics, research studies or conceptual models) or observational predictors like clouds on IR satellite images, boundary layer winds from surface, upper air and Doppler velocity observations and even convergence lines as diagnosed by meteorologists. Nevertheless, the accuracy of these nowcasting methods degrades rapidly with time (Elvander (1977); Dixon and Wiener (1993); Wilson et al. (1998)) as qualitatively illustrated in figure 1.1, primarily as a consequence of individual convective cells and thunderstorms only having a mean lifetime of approximately 20 minutes (Battan (1953); Foote and Mohr (1979)), making extrapolation based methods increasingly more unreliable beyond this time scale in a forecast. Exceptions are cases where cells merge or are otherwise organised as a larger mesoscale system (Henry (1993)) or when precipitation results from strong and extensive front (Hill et al. (1977)). Once individual convective cells have dissipated, a correct nowcast relies on the ability to forecast initiation, growth and decay of new cells, which requires a detailed and accurate description of atmospheric conditions as well as a model that is able to explicitly describe the evolution of processes that lead to the formation of precipitation.

For that reason, nowcasting for lead times > 1 hours has shifted to NWP based methods. The development of NWP methods combined with increasing computational power has enabled employing high resolution models to forecast precipitation. Only modelling a regional area for a limited time (e.g. +24 or +36 hours) allows defining resolutions in the order of 1-4 km, generally referred to as "convection permitting models" for their ability to explicitly describe large scale, organized, deep convection, obviating the need for convective parametrization (CP) schemes. Numerous studies (Done et al. (2004); Kain et al. (2006); Clark

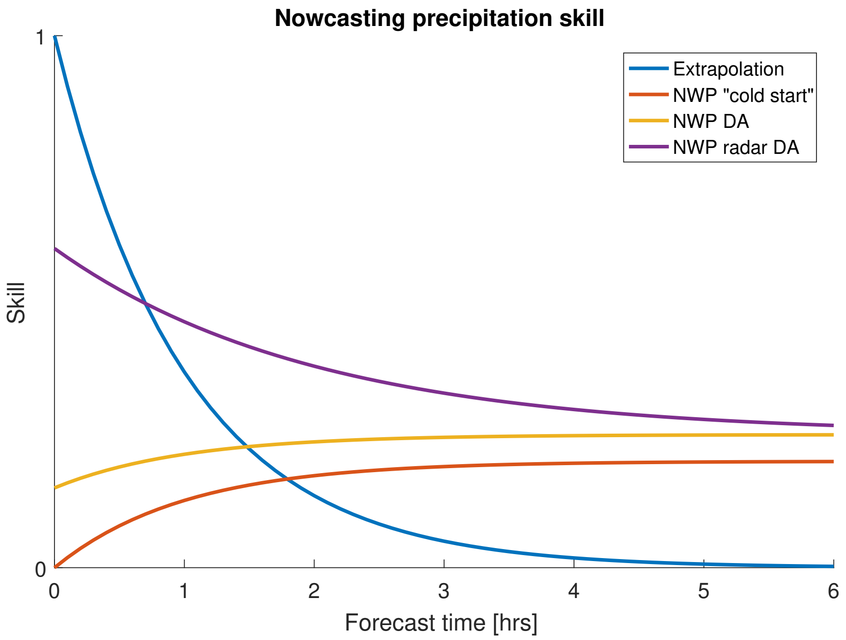


Figure 1.1: Qualitative illustration of forecast skill of various nowcasting methods such as radar extrapolation and NWP based methods.

et al. (2009)) have proved that these high resolution models without CP have a better skill in forecasting precipitation than CP employing models with a coarser resolution. Despite the improved ability of these high resolution NWP models to forecast initiation and mode of convection, accurately forecasting timing, location and intensity as required to meet nowcasting demands remains problematic, as shown by Weisman et al. (2008) who compared a 4-km resolution model with the 12-km CP employing ETA model. Moreover, high resolution regional NWP models require a certain time to establish a dynamical balance and start physical processes relevant for convection after they have been initialized from coarse resolution global models. This period, generally referred to as spinup time, can take up to 3-6 hours, rendering these first hours of a so called "cold start" forecast worthless from a nowcasting perspective. In figure 1.1 this spinup time time is qualitatively shown with the cold start NWP taking some time to reach its forecasting skill potential. To reduce the spinup time and improve model accuracy, the model description of the atmospheric state is updated with high resolution observations by means of data assimilation, of which the qualitative skill score has also been plot in figure 1.1. Several implementations exist, but in the last decade however **variational data assimilation** has dominated in NWP research and has now been implemented in some form in all of the current operational high resolution models from leading NWP centres like European Centre for Medium-Range Weather Forecasts (ECMWF), National Centers for Environmental Prediction (NCEP), United Kingdom Meteorological Office (UKMO), Deutsche Wetterdienst (DWD) and Environment Canada (EC). While most of these centres now employ really advanced forms of data assimilation (see Bannister (2016) for an overview) combining the hybrid approach to describe model error covariance and 4D methods that can use the evolution of the NWP as an additional constraint for observations, these methods require vast computational resources to generate large ensemble forecasts for hybrid data assimilation and to evolve and linearise the NWP model in every iteration of the minimization procedure that is central in any four dimensional variational assimilation (4DVar) system. Consequently, computationally less demanding 3DVar systems in combination with rapid update cycles of < 1 hour remain a popular alternative in operational setups for these expensive methods.

Essential in improving forecast skill of NWP models has proven to be the assimilation of radar observations. It is currently the only observation technique that is able to detect the size and extent of precipitation with high spatial and temporal resolution that is required for nowcasting. So, along with the development of variational data assimilation, efforts have been made to assimilate radar observations into NWP models. Assimilation of radar reflectivity is not trivial since it concerns an indirect measurement of precipitation and has to be linked to model variables using theoretical or empirical relations and several assumptions. Besides relating radar reflectivity to hydrometeors, often reflectivity is also used to employ some form of diabatic initialisation in which favourable conditions for precipitation are created by including the latent heat release associated with the formation of hydrometeors as well as assuming saturation of humidity in areas where significant radar reflections are measured or otherwise analysing 3D cloud cover with aid from other observations. Research has shown that the diabatic initialization techniques based on reflectivity reduces the spinup time (Ducrocq et al. (2002); Hu et al. (2006b); Weygandt et al. (2008); Dixon et al. (2009); Caumont et al. (2010); Schenkman et al. (2011); Wang et al. (2013)), thereby improving forecast skill in the first hours of the forecast that are so relevant for nowcasting. In addition to reflectivity Doppler weather radars are also able to retrieve the radial velocity of a reflection, which can also be assimilated to improve model wind estimates.

This study has been set up with a number of objectives in mind, the first one being to improve the accuracy of the precipitation forecast from a regional Weather Research and Forecasting model (WRF) (Skamarock et al. (2008)) covering west-central Europe. To that end, observations from various sources will be assimilated into the WRF model using the WRFDA (Barker et al. (2012)) with a 15 minute rapid update 3DVar cycling strategy. This rapid 3DVar approach is adopted with the idea that most of the observations can be assimilated close to their respective times, similar to a 4DVar approach but far less computationally demanding. As noted before, amongst the assimilated observations radar reflectivity is most prominent, for its ability to provide the necessary detail regarding precipitation to the initial model state. Radar reflectivity observations from Dutch and Belgian weather radars are assimilated using the indirect approach from Wang et al. (2013), which also includes a form of diabatic initialization where near saturation humidity is assimilated when significant reflections are measured. Moreover, the absence of a significant radar reflection is used to penalize spurious model precipitation, employing null-echo assimilation similar to Min et al. (2016). Besides radar, observations from other sources like surface stations, airports, buoys are assimilated. In addition, GPS ZTD relating to IWV are assimilated to provide some constraints on the humidity assimilation from radar reflectivity. Logically, a second objective is to assess how assimilation of each observation type affects the skill of forecasting precipitation. A case with a MCS squall line passing the Netherlands is examined with a set of data denial experiments where in every scenario only a single or subset of observation types is assimilated to isolate their effect on the precipitation forecast. Finally, besides the MCS case, as a third objective two additional cases are investigated to determine the ability of data assimilation to improve forecasting skill in the case of more local convection and additionally in a case with stratiform rain of really large extent and moderate intensity.

To reach these objectives, this thesis report is structured as follows. Chapter 2 will discuss the fundamentals of assimilating observations into a NWP model. First a brief motivation for data assimilation is given, followed by an overview of (variational) data assimilation methods. Then the mathematical framework for data assimilation is examined, with the definition of a cost function to quantify the difference between the model and observations. Computation of the cost function requires modelling and estimation of the errors associated with the model and observations. Another important component of variational data assimilation is how exactly observations are related to model variables. Especially remote sensing instruments provide indirect measurements of atmospheric variables, so for radar and GPS measurements it is discussed separately how they are assimilated. The chapter on data assimilation is concluded with range of pseudo single observation test (PSOT), that shows how a single observation influences the analysis and at the same time provides insight in the way error covariances are modelled. In data assimilation it is critical that the observations that are assimilated are of good quality. In chapter 3 the preprocessing steps that are carried out before observations are assimilated into the model are addressed. Especially radar reflectivity requires several operations to obtain quality observations ready for assimilation. For radar radial velocity also numerous processing steps have to be applied to prevent assimilation of aliased observations. Finally, for GPS it is described how an important height correction for model and receiver height discrepancy is performed. Several experiments are defined where the assimilation of observations is investigated. A detailed description of each experiment together with the choice for a number of precipitation cases can be found in chapter 4. Before the experiments are described the model setup is discussed, with a focus on the important parametrizations that are employed. The chapter is concluded with an explanation of the method that is used to verify the precipitation forecast. The description of the experimental setup is followed by a discussion of the results that the assimilation of the various observations yields. Every experiment is examined in detail to assess the influence of data assimilation on the precipitation forecast and is verified using the method described in chapter 4. Performing the data assimilation experiments enables to draw a conclusion regarding the added value of data assimilation to the precipitation forecast, which is summarized in chapter 6. The conclusion is followed by a set of recommendations for future research.

2

Variational data assimilation

This chapter will discuss the fundamentals of data assimilation. Firstly, the incentive for data assimilation is explained in section 2.1 together with a brief overview of data assimilation methods in section 2.2. Following that the mathematical framework that forms the basis of variational data assimilation is addressed (section 2.3). Central is a cost function that quantifies the difference between model and observation. In order to evaluate the cost function, the uncertainty of the model and observations have to be quantified. How their respective errors are defined, modelled and estimated will be discussed in section 2.4. Throughout the development of data assimilation methods this cost function has been adapted to be able to accommodate assimilation of other (remotely sensed) observations like reflectivity (section 2.5) and radial velocity from Doppler weather radars (section 2.6) that indirectly relate to model quantities. GPS retrieved estimations of signal delay by water vapour is another example remotely sensed observations that will be assimilated in this study. In section 2.7 it will be shown what method exists to achieve this. This chapter concludes with a series of PSOTs using different kinds of observations to assess the impact that assimilating a single observation has on the model state.

2.1. Introduction

Ever since meteorologists started using numerical weather prediction models to aid in weather forecasting, the desire to improve the accuracy of these models existed. On the one hand this can be achieved by improving the way the physical processes that govern the atmosphere are described in the model. However the equations can already describe these processes reasonably well, although subgrid scale processes like turbulence and droplet formation still have to be parameterized. To be able to resolve and describe small scale processes like localized convection usually the resolution is increased, but naturally that requires a lot more computing power. E.g. using half the grid scales will result in 4 times as much calculations that have to be made, only considering the enhanced areal resolution. However even if computing power would allow refining resolution to sub kilometer level it is not necessarily more accurate since the information with which the model was initialized might have come from a source which did not have that level of detail. In fact, it is common to initialize a regional weather model with a global model like the Global Forecast System (GFS) or ECMWF model that produce forecasts at much coarser resolution (e.g. approximately 28 and 56 km for Global Forecast System (GFS)). Moreover, these models themselves have to be initialized from observations. There are not nearly enough observations to initialise every variable for every model grid box, so interpolation is used to obtain values for the model grid. Combined with the fact that these global models require at least several hours in state of the art facilities to perform their computations, a regional model initialized with this data first needs to catch up with reality to actually become a forecast. For example it would not be uncommon for a regional weather model forecast for central Europe to be initialized with global model data which itself was initialized from observations e.g. 12 hours ago. Naturally, in that time differences arise between the modelled and real atmosphere, especially in turbulent weather regimes like heavy storms. To make the model more accurately represent the current state of the atmosphere, observations are being integrated into the model which is commonly referred to as **data assimilation**.

Over the years many data assimilation methods have been developed, all with the purpose of using observations to improve the accuracy of a model. An overview of current, widely employed data assimilation methods is given in figure 2.1. The most basic approach is actually not to use a model reference or first guess at all and simply interpolate the observations to a model grid. However this approach is quite limited, in the first place because there are not enough observations to obtain a full model description of the atmosphere. Cressman (1959) developed one of the first methods to combine model and observations. In a so called Cressman analysis the model state is set equal to observations at the observation locations. Up to an arbitrary range the model state is updated by an observation:

$$\mathbf{x}_a(i) = \mathbf{x}_b(i) + \frac{\sum_{j=1}^{k} w(i,j)\left[\mathbf{y}(j) - \mathbf{x}_b(j)\right]}{\sum_{j=1}^{k} w(i,j)} \tag{2.1}$$

Where $\mathbf{x}_a(i)$ and $\mathbf{x}_b(i)$ represent element i of the analysis and background model state, $\mathbf{y}(j)$ the j-th observation $w(i,j)$ a weight that is given, based on the distance $d_{i,j}$ between model grid point i and observation j:

$$w(i,j) = \left\{ \begin{array}{cc} \frac{R^2 - d_{i,j}^2}{R^2 + d_{i,j}^2} & d_{i,j}^2 \leq R^2 \\ 0 & d_{i,j}^2 > R^2 \end{array} \right\} \tag{2.2}$$

With R an arbitrary range of influence. Cressman analysis is usually classified under the successive correction methods, as it is often applied iteratively. The analysis is updated several times to improve the smoothness of the analysis. However, interpolation and the successive correction methods have their disadvantages (Bouttier and Courtier, 2002):

- Inability to make **physically realistic** changes to a model. Setting a model variable equal to observations might produce significant artefacts that are not consistent with the laws of physics if observations are inaccurate. The analysis should follow physical constraints like the hydrostatic balance (the upward force on a slab of air resulting from the vertical pressure gradient is equal to the downward gravitational force of the slab), geostrophic wind (a balance between the pressure gradient and Coriolis force) or a relative humidity that does not exceed 100%. A pressure observation should also lead to an alteration of the wind in the analysis. Also, the length scale of the increments should be realistically smooth and not produce any high frequency jumps in a model field.

- Inability to take into account **uncertainty**. This is a major drawback since the interpolation and successive correction methods assume that observations are true: they contain no error. Therefore only assimilating few erroneous observations can significantly degrade the analysis. Even accurate observations have a certain error and moreover if the background is already of good quality we would not like it to be impacted easily.

- Inability to directly integrate **remotely sensed observations**. Remote sensing measurements are most often indirect retrievals of model variables. Nowadays the major part of observations that are readily available come from remote sensing instruments like weather radars, satellites, lidar etc. Model quantities have to be retrieved separately before the observations can be assimilated into the model.

All these inadequacies have given rise to the development of a statistical data assimilation method, known as **variational data assimilation**. The primary advantage of variational data assimilation is that it can take into account the uncertainties in both model and observation to produce an optimal estimate. It is optimal in the sense that is the most likely estimate given the uncertainties of the model and observation, as demonstrated in section 2.3.2. Under various assumptions it also minimizes the variance of the solution.

As mentioned in the introduction there is an increasing need for accurate forecasts, especially for precipitation. For accurate forecasting of precipitation it is vital that models are initialized with a detailed and accurate description of the atmosphere. Especially moisture, which is crucial for formation of clouds and precipitation, is not as homogeneous as temperature or pressure which makes it difficult to describe detailed precipitation from model states initialized from coarsely sampled information. There lies the added value of data assimilation: supplying the model with high resolution and frequently sampled observations to initialise a model with more detail regarding the current state of the atmosphere. Remote sensing instruments have opened a new chapter in data assimilation as they sample very frequently and moreover provide information at higher altitudes in the atmosphere, contrary to conventional surface weather stations at the surface. For example Doppler weather radars sample the atmosphere with both a high temporal (e.g. 5 minutes) and spatial (e.g. 1 km) resolution. Also, the different scan angles from the radar can provide information at various elevations. Measured echoes from weather radar can provide a model with frequent and detailed information on the presence of precipitation in the atmosphere.

2.2. Overview

In general, variational data assimilation can be defined as a statistical method to obtain the best estimate of model state variables. Applied to atmospheric science, data assimilation combines observations and a model forecast commonly referred to as first guess together with an estimate of the associated errors to provide an analysis which is closer to the real atmospheric state. This analysis is used to initialize a subsequent numerical weather prediction model run. In essence it can be regarded as a model update at a certain time step by comparing the model state to observations.

Nearly all meteorological institutions run operational models that include variational data assimilation. These variational data assimilation methods have become increasingly more complex over the years. See figure 2.1 again for an overview of the various types of data assimilation methods that are commonly used in these institutions. Statistical data assimilation methods now dominate, for the reasons mentioned in the previous section. At the basis stands the three dimensional variational assimilation, where all variables in a three dimensional model can be updated simultaneously. Correlations between model variables, e.g. same physical variable but other grid points or different physical variables through their balance property, ensure a physically sound analysis. These correlations are represented by the error covariance between the model variables. The error characterization of both model and observation is used for the uncertainty based weighting of model and observation. First implementations of three dimensional variational assimilation were developed by European Centre for Medium-Range Weather Forecasts (ECMWF) (Courtier et al., 1998; Mcnally et al., 1998), High Resolution Limited Area Model (HIRLAM) (Lindskog et al., 2001), National Centers for Environmental Prediction (NCEP) (Parrish and Derber, 1992), United Kingdom Meteorological Office (UKMO) (Lorenc et al., 2000) and by National Center for Atmospheric Research (NCAR) (Barker et al., 2004). The latter, Community fifth generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) Three Dimensional Variational Assimilation (3DVAR) system, has formed the basis for the release of a community data assimilation system known as Weather Research and Forecasting model's Community Variational/Ensemble Data Assimilation System (WRFDA) (Barker et al., 2012). It is designed to be compatible with the Weather Research and Forecasting model (WRF) that is distributed by the same multi-agency collaboration including National Center for Atmospheric Research (NCAR), National Oceanic and Atmospheric Administration (NOAA), Air Force Weather Agency (AFWA), Oklahoma University (OU) and Naval Research Laboratory (NRL) that maintains the WRFDA system. WRFDA has been developed to cater the growing demand for a state of the art data assimilation system from universities, federal offices, private companies who lack the resources to develop these systems. In this thesis all data assimilation experiments are carried out using the latest version of WRFDA (V3.9).

The principle of three dimensional variational assimilation has been extended to four dimensional variational assimilation which uses the evolution of a model through time as an additional constraint. That way the **analysis** is the optimal model state for which its linearised model trajectory fits the assimilated observations in the best way. The benefit of 4D over 3D is that the observations can be used at their appropriate time, or at least closer to their valid time since for practical reasons observations are still binned in time slots. Figure 2.2 illustrates this principle. In three dimensional variational assimilation, the data assimilation algorithm assumes all observations within a predefined time window are valid at the time of assimilation. Instead, four dimensional variational assimilation compares observations against the model trajectory valid at the time of observation.

As will be discussed in section 2.4, the description of the precision of model and observations is essential for the quality of data assimilation. Accurately estimating model variable covariance is one of the most fundamental challenges in data assimilation. Because in three dimensional variational assimilation and four dimensional variational assimilation usually a time invariant estimate of model covariance is used, alternative assimilation methods have been developed that use an ensemble of perturbed forecasts to obtain a more accurate representation of model errors specific to a single forecast. These flow dependent error covariance estimation methods are referred to as **ensemble methods**, see figure 2.1. Pure ensemble methods like the Ensemble Kalman Filter and Ensemble Kalman Smoother assimilate observations into each ensemble member, with the model error covariance estimated from the spread in the ensembles. Later, **hybrid methods**

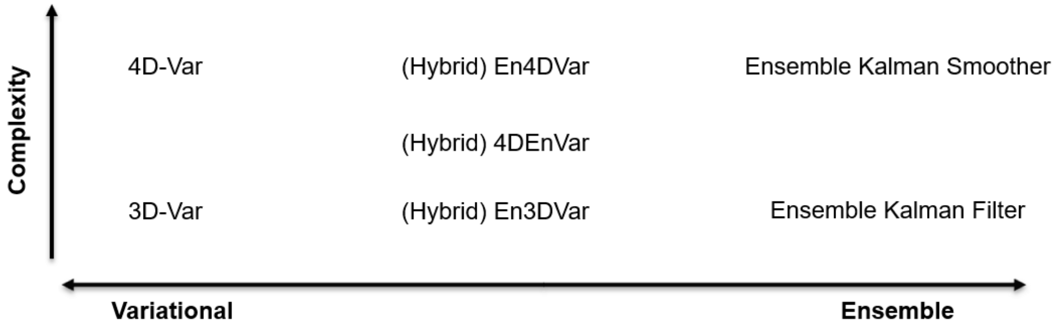


Figure 2.1: Overview of the types data assimilation algorithms currently employed by leading weather forecasting institutions. The horizontal axis shows whether the algorithm is variational, which uses a prescribed background error covariance matrix, or utilizes an an ensemble of forecasts to describe model covariance. The vertical axis shows the increasing complexity and computational costs of these algorithms. 3D-Var and 4D-Var are abbreviations for three dimensional variational assimilation and four dimensional variational assimilation respectively. The naming conventions have been adopted from Bannister (2016), who provided a comprehensive overview of current data assimilation algorithms. The term hybrid refers to the combination of a static covariance as used in 3D/4D-Var and an ensemble based estimate of covariance as used in the pure ensemble methods. The primary difference between En4DVar and 4DEnVar is that En4DVar uses the tangent linear and adjoint of the NWP model to evolve the background error covariance matrix from the start of the forecast through time just like 4D-Var does, while 4DEnVar estimates the development of the background error covariance matrix just from the ensembles and therefore does not require linearisation of the NWP model. The reader is referred to Bannister (2016) for a more thorough description of the differences between these algorithms.

have been developed that combine the merits of static and flow dependent description of model error covariance. A brief discussion regarding these methods to estimate model covariance can be found in section 2.4.5. The technical differences between these methods and full implementation are not a focus of this study, thus the reader is referred to Bannister (2016) who provides an elaborate description and comparison of these methods.

2.3. Cost function

2.3.1. Definition

Suppose a numerical weather prediction model describes the atmosphere with variables like temperature T, pressure p, density ρ, wind components u, v and w and specific humidity q at every model grid box. These variables are concatenated into the state vector $\mathbf{x}$ which describes the total model state.

$$\mathbf{x} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \\ \mathbf{t} \\ \mathbf{p} \\ \rho \\ \mathbf{q} \end{bmatrix} \tag{2.3}$$

In 2.3 $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, $\mathbf{t}$, $\mathbf{p}$, ρ and $\mathbf{q}$ are the model vectors containing all grid values for the model variables mentioned above. For example, the vector $\mathbf{u}$ of the horizontal west-east wind component u contains the wind speed for all k grid points:

$$\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_k \end{bmatrix} \tag{2.4}$$

In standard data assimilation, a subset of the model variables is used, namely $\mathbf{u}$, $\mathbf{v}$, $\mathbf{t}$, $\mathbf{p}_s$, and $\mathbf{q}$ because most observations relate to these quantities, although as we will see later in the assimilation of radar data, more variables may be added to the state vector if the observations cannot be linked to the standard data assimilation model variables. Taking a subset of 2.3, $\mathbf{x}$ is redefined as the model vector for data assimilation:

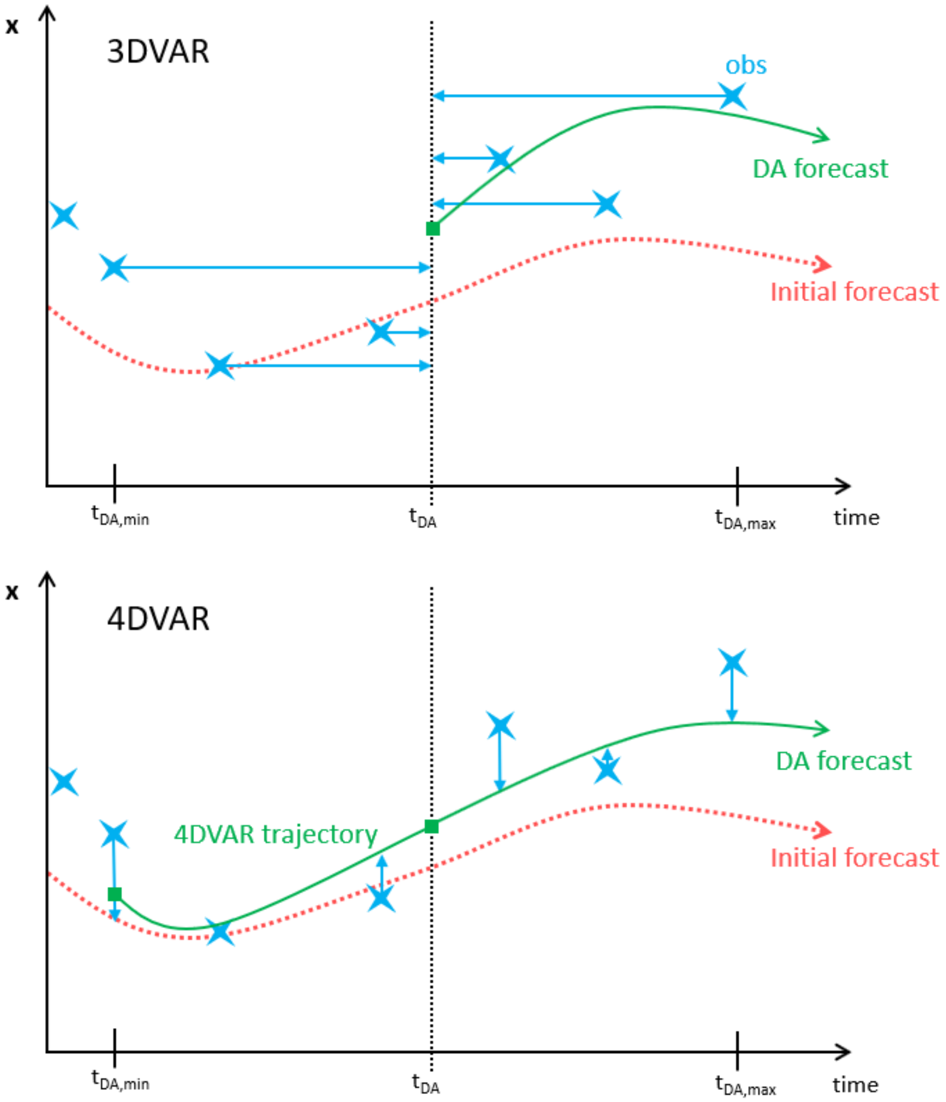


Figure 2.2: Schematic graphs illustrating the principle of three dimensional variational assimilation and four dimensional variational assimilation. The top figure shows the principle of three dimensional variational assimilation, where all observations within a limited time window (e.g. 1 hour, in the graph denoted with $t_{DA,min}$ and $t_{DA,max}$) are valid for time t_{DA} at which the assimilation is carried out. This results in an analysis (indicated with the green square) which is then used to initialize a new forecast, depicted with the green arrow. Four dimensional variational assimilation however takes into account the time at which the observations were acquired, by finding an initial condition of which the model trajectory best fits the observations at their respective time. That trajectory is indicated with the green, dashed line in the bottom graph.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{t} \\ \mathbf{p}_s \\ \mathbf{q} \end{bmatrix} \tag{2.5}$$

$\mathbf{p}_s$ represents the surface pressure. Typically, this will result in a vector with n entries in the order of $10^6 - 10^7$, corresponding to the number of grid points k times the number of model variables. Besides the model vector $\mathbf{x}$ a vector $\mathbf{y}$ can be constructed with a number of m observations that are to be assimilated into the model:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \tag{2.6}$$

Any number of observations could be assimilated, but usually however the number will not exceed $10^5 - 10^6$, making the system where the model state would be inferred purely from observations underdetermined. This explains the mathematical need for using a first guess in the estimation of the optimal state for **x**. Of course the first guess is also used because it is assumed that is already a relatively good approximation of the true atmospheric state. The terms first guess and background are used interchangeably, although as will become clear later they are technically not the same when multiple outer loops are carried out during numerical minimization. To quantify the difference between model and observations, a cost function J_x (2.7) is defined. The cost function is central to data assimilation and will be found in any scheme that implements variational data assimilation, like WRFDA. It can be derived from the maximum likelihood principle as shown in appendix A.1. The data assimilation analysis is the most likely model representation of the atmosphere given the background state $\mathbf{x_b}$ and observations **y**. It consists of a term that measures the offset of a model state vector **x** with the background forecast $\mathbf{x_b}$ and with model variables **x** mapped to observation space by observation operator H:

$$\begin{aligned} J(\mathbf{x}) &= J_b + J_o \\ J(\mathbf{x}) &= \frac{1}{2}(\mathbf{x}-\mathbf{x_b})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x_b}) + \frac{1}{2}\big(H(\mathbf{x})-\mathbf{y}\big)^T\mathbf{R}^{-1}\big(H(\mathbf{x})-\mathbf{y}\big) \end{aligned} \tag{2.7}$$

A simple example of the observation operator would be the interpolation of model temperature to the observation location situated in between the model grid points. For remote sensing instruments the observation operator is usually more complex and non linear. E.g. for satellite radiance assimilation the observation operator includes estimating the energy flux at the top of the atmosphere from model temperature at various levels using Stefan-Boltzmann's law that is strongly non linear. In 2.7 the squared vectors are weighted with the inverse of the model variable error covariance matrix **B** and observation error covariance matrix **R** to obtain a measure of dissimilarity between the the updated model state and the background and observations. A description of these error covariance matrices is given in section 2.4. When using equation 2.7 it is assumed that the described errors are unbiased and their distribution is described by a Gaussian probability density function. Moreover, the errors of model variables and observations are assumed to be uncorrelated so that no cross product of the two exists in the cost function. Generally that is a valid assumption, although e.g. in the case of satellite radiance assimilation the model background is used to perform a bias correction on the observations, so that errors in the model background can propagate to the observations.

Logically, the cost function is minimized to obtain an **analysis** state referred to as $\mathbf{x_a}$ that is the best estimate for the atmospheric state described by the model. It is best in the sense that it maximizes the likelihood of the analysis given the likelihood of the first guess and observations. Because in the definition of the cost function it is assumed that model and observation errors have a Gaussian distribution, the analysis also minimizes the variance of the analysis when the observation operator is linear, as will be shown in the next section. Minimizing the cost function is equivalent to finding a model state **x** for which the gradient of J equals zero:

$$\nabla J(\mathbf{x_a}) = 0 \tag{2.8}$$

2.3.2. Exact solution and precision of the analysis

First it is shown how to obtain an analytical solution of the minimization problem. This will aid in understanding how variational assimilation updates the model state with observations based on their error characterization. The second part of the cost function J_o that measures the departure from observations is linearised with a Taylor expansion around $\mathbf{x_b}$.

$$H(\mathbf{x}) - \mathbf{y} = H(\mathbf{x}_b + (\mathbf{x}-\mathbf{x}_b)) - \mathbf{y} \approx H(\mathbf{x}_b) + \mathbf{H}(\mathbf{x}-\mathbf{x}_b) - \mathbf{y} \tag{2.9}$$

Where **H** represents the linearised version (Jacobian) of the non linear observation operator H in the vicinity of $\mathbf{x_b}$. By performing this linearisation it is assumed that the first guess is close enough to the analysis such that a linear approximation is valid. Rewriting the cost function 2.7 using 2.9 results in the following:

$$
\begin{aligned}
J(\mathbf{x}) = & \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \left(\mathbf{B}^{-1}+\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)(\mathbf{x}-\mathbf{x}_b) \\
& + \left[H(\mathbf{x}_b)-\mathbf{y}\right]^T \mathbf{R}^{-1}\mathbf{H}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\left[H(\mathbf{x}_b)-\mathbf{y}\right]^T \mathbf{R}^{-1}\left[H(\mathbf{x}_b)-\mathbf{y}\right]
\end{aligned} \tag{2.10}
$$

See appendix A.2 for the intermediate steps taken to obtain 2.10. Instead of computing J as a function of $\mathbf{x}$ which is computationally very demanding, the cost function is expressed as a function of the analysis increment $\delta\mathbf{x} = \mathbf{x}-\mathbf{x}_b$. This is known as the incremental formulation and is widely adopted as a starting point for data assimilation algorithms.

$$
J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\left(\mathbf{B}^{-1}+\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\delta\mathbf{x} + \left[H(\mathbf{x}_b)-\mathbf{y}\right]^T\mathbf{R}^{-1}\mathbf{H}\delta\mathbf{x} + \frac{1}{2}\left[H(\mathbf{x}_b)-\mathbf{y}\right]^T\mathbf{R}^{-1}\left[H(\mathbf{x}_b)-\mathbf{y}\right] \tag{2.11}
$$

Evidently, the cost function in 2.11 is a quadratic function of $\delta\mathbf{x}$, in the general form of:

$$
J(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c \tag{2.12}
$$

With a weighting matrix $\mathbf{A}$, a vector $\mathbf{b}$ and a scalar c. The gradient of a quadratic function is given by:

$$
\nabla J = \left(\mathbf{A}+\mathbf{A}^T\right)\mathbf{x} + \mathbf{b} = 2\mathbf{A}\mathbf{x} + \mathbf{b} \tag{2.13}
$$

The last equality is only valid if $\mathbf{A}$ is symmetric. That is true for this case since $\left(\mathbf{B}^{-1}+\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)$ is a symmetric matrix (see appendix A.3). Applying 2.13 to the incremental cost function 2.11, an expression for the gradient is obtained:

$$
\nabla J(\delta\mathbf{x}) = \left(\mathbf{B}^{-1}+\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\delta\mathbf{x} + \mathbf{H}^T\mathbf{R}^{-1}\left[H(\mathbf{x}_b)-\mathbf{y}\right] \tag{2.14}
$$

Setting the gradient equal to zero and solving for $\delta\mathbf{x}$:

$$
\delta\mathbf{x} = \left(\mathbf{B}^{-1}+\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\left[\mathbf{y}-H(\mathbf{x}_b)\right] \tag{2.15}
$$

This is equivalent to the analytical solution for the so called Optimal Interpolation method (see e.g. A.C. Lorenc (1986, p. 1180)), where the expression for an analysis increment is:

$$
\delta\mathbf{x} = \mathbf{B}\mathbf{H}^T\left(\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^T\right)^{-1}\left[\mathbf{y}-H(\mathbf{x}_b)\right] \tag{2.16}
$$

For a proof, see appendix A.4. Note that the equivalence is only valid if the observation operator H is linear. The matrix product $\mathbf{B}\mathbf{H}^T\left(\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^T\right)^{-1}$ is a matrix that supplies the weights corresponding to each virtual observation increment $\delta\mathbf{y}_i$ for each state vector entry $\delta\mathbf{x}_i$. Formally it is known as the optimal Kalman gain matrix $\mathbf{K}$ in the widely used Kalman filter:

$$
\mathbf{K} = \mathbf{B}\mathbf{H}^T\left(\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^T\right)^{-1} \tag{2.17}
$$

Let us define the covariance matrix $\mathbf{A}$ which describes the precision of the analysis $\delta\mathbf{x}_\mathbf{a}$. As noted before, the solution $\delta\mathbf{x}_a$ minimizes the variance and is an improvement of the original background error $\mathbf{B}$. The precision of the analysis is given by:

$$
\mathbf{A} = E\left[\varepsilon_a{\varepsilon_a}^T\right] = (\mathbf{I}-\mathbf{K}\mathbf{H})\mathbf{B} \tag{2.18}
$$

This expression is obtained by defining the analysis error as

$$
\varepsilon_a = \varepsilon_b + \mathbf{K}(\varepsilon_o - \mathbf{H}\varepsilon_b) = (\mathbf{I}-\mathbf{K}\mathbf{H})\varepsilon_b + \mathbf{K}\varepsilon_o \tag{2.19}
$$

and then multiplying the analysis error with its transpose and then taking the expected value, also noting that the background and observation errors are uncorrelated. Additionally the expression for the optimal Kalman gain matrix (2.17) that is found by minimizing the variance has to be substituted into the general expression for the analysis precision that is valid for any gain matrix $\mathbf{K}$ to arrive at the simplified form of equation 2.18. See also Brown and Hwang (1997, chapter 4) or Haykin (2001, chapter 1.3) for a full derivation of the analysis precision. Perhaps a more intuitive, equivalent formulation of the analysis precision is:

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \tag{2.20}$$

This expression shows that for any nonzero background **B** and observation **R** variance, the variance of **A** is smaller than the variance of the background **B**. Again, note that $\mathbf{x_a}$ only exactly minimizes the variance when errors have a Gaussian distribution and the observation operator is linear.

2.3.3. Numerical minimization

Unfortunately, the analysis is increment is not as easily computed as the analytical expression that was found in 2.15 suggests. In the first place because of the sheer size of the matrices involved, for example with matrix **B** having $O(10^{14})$ entries which would require immense storage space. To reduce matrix sizes a variable transformation is performed to make the matrix near diagonal, requiring less storage and the same time improving the condition number. This will be discussed more in detail in section 2.4.4. Secondly, computing the inverse of these large matrices is impossible. Optimal Interpolation circumvents this problem by solving 2.15 separately for subsets of **x** with a limited number of observations so that the inverse matrix is small enough to actually compute directly. However the regional approach of this solution lead to physically unrealistic features or discontinuities at the region boundaries, giving rise to the development of 3DVar, which is able to incorporate all observations simultaneously.

In 3DVar, instead of sticking to the analytical solution where the gain matrix **K** is evaluated as with Optimal Interpolation, the cost function is solved iteratively using numerical minimization algorithms. Iterative methods like steepest descent, Quasi-Newton and conjugate gradient for instance are used to minimize the cost function. Their main principles are shortly addressed. Most of these methods are based on finding the local gradient to provide guidance in making an analysis increment for the next iteration. See figure 2.3 for a flowchart describing the general steps of these algorithms. The difference between the methods lies in how they exactly determine the increment $\delta\mathbf{x}$ as a function of the gradient. Apart from that the general procedure is the same. Using a first guess the cost function and its gradient are evaluated. Then, employing the method specific way of computing the analysis increment, the next point is calculated at which the convergence criteria are evaluated. Usually it is said that convergence is reached when the cost function or gradient reaches a certain fraction ϵ of its initial value, e.g. 0.01. Or alternatively when the analysis increment becomes smaller than a certain absolute threshold.

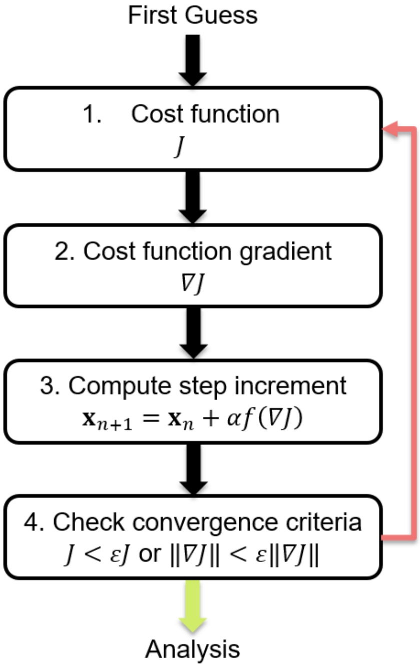


Figure 2.3: Flowchart illustrating the steps performed in minimizing the cost function.

A visual example of iteratively minimizing the cost function using a gradient algorithm is shown in figure 2.4. In every step the gradient at the current position is calculated to determine the direction for the next minimization increment. For the sake of simplicity the cost function is defined in two dimensional space. Note that the cost function is a paraboloid, the two dimensional equivalent of a parabolic function. Again, the point $\mathbf{x}$ for which the cost function is minimal is the optimal point $\mathbf{x_a}$.

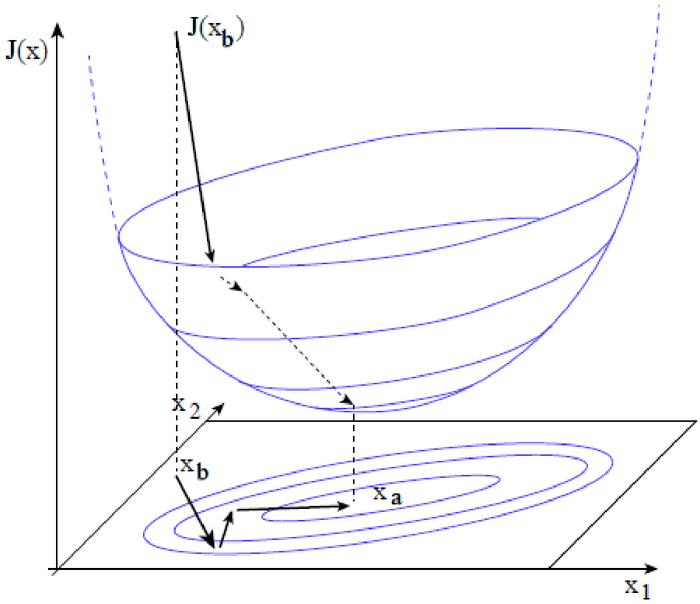


Figure 2.4: Schematic representation of an iterative numerical minimization algorithm applied to the data assimilation cost function J using the gradient. For simplicity a two dimensional state vector space is used. Figure from Bouttier and Courtier (2002)

The most simple algorithm used for minimization is called the steepest descent method. It uses the gradient of the cost function in a straightforward way: the gradient is calculated as the direction of maximal increase and thus the steepest descent in the direction of the negative gradient. The term $f(\nabla J)$ then simply reduces to $-\nabla J$ and step size α is found using the constraint that the next gradient has to be orthogonal to the current one. However this means that algorithm will converge quite inefficiently for ill conditioned matrices with large condition numbers. The condition number $\kappa = \lambda_{\max}/\lambda_{\min}$ with eigenvalues λ_i is a measure of the sensitivity of a function to small changes in its input. Applied to the example two dimensional case the eigenvalues are the principal axes of the ellipsoid and measure the degree of ellipticity. If the condition number is big then the steepest gradient algorithm is inefficient since the step size α is limited. Sometimes this is referred to as the *narrow valley* effect (see also figure 2.5), since the minimization algorithm keeps bouncing between the steep walls (direction with large eigenvalue) before it finds the minimum.

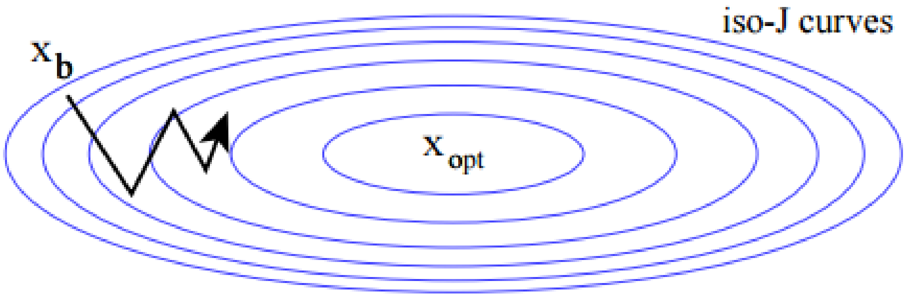


Figure 2.5: Graphical example of the narrow valley effect. The highly elliptic isosurfaces of the cost function imply a high condition number. The steep cost function makes the minimization perform a zigzag pattern which is inefficient. Figure from Bouttier and Courtier (2002)

The conjugate gradient algorithm aims to solve that problem by finding a search direction that instead is A-orthogonal to the residual. A-orthogonal refers to two vectors being orthogonal with an A matrix multiplication in between, with A being the matrix in a quadratic function like 2.12. Geometrically it can be interpreted as finding the orthogonal vector in a space stretched by matrix A. The matrix A is chosen such that the A-stretched space becomes approximately spherical. For the example in figure 2.5 it would mean the iso-J curves are transformed so that they become circles. The direction of the orthogonal vector in A-stretched space is then used to make the next increment, resulting in a significantly more efficient algorithm for ill-conditioned matrices. The conjugate gradient algorithm has been implemented in WRFDA for minimization of the cost function. For a proper introduction, full derivation and convergence analysis of the conjugate gradient algorithm the reader is referred to Shewchuk (1994).

Another family of optimization algorithms that is commonly used for cost function minimization is called quasi-Newton method. They have been named 'quasi' since they approximate the full Hessian matrix that is used in the original Newton method. The full Newton method employs a second order approximation to a multivariate non linear problem in order to evaluate the cost function at a point $\mathbf{x}+\delta\mathbf{x}$:

$$J(\mathbf{x}+\delta\mathbf{x}) \approx J(\mathbf{x}) + \delta\mathbf{x}^T \nabla J(\mathbf{x}) + \frac{1}{2}\delta\mathbf{x}^T J'' \delta\mathbf{x} \tag{2.21}$$

And consequently with a gradient:

$$\nabla J(\mathbf{x}+\delta\mathbf{x}) \approx \nabla J(\mathbf{x}) + J'' \delta\mathbf{x} \tag{2.22}$$

Where J'' represents the Hessian matrix of the cost function that contains all second order partial derivatives of the cost function evaluated around $\mathbf{x}$. For every iteration the approximated gradient 2.22 is set to zero and is solved for $\delta\mathbf{x}$. However this requires the full Hessian matrix, which is impossible to build and invert since its dimensions are in the order of $O(10^{14})$. With quasi-Newton methods attempts are made to approximate a the Hessian at every iteration. Amongst others a widely adopted quasi-Newton algorithm is the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm, named after Broyden (1970), Fletcher (1970), Goldfarb (1970), Shanno (1970). A limited memory BFGS method was developed by Liu and Nocedal (1989) that saves memory by using the history of updates of $\mathbf{x}$ and the gradient $\nabla J(\mathbf{x})$. This is particularly suited for data assimilation since there are so many variables contained in the state vector.

It has to be noted that before any of the mentioned optimization algorithms are employed the cost function has to be *preconditioned*. With preconditioning the condition number and equivalently the ellipticity of a matrix are reduced. Usually a control variable transform is performed to approximate a spherical cost function (with condition number close to 1). A more elaborate description is given in the next section (2.4.4).

2.3.4. Non linearity

These minimization algorithms assume that the cost function is quadratically related to $\mathbf{x}$. As noted before, this is the case if the observation operator H is linear. If H however is not linear, it is assumed that the process can still be approximated by a linear function, provided the first guess is already not too far from the analysis.

Often, especially with remote sensing observations, the observation operator is not linear. Since then the cost function is not convex anymore as in figure 2.4, it can have multiple local minimum values. Moreover, the non linear shape of the cost function causes the minimization of the approximated quadratic function to be suboptimal. This is illustrated by figure 2.6, where the first minimization of a quadratic function tangent at the starting point leads to an analysis increment $\delta\mathbf{x}_a^{(0)}$ that is still relatively far from the actual minimum of the non linear cost function. To mitigate the problem of non linearity, multiple *outer loops* may be performed during minimization. In these outer loops the cost function is re-linearised around the analysis $\delta\mathbf{x}_a$ and then the minimization of the cost function (inner loop) as illustrated in 2.3 is repeated using the analysis from the previous minimization as the new background, resulting in a new analysis $\delta\mathbf{x}_a^{(1)}$. The amount of outer loops that are appropriate naturally depends on the degree of non linearity. In this study whenever observations are being assimilated, two outer loops are carried out during the minimization to be able to account for some non linearity.

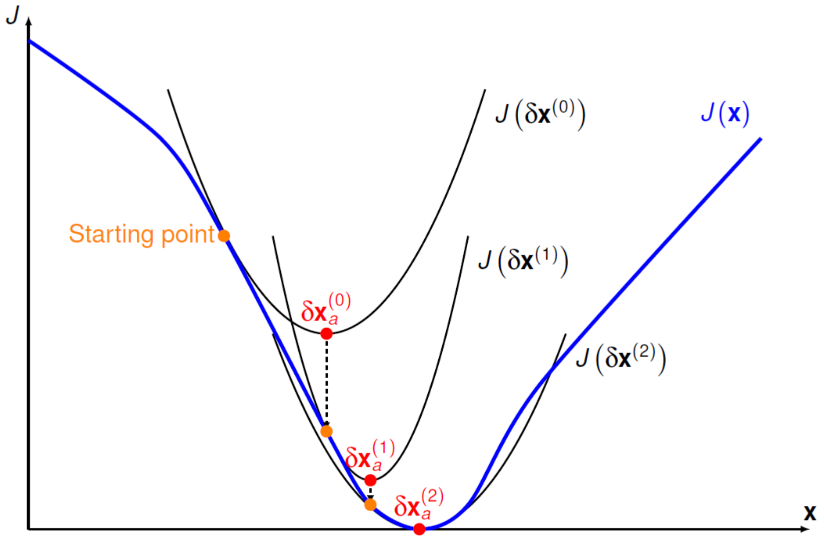


Figure 2.6: Schematic example of performing several outer loops during the minimization of the cost function $J(\mathbf{x})$. In each outer loop the analysis $\delta\mathbf{x}_a$ from the minimization in the previous outer loop is used as the background/starting point and the observation operator is linearised around the new background. Illustration by Massart (2017).

2.4. Error modelling

2.4.1. Uncertainty

In order to combine model and observations in variational assimilation their uncertainties should be known. That way the assimilation can apply appropriate weighting to the model and observations (through their error covariance matrices **B** and **R** respectively) to produce an optimal analysis. The true uncertainties are unknown and can only be estimated. It is vital that this estimate represents the uncertainty in model variables as accurately as possible. For example if the uncertainty of the model variables is overestimated, the model tends to the observations although these observations actually have a similar or worse precision than the model. On the other hand, using a model uncertainty that is really small will practically result in no observations being used to update the model.

In general an error ϵ can be defined as the difference between an estimate $\hat{x}$ and the true value x of a variable:

$$\varepsilon = \hat{x} - x \tag{2.23}$$

Every single time the variable x is estimated, a realization of the error is obtained. However almost always the true value of x is unknown. We can however make an assumption that if the estimation is repeated often enough, the mean value of all the realizations will converge to the true value of the desired variable. Also, the distribution of the error realization characterises the uncertainty of the estimate. This distribution is described mathematically with a probability density function. The probability density function (PDF) can be thought of a histogram with infinitely small bins that is normalized so that the integral of the distribution equals unity. Using the PDF a next realization of the estimate and its error can be interpreted to make a statement about the likelihood of the estimated value for the variable.

Often a probability density function is assumed to be *Gaussian*. The Gaussian can be characterized by a mean $\bar{x}$ and a variance σ_x^2.

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left\{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma_x}\right)^2\right\} \tag{2.24}$$

In the case of vectors the scalar principle from 2.24 can be extended so that a vector equivalently follows a Gaussian *multivariate* distribution:

$$f_x(\mathbf{x}) = \frac{1}{\sqrt{\det(2\pi\mathbf{Q}_{xx})}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{Q}_{xx}(\mathbf{x}-\bar{\mathbf{x}})\right\} \tag{2.25}$$

with $\bar{\mathbf{x}}$ the vector mean and $\mathbf{Q}_{xx}$ the covariance matrix of $\mathbf{x}$

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_n \end{bmatrix} \tag{2.26}$$

$$\mathbf{Q}_{xx} = \begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \cdots & \rho_{1,n}\sigma_1\sigma_n \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1,n}\sigma_1\sigma_n & \cdots & \cdots & \sigma_n^2 \end{bmatrix} \tag{2.27}$$

In 2.27 σ_n represents the standard deviation of variable x_n and $\rho_{n,m}$ the correlation between two variables. The diagonal elements in $\mathbf{Q}_{xx}$ are the variances and the off diagonal elements represent the covariances. Recall the definition of covariance:

$$Cov(x_1, x_2) = E\left[(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)\right] = \rho_{1,2}\sigma_1\sigma_2 \tag{2.28}$$

And the definition of variance that is equal to the covariance of a variable with itself.

$$Var(x_1) = Cov(x_1, x_1) = E\left[(x_1 - \bar{x}_1)^2\right] = \sigma_1^2 \tag{2.29}$$

2.4.2. Definition of errors

In data assimilation the Gaussian multivariate distribution 2.25 is used to model the error uncertainty of the model and observations. Several relevant errors can be defined. Similarly to 2.23, firstly the **background error** is defined as the difference of a background forecast state vector with respect to the true state of the atmosphere:

$$\varepsilon_b = \mathbf{x}_b - \mathbf{x}_t \tag{2.30}$$

Note that the *true state* denoted by $\mathbf{x}_t$ is not actually a state that describes the atmosphere perfectly, because a model is a discretization of reality and moreover because several parametrizations and assumptions are made to simplify real physical processes. The errors resulting from this are called *representativeness* errors. In this context the true state $\mathbf{x}_t$ refers to the best possible representation of the atmosphere by the model. Obviously the true state is not known. We can only assume that with enough realizations the average error is close to 0 (*unbiasedness*) and the covariance is described by the covariance matrix **B**, which describes the covariance between each element of the background state vector $\mathbf{x}_b$ as follows:

$$\mathbf{B} = \overline{\varepsilon_b \varepsilon_b^T} = \begin{bmatrix} \sigma_{b,1}^2 & \cdots & \rho_{1,n}\sigma_{b,1}\sigma_{b,n} \\ \vdots & \ddots & \vdots \\ \rho_{1,n}\sigma_{b,1}\sigma_{b,n} & \cdots & \sigma_{b,n}^2 \end{bmatrix} \tag{2.31}$$

Which is the equivalent of 2.27 applied to the background state vector $\mathbf{x}_b$. Besides the background error also the **analysis error** can be defined that measures the offset of the data assimilation analysis and the true state:

$$\varepsilon_a = \mathbf{x}_a - \mathbf{x}_t \tag{2.32}$$

And correspondingly the analysis covariance matrix **A**:

$$\mathbf{A} = \overline{\varepsilon_a \varepsilon_a^T} = \begin{bmatrix} \sigma_{a,1}^2 & \cdots & \rho_{1,n}\sigma_{a,1}\sigma_{a,n} \\ \vdots & \ddots & \vdots \\ \rho_{1,n}\sigma_{a,1}\sigma_{a,n} & \cdots & \sigma_{a,n}^2 \end{bmatrix} \tag{2.33}$$

The trace (diagonal) of matrix **A** contains the analysis variables' precision. It has been shown in section 2.3.2 that in the case of Gaussian distributed errors and a linear observation operator the most likely state or analysis also minimizes the variance of the analysis. An analytical expression for the analysis precision **A** has been derived in 2.18 for the case of explicitly calculating the Kalman gain matrix. Recall that is infeasible considering the matrix sizes and 3DVar instead minimizes the cost function iteratively.

Last but not least the observations that are to be assimilated also contain errors. The **observation error** is defined as the difference between observation $\mathbf{y}$ and the model equivalent observation $H(\mathbf{x}_t)$:

$$\varepsilon_o = \mathbf{y} - H(\mathbf{x}_t) \tag{2.34}$$

The observation error consists of an error that rises from the observation itself, the **instrumental error**. For example a thermometer or a pressure sensor have a certain instrumental precision that is often specified by the manufacturer or obtained from verification studies. Secondly the observation error in 2.34 contains errors produced by invoking the observation operator H, often referred to as the **forward model error**. If a temperature measurement is observed at a certain level in the atmosphere that is in between the model levels the model equivalent temperature has to interpolated from nearby model points, which introduces errors since the interpolation is based on certain assumptions. Also, **processing errors** can be introduced when pre-processing observations, for example when hydrometeors are retrieved from radar reflectivity using a certain empirical relation. Finally, **representativeness errors** generated for example in the discretization of physical processes contribute to the observation error. Note that this was not the case for the background and analysis errors, because two model states are being subtracted. Analogous to the background and analysis, the observation covariance matrix is:

$$\mathbf{R} = \overline{\varepsilon_o \varepsilon_o^T} = \begin{bmatrix} \sigma_{o,1}^2 & \cdots & \rho_{1,n}\sigma_{o,1}\sigma_{o,n} \\ \vdots & \ddots & \vdots \\ \rho_{1,n}\sigma_{o,1}\sigma_{o,n} & \cdots & \sigma_{o,n}^2 \end{bmatrix} \tag{2.35}$$

Often the observation errors are assumed to be uncorrelated. That is a valid assumption for example for observations from individual weather stations. For other observations this might not be the case. E.g. for remote sensing observations like in case of satellite images the brightness temperature of one pixel is correlated with the adjacent pixels. To mitigate the correlation between observations, they are spatially thinned in a pre-processing step or quality control step before they are being assimilated. If the observations are uncorrelated then $\mathbf{R}$ reduces to a diagonal matrix:

$$\mathbf{R} = \begin{bmatrix} \sigma_{o,1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{o,n}^2 \end{bmatrix} \tag{2.36}$$

2.4.3. Modelling of background covariance

Representing real background error covariance in the matrix $\mathbf{B}$ accurately is maybe the most challenging part of data assimilation. As the magnitude and the shape of the covariance matrix determine how much the background forecast is altered, it is crucial that the matrix is modelled as accurately as possible. More specifically, background error correlations are in important for multiple reasons:

1. **Observation spreading**
   The cross correlation between variables determines how much one variable should be changed when another is updated. When assimilating a single observation, the information is spread to neighbouring areas that are correlated. The extent of these areas are governed by the correlation length. Also in the case that only few observations are assimilated, still a large part of the background forecast can be altered because of this property.

2. **Observation smoothing**
   Contrary to the case where only few observations are assimilated, in areas where there is a dense network of observations available the degree to which observations are smoothed is governed by the covariance in $\mathbf{B}$. If there would be no cross correlation then the analysis will contain steep gradients that are physically very unlikely. The smoothing makes sure the length scales of the increments are consistent with the realistic length scales of physical quantities. For example if two nearby but different temperature measurements are assimilated with no cross correlation between the temperature of the model grid points that are close to the observations, the assimilation will produce a physically unrealistic jump in temperature. Taking into account the cross correlation both model grid points will also

be updated according to the other observation that is slightly further away. Of course this smoothing effect is a trade-off between having more detail in the analysis that might produce high frequency and physically unrealistic features versus physically consistent analysis increments which lack the detailed description of the atmosphere.

3. **Physical constraints**
   Another advantage of three dimensional variational assimilation is that additional physical constraints can be added so that analysis increments are made physically consistent with other quantities. For example if it is assumed that the atmosphere is approximately hydrostatic, then a pressure increment should lead to increments in the whole column. This is reflected by the covariances between vertical grid points for pressure. Also for example under the assumption that approximately the geostrophic balance holds a change to wind speed should be accompanied by a change in pressure in the direction orthogonal to the wind. These examples illustrate that changes can be made to any model variable that is balanced with the primary variable that is being updated by a measurement. The balance is expressed in the correlations between these variables.

2.4.4. Preconditioning: Transformation of control variables

To be able to compute the cost function and its gradient (2.14) the background error covariance matrix has to be inverted. As discussed earlier in section 2.3.3, explicitly constructing $\mathbf{B}^{-1}$ is impossible since the dimensions are just too large and inverting large matrices is really inefficient. To bypass the problem of having to invert a large matrix a **control variable transform** is applied in WRFDA. That is, transforming the control variable $\delta\mathbf{x}$ such that the matrix $\mathbf{B}^{-1}$ disappears from the cost function. The control variable transform is defined as:

$$\delta\mathbf{x} = \mathbf{U}\mathbf{v} \tag{2.37}$$

Where $\mathbf{v}$ is now the **control variable**. The matrix $\mathbf{U}$ is called the *preconditioner*. The preconditioner is the square root of the covariance matrix $\mathbf{B}$:

$$\begin{aligned} \mathbf{B} &= \mathbf{B}^{1/2}\mathbf{B}^{T/2} = \mathbf{U}\mathbf{U}^T \\ \mathbf{U} &= \mathbf{B}^{1/2} \end{aligned} \tag{2.38}$$

To demonstrate that transforming the control variable will result in the elimination of $\mathbf{B}^{-1}$, the incremental formulation of the linearised cost function is recalled:

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{c}+\mathbf{H}\delta\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{c}+\mathbf{H}\delta\mathbf{x}) \tag{2.39}$$

with

$$\mathbf{c} = H(\mathbf{x}_b) - \mathbf{y} \tag{2.40}$$

Then the control variable transform in 2.37 can be applied to the cost function 2.39, which results in

$$J(\mathbf{v}) = \frac{1}{2}(\mathbf{U}\mathbf{v})^T\mathbf{U}^{-T}\mathbf{U}^{-1}\mathbf{U}\mathbf{v} + \frac{1}{2}(\mathbf{c}+\mathbf{H}\mathbf{U}\mathbf{v})^T\mathbf{R}^{-1}(\mathbf{c}+\mathbf{H}\mathbf{U}\mathbf{v}) \tag{2.41}$$

Now it is obvious that the first term can be reduced to

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}(\mathbf{c}+\mathbf{H}\mathbf{U}\mathbf{v})^T\mathbf{R}^{-1}(\mathbf{c}+\mathbf{H}\mathbf{U}\mathbf{v}) \tag{2.42}$$

The gradient of the cost function becomes

$$\nabla J(\mathbf{v}) = \mathbf{v} + \mathbf{U}^T\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{c}+\mathbf{H}\mathbf{U}\mathbf{v}) \tag{2.43}$$

Equations 2.42 and 2.43 are used in every iteration of the minimization algorithm. That requires transforming from model to control variable space before the minimization algorithm is called and then transforming the analysis back to model variable space:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{U}\mathbf{v}_a \tag{2.44}$$

In which $\mathbf{v}_a$ represents the final analysis increment in control variable space.

The inverse background error covariance matrix $\mathbf{B}^{-1}$ has been absorbed by $\mathbf{v}$. In essence the first term has been diagonalized and the inverse error covariance matrix is equal to the identity matrix when transformed to control variable space. That way the condition number of the background term is reduced to 1. The control variable transform is designed with the idea to change to a basis which components are uncorrelated. That is realised by projecting the background error covariances onto an orthogonal basis of vectors that are uncorrelated by definition.

Then of course the challenge of modelling $\mathbf{B}$ is shifted to how $\mathbf{U}$ is constructed. $\mathbf{U}$ in WRFDA is built up using a series of transformations:

$$\mathbf{U} = \mathbf{U}_p \mathbf{U}_v \mathbf{U}_h \tag{2.45}$$

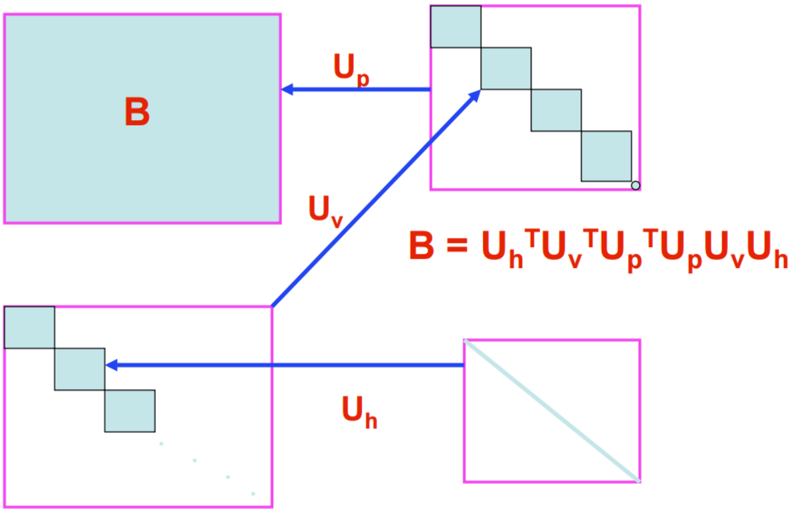


Figure 2.7: Schematic illustration of how the background error covariance matrix $\mathbf{B}$ is diagonalized by employing the physical $\mathbf{U}_p$, vertical $\mathbf{U}_v$ and horizontal $\mathbf{U}_h$ transformations. By performing these transformations $\mathbf{B}$ is diagonalized step by step. In the top right rectangle 4 bigger blocks and 1 smaller block illustrate that the surface pressure p_s (which is often both a model and control variable) is a smaller array since it only represents a 2D field instead of the other control variables which are 3 dimensional. Image from Rizvi (2012).

Where $\mathbf{U}_p$, $\mathbf{U}_v$ and $\mathbf{U}_h$ represent a physical, vertical and horizontal transformation. These operations transform the model variables used in data assimilation (u, v, T, q, p_s) contained by model vector $\mathbf{x}$ to the eventual amplitudes of the empirical orthogonal functions (EOF's) of control variables in $\mathbf{v}$. First, (the inverse of) $\mathbf{U}_p$ converts from model space (u, v, t, q, p_s) to a set of control variables. These control variables are assumed to be independent and thus no correlation with each other. In algebraic terms, this equates to block diagonalization of matrix $\mathbf{U}$. The blocks that remain on the diagonal represent the spatial autocorrelation of each variable that is left. This is illustrated in figure 2.7. Applying the vertical transformation $\mathbf{U}_v$ further reduces the off diagonal elements, namely the correlation of all elements that are not both on the same model level, for example the correlation of model humidity variable q at two grid points on model level 1 and 2. Finally, to arrive at a diagonal matrix as illustrated in the last box of figure 2.7, a horizontal transformation is applied so that the horizontal autocorrelation is compressed.

Physical transformation

Different sets of control variables may be chosen. A commonly used combination of control variables in data assimilation and especially for WRFDA are the stream function ψ, unbalanced velocity potential χ_u, unbalanced temperature T_u, pseudo-relative humidity RH_s and unbalanced surface pressure $P_{s,u}$. It is pseudo in the sense that the relative humidity is with respect to the saturation of the background forecast. Also, the term “unbalanced” refers to the residual of each variable after a statistical balance to the stream function is removed. The first part of the physical transformation exists of transforming from model variables (u, v, T, q, p_s) to control variables. In case (ψ, χ_u, T_u, RH_s, $P_{s,u}$) are used, the transformation matrix is:

$$\begin{bmatrix} u \\ v \\ T \\ q \\ P_s \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{u,\psi} & \mathbf{C}_{u,\chi} & 0 & 0 & 0 \\ \mathbf{C}_{v,\psi} & \mathbf{C}_{v,\chi} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{C}_{q,RH} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \psi \\ \chi \\ T \\ RH_s \\ P_s \end{bmatrix} \tag{2.46}$$

Where $\mathbf{C}_{u,\psi}$, $\mathbf{C}_{u,\chi}$, $\mathbf{C}_{v,\psi}$ and $\mathbf{C}_{v,\chi}$ map the variables indicated by their subscripts to each other. $\mathbf{I}$ represents the identity matrix. How the mapping is performed exactly is beyond the interest of this research. For a full description of the mapping to stream function and velocity potential the reader is referred to Barker et al. (2003). The unbalanced form of the control variables is obtained by the following transformation:

$$\begin{bmatrix} \psi \\ \chi \\ T \\ RH_s \\ P_s \end{bmatrix} = \begin{bmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ \mathbf{C}_{\chi,\psi} & \mathbf{I} & 0 & 0 & 0 \\ \mathbf{C}_{T,\psi} & 0 & \mathbf{I} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{I} & 0 \\ \mathbf{C}_{Ps,\psi} & 0 & 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \psi_u \\ \chi_u \\ T_u \\ RH_s \\ P_{s,u} \end{bmatrix} \tag{2.47}$$

Matrices $\mathbf{C}_{\chi,\psi}$, $\mathbf{C}_{T,\psi}$ and $\mathbf{C}_{Ps,\psi}$ are regression matrices of unbalanced velocity potential, temperature, and surface pressure with the stream function. Equation 2.47 shows how for example the temperature is modelled as a combination of a balanced part with the streamfunction and an independent, unbalanced part. Additional correlations might be added to this matrix if there are reasons to assume that variables are correlated. It is important to recognise that adding any correlation between variables means that an increment in one variable will lead to changes in other variables. For example, if the transformation matrix in 2.47 is used, an innovation in the wind component u will result in a change in the streamfunction and velocity potential (2.46), while the change in the streamfunction in turn will result in a change in surface pressure as well since they are correlated according to $\mathbf{C}_{Ps,\psi}$ in 2.47. This can be regarded as the implementation of imposing the geostrophic balance on any changes applied to the model state. Remember that this is one of the added values of variational data assimilation mentioned in section 2.1 and 2.4.3.

A recent study by Sun et al. (2016) investigated the effects of using different control variables on data assimilation using WRFDA and assessed the impact on precipitation forecasting. The control variables commonly used in data assimilation for momentum are the pairs streamfunction/velocity potential (ψ, χ), vorticity/divergence (ζ,δ) and the east and west horizontal wind components (u, v). Often times (ψ, χ) and (ζ,δ) have been used for data assimilation in global and synoptic scale models, primarily for their ability to impose physical balances (mainly the geostrophic one) mentioned earlier. In their study they refer to Xie and MacDonald (2012) who conducted numerical experiments to show that using the (ψ, χ) momentum control variable produces errors since (ψ, χ) preserves the integral values of the wind so that unrealistic increments are made when the observation network is sparse or inhomogeneous. Also, they note that for smaller, regional domains the lateral boundary condition specification for (ψ, χ) in the estimation of model background error covariance matrix adds to these errors. Moreover, it tends to capture long wave information from the background forecast, while short wave information from observations is taken. The opposite is the case for the (u, v) pair, which suggest observations in a (u, v) framework have a longer lasting effect. Based on these findings from Xie and MacDonald (2012), Sun et al. (2016) question if using the (ψ, χ) still adds value to the data assimilation, when applied in a smaller scale regional, high resolution convection permitting domains when a dense set of observations is assimilated. Large scale balances like the geostrophic assumption might be less valid for these small domain models that focus on enhancing short term precipitation estimation in which more vertical motions are involved. These regional models rely on high resolution observations from e.g. radars to provide information on small scale processes which might not come forward in the data assimilation analysis when using the stream function and velocity potential momentum control variables. To investigate these hypotheses Sun et al. (2016) conducted experiments comparing the use of horizontal wind components (u, v) and the (ψ, χ) pair momentum control variables for 7 cases of forecasting convective events of which one also included data assimilation of radar reflectivity and radial velocity of eight NEXRAD radars. They found that the correlation between (u, v) is minimal compared to (ψ, χ) and therefore may be used as independent control variables. Also, the correlation of (u, v) with other variables like the surface pressure was found to be insignificant, except for some correlation with temperature and relative humidity on lower levels which was chosen to be neglected. On the contrary, the (ψ, χ) momentum control variables did show correlation with other variables, justifying the use of multivariate correlation as in 2.47 when (ψ, χ) are used as momentum control variables. From their experiments they concluded that (ψ, χ) increased length scale and decreased variance for u and v, having a negative impact on the analysis velocity field. Because of this smoothing behaviour of (ψ, χ) they also found a closer fit to radar wind velocity observations using (u, v) as control variables, as they are better in representing small scale features. Finally, the (u, v) control variables improved precipitation forecasting skill, also for the cases without radar data assimilation.

From the studies conducted by Xie and MacDonald (2012) and Sun et al. (2016), the choice is made to use the set (u, v, t, RH_s, p_s) as control variables for data assimilation, since the focus of this thesis is on short term precipitation forecasting on a small regional domain rather than producing a synoptic data assimilation system which captures mainly large scale motions. By taking (u, v) as momentum control variables, the first part of the physical transform (compare to 2.46) is reduced to:

$$\begin{bmatrix} u \\ v \\ T \\ q \\ P_s \end{bmatrix} = \begin{bmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{C}_{q,RH} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} u \\ v \\ T \\ RH_s \\ P_s \end{bmatrix} \tag{2.48}$$

And the second part of $\mathbf{U}_p$ reduces to the identity matrix, because no cross correlation exists between the control variables:

$$\begin{bmatrix} u \\ v \\ T \\ RH_s \\ P_s \end{bmatrix} = \begin{bmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} u \\ v \\ T \\ RH_s \\ P_s \end{bmatrix} \tag{2.49}$$

This is how the physical transformation $\mathbf{U}_p$ is performed in the evaluation of the cost function and the computation of the analysis increments in physical space. Although no thorough study was conducted comparing different sets of control variables, one experiment has been carried out to verify the choice of (u, v) over (ψ, χ). In this experiment radar data has been assimilated using the two sets of control variables. The forecast after assimilation with control variable set (u, v) resulted in a higher forecasting skill, producing a more accurate structure of rainfall when compared to actual radar images. Therefore, in the remainder of this report only the (u, v) momentum control variable pair will be considered in any data assimilation experiment.

Vertical transformation

The vertical transformation is carried out by decomposing the vertical error covariances into empirical orthogonal functions (EOF's). Suppose for a single vertical column a matrix $\mathbf{B}_v = \varepsilon_b {\varepsilon_b}^T$ describes the error covariance of k vertical levels, where the vector $\varepsilon_b = (\varepsilon_1, \cdots, \varepsilon_k)$ contains the error at every model level. Then a decomposition can be performed into eigenvectors $\mathbf{E}$ and eigenvalues Λ:

$$\mathbf{B}_v = \mathbf{E}\Lambda\mathbf{E}^T \tag{2.50}$$

EOF's are essentially the k eigenvectors (columns of $\mathbf{E}$) of the vertical error covariance matrix $\mathbf{B}_v$ and they are determined so that the first eigenvector explains maximum total error variance, the second the maximal of the residual variance and so on for the remaining eigenvectors, of which there are a total of k. In general it is difficult to physically interpret these eigenvectors. They can be considered natural modes of variation, however physical modes are not necessarily orthogonal. The eigenvectors are just geometrical structures aimed to provide a more efficient basis of vectors to describe data and do not need to have a physical meaning. They merely show patterns explaining the observed variation. The determination of the eigenvectors is obviously dependent on the observations of variability. So even for the same geographical area using different statistics (e.g. from different times) will result in different eigenvectors, although it is assumed that this effect is somewhat mitigated by averaging the variability over an extensive time period like a month. The matrix Λ in 2.50 is a diagonal matrix containing the k eigenvalues λ_i which are proportional to the amount of variance that the corresponding eigenvector explains, that is:

$$Var_{\text{exp}} = \frac{\lambda_i}{\sum\limits_{i=1}^{k} \lambda_i} \cdot 100\% \tag{2.51}$$

Any increment $\delta\mathbf{x}$ in the data assimilation variables introduced by observations will be projected onto the eigenvectors. The corresponding eigenvalues determine the relative weight of the increments from every eigenvector when the cost function 2.41 is computed. In the next section it is explained how the eigenvectors are estimated along with some plots of the eventual estimates.

Horizontal transformation

The horizontal transformation $\mathbf{U}_h$ models the horizontal autocorrelation of a model variable. It is assumed that the autocorrelation is isotropic, meaning that the change in autocorrelation is the same for every direction and just depends on distance. Another assumption is that the autocorrelation is homogeneous, that is being constant throughout the domain and not location dependent. A visual representation of these assumptions will be given later in section 2.8 where single observation tests are performed to assess the spatial pattern of increments resulting from assimilation of a single observation. The modelling of autocorrelation is implemented in WRFDA by means of a recursive filter. It involves specifying the covariance of the coefficients of the eigenvectors/EOF's. The covariance is said to decay exponentially with distance:

$$B(r) = B(0)\exp\left(-\frac{r^2}{8s^2}\right) \tag{2.52}$$

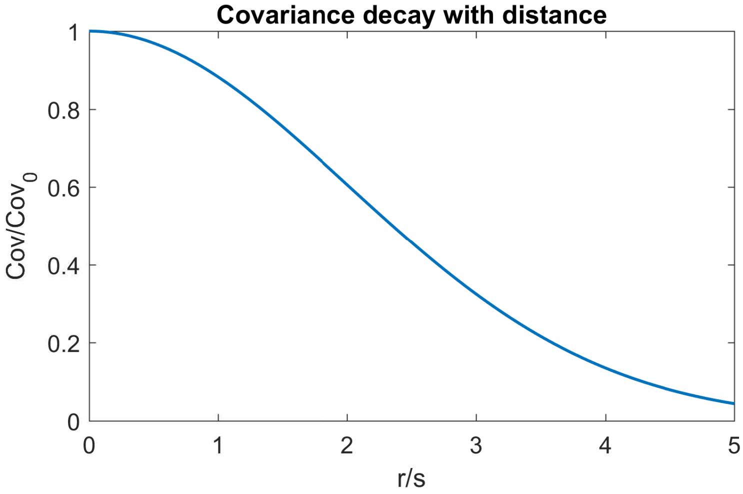


Figure 2.8: Exponential decaying function that is used to estimate horizontal length scales of the amplitudes of the EOF's by fitting the function to computed covariances binned in distance intervals. The ratio of covariance at distance r and and distance 0 is shown as a function of the ratio of distance to length scale s.

Where $z(r)$ is the autocovariance of the coefficient of the eigenvector that depends on distance r only, and is thus isotropic. $z(0)$ is the autocovariance at zero distance, i.e. the variance of the coefficient of the EOF's. The characteristic horizontal length scale s is estimated by fitting the exponential function 2.52 to the experimental covariances. This length scale determines how quickly a variable decorrelates with distance. At $r = s$, the covariance has reduced to 88% of $z(0)$ and 50% reduction occurs at $r = 2.35s$. Refer to figure 2.8 to see how the covariance decays with ratio r/s. When the actual data assimilation is employed, the length scale can be adjusted to allow some liberty in determining the extent of increments resulting from assimilation of observations. This will be discussed in more detail in section 2.4.5 and 2.8.

For a detailed description of the recursive filter in employed in the horizontal transform and also of the practical implementation of the physical and vertical transforms, refer to the technical notes of the control variable transformation in WRFDA from Barker et al. (2003).

2.4.5. Estimating background error covariance: the NMC method

A widely used technique throughout the variational data assimilation community to estimate background forecast error covariance is the so called National Meteorological Center (NMC) method, originally developed by Parrish and Derber (1992). It owes its name to the National Meteorological Center that is now known as the National Centers for Environmental Prediction. This method uses an ensemble of forecast differences between forecasts of different length, however valid for the same time to estimate the model error:

$$\varepsilon = \hat{\mathbf{x}} - \mathbf{x} = \mathbf{x}_{t_1} - \mathbf{x}_{t_2} \tag{2.53}$$

It is common to take the difference between a 12 hour forecast and a 24 hour forecast valid for the same time and domain. For example, at 30-08-2016 12 UTC a 24 hour forecast initialized at 29-08-2016 12 UTC and a 12 hour forecast initialized at 30-08-2016 00 UTC are subtracted for one realization of the model error. The error covariance is then estimated from the statistics of these differences:

$$\mathbf{B} = \overline{(\mathbf{x}_{24} - \mathbf{x}_{12})\,(\mathbf{x}_{24} - \mathbf{x}_{12})^T} \tag{2.54}$$

Of course one could argue for other forecast lengths, but setting the length much shorter to 6 hours or less might result in non representative errors since the spinup time (see glossary) of a model can also be several hours. Therefore, by using a 12 hour forecast length it is implicitly assumed that these forecasts are fully representative of the model's capabilities. Other common values are 24 and 48 hour forecasts, however bearing in mind that for most data assimilation applications and nowcasting focus on short term weather prediction (e.g. the 0-12 hour range), statistics generated from forecast differences between 24 and 48 hour forecasts are not that representative anymore. In general uncertainty increases with forecast time and thus so will the error variance of the model variables, leading to an overestimation of background error covariance when the covariances are used for data assimilation at an earlier point in the forecast (e.g. +3h or +6h). Of course this also a consequence of using a static, climatological background error covariance matrix to describe the error covariance. To somewhat mitigate the misrepresentation error WRFDA allows scaling both the magnitude and correlation length of the covariances. That way the background error covariance matrix can be tuned to what is appropriate according to the users judgement. Some examples of this scaling are given in 2.8, where single observation assimilation tests are used to illustrate how background error covariance matrix affects the changes in model variables.

It is common to generate the climatological statistics over quite an extensive time period, preferably a month or longer, to obtain an error sample large enough to produce a valid estimate for **B**. In this study the background error covariance is estimated from 60 forecast differences, with the 12 and 24 hour forecasts valid for 00 and 12 UTC during 30 days ranging from May 24^{th} to June 23^{rd} in 2016. The forecast differences at each time are then transformed to control variable space using the inverse of the physical transformation defined in 2.48 and 2.49, described in the previous section 2.4.4. After the mean for every vertical model level is removed, the errors are averaged horizontally for every model level k to obtain a domain averaged vertical background error, for every time step. These vertical background errors are then averaged over time to obtain the final time and domain averaged vertical background error covariance matrix $\mathbf{B}_v$. As discussed in the previous section, EOF's are computed from an eigenvalue decomposition according to 2.50. The first five (of the k total) eigenvectors resulting from this decomposition have been plot in figure 2.9. These vectors explain the largest part of the variances in the vertical background error covariance. The amount of variance that is explained is based on the magnitude of the eigenvalues corresponding to the eigenvectors (see 2.51). The eigenvalue magnitude of each eigenvector has also been plot in figure 2.10. The eigenvalues decrease with vertical mode for every control variable, as expected from the definition of the eigendecomposition. The fractions of the first five eigenvalues of the total have also been included in table 2.1. The table shows the percentage of variance that is explained by the first five eigenvectors for each control variable, except surface pressure since it is not three dimensional. For all control variables, about 75% of the total variance in $\mathbf{B}_v$ is explained by the first five eigenvectors in figure 2.9, except for temperature for which however the first five eigenvectors still account for a large part (66%) of the error variance. This also proves the added value of an eigendecomposition: a set of five vectors is able to describe approximately 75% of the variation that would normally be described by 40 vectors.

From the top two graphs in figure 2.9 showing the eigenvectors for wind components u and v, it is observed that a reasonable part of the error variance is quite evenly distributed over the model levels, since the magnitude of the first (black) eigenvector that explains most of the vertical error variance is relatively stable over the vertical levels, with a slight peak around level 31. Most likely differences in positioning of the jet stream that corresponds to these higher model levels causes the variability in horizontal wind speeds to be largest here. The second mode (red) represents most of the variation at lower model levels for u and v. The remaining (third to fifth) eigenvectors have their peak at higher model levels, indicating a large part of total horizontal wind speed variability is in the upper areas of the troposphere. For temperature the pattern is different, see the lower left figure in 2.9. Most variability in the error for temperature is close to the surface, given the first mode only has significant magnitude at the lowest model levels. This is natural since the

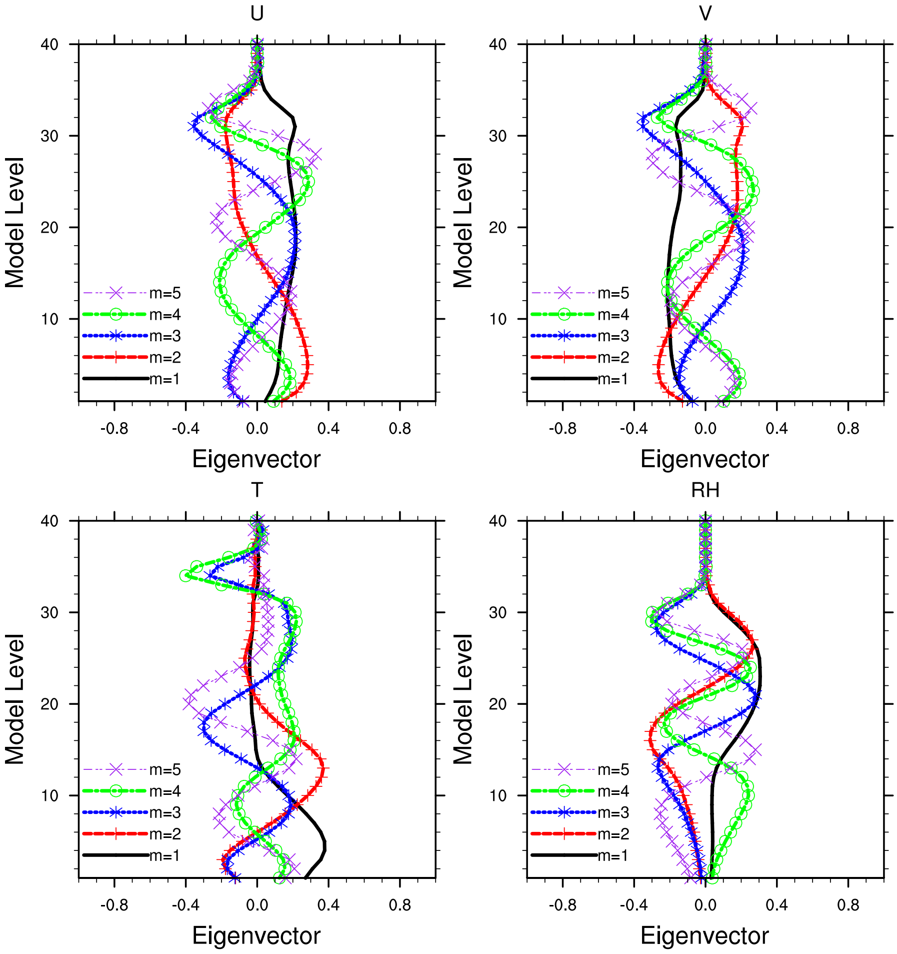


Figure 2.9: First five eigenvectors corresponding to the five largest eigenvalues (see figure 2.10). Top left shows the eigenvectors of the horizontal west-east wind component u, top right of horizontal north-south wind component v, bottom left of temperature T and bottom right the relative humidity RH.

surface is heated by incoming radiation which might vary considerably between model forecasts, for example because of varying cloud cover. The next eigenvector (second, red) has a peak just slightly above the first eigenvector. The same story applies to the following eigenvectors, illustrating that the temperature variability generally decreases with height. A sign change (positive/negative) for example for the second (red) eigenvector of temperature around the sixth model level indicates some negative correlation between the levels above and below exists for this eigenvector. Thus, if for example only the second temperature eigenvector makes a positive change at model level 12, it will at the same time cause a negative temperature increment near the surface. Finally, the five eigenvectors for relative humidity show that most of the uncertainty is in the mid to higher regions in the troposphere. The first mode has its peak around level 25, while showing little to no variability at model levels below 11. The same is true to a lesser extent for the eigenvectors that follow. The high variability in mid to upper model levels is likely amplified in the summer when more and heavier convective events occur, transporting heat and moisture from the surface to high levels in the troposphere. If forecasts from a winter period would have been used to generate the background error statistics, possibly the moisture variability will be slightly smaller at these higher levels, but still significantly larger than at the surface. At the surface the model is better able to forecast humidity, partly because nearly all humidity observations are made at the surface while they are sparse at higher levels in the atmosphere. The eigenvalues and eigenvectors estimated from the summer month forecast differences are stored and used in the computation of the cost function (2.41) and its gradient (2.43).

The last part in the estimation of the background error covariance matrix involves the estimation of the horizontal covariance. Again, refer to Barker et al. (2003) for a detailed description of how the horizontal covariance is estimated. Similar as for the estimation of the eigenvalues for the vertical background error covari-

	U	**V**	**T**	**RH**
1	24.1	24.7	32.0	24.8
2	19.6	19.2	11.8	17.4
3	15.2	15.2	8.5	12.2
4	9.4	9.3	7.6	9.2
5	6.9	6.7	5.9	7.3
Sum	75.2	75.1	65.8	70.9

Table 2.1: Percentage [%] of variance explained of the first 5 eigenvectors determined from eigendecomposition of the vertical background error covariance matrix. Top left shows the eigenvalues of the horizontal west-east wind component u, top right of horizontal north-south wind component v, bottom left of temperature T and bottom right the relative humidity.

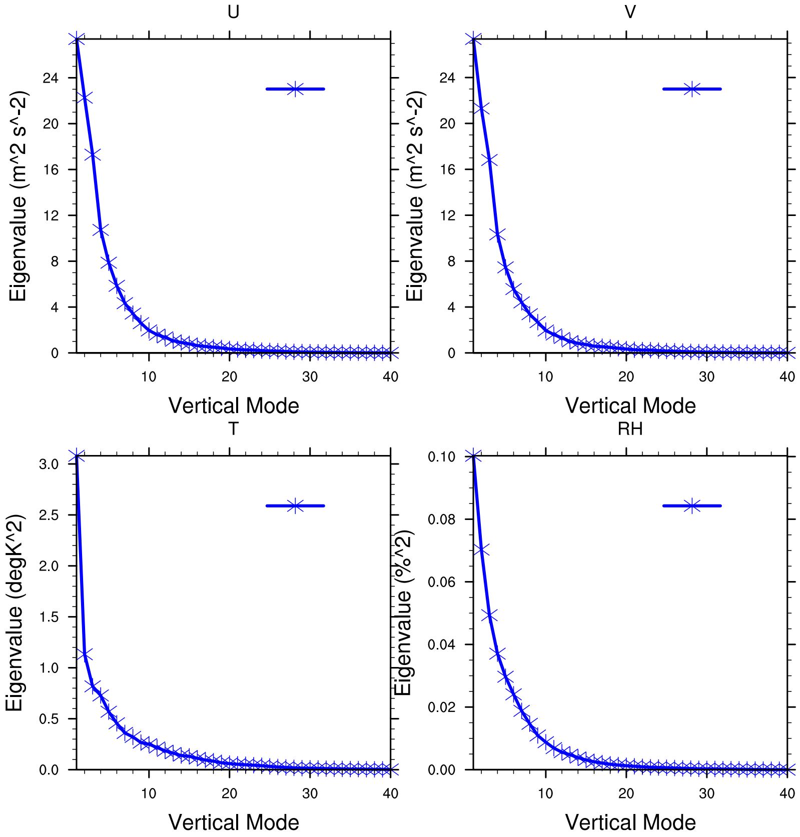


Figure 2.10: Eigenvalues corresponding to all the k vertical modes resulting from the eigendecomposition of the vertical background error covariance matrix. Top left shows the eigenvalues of the horizontal west-east wind component u, top right of horizontal north-south wind component v, bottom left of temperature T and bottom right the relative humidity RH.

ance, the forecast differences are computed for every time step and transformed to control variables via $\mathbf{U}_p{}^{-1}$, so that the forecast differences now have a vertical coordinate in terms of the eigenvector number rather than vertical model level. The control variable differences are then projected onto the vertical eigenvectors using the previously determined vertical transformation $\mathbf{U}_v{}^{-1}$. Afterwards, the two dimensional average is subtracted for the projected control variable differences, for every horizontal field. The product of the projected and zero centralized control variable differences is taken to obtain an estimate of the covariances which are then binned based on the 2D-distance separating two points, essentially building a variogram of horizontal covariances for every eigenvector. For every range bin of the horizontal covariances, the mean is calculated to obtain the mean covariance of the control variable differences. Then, the exponentially decaying function defined in 2.52 is fit to the mean covariances in the variogram, resulting in an estimated horizontal covariance length scale s for every eigenvector. These length scales have been plot for u, v, T and RH again in figure 2.11. Generally speaking, the covariance length scale is largest for the first eigenvectors for all control variables, except for the first two eigenvectors of the temperature. That means the magnitude of the first eigenvectors

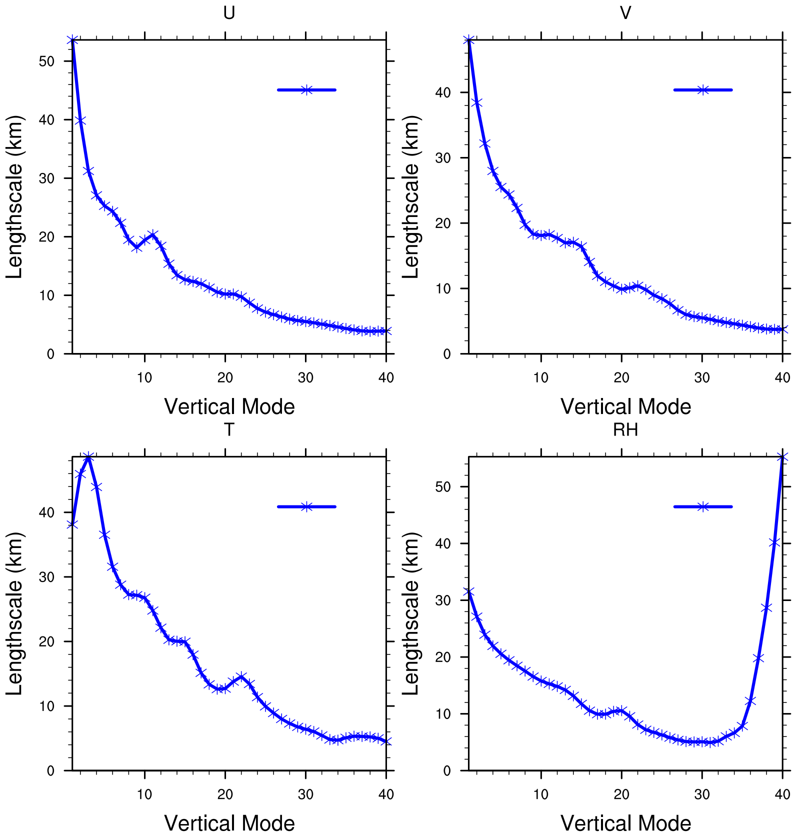


Figure 2.11: Horizontal length scales for every vertical eigenvector k. Top left shows the length scales of the horizontal west-east wind component u, top right of horizontal north-south wind component v, bottom left of temperature T and bottom right the relative humidity RH.

will propagate to a larger area than following eigenvectors. Only for the first two eigenvectors of temperature length scale increases, although not significantly. Physically it means that temperature increments from the second and third eigenvector of the temperature, that correspond to vertical model levels just above the vertical model levels of the first eigenvector, will spread over a larger distance than increments resulting from the first eigenvector. For relative humidity (bottom right) a sudden increase of the length scale is observed for the last eigenvectors. There is no straightforward explanation at hand, other than that the last part of residual variance (noise) apparently has a large correlation length. It has to be noted however that these last five eigenvectors combined explain 0.3% of the total variance, which means the effect of these eigenvectors and corresponding length scales on any model increment is negligible. Interestingly enough, the length scales of all variables in figure 2.11 are quite similar, in the order of tenths of kilometers. This is still quite big, bearing in mind that the covariance is reduced with 50% only at $2.35s$. As noted earlier, WRFDA enables scaling these length scales in order to obtain more realistic increments. In section 2.8 some PSOTs will be conducted to show the magnitude and extent of assimilating a single observation and the effects of adjusting the length scale.

In reality it remains a challenge to obtain a valid estimate for the background covariance. In the first place because the uncertainty is unique for every forecast, governed by the weather regime that is present at the time of assimilation. In the summer for example convective events can cause more variability between model runs. However from day to day or even between subsequent model runs the spatial pattern of error covariance can differ. Consider for example a squall line where in the direction of propagation of the squall line the variability of temperature and pressure at some distance from the front is quite small. However along the boundary of the front locally large variations can exist in humidity, temperature et cetera, depending on the timing of the boundary and how well that is captured by the model. Ideally this should be reflected in the covariance matrix so that observations along the squall line have a bigger impact in the assimilation. This requires a case specific or *flow-dependent* background error description. The so called 'error of the day' approach can be realised by using an **ensemble** of forecasts to estimate the variability in this specific forecast. The ensembles are generated by making small perturbations to a reference forecast, letting the model take its course for every ensemble member and then estimating the variance at a certain time from the spread between the ensembles. This way the actual weather regime will govern the pattern of the error variance. Besides a perturbation of the model it is also common to initialize ensemble members using different parametrizations, e.g. different microphysics or boundary layer schemes, to represent the uncertainty of the model as a result of parametrization. Or in some cases even an ensemble of various NWP models could be used. Regardless of how the ensembles are generated, a sufficient amount of ensembles is required to obtain a valid estimate of the covariance. Logically, there has to be a trade-off between ensemble size (the representativeness of the statistics) and computing power as possibly a lot more forecasts have to be made. In general, the institutions that operationally employ ensemble data assimilation generate ensembles with at least 20 members (see Bannister (2016) for an overview of (ensemble) data assimilation methods currently employed by leading NWP centres). Still, the limited amount of ensemble members to estimate model covariance will lead to sampling errors, so that the statistics derived from the ensembles are not necessarily representative for the actual covariance. For example sampling errors include unrealistically strong covariance between points separated by a large distance. This problem can be mitigated by applying **localization** (Gaspari and Cohn (1999)), where covariances are damped more with increasing separation distance. In essence, this follows the same rationale as the horizontal transformation $\mathbf{U}_h$ that is applied to any control variable increment, where it assumed that covariance decreases with distance according to the exponential decaying function in 2.52. Another sampling error occurs when covariances are underestimated (Houtekamer et al. (2001)) in case the spread in the ensembles is not big enough, as a result of e.g. too small perturbations or no variation in model settings. The ensemble spread can be artificially increased by **ensemble inflation** (e.g. Anderson et al. (1999)). However, because of the ad hoc nature of these approaches it is difficult to obtain consistently good estimates of ensemble covariance, as for instance the inflation factor is unique for every forecast.

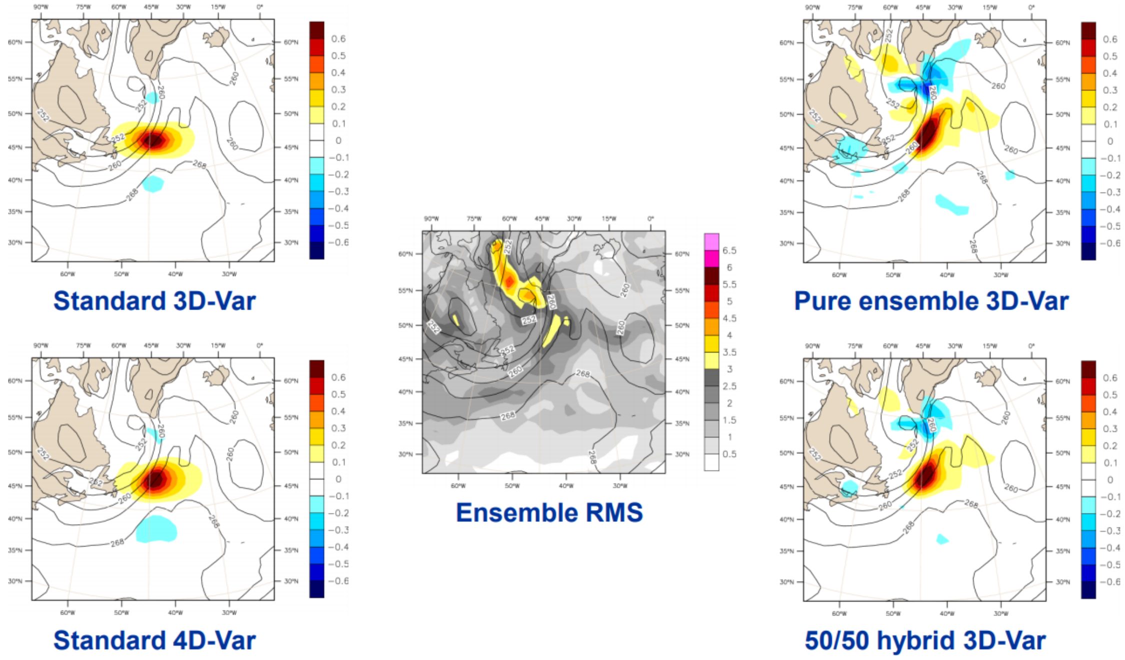


Figure 2.12: Horizontal wind component u increments in the data assimilation analysis resulting from a single wind speed observation, using the various types of data assimilation. In standard variational data assimilation, increments are based only on the climatological covariance matrix. Pure ensemble methods use an ensemble derived covariance matrix, resulting in increments that better follow the forecast specific variance (as illustrated by the ensemble RMS). In hybrid data assimilation the increments a combination of the climatological and ensemble based covariance is used, so that the analysis increments show features of both methods. Illustration from Clayton et al. (2011).

Hybrid methods have been developed that combine the static and flow-dependent description of background error. That way the data assimilation can incorporate forecast specific variability related to the current weather situation, while some robustness is provided by the static covariance matrix derived from climatology that is used in pure variational data assimilation. Usually, the covariance matrix used in hybrid methods is a weighted combination of the static covariance matrix $\mathbf{B}_s$ and ensemble based covariance matrix $\mathbf{B}_{ens}$:

$$\mathbf{B} = a\mathbf{B}_s + (1-a)\,\mathbf{B}_{ens} \circ \mathbf{L} \tag{2.55}$$

The factor a determines the weights and $\mathbf{L}$ is the localization matrix that relaxes the ensemble covariances to zero for increasing distances. The way that hybrid data assimilation is actually implemented is slightly different from 2.55 and is usually done by extending the control variables or in other words, adding new terms to the cost function that represent the ensemble part in the assimilation. A visual example of employing hybrid data assimilation is given in figure 2.12, where data assimilation analysis have been plot resulting from assimilation of a single wind observation, for different types of variational data assimilation. Note how hybrid data assimilation combines the structure of analysis increments from the static (3DVar/4DVar) and flow-dependent (ensemble derived) error description. Hybrid data assimilation is now the main type of data assimilation that is employed operationally by leading NWP institutions (Bannister (2016)), for its ability to provide an adaptive yet robust representation of model covariance.

Finally, in 4DVar, the increasing uncertainty of a NWP model with time can be accounted for. Observations assimilated at later times in the forecast then obtain a relatively higher influence, as the model precision reduces with forecast length. In the most basic 4DVar framework the static background error at the beginning of the assimilation window is propagated through time by the linearised NWP model. So called hybrid En4DVar applies the same algorithm, however with the difference that the background error is now a combination of $\mathbf{B}_s$ and $\mathbf{B}_{ens}$ as in 2.55. Hybrid 4DEnVar on the other hand does not estimate the error covariance at later times through error propagation of the covariance at the start, but directly from the ensemble. This avoids the need to linearise the forecast model so that this form of 4D assimilation can be an order of magnitude cheaper than ordinary 4DVar (Caron et al. (2015)). For a more in depth analysis of the various combinations of hybrid and 4D assimilation methods, refer to Bannister (2016).

2.5. Radar reflectivity assimilation

In data assimilation, any observation can be assimilated as long as it can be related to a model variable. In this section it will be discussed how a reflectivity observation from a weather radar can be integrated into a weather model. First it will be explained what a radar reflectivity measurement is and how it links to rainfall rate, which is the most commonly associated quantity with radar reflectivity. Moreover, it will be used in the verification of the model rainfall prediction as well. Then it is discussed how radar reflectivity can actually be linked to model variables, requiring some additions to the cost function. Also, the assimilation of humidity observations derived from radar will be addressed.

2.5.1. Definition of radar reflectivity and rainfall rate

Weather radars send electromagnetic signals into the atmosphere and measure the backscattered radiation. The amount of radiation that is backscattered is called the reflectivity factor z and it is defined as the summation of the diameters to the sixth power of all droplets within a unit volume of $1m^3$, called the reflectivity factor:

$$z = \sum_{volume} D_i^6 = \int_0^\infty D^6 N(D)\,dD \quad \left[mm^6 m^{-3}\right] \tag{2.56}$$

Where D is the diameter of a droplet contained in the volume. The definition in 2.56 assumes the particles causing a reflection are in the Rayleigh scattering regime, where the amount of scattering is strongly wavelength dependent. In this regime the size of the particles are considerably smaller than the wavelength of the radar. For example the Dutch weather radars from KNMI operate at a wavelength $\lambda = 5.3\ cm$ that is significantly bigger than the common hydrometeors present in the atmosphere, except for some extreme cases like very large hailstones. A logarithmic scale is preferred because z spans such a big range. Therefore, Z, known as the radar **reflectivity**, is defined:

$$Z = 10\log_{10}\left(\frac{z}{z_0}\right) \quad [dBZ] \tag{2.57}$$

Where $z_0 = 1\ mm^6 m^{-3}$ is the reflection of a single droplet with $1mm$ diameter which is used as a reference. $N(D)$ in 2.56 is the drop size distribution (DSD), expressed in the number of droplets per unit volume

$\left[m^{-3} mm^{-1}\right]$. The DSD is unique for every rain event, but is often modelled for example with an exponential distribution as formulated by (Marshall and Palmer, 1948):

$$N(D) = N_0 \exp(-\Lambda D) \tag{2.58}$$

In equation 2.58, N_0 represents the intercept parameter and Λ the slope parameter, respectively. In the case of single (horizontal) polarization reflectivity measurements, the intercept parameter is taken to be constant so that rainfall rate can be determined by a single reflectivity measurement. A value of $N_0 = 8000\ m^{-3} mm^{-1}$ has been found by the same Marshall and Palmer (1948) by measurement fit. The empirical expression for the slope parameter varies with rainfall rate R $\left[mm\, h^{-1}\right]$:

$$\Lambda = 4.1 R^{-0.21} \quad \left[mm^{-1}\right] \tag{2.59}$$

Figure 2.13 shows how the DSD looks like for different rainfall rates. The graph gives an example of how the DSD varies for situations with different rainfall rates.

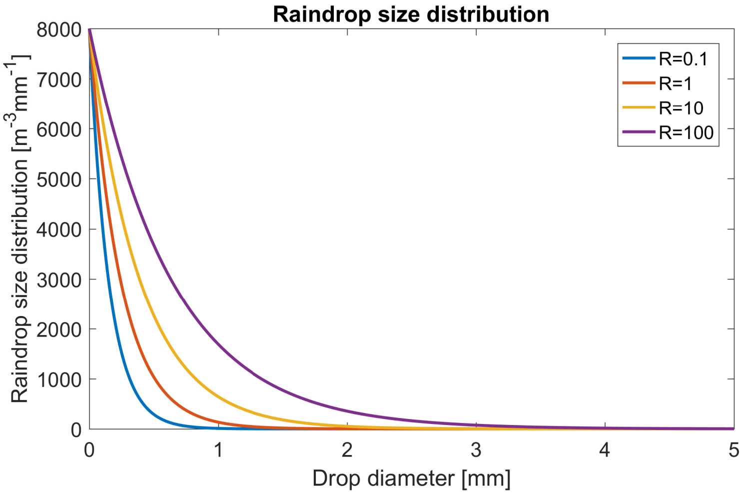


Figure 2.13: Drop size distribution for different rain rates using the exponential distribution of Marshall and Palmer (1948).

The other way round, the **rainfall rate** is a function of the DSD. By definition, the rainfall rate is the total volume of water passing a unit area per second:

$$R = 6\pi \cdot 10^{-4} \int_0^\infty D^3 v(D)\, N(D)\, dD \quad \left[mm\, h^{-1}\right] \tag{2.60}$$

This expression for rainfall rate assumes spherical droplets and that no vertical wind or turbulence is present that influences the terminal fall velocity v, so that it only depends on the diameter. By taking a certain drop size distribution and terminal fall velocity, the relationship between rainfall rate and reflectivity is generalized with a power law relationship (also referred to as $Z-R$ relation) with prefactor a and exponent b:

$$Z = aR^b \tag{2.61}$$

These constants a and b (and consequently the DSD) are actually different for every precipitation event. The type of precipitation that occurs (like convective, orographic or stratiform) governs the drop sizes and their distribution, for which a certain pair of constants exist. Battan (1973) provides a comprehensive list for different precipitation types as a result of several studies that have used regression analysis fitting equation 2.61 to measured rainfall rates and reflectivity. One that is widely used throughout meteorology is the $Z-R$ relation from Marshall et al. (1955):

$$Z = 200 R^{1.6} \tag{2.62}$$

Typical values for Z and the corresponding rainfall rates are shown in table 2.2. Additionally, an attempt is made to classify the rainfall intensities associated with the numerical values for reflectivity and rainfall rate.

Reflectivity [dBZ]	**Rain rate** [$mm\ h^{-1}$]	**Intensity**
0	0.04	Not visible
5	0.07	Hardly visible
10	0.15	Light mist
15	0.32	Mist
20	0.65	Very light rain
25	1.33	Light rain
30	2.73	Light to moderate rain
35	5.62	Moderate rain
40	11.53	Strong rain
45	23.68	Strong to heavy rain
50	48.62	Heavy rain
55	99.85	Very heavy rain
60	205.05	Graupel or extreme rain
65	421.07	Large graupel
70	864.68	Extreme graupel

Table 2.2: Reflectivity with corresponding rainfall using the Marshall-Palmer Z-R relation (2.62). Note that for high 55+ reflectivity the equivalent rain rates can be slightly misleading since usually only graupel or similar icy hydrometeors can cause such high reflections. Additionally, the experienced intensity corresponding to rain rates is included for reference.

The Marshall-Palmer relation (2.62) is close to the average of the 69 $Z-R$ relations studied in Battan (1973) for rainfall rates between 1 and 50 $mm\ h^{-1}$. This relationship has shown to fit $Z-R$ measurements of stratiform precipitation quite decently, as well as reasonable representation up to higher rainfall rates for other precipitation types like convection. For the most accurate estimation of rainfall rate given a reflectivity measurement, one would have to calibrate the $Z-R$ relation for every single rain event with rain gauge measurements. Since it is not a focus of this study to derive the optimal $Z-R$ relation and moreover because the necessary resources both in time and in terms of data availability are lacking, the Marshall-Palmer (MP) $Z-R$ relation (2.62) will be used to deduce rainfall rates from reflectivity measurements in this study, most prominently in the verification.

2.5.2. Direct radar reflectivity assimilation

Any observation to be assimilated into a forecast requires an observation operator to be able to quantify the difference between observation and model in the cost function (2.7). This observation operator converts the model to observation space, or in other words simulates an observation from the model state at the observation location. That implies a radar reflectivity has to be linked to model variables. A basic approach is to link reflectivity to the liquid water or equivalently rain water mixing ratio, defined as the ratio of the weight of liquid water to dry air contained in a unit volume:

$$r_r = \frac{m_r}{m_d} \tag{2.63}$$

The specific humidity for rain water is similarly defined as the ratio of the mass of liquid water to the total mass of the unit volume m_{tot} that may contain water vapour, liquid water, snow, ice, etcetera:

$$q_r = \frac{m_r}{m_{tot}} \tag{2.64}$$

In general atmospheric situations without extreme humidity, the mass of dry air is dominant over the contributions from vapour and hydrometeors ($m_v, m_r, m_s \ll m_d$) so that a mixing ratio is approximately equal to the specific humidity, e.g. for rain:

$$r_r \approx q_r \tag{2.65}$$

Throughout this report this assumption of 2.65 is made unless specified otherwise. Coming back to the objective to link reflectivity to rain mixing ratio, it is noted that both reflectivity and rain mixing ratio are related to the liquid water content (LWC), the total amount of liquid water in a unit volume of dry air. Assuming only spherical droplets of water exist, the LWC is defined as:

$$LWC = \rho_w 10^{-9} \int_0^\infty \frac{\pi}{6} D^3 N(D)\, dD \quad \left[g\ m^{-3}\right] \tag{2.66}$$

With water density $\rho_w = 10^6 g\ m^{-3}$ and $D\,[mm]$ diameter, $N(D)\left[m^{-3} mm^{-1}\right]$ the DSD again, respectively. In essence this is the total volume of all droplets multiplied by the density to obtain the total weight of the water droplets in a unit volume of air. By assuming the exponential distribution of Marshall and Palmer (1948) for droplets as in 2.58, it can be shown analytically that reflectivity factor z and LWC, analogous to the reflectivity and rainfall rate, relate to each other also via a power law relation (2.61) as follows:

$$z = 2.04 \cdot 10^4 (LWC)^{1.75} \quad \left[mm^6 m^{-3}\right] \tag{2.67}$$

Or equivalently when the LWC is specified in units of $kg\ kg^{-1}$ instead of $g\ kg^{-1}$

$$z = 1.15 \cdot 10^8 LWC^{1.75} \tag{2.68}$$

By recognising that $LWC = \rho r_r$, taking the logarithm and multiplying by 10 as in 2.57, the $Z - r_r$ relation is obtained:

$$Z = 43.1 + 17.5 \log_{10}\left(\rho r_r\right) \quad [dBZ] \tag{2.69}$$

Where ρ is the density of dry air. This $Z - r_r$ relation can be used as the observation operator H that links model's rain mixing ratio r_r to a reflectivity measurement Z. Do note that applying this relation assumes that the reflectivity is solely caused by reflection from rain. That is a valid assumption for most stratiform precipitation especially in the summer, but often significant reflections are caused by for example snow, ice and graupel in for example the melting layer in a convective system. The observation operator in 2.69 has been implemented by Xiao et al. (2007) to be radar reflectivity observations into the MM5-3DVAR system and consequently into the WRFDA later. To be able to compute analysis increments at every iteration in the minimization of the cost function, the rain water mixing ratio has to be added to the control variables. Because the NMC method is not able to properly generate background error statistics for r_r (since often it does not rain and subsequently the error would be zero), the total water mixing ratio (the sum of mixing ratios of water vapour, cloud water and liquid water) was used instead. Background error statistics could be generated for this and therefore the total water mixing ratio was added as a control variable to the cost function. A warm rain partitioning scheme as described by Xiao et al. (2007) was used to split total water mixing ratio increments found in every iteration step into the individual mixing ratios of vapour, cloud and liquid water. Xiao et al. (2007) showed that by assimilating both reflectivity and radial velocity measured by Doppler weather radars the number of correct rain forecasts improved and that the added value of assimilating reflectivity was most evident in the first 3 hours of the forecast after assimilation.

2.5.3. Indirect radar reflectivity assimilation

The direct assimilation of radar reflectivity through operator 2.69 enables to link model to observations. However, as shown in section 2.3.2, the observation operator is linearised so that the cost function (2.11), incremental) or equivalently the preconditioned cost function (2.42) is quadratic and can be minimized using the optimization algorithms described in section 2.3.3. Since the observation operator is logarithmic, large linearisation errors can be introduced if the optimal model state is far away from the background.

Wang et al. (2013) have showed that if the power law relationship for $Z - r_r$ in the logarithmic form of

$$Z = c_1 + c_2 \log_{10}\left(\rho r_r\right) \tag{2.70}$$

is used as the observation operator, when linearised to

$$dZ = \frac{c_2 dr_r}{r_r \ln(10)} \tag{2.71}$$

compared to an increment dZ_{new} using the full non linear relation from 2.70

$$dZ_{new} = Z_{new} - Z = c_2 \log_{10}\left[\frac{r_r + dr_r}{r_r}\right] \tag{2.72}$$

will lead to a linearisation error that is dependant only on the ratio $k = \frac{dr_r}{r_r}$ of rain water increment to background rain water:

$$\varepsilon_{lin} = \frac{c_2 k}{\ln(10)} - c_2 \log_{10}(1+k) \tag{2.73}$$

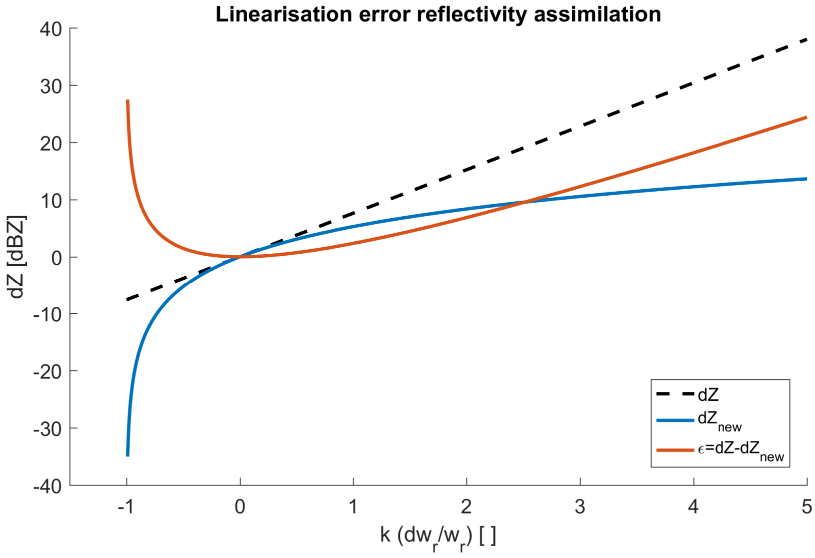


Figure 2.14: Graph showing the how the linearisation error ε_{lin} varies with ratio k. The dashed line shows the increment dZ from the linearisation. The blue line shows the magnitude of dZ when the full non linear operator is used. Finally, the red line shows the difference between the two. This figure has been based on the original illustration from Wang et al. (2013).

This linearisation error is illustrated in figure 2.14. The blue curve shows how reflectivity is affected by changes in the rain mixing ratio using the full non linear relation 2.70. The black, dashed line shows the increment dZ corresponding to a change r_r (and consequently k) when using a linear approximation. Finally, the red curve shows the error between the non linear and linearly approximated reflectivity increments as formulated by 2.73. It is evident that when the increment dr_r is several times the size as the background rain mixing ratio r_r, e.g. at $k = 5$ a significant linearisation error of 24 dBZ is introduced. When forecasting convection or even precipitation in general it is quite common for a model to incorrectly forecast the position of an area of rain. Correspondingly, the rain water mixing ratio in the model at an area where no rain is present will be (nearly) zero so that any non zero increment resulting from a radar reflectivity observation will lead to factor k blowing up and consequently cause a huge linearisation error. Then of course the linearisation is by far not valid anymore. Also, Wang et al. (2013) note that the linearisation error is always positive, meaning that any reflectivity innovation using the linear approximation will yield an increment dr_r that is always smaller than an increment dr_r corresponding to the same reflectivity innovation that would be obtained by the non linear $Z - r_r$ relation. In other words, the linearisation causes a *dry bias* in the data assimilation analysis.

To circumvent the significant errors introduced by linearisation of the reflectivity operator when there is a large discrepancy between model and observation (rain/no rain), the same Wang et al. (2013) proposed to first reconstruct hydrometeor quantities from radar reflectivity observations and then assimilate those instead. So the reflectivity measurements are replaced by mixing ratio 'observations'. That way it is avoided that the reflectivity observation operator is linearised. Besides the mixing ratio for rain WRFDA is also capable of assimilating other hydrometeors from reflectivity, namely snow and graupel. It is assumed that a reflectivity measurement consists of contributions from rain, dry snow, wet snow and graupel. The equivalent reflectivity factor is therefore a sum of the contributions to reflectivity factor of the model mixing ratios for rain r_r, snow r_{sn} and graupel r_g:

$$z_e = z(r_r) + z(r_{sn}) + z(r_g) \tag{2.74}$$

For every $z - r$ relation, empirical expressions obtained from regression analysis are used in WRFDA and are described in Gao and Stensrud (2012). The equivalent reflectivity factor for rain is given by the relation

$$z(r_r) = c_r\left(\rho r_r\right)^{1.75} = 3.63 \cdot 10^9 \left(\rho r_r\right)^{1.75} \tag{2.75}$$

Where ρ represents the atmospheric density and c_r the prefactor for rain. For snow the distinction is made between dry and wet snow. If snow is present, then for temperatures below $0°C$ the snow is regarded as dry snow and the first expression in 2.76 is used. The prefactor here is smaller than for rain since dry ice crystals don't reflect as strong compared to water droplets. For wet snow at temperatures larger than $0°C$ associated with the melting layer a different relation is used with a bigger prefactor to account for the large, melting ice crystals that now have a surface of liquid water.

$$z(r_{sn}) = \begin{cases} c_{dsn}\left(\rho r_{sn}\right)^{1.75} = 9.80 \cdot 10^8 \left(\rho r_{sn}\right)^{1.75} & t \leqslant 0°C \\ c_{rsn}\left(\rho r_{sn}\right)^{1.75} = 4.26 \cdot 10^{11} \left(\rho r_{sn}\right)^{1.75} & t > 0°C \end{cases} \tag{2.76}$$

Finally, for graupel, reflectivity is related to its mixing ratio as follows:

$$z\left(r_g\right) = c_g\left(\rho r_g\right)^{1.75} = 4.33 \cdot 10^{10} \left(\rho r_g\right)^{1.75} \tag{2.77}$$

Again, a relatively high prefactor is used. In the case of graupel it reflects the size of the hailstone rather than the strong reflection caused by water. Although the above way of assimilating reflectivity is a better representation of reality, it is impossible to determine the three mixing ratio variables from a single reflectivity measurement. Therefore, the hydrometeor variables have to be partitioned according to some assumptions. The partitioning in WRFDA is done in line with Gao and Stensrud (2012) who proposed to partition the hydrometeors based on model temperature. First it is assumed that for temperatures above $5°C$ only rain leads to a radar reflection, while at temperatures lower than $-5°C$ a combination of snow and graupel make up the equivalent reflectivity factor. In between these temperatures, all hydrometeors in 2.74 exist:

$$z_e = \begin{cases} z(r_{sn}) + z\left(r_g\right) & t < -5°C \\ z(r_{sn}) + z\left(r_g\right) + z(r_r) & -5° \leqslant t \leqslant 5°C \\ z(r_r) & t \geqslant 5°C \end{cases} \tag{2.78}$$

This reduces the complexity of the radar reflectivity assimilation, though a partitioning still needs to be applied. In the transition zone from solid to liquid phase ($-5° \leqslant t \leqslant 5°C$) it is assumed that the fraction α of the equivalent reflectivity caused by rain linearly increases from 0 at $T = -5°C$ to 1 at $T = 5°C$. Secondly, at temperatures where snow and graupel exist, the reflectivity factor is distributed according to their respective prefactors c_{sn}, c_g, with an additional distinction for dry and wet snow as defined in 2.76. The prefactor partitioning for snow and graupel implies that an equal observation is made for both r_{sn} and r_g. The partitioning leads to the following values for $z(r_r)$, $z(r_{sn})$ and $z\left(r_g\right)$ at certain temperature intervals:

$$z_e = \begin{cases} \begin{array}{cccc} \mathbf{z(r_{sn})} & \mathbf{z\left(r_g\right)} & \mathbf{z(r_r)} & \mathbf{T} \\ \frac{c_{dsn}}{c_{dsn}+c_g} z_e & \frac{c_g}{c_{dsn}+c_g} z_e & 0 & T < -5°C \\ (1-\alpha)\frac{c_{dsn}}{c_{dsn}+c_g} z_e & (1-\alpha)\frac{c_g}{c_{dsn}+c_g} z_e & \alpha z_e & -5° \leqslant T < 0°C \\ (1-\alpha)\frac{c_{rsn}}{c_{rsn}+c_g} z_e & (1-\alpha)\frac{c_g}{c_{rsn}+c_g} z_e & \alpha z_e & 0° \leqslant T < 5°C \\ 0 & 0 & z_e & T \geqslant 5°C \end{array} \end{cases} \tag{2.79}$$

Using the values for the prefactors from 2.75, 2.76 and 2.77, the magnitude of the fractions defined in 2.79 have been plot in figure 2.15. Evidently, at temperatures below $T = -5°C$ the reflection is dominated by graupel. With increasing temperature the contribution from snow and graupel decreases until the melting point where now the high prefactor for wet snow is causing the bulk of the remaining fraction of reflection from snow and graupel. At $T = -5°C$, only rain is assumed to be present.

The hydrometeors that are to be assimilated have to be included in the control variables. For example for the rain mixing ratio q_r, an extra term is added to the cost function:

$$J_{r_r} = \frac{1}{2}\left(\mathbf{r}_r - \mathbf{r}_r^b\right)\mathbf{B}_{r_r}^{-1}\left(\mathbf{r}_r - \mathbf{r}_r^b\right) + \frac{1}{2}\left(\mathbf{r}_r - \mathbf{r}_r^o\right)\mathbf{R}_{r_r}^{-1}\left(\mathbf{r}_r - \mathbf{r}_r^o\right) \tag{2.80}$$

With $\mathbf{r}_r$, $\mathbf{r}_r^b$ and $\mathbf{r}_r^o$ the model, background and observation vector for the rainwater mixing ratio. $\mathbf{B}_{r_r}$ and $\mathbf{R}_{r_r}$ represent the background and observation error covariance matrix. As noted earlier, the background error covariance matrix cannot be properly estimated using the NMC method, because most often there is no rain present in the model. Instead, $\mathbf{R}_{r_r}$ is hard coded using a variance of $\sigma_{r_r}^2 = 1.0 \cdot 10^{-6}\ kg^2\ kg^{-2}$ or equivalently,

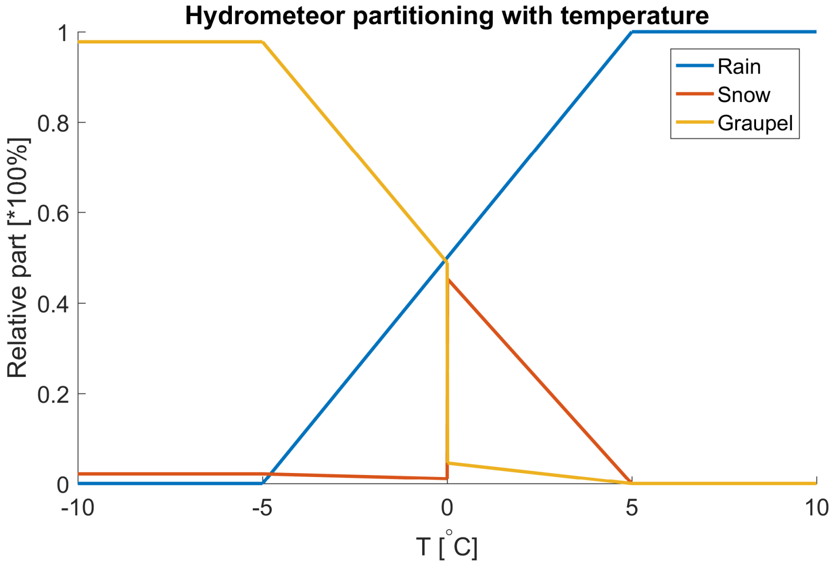


Figure 2.15: Partitioning of hydrometeors as based on Gao and Stensrud (2012)

a standard deviation of $\sigma_{r_r} = 1.0\ g\ kg^{-1}$. The covariance length scale is equal to the resolution of the domain, 4.0 km. Also, r_r is assumed to be only isotropically horizontally correlated. So, any radar observed mixing ratio r_r will only lead to an increment dr_r on horizontal model levels. The main reason for assuming only horizontal correlation is that the vertical layering in the WRF model is domain dependent and that is is not possible to derive the vertical correlation using the NMC method as the mixing ratio is often zero.

Similar to the rain cost function term 2.80, additional cost function terms are added for the snow and graupel mixing ratios, see 2.81 and 2.82. The variance and length scales are the same, making the background error covariance matrices $\mathbf{B}_{r_{sn}}$ and $\mathbf{B}_{r_g}$ equivalent. In WRFDA however any scaling factor for both variance and length scale can be applied to the control variables if the analysis increments are unrealistic according to the user's judgement.

$$J_{r_{sn}} = \frac{1}{2}\left(\mathbf{r}_{sn} - \mathbf{r}_{sn}^{b}\right)\mathbf{B}_{r_{sn}}^{-1}\left(\mathbf{r}_{sn} - \mathbf{r}_{sn}^{b}\right) + \frac{1}{2}\left(\mathbf{r}_{sn} - \mathbf{r}_{sn}^{o}\right)\mathbf{R}_{r_{sn}}^{-1}\left(\mathbf{r}_{sn} - \mathbf{r}_{sn}^{o}\right) \tag{2.81}$$

$$J_{r_g} = \frac{1}{2}\left(\mathbf{r}_g - \mathbf{r}_g^{b}\right)\mathbf{B}_{r_g}^{-1}\left(\mathbf{r}_g - \mathbf{r}_g^{b}\right) + \frac{1}{2}\left(\mathbf{r}_g - \mathbf{r}_g^{o}\right)\mathbf{R}_{r_g}^{-1}\left(\mathbf{r}_g - \mathbf{r}_g^{o}\right) \tag{2.82}$$

To construct the observation error covariance matrix **R** for the hydrometeor variables, each radar observation requires an estimate of the associated error. This is not a trivial task, in the first place because the instrumental error of radar measurements is difficult to quantify because various error sources contribute to the error. From correspondence with H. Leijnse, the custodian for radar data at KNMI, the error variance for a reflectivity measurement is said to be approximately 2 dB^2Z. However, as pointed out in section 2.4.2, besides the measurement error also the observation operator error and representativeness have to be accounted for. For example even if a measured reflectivity would be perfect in terms of representing reality, in the model parametrizations and simplifications are made so that a different reflectivity is obtained and thus an error between model and observation exist, simply because the model cannot represent the full state of the atmosphere perfectly. Although a solid estimation technique for all these errors is lacking, an attempt is made to account for them by setting the variance of a reflectivity measurement to 5 dB^2Z. For a single reflectivity measurement 2.71 is used to estimate the error variance for a hydrometeor mixing ratio from a reflectivity error variance.

2.5.4. Radar reflectivity humidity assimilation

In the previous section it was discussed how radar reflectivity can be related to model hydrometeor quantities. Radar reflectivity observations are transformed into 'observations' for rain, snow and graupel mixing ratios which are then assimilated. However, when only this information is assimilated, generally the model is unable to sustain the precipitation and it dissipates quite quickly (Sun and Wang (2013)). Assimilation of the hydrometeor mixing ratios in essence is equivalent to assimilating only the last stage of a precipitation event, rather than the mechanism driving it.

To tackle this problem so called cloud analysis schemes have been developed. These schemes adjust model humidity and temperature to create favourable conditions for the initiation of new precipitation. For example Hu et al. (2006a) indicated an enhanced storm forecast when using a cloud analysis scheme to change heat and humidity based on radar reflections. Also, Zhao and Xue (2009) reported better representation of intensity and structure of forecasts for hurricane Ike when a cloud analysis scheme was used in combination with 30 minute radar assimilation in the Advanced Regional Prediction System (ARPS) 3DVar system. In WRFDA, a basic adoption of such a cloud analysis scheme was proposed by Wang et al. (2013) based on the assumption that the water vapour pressure is close to saturation vapour pressure at a point where the radar reflection is obtained. Basically, where a radar measures a reflection higher than a certain reflectivity threshold at a height above cloud base, a relative humidity 'observation' close to saturation is made which is then assimilated. The cloud base is assumed to be at the lifting condensation level (LCL) which is estimated as follows:

$$LCL = h_{sfc} + 125\left(T_{sfc} - T_{d,sfc}\right) \tag{2.83}$$

Where h_{sfc} is the model surface height, T_{sfc} and $T_{d,sfc}$ the model surface and surface dewpoint temperature, respectively. Evidently, a lapse rate of $\Gamma = 8°\ C\ km^{-1}$ is assumed. Also, the LCL is bounded by:

$$h_{sfc} + 500 < LCL < 3000 \tag{2.84}$$

To be able to assimilate a water vapour observation, an addition to the cost function is made similar as for rain water in 2.80.

$$J_{r_v} = \frac{1}{2}\left(\mathbf{r}_v - \mathbf{r}_v^o\right)^T \mathbf{R}_{r_v}^{-1}\left(\mathbf{r}_v - \mathbf{r}_v^o\right) \tag{2.85}$$

With $\mathbf{r}_v$ the model vector of water vapour mixing ratio and $\mathbf{R}_{r_v}$ the water vapour observation error covariance matrix. In 2.85 there is no background term as was the case for hydrometeors like in 2.80, 2.81 and 2.82, since water vapour is already accounted for in the the standard 3DVar control variables without radar assimilation through the pseudo RH control variable. The model water vapour mixing ratio r_v at a specific point is converted from the relative humidity RH using the model saturation vapour mixing ratio r_s:

$$r_v = RH \cdot r_s \tag{2.86}$$

Equation 2.86 is the observation operator H for r_v. It links the relative humidity control variable to a water vapour mixing ratio. When the contribution to weight from water (present in any phase) is neglected, the water vapour saturation mixing ratio is approximately equal to the saturation specific humidity q_s, which in turn is defined as:

$$r_s \approx q_s = \frac{\varepsilon e_s}{p - (1-\varepsilon)\, e_s} \tag{2.87}$$

The saturation water vapour mixing ratio depends on the saturation vapour pressure e_s, total pressure p and the ratio between the gas constants for dry air and water vapour $\varepsilon = \frac{R_d}{R_v} \approx 0.622$. The water vapour saturation pressure e_s varies with temperature. Therefore the the estimated r_s is a function of both model pressure and temperature and thus the quality of the r_s estimate depends on these model fields.

Recall that for minimization of the preconditioned cost function (2.42) the observation operator 2.86 has to be linearised. Wang et al. (2013) showed that the linearised observation operator

$$dr_v = dRH \cdot r_s + RH \cdot dr_s \tag{2.88}$$

can be approximated by:

$$dr_v \approx dRH \cdot r_s + r_v \frac{c_2 c_3}{(T + c_3)^2} dT \tag{2.89}$$

With increments dr_v, dRH, dT in water vapour mixing ratio, relative humidity and temperature respectively. Also T represents temperature and constants $c_2 = 17.67$ and $c_3 = 243.5$ that originate from the expression for the water vapour saturation pressure e_s A.37. See also appendix A.6 for an elaborate description of how to arrive at this approximation. Also, Wang et al. (2013) note that any positive increment in water vapour will lead to increments in relative humidity and temperature that are positive as well, since $\frac{c_2 c_3}{(T+c_3)^2} > 0$. Both moisture and heat are added to areas where a radar reflection is obtained and where the relative humidity is not close to saturation. This approach should catalyse the formation of new convective areas where it does not exist yet in the model.

As mentioned, a water vapour mixing ratio 'observation' r_v^o is made at a radar reflection point. It seems logical to base the magnitude of r_v^o on the measured reflectivity from the radar. r_v^o is constructed by assuming a relative humidity close to saturation and using 2.86 in combination with the model saturation water vapour mixing ratio which in turn is a function of model pressure and temperature, see A.36 and A.37. Consequently, the magnitude of the relative humidity observation that is assimilated is based on the reflectivity as follows:

$$RH_o = \begin{cases} RH_1 & Z_1 \leqslant Z < Z_2 \\ RH_2 & Z > Z_2 \end{cases} \tag{2.90}$$

Of course values for RH_1, RH_2, Z_1 and Z_2 are arbitrary. A straightforward value for Z_1 would be a reflectivity corresponding to light precipitation. Looking back at table 2.2, a value of 20 dBZ seems to be appropriate. Setting Z_1 to a lower threshold value means that a larger area of humidity observations will be assimilated. Z_2 is set to 35 dBZ, corresponding to moderate rain intensity. RH_1 in between these thresholds is set to 88%. RH_1 mainly controls how much water vapour is assimilated, since often most precipitation is contained within this interval. Above 35 dBZ, associated with heavy rain, a relative humidity of 95% is assimilated. Areas in this reflectivity range are often restricted to (cores of) convective regions. Updating 2.90 with these values:

$$RH_o = \begin{cases} 88\% & 20 \leqslant Z < 35\,dBZ \\ 95\% & Z > 35\,dBZ \end{cases} \tag{2.91}$$

For a reflectivity smaller than Z_1, no humidity observation is assimilated. The RH values of 88% and 95% have been set after some experiments testing the sensitivity to these values. Setting RH_1 and RH_2 to larger values resulted in significant overestimation of precipitation in the subsequent forecast, while setting lower RH assimilation values resulted in too few new areas of precipitation.

Additionally, it is important to note that when a radar derived humidity observation is assimilated, the innovation cannot be negative. That is, if a reflection stronger than 20 dBZ is measured and the humidity is already larger than RH_1 or RH_2 depending on the reflectivity strength, there is no assimilation of a humidity observation. The rationale is that an area with such a near saturation RH is already likely to produce precipitation and it's RH should not be reduced.

As for any observations, the error of a humidity observation deduced from radar reflectivity has to be quantified. In this case, a constant fraction of 20% of the humidity observation is used to specify the standard deviation of the error. The minimum value that the error may attain is $1.0 \cdot 10^{-3}\ kg\ kg^{-1}$. So for any humidity observation below $5.0 \cdot 10^{-3}\ kg\ kg^{-1}$, σ_{r_r} is set to this minimum value.

2.5.5. Null echo assimilation

In the assimilation of radar reflectivity up until now only the case where the radar indicates precipitation has been considered. Equally important however is that precipitating areas that exist in the model where the radar does not measure a reflection (called a *null echo*) are suppressed. In other words, precipitating model points have to be penalized for indicating precipitation while it is not true, at least according to the radar reflection. Therefore, at places where the model is indicating precipitation but the radar does not measure any significant reflection, a humidity observation just under saturation is assimilated to suppress the model formation of precipitation. The reflectivity threshold for model precipitation is again taken to be 20 dBZ. If the model equivalent reflectivity is above this threshold, the humidity observation that is assimilated is constructed using the model saturation mixing ratio, analogous to 2.86:

$$r_v^o = RH_{np} r_s \tag{2.92}$$

The innovation from r_v^o ought to be negative, otherwise it is not assimilated. In other words, if the model water vapour mixing ratio is already smaller than RH_{np}, no observation is assimilated. The relative humidity value RH_{np} used to downscale the humidity in non precipitating areas is set according to temperature:

$$RH_{np} = \begin{cases} 85\% & T < -5^\circ C \\ \alpha \cdot 85\% + (1-\alpha) \cdot 90\% & -5^\circ \leqslant T < 5^\circ C \\ 90\% & T \geqslant 5^\circ C \end{cases} \tag{2.93}$$

In the transition $-5° \leqslant T < 5°C$ from ice to water, RH_{np} is a combination of both thresholds, where the fraction of $RH_{np,rain}$ of 90% linearly increases from 0 to 1, similar as for the hydrometeor assimilation in 2.79. 90% has been taken as threshold for rain since it is not too close to saturation yet it is not so low that no precipitation can form anymore at a later stage because the relative humidity is so far away from saturation. Additionally, a lower RH is used for freezing temperatures as it is assumed that precipitation forms faster at equal RH in freezing clouds than in conditions where only liquid water is present in clouds. Also, if the hydrometeor assimilation as described in section 2.5.3 is employed, near zero mixing ratios are assimilated for rain, snow and graupel, effectively removing erroneous model hydrometeors. This is realised by assimilating a reflectivity observation of $-15\ dBz$ using 2.78. Assuming that only liquid water causes a reflection and thus, using 2.75, this would correspond to an observation $r_r^o = 4.8 \cdot 10^{-7} kg\ kg^{-1}$, using a dry air density of $\rho = 1.0\ kg\ m^{-3}$ that is common at 1.5 km altitude.

A radar is said to measure a null echo if the measured reflectivity is smaller than $-10\ dBz$. The threshold $-10\ dBz$ is chosen so that areas with a small measurable reflection ($-10 \leqslant Z < 20\ dBz$) which might serve as a favourable area for new precipitation are not affected. Only the areas with erroneous model precipitation where the radar does not measure a significant reflection are flagged as having a null echo. Erroneous model precipitation is indicated by a model simulated reflectivity bigger than 20 dBz, which is required for the non precipitating observation to be assimilated. Summarizing, the following conditions have to be satisfied in order for the hydrometeor suppression:

$$\begin{aligned} Z_{radar} &< 0\ dbZ \\ Z_{model} &> 20\ dbZ \end{aligned} \tag{2.94}$$

Similar to the humidity assimilation for precipitation, these reflectivity thresholds are defined arbitrarily. Usually, for a lot of radar measurements there is no detectable reflection above a certain signal to noise ratio threshold, which are classified as *no data* points in a radar image. Technically it is not correct to assume that at these no data points no precipitation can be present. For example attenuation of radar beams by heavy precipitation can result in the radar not being able to detect precipitation at large ranges, as there is there is no radar power to scatter back anymore. However, this is rarely the case for the C-band radars employed at KNMI. Most importantly however, the assumption that the no data points contain no precipitation is required in order to penalize erroneous model precipitation. The ability to suppress spurious model precipitation is favoured over the drawback of the sporadic cases where model precipitation is suppressed while it is present in reality but just not measured by the radar.

2.6. Radar radial velocity assimilation

Besides measuring the strength of the backscattered signal from hydrometeors, radars are also capable of detecting a frequency shift that results from a moving target, known as the Doppler shift. This Doppler shift is directly related to the radial velocity v_r:

$$f_D = -\frac{2v_r}{\lambda} \tag{2.95}$$

Where f_D is the Doppler frequency shift, v_r the radial velocity and λ the wavelength of the electromagnetic wave. When a the target is moving towards the radar, the phase of the received pulse changes, causing an apparent higher frequency and thus a positive Doppler frequency shift. In that case the radial velocity is negative. Vice versa, if an object moves away from the radar, the distance between wave fronts increase which is perceived by the radar as a wave with lower frequency (i.e. a negative Doppler frequency shift). It is emphasized that a radar is only able to measure the radial velocity, that is the velocity component in the line of sight of the radar. This concept is illustrated in figure 2.16. Given a constant flow of westerly winds, the coloured arrows show the magnitude of the radial velocity component depending on the angle of the radar beam with the wind field. Note that the radial velocity is equal to the actual wind magnitude when the radar beam is parallel to the wind vector, however is zero when the radar beam is tangent to the wind vector. While figure 2.16 is a two dimensional illustration, the same principle applies to a three dimensional wind: the radial velocity is maximal when it aligns exactly with the 3D wind vector. So at higher elevation scans the radial velocity also contains a significant vertical component of the wind speed.

Contrary to the previous observations derived from radar reflectivity, the assimilation of radar radial velocity is quite straightforward in the sense that it is assimilated in a conventional way using the observation

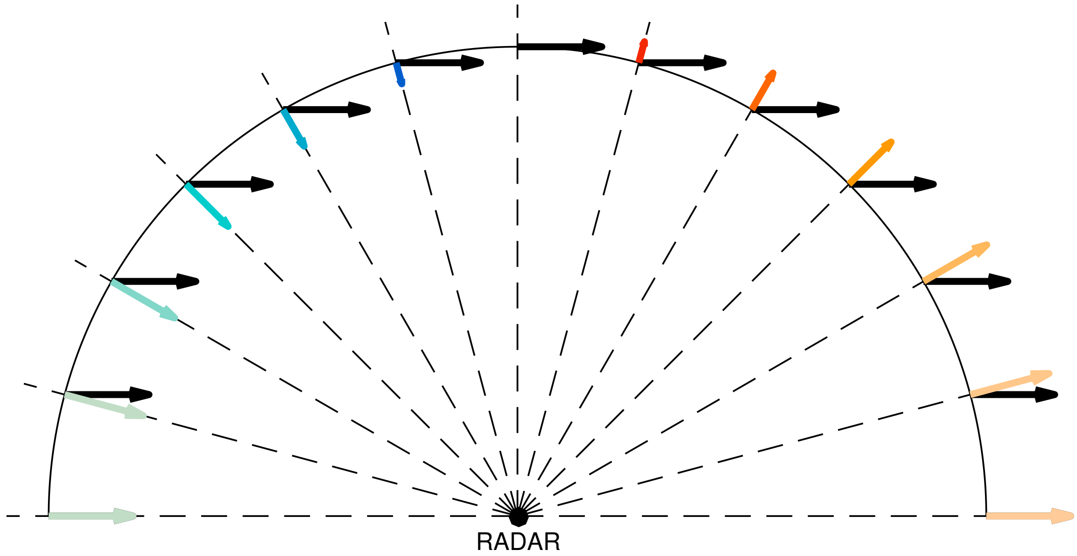


Figure 2.16: Radial velocity component of a radar for various angles with respect to a constant westerly wind field.

operator, without requiring constructing derived observations under certain assumptions or defining additional control variables. Similar as to wind speeds retrieved from other observation sources, assimilation of radar radial velocities just requires the definition of an observation operator H that is used in the cost function (2.7) to compute a model equivalent of the observation so that the difference with respect to the observation $H(\mathbf{x}) - \mathbf{y}$ can be quantified. The observation operator for radial velocity is defined as follows:

$$v_r = \frac{x - x_r}{r} u + \frac{y - y_r}{r} v + \frac{z - z_r}{r} (w - v_T(r_r)) \tag{2.96}$$

Where x, y, z and x_r, y_r, z_r denote the Cartesian components of the positions on the model grid of a certain location and of the radar, respectively. Similarly u, v, w denote the three dimensional components of the wind velocity vector at this certain location. Note that a correction to the vertical velocity w is made to account for the terminal fall speed v_T of rain, that is defined as (Sun and Crook (1997),Sun and Crook (1998)):

$$v_T = 5.4a(\rho r_r)^{0.125} \tag{2.97}$$

The terminal fall speed depends on the air density ρ, rainwater mixing ratio r_r and a pressure correction factor a.

$$a = \left(\frac{p_0}{\bar{p}}\right)^{0.4} \tag{2.98}$$

With p_0 the surface pressure and $\bar{p}$ the base state pressure.

The determination of the Doppler frequency shift is limited by the frequency of the radar pulses. The maximum unambiguous Doppler frequency is equal to half of the frequency at which the phase is sampled (known as the Nyquist frequency), in this case the pulse repetition frequency (PRF):

$$f_D \leq \frac{1}{2} PRF \tag{2.99}$$

At frequencies $|f_D| > \frac{1}{2} PRF$ the sampling is not dense enough to unambiguously determine the frequency and thus they are mapped into the unambiguous range $-\frac{1}{2} PRF \leq f_D \leq \frac{1}{2} PRF$, so called frequency **aliasing**. Often, the maximum unambiguous Doppler frequency shift is expressed in terms of radial velocity, using 2.95:

$$|v_r| < \frac{\lambda \cdot PRF}{4} = v_{\max} \tag{2.100}$$

And consequently all radial velocity observations are contained within the range of maximum unambiguous Doppler or Nyquist velocity $-v_{\max} \leq v_r \leq v_{\max}$. Any actual wind velocity outside this interval is aliased into this interval according to:

$$v_r + k \cdot 2v_{\max} \tag{2.101}$$

With k being an integer. The aliasing of Doppler radial velocity is illustrated in figure 2.17. In this example $v_{\max} = 10\ m\,s^{-1}$ and thus an actual wind velocity of 15 $m\,s^{-1}$ is measured as being $-5\ m\,s^{-1}$. This implies radar radial velocities should be de-aliased before they are to be assimilated.

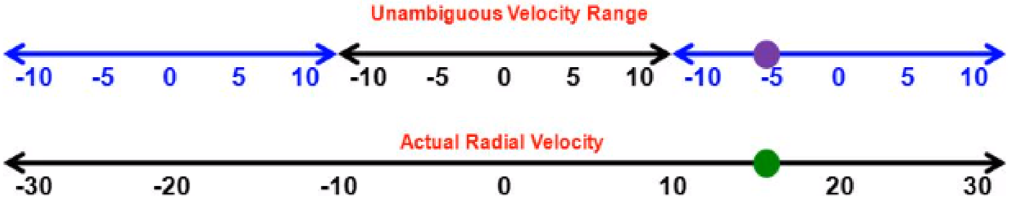


Figure 2.17: Illustration of radial velocity aliasing.

The precision of a radial velocity is proportional to the Doppler spectrum width, which describes the spread in Doppler frequency measurements. The spectrum width is measured during the radar scanning and depends on the variability of the motions inside a radar reflection volume. E.g. turbulence or strong updrafts can cause significant variability in retrieved Doppler frequencies, thus increasing the spectrum width.

2.7. GPS zenith delay assimilation

Global Positioning System (GPS) uses a constellation of satellites to determine the position of a user on Earth. A receiver measures the time it takes for electromagnetic waves to travel back and forth to at least four satellites to estimate a three dimensional position plus a clock bias of the receiver. To obtain the desired accuracy in the position estimate, precise knowledge of the signal travel time is required. However, a GPS signal is delayed by interaction with the atmosphere, see figure 2.18. For example free electrons in the ionosphere slow down the propagation speed of the waves, causing some of the biggest delay of several meters in the GPS signal. Additionally, the gases present in the troposphere also cause a delay in the travel time of a GPS signal. These delays have to be accurately modelled in order to be able to make accurate position estimates. The total vertical delay from the troposphere is also referred to as zenith total delay (ZTD). Generally, the ZTD is modelled as being a combination of delay caused by the gases in the 'dry' hydrostatic atmosphere (zenith hydrostatic delay (ZHD)) and the 'wet' part (zenith wet delay (ZWD)) from the water vapour present in the atmosphere:

$$ZTD = ZHD + ZWD \tag{2.102}$$

The ZWD is of primary interest for meteorology since it contains information about the water vapour in a vertical column the atmosphere. Assimilation of this information in a NWP model should lead to an improved estimate of water vapour, affecting areas where precipitation can form. In an actual GPS measurement the path of a GPS signal is slant and thus the slant paths are converted to zenith angle (locally orthogonal to the surface) to obtain the vertical path using a mapping function, which adds uncertainty to the retrieval. Given a ZTD measurement, a valuable estimate for ZWD can be obtained by subtracting the ZHD from the total delay. The hydrostatic delay caused by the gases other than water vapour is usually modelled as a function of the surface pressure measured at a ground based GPS receiver. As for any observation that is to be assimilated, it has to be linked to model variables through an observation operator to be able to compute the cost function. In WRFDA the observation operator H for ZHD is based on Saastamoinen (1972) and Vedel and Huang (2004):

$$ZHD = \frac{k_1 p_s}{1 - 0.0026\cos\left(2\varphi\right) - 0.000028h} \tag{2.103}$$

Where $k_1 = 2.2768 \cdot 10^{-5}\ m\,Pa^{-1}$ an empirical constant, p_s $[Pa]$ model surface pressure, φ [deg] the observation latitude and h $[km]$ the antenna altitude above the geoid. The surface pressure can be measured quite accurately, enabling a decently accurate estimate of the delay caused by water vapour.

The delay caused by water vapour is described by the operator for ZWD:

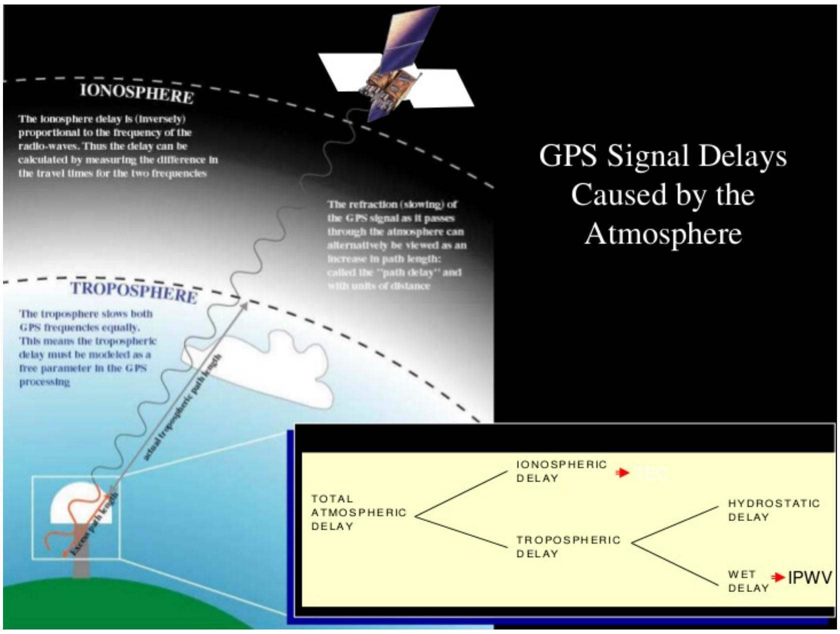


Figure 2.18: Schematic illustration of how the troposphere delays the GPS signal. Illustration from Geophysics (2009).

$$ZWD = 10^{-6}\int_0^\infty N(z)\,dz \approx 10^{-6}\sum_{k=1}^{k=k_{top}}\left[\frac{p_k q_k}{T_k \varepsilon}\left(k_2 + \frac{k_3}{T_k}\right)\right]dz_k \tag{2.104}$$

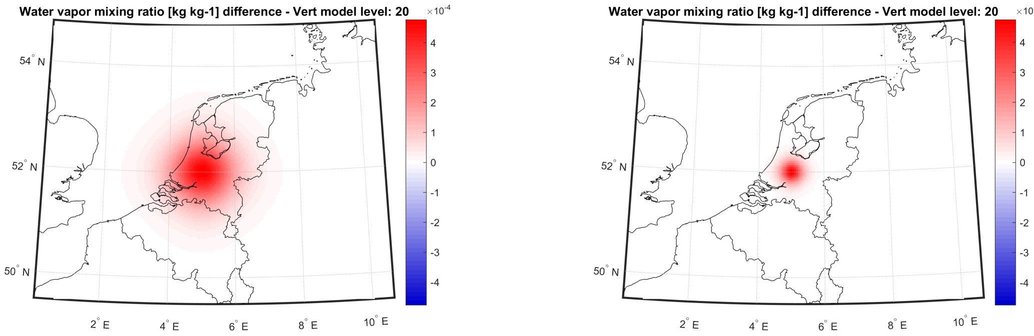
Obviously, the total delay is a sum of the delays caused by water vapour at every model level k. Variables p_k, q_k and T_k represent the model pressure, specific humidity and temperature at model level k that determine the refractivity N. The water vapour specific humidity is assumed to be approximately equal to the mixing ratio of water vapour: $q_k \approx r_k$. dz_k is the thickness of every model layer in terms of height. Additionally, empirical constants $k_2 = 2.21 \cdot 10^{-7}\ KPa^{-1}$ and $k_3 = 3.73 \cdot 10^{-3}\ K^2Pa^{-1}$ are used that were found by Bevis et al. (1994).

2.8. Pseudo single observation tests

To see how observations that are being assimilated into a model propagate in space, a *pseudo single observation test (PSOT)* can be carried out. In a PSOT only a single observation is assimilated, which has numeral benefits:

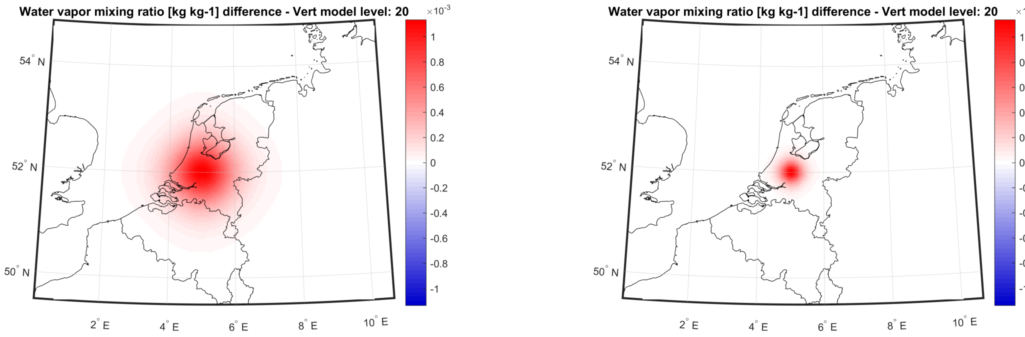
- Assessing the impact of a single observation. It is important to know to what extent an observation can change the model state and moreover how that increment spreads in space.
- Understanding the structure of the background error covariance matrix. The background error covariance matrix is too big and complex too look at directly and the transformed version with EOF's amplitudes contains quantities that are difficult to interpret directly. A PSOT shows how a single observation affects other model points or even other variables through balance and correlation. Thus it shows the structure of the background error covariance matrix in model space so that its strengths and shortfalls can be identified.
- Offering a guideline for local tuning of the background error covariance matrix. If the radius of influence determined from the error estimation is unrealistically large, the length scale of correlation can be reduced to limit the extent to which increments to the forecast are made. Similarly, the variance might need rescaling to produce a more impactful analysis.

An example of the increments in model variables resulting from a PSOT using a single reflectivity measurement is given in figure 2.19. In this theoretical case a single reflectivity observation of 40 dBZ is assimilated at an altitude of 3 km at 52°N, 5°E, with the sole purpose of assessing the impact a single observation



(a) No length and variance scaling is applied.

(b) Length scaling with 0.35. No variance scaling.



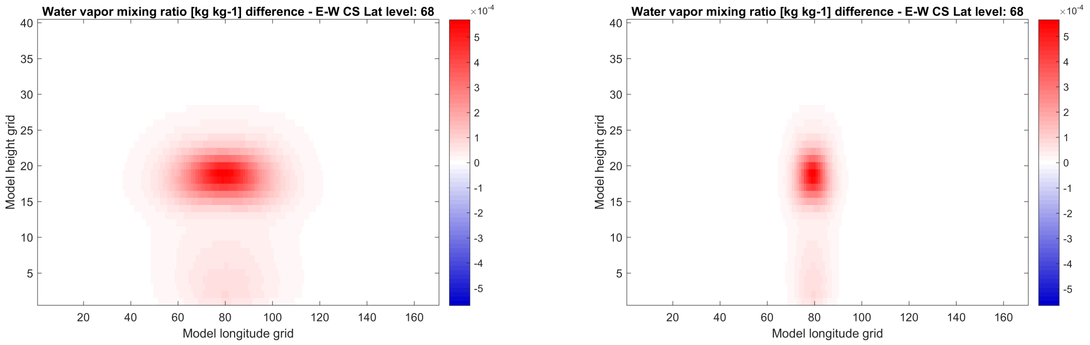
(c) No length scaling. 1.5 variance scaling.

(d) Length scaling with 0.35 and 1.5 variance scaling.

Figure 2.19: Change in model water vapour resulting from a radar reflectivity PSOT of 40 dBZ. Horizontal cross section at vertical model level 20, with a geopotential height equal to 3063 m and pressure 695 hPa at the model latitude/longitude closest to the observation.

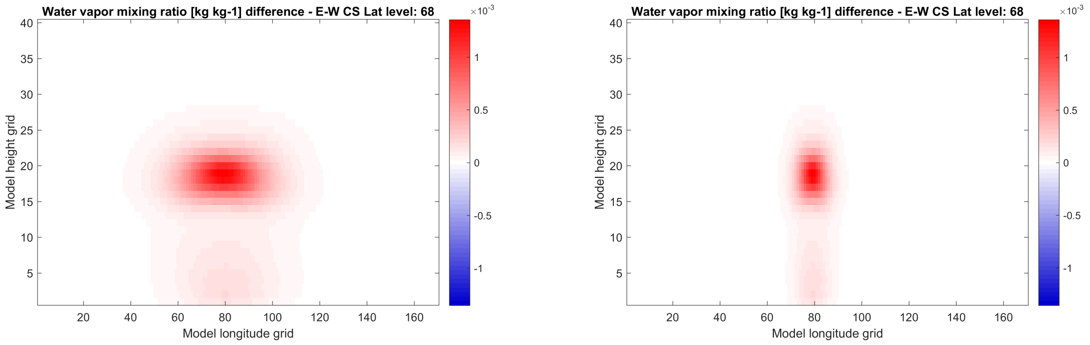
has on the model state. The background forecasts is a three hour WRF model simulation for the Benelux area with 4x4 km resolution and 40 vertical model levels, valid for August 30th, 2015. Figure 2.19a shows the changes on vertical model level 20 (corresponding to a geopotential height of 3063 m and pressure 695 hPa at the nearest model latitude/longitude) water vapour as a result of the assimilated water vapour observation that was deduced from a reflectivity measurement. The difference with respect to the background forecast, also known as *innovation* and O-B, of the reflectivity observation is 1.7 $g\ kg^{-1}$. After assimilation which adds water vapour to the forecast the analysis residual or O-A has been reduced to 1.2 $g\ kg^{-1}$. This corresponds with the maximum analysis increment of approximately 0.5 $g\ kg^{-1}$ as shown in figure 2.19a. The radius of influence in which increments are made is quite large. Humidity is not expected to be correlated for differences of several hundreds of kilometers. Besides, the goal of the radar assimilation is to provide the model with high resolution information that contains small scale features. That requires a smaller length scale. If a scaling of 0.35 is applied, this results in an increment of the same magnitude but with a smaller radius of influence, see 2.19b. This is considered to be a more reasonable radius of influence. The humidity observation has some influence on the surrounding model fields so that some spatial smoothness is maintained, but not as extensively as when no scaling is applied. Similarly to length scaling, WRFDA options also permit scaling the variance of the model background error variances, basically changing the weight of the model background relative to the observations. Assuming a larger variance, e.g. by scaling with 2.5, will reduce the weight of the model background because it is now considered to be less precise. The variance scaling results in larger analysis increment, as seen in figure 2.19c. The water vapour analysis increment is now equal to 1.2 $g\ kg^{-1}$ and O-A 0.5 $g\ kg^{-1}$, compared to 0.5 $g\ kg^{-1}$ and O-A 1.2 $g\ kg^{-1}$ for no variance scaling. Evidently the radar humidity observation now has a bigger impact on the model state. Applying both length and variance scaling results in the same increment but then with the smaller radius of influence from figure 2.19b, shown in figure 2.19d. This way an observation has a decent impact on the analysis and at the same time a realistic scale of influence.

In general on a horizontal level the vapour increments are spread isotropically in space and with a decreasing magnitude when distance from the observation increases, as was modelled using the exponentially decaying covariance (2.52). This pattern of how the information of an observation is spread is solely based on how the background error covariance matrix is constructed or equivalently which model points are correlated.



(a) No length and variance scaling is applied.

(b) Length scaling with 0.35. No variance scaling.



(c) No length scaling. 2.5 variance scaling.

(d) Length scaling with 0.35 and 2.5 variance scaling.

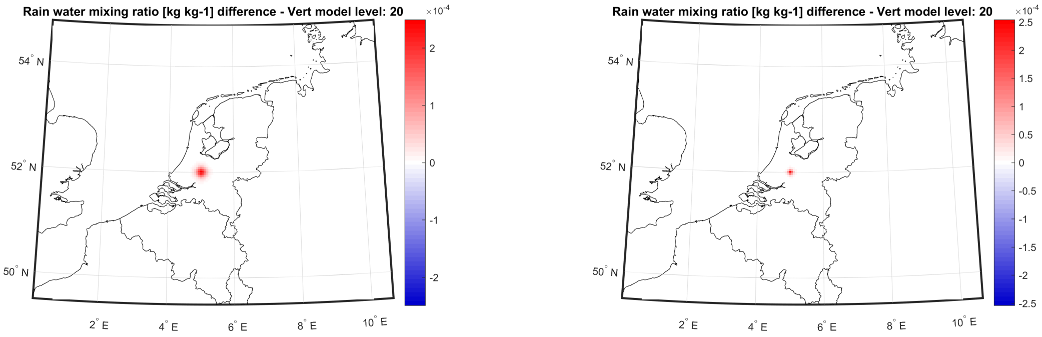
Figure 2.20: Change in water vapour mixing ratio resulting from a radar reflectivity PSOT of 40 dBZ at 3 km altitude. Vertical West-East cross section at the model latitude closest to the observation latitude: 52.01° N.

Besides looking at model increments on a horizontal plane, it also interesting to examine the vertical structure of any increments resulting from assimilation of a single observation. Figure 2.20 shows the vertical cross section from West to East that has been made at a model latitude that is closest tot the observation latitude. Again, the four combinations of length and variance scaling of model background error covariance have been applied in the assimilation of the single reflectivity observation. The big radius of increments also applies to model levels below and above the levels corresponding to the observation height (19 and 20). Considering the increments have a range of influence of nearly 80 model grid points which equals 320 km, it makes sense to apply a length scaling as shown in figure 2.20b to reduce the horizontal spreading of the increment introduced by the observation. Again the increased magnitude of the analysis increment can be recognised in the case of variance scaling (figures 2.20c and combined with reduced lengthscale 2.20d). Of course the amount of scaling remains arbitrary. After various PSOTs with different scaling values however the previously mentioned length and variance scaling of 0.35 and 2.5 were found to produce what could be considered decent and realistic changes to the model. Table 2.3 summarizes the length and variance scale factors for every control variable that are used. These scalings will be applied to any data assimilation experiment if not indicated otherwise. The scaling factor for the surface pressure is not as big as for the other primary control variables u, v, T and RH, because it is a surface variable and considered less uncertain. For the hydrometeor control variables that include mixing ratios for cloud r_c, rain r_r, ice r_i, snow r_{sn} and r_g only a length scaling of 0.5 is applied. The effect of this scaling is not as big as for the primary control variables since the hydrometeor control variables already have a small length scale (equal to the model resolution, 4km in this case). The length scaling is applied so that really only a couple of neighbouring model grid points are affected. Hydrometeor concentration and corresponding mixing ratios vary greatly on a local scale. Of course this varies with the weather regime that is causing the precipitation, but mixing ratios in small convective cells are not likely to be correlated over distances bigger than several tenths of kilometers. That should be reflected in the length scale of the hydrometeor mixing ratio increments. For consistency the same length scaling factor of 0.5 is applied to all hydrometeor variables. The result is shown in figure 2.21 for rain, where at

Nr	Control Variable	Length scaling	Variance Scaling
1	u	0.35	2.5
2	v	0.35	2.5
3	T	0.35	2.5
4	RH	0.35	2.5
5	P_s	0.35	1.5
6	r_c	0.5	1.0
7	r_r	0.5	1.0
8	r_i	0.5	1.0
9	r_{sn}	0.5	1.0
10	r_g	0.5	1.0
11	w	1.0	1.0

Table 2.3: Control variable length and variance scaling used throughout the data assimilation experiments conducted in this study.

the same 20th model level model increments in rainwater mixing ratio induced by the radar observation have been plot. The area of increments has again been reduced. A critical reader may have noticed the additional control variables cloud water mixing ratio r_c, ice mixing ratio r_i and vertical velocity w in table 2.3. The first two are not really important in this study but they are used for assimilation of other observations that link to these quantities like satellite radiances from infra-red channels. Updraft velocity w is employed when radar radial velocities are assimilated, which contain a significant vertical component when the higher elevation scans are used.

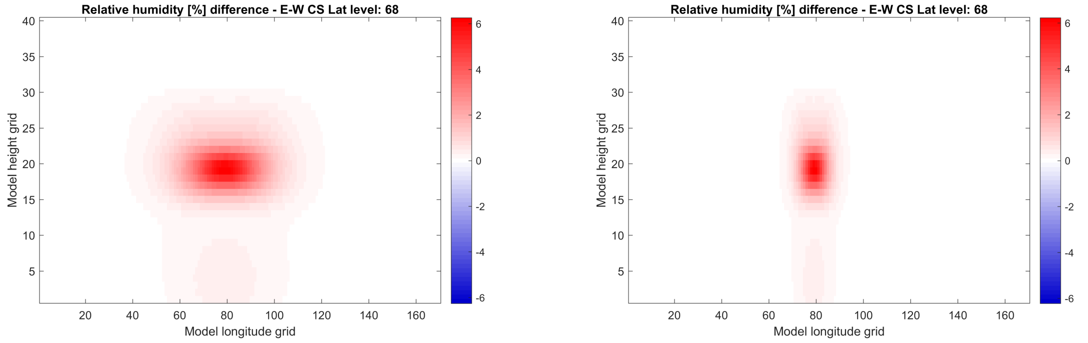


(a) No length scaling is applied.

(b) Length scaling with 0.5.

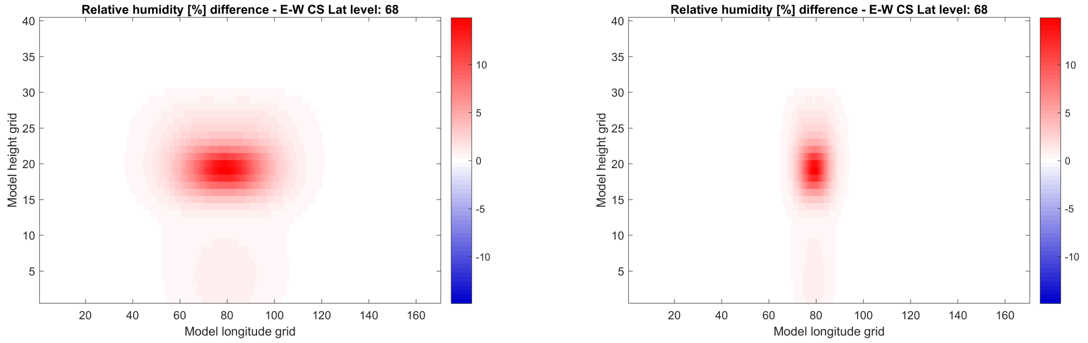
Figure 2.21: Change in model rain water mixing ratio resulting from a radar reflectivity PSOT of 40 dBZ. Horizontal cross section at vertical model level 20, with a geopotential height equal to 3063 m and pressure 695 hPa at the model latitude/longitude closest to the observation.

From figure 2.21b it might seem that the observation spreads the increment more in the vertical but this is misleading since the model grid resolution is much coarser in the horizontal than vertically (kilometers versus pressure levels, approximately 100−200 m for lower model levels), making these cross section plots not to scale. Interestingly, the humidity observation at 3 km also causes increments in water vapour down to the surface, however the relative impact is not as big as at the observation location since the water vapour mixing ratio at the surface is usually bigger than at higher altitudes. That is shown in figure 2.22, where an increase in relative humidity near the surface can be observed of approximately 1% for all scaling combinations. The shape of vertical increments is the same for every scenario, only the horizontal extent varies because of the different horizontal length scales and the magnitude of the increments because of the variance scaling. This is to be expected, since the total vertical shape of the increments is the combination of the eigenvectors (see section 2.4.5 and in particular figure 2.9) that minimizes the cost function. In essence the data assimilation increments are a linear combination of the natural modes of variation, described by the eigenvectors. An increment made at lower levels away from the observation therefore is not always easily physically explained. Rather, it is a consequence of the fact that the eigenvectors, which determine the vertical shape of increments, are based on error statistics that contain correlation between vertical model points. In the case for humidity, the correlation is likely due to surface convection, as the background error statistics were generated using



(a) No length and variance scaling is applied.

(b) Length scaling with 0.35. No variance scaling.



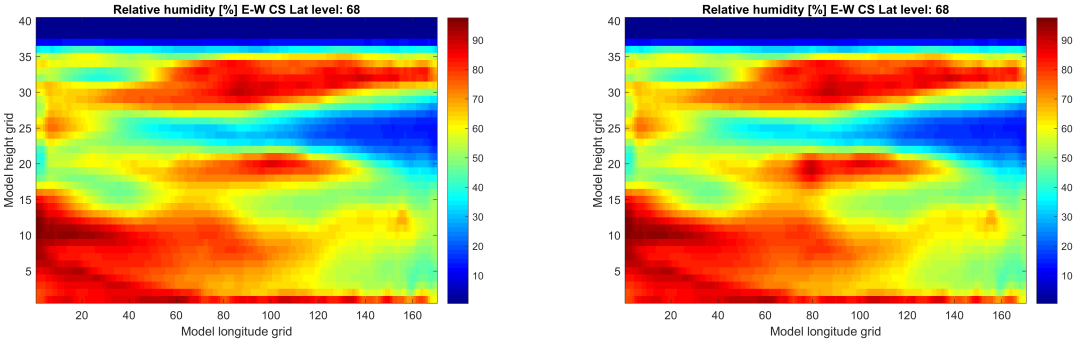
(c) No length scaling. 2.5 variance scaling.

(d) Length scaling with 0.35 and 2.5 variance scaling.

Figure 2.22: Change in relative humidity resulting from a radar reflectivity PSOT of 40 dBZ at 3 km altitude. Vertical West-East cross section at the model latitude closest to the observation latitude: 52.01° N.

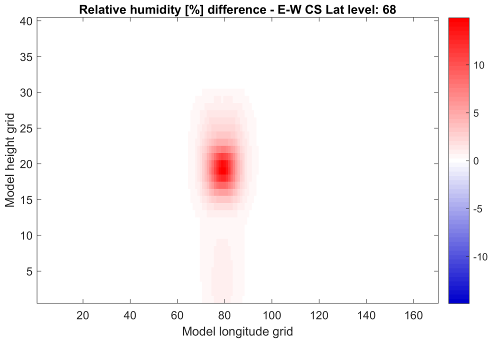
forecasts with spring/summer conditions in May and June where generally stronger convection occurs than in winter. The correlation with lower model levels associated with the boundary layer is actually important since most convective events are surface driven. Because of the correlation a radar observation at a height not that close to the surface, as is the case for the PSOT, can still help to trigger new convection, albeit the relatively small impact on relative humidity on these lower levels. Figure 2.23c shows the same vertical cross section for relative humidity as 2.22d where the scaling is applied, but additionally the model state before (2.23a) and after (2.23b) have been plot. It shows how the radar based humidity observation creates a more favourable environment for clouds and precipitation to form.

Besides the moisture increment made by the PSOT also the temperature is affected, as was pointed out in 2.88. Because a positive vapour mixing ratio innovation was introduced by the radar derived humidity observation, also a positive increment in temperature is made. Figure 2.24a shows how the temperature on the same model level 20 increases as a result of the single reflectivity observation. Looking at figure 2.24a and comparing to 2.19, again the changes are isotropic on the horizontal model level because of the distance based variance decay. However the magnitude of the temperature increments are relatively small, given the maximal temperature increase of approximately 0.02 K in the vicinity of the observation. Apparently the temperature term in the linearised observation operator for r_v 2.89 is not big enough to cause significant changes in temperature. The temperature increase is solely a consequence of 2.89, since only a humidity observation is assimilated and no correlation between the control variables is modelled in the background error covariance matrix. Similarly as for water vapour, the same vertical cross section is made for the temperature increase and is shown in figure 2.24b. The extent of the positive temperature increment is bigger than as for humidity, showing that the correlation length of temperature determined by the background error estimation is larger. It is striking however that also some small negative increments are made for the temperatures above and below the observation. This is likely due to the relatively high altitude at which the observation is made: 3 km, halfway of all model levels. The eigenvectors for temperature with largest amplitudes around model level 20 (refer to figure 2.9), are the third and fifth. These eigenvectors change sign at the model levels below and above level 20, indicating negative correlation between level 20 and these surrounding levels, thus resulting in negative increments. In contrast, if a temperature observation is assimilated at a lower level, the



(a) Model RH before DA.

(b) Model RH after radar PSOT.



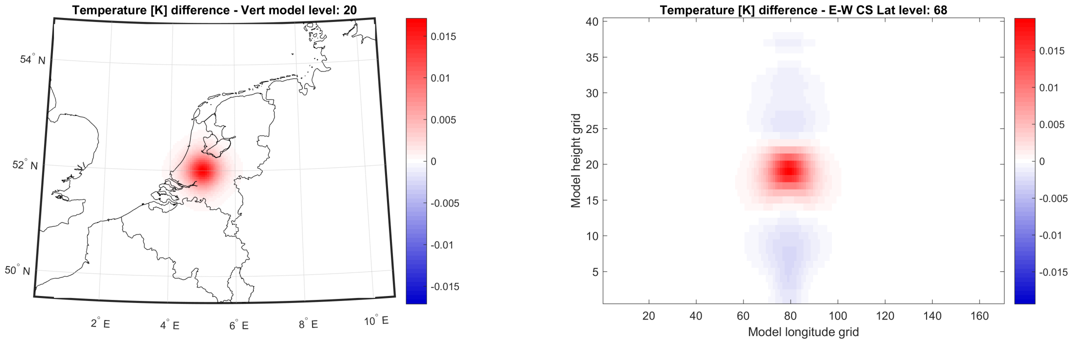
(c) RH difference as a result of radar PSOT.

Figure 2.23: Change in relative humidity resulting from a radar reflectivity PSOT of 40 dBZ at 3 km altitude. Vertical West-East cross section at the model latitude closest to the observation latitude: 52.01°N.

first eigenvector of temperature will largely determine the vertical shape of increments. Figure 2.25b shows the vertical structure of increments made by a temperature PSOT at the fifth model level with O-B 2 K and final O-A equal to 1.5 K. The increments follow the vertical structure of the first eigenvector with the large increments near the surface and relatively small (compared to the near surface magnitudes) increments at model levels above 13.

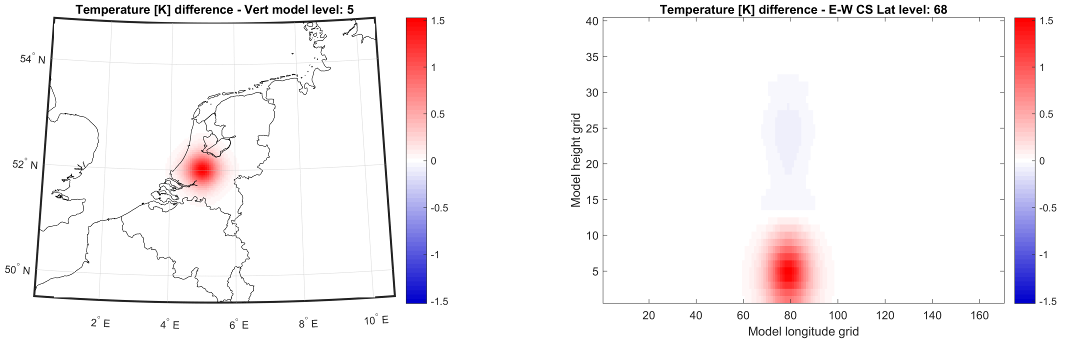
In figure 2.21b the increment in rainwater mixing ratio resulting from the radar reflectivity observation can be observed. However as demonstrated in section 2.5.3 and more specifically in the reflectivity partitioning of 2.78, a radar reflectivity observation can also lead to increments for snow and graupel if the temperature is below 5 °C. To demonstrate this, a new radar reflectivity is assimilated, but now at an altitude of 5 km. The temperature at the nearest model level (25, 583 hPa) is below freezing level. The increments in snow and graupel mixing ratio, respectively, have been plot in figure 2.26, both for horizontal model level 25 and the same vertical E-W cross section closest to the observation as before. The model r_{sn} increases from 0 to 0.2 $g\ kg^{-1}$. The same is true for r_g since the partitioning of reflectivity in 2.79 was constructed so that a reflectivity observation is caused by observed mixing ratios equal in magnitude. In this case both background fields are zero, so also the increments in r_{sn} and r_g are the same. Also, the fact that no vertical correlation is modelled in the background error covariance matrix for snow and rain can be recognized in the vertical cross section plot in figures 2.26c and 2.26c. Only the horizontal model levels that the observation is in between are affected. The observation increment then decays exponentially as normal, with the scaled length scale, although that is difficult to recognize since the length scale is so small.

Lastly, it is interesting to see how PSOT with a GPS ZTD affects model humidity. The ZTD is related to the total integrated amount of water vapour in a column. A PSOT is carried out using a ZTD innovation of 2 cm and an error of 0.4 cm. This leads to a profile of vertical increments shown in figure 2.27c, where the vertical cross section of relative humidity has been plot. Again, the vertical profile is governed by the eigenvectors for RH as shown in figure 2.9. The largest increments can be found at the mid model levels 15 to 25, the same model heights as where the first eigenvectors for RH are at their maximum magnitudes. Contrary to the radar PSOT which only assimilates an observation at a certain altitude, a GPS ZTD observation is a measure of the vertically integrated water vapour column. In the case of a radar PSOT the increments were mainly deter-



(a) Horizontal cross section (b) Vertical cross section

Figure 2.24: Change in model temperature resulting from a radar reflectivity PSOT of 40 dBZ at 3 km altitude. Both cross section at levels closest to the observation location latitude: 52.01°N.

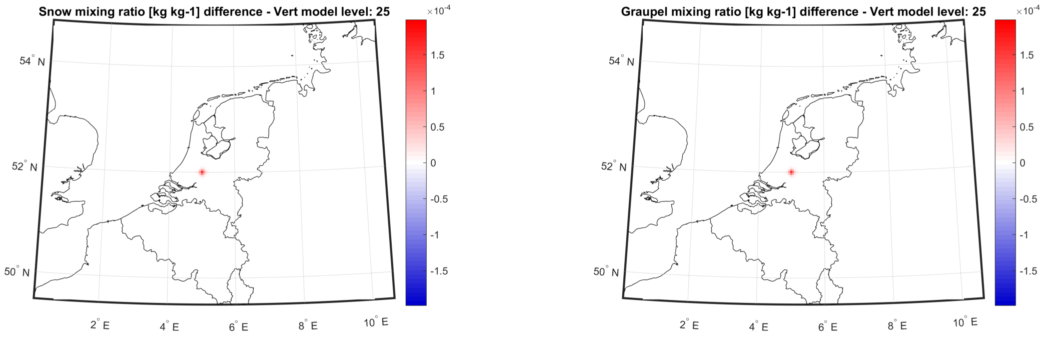


(a) Horizontal cross section (b) Vertical cross section

Figure 2.25: Change in model temperature resulting from a temperature PSOT using an innovation of 2 K at model level 5.

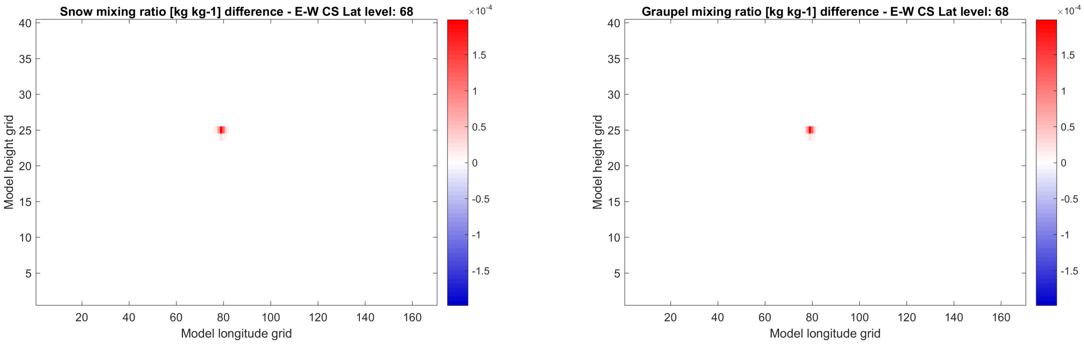
mined by a limited amount of eigenvectors which have their peak close to observation altitude/pressure. The model RH increment resulting from the GPS ZTD PSOT has more eigenvectors contributing to its shape since there are no additional constraints next to the fact that the total integrated water vapour has to be increased. Thus, the vertical increment magnitude profile is similar to the combination of all of the first eigenvectors which explained most variation in humidity in the background error statistics. It could have been more beneficial if relatively more humidity is added to lower levels, assuming most of the summer convective events initiate from the boundary layer. On the other hand, since only a measure of the total water vapour in a vertical column is given, no information is available on the profile in this column. It is not a bad assumption to have the biggest increments in the area where the model humidity estimation is most uncertain, as determined by the eigenvector decomposition of the background error covariance matrix.

The absolute magnitude of RH increments is limited for the PSOT of the GPS ZTD, considering the maximum RH increment is approximately 8%. However, as mentioned above, the ZTD observation can only cause an increment for the whole column and not locally as the radar. An observation that would increase the relative humidity throughout the column with more than 10% is also quite unrealistic. Moreover, we should rely on radar data assimilation for local increments and consider the ZTD assimilation more as an additional constraint for the humidity increments induced from radar.



(a) Horizontal, snow r_{sn}

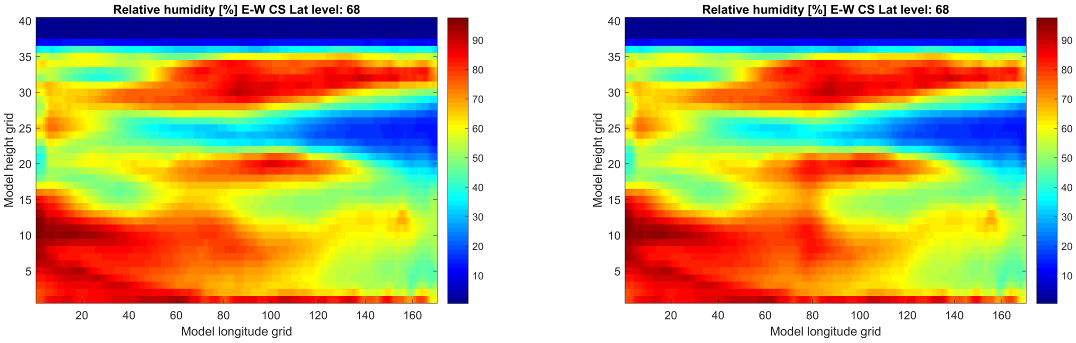
(b) Horizontal, graupel r_g



(c) Vertical, snow r_{sn}

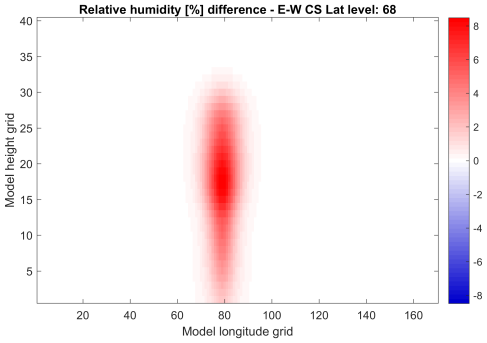
(d) Vertical, graupel r_g

Figure 2.26: Change in mixing ratios resulting from a radar reflectivity PSOT of 40 dBZ at 5 km altitude, above freezing level. Vertical West-East cross section at the model latitude closest to the observation latitude: 52.01°N.



(a) Model RH before DA.

(b) Model RH after ZTD PSOT.



(c) RH difference as a result of ZTD PSOT.

Figure 2.27: Change in relative humidity resulting from a GPS ZTD PSOT with O-B 5 cm. Vertical West-East cross section at the model latitude closest to the observation latitude: 52.01°N.

3

Data preprocessing

In data assimilation it is essential that the observations that are to be assimilated into a NWP model are of good quality. In the first place in the sense of precision because one would like the assimilation analysis to be as precise as possible, however possibly more importantly because inaccurate observations add unnecessary weight to the cost function (2.7). A biased observation will namely result in a large difference between model and observation $H(\mathbf{x}) - \mathbf{y}$ and consequently increase the second term of the cost function that quantifies the difference between model and observations. In the subsequent minimization of the cost function the analysis will tend to fit the model to these inaccurate observations, as they make a relatively large contribution to the cost function compared to the good observations. This way only a limited amount of biased observations can really degrade the analysis. To ensure that observations which are ingested by the data assimilation are of proper quality, several preprocessing steps have to be applied different per instrument or observation type. The observations that require most preprocessing are often the ones obtained through remote sensing, as they can contain significant noise or the data contains artefacts. Since this study is primarily concerned with the assimilation of radar data, the focus in the description of preprocessing steps will be on radar reflectivity and radial velocity. Most surface observations that directly measure atmospheric variables like temperature, pressure, humidity and wind speed are straightforward and don't require preprocessing. These observations often come from permanent stations from meteorological agencies or airports who also use these observations for climate analysis and thus have already applied several quality control procedures before the data is released. In the data assimilation experiments performed in this study no additional preprocessing steps for these so called conventional observations have been applied. The only quality control that remains present for every type of observation is the **maximum innovation check**. This test rejects any observation that fails the following test:

$$O - B < c\sigma_{obs} \tag{3.1}$$

Where $O - B$ denotes the difference between an observation O and equivalent observation from the model background B. σ_{obs} is the standard deviation of the observation and c an arbitrary constant that may differ per variable or observation type. By default $c = 5$, which will also be used in this study unless indicated otherwise.

3.1. Radar reflectivity

As mentioned, data from remote sensing instruments often requires some additional processing procedures before the desired degree of quality is obtained. In this section the preprocessing steps for radar will be discussed. First radar reflectivity will be examined with a description of the steps that are applied to the raw scan acquisitions to obtain data that is ready to be ingested by the data assimilation algorithm. Secondly, the procedure to obtain radial velocity observations that are ready for assimilation are discussed.

3.1.1. Overview

The radar data used in this study has been provided by the Koninklijk Nederlands Meteorologisch Instituut (KNMI) and Koninklijk Meteorologisch Instituut (KMI), the national meteorological services of the Netherlands and Belgium respectively. Raw volumetric radar scan data can be obtained through the KNMI website

(Lijense (2016)) for the (former) locations of Dutch weather radars at De Bilt (DB) and Den Helder (DH). These datasets contain both radar reflectivity and radial velocity observations for a total of 14 scan elevations. The exact elevation angles (α) together with maximum unambiguous velocity and radar range are summarized in table 3.1.

Scan	**Elevation [deg]**	**V_{max} [m s^{-1}]**	**Range [km]**	**Resolution [m]**
1	0.3	-	320	1000
2	0.4	23.85	240	1000
3	0.8	23.85	240	1000
4	1.1	23.85	240	1000
5	2	23.85	240	1000
6	3	31.8	170	500
7	4.5	31.8	170	500
8	6	39.75	150	500
9	8	39.75	150	500
10	10	47.7	120	500
11	12	47.7	120	500
12	15	47.7	120	500
13	20	47.7	120	500
14	25	47.7	120	500

Table 3.1: Operational radar settings at KNMI for radars at De Bilt and Den Helder. 14 scan angles are listed together with the corresponding elevation angle, maxmimum unambiguous velocity V_u, maximum unambiguous range and resolution.

Besides raw volumetric data from the Dutch radars also data from Belgian weather radars at Jabbeke (JAB) and Wideumont (WID) has been provided upon request. In this case only reflectivity data was available, for a total of 11 elevations for both radars. Details of each scan can be found in table 3.2.

		Jabbeke		**Wideumont**	
Scan	**Elevation [deg]**	**Range [km]**	**Resolution [m]**	**Range [km]**	**Resolution [m]**
1	0.3	300	500	250	250
2	0.9	300	500	250	250
3	1.5	300	500	250	250
4	2.2	300	500	250	250
5	2.9	300	500	250	250
6	3.8	300	500	250	250
7	4.8	150	500	250	250
8	6.5	150	500	250	250
9	9	150	500	250	250
10	13	150	500	250	250
11	25	150	500	250	250

Table 3.2: Metadata of the Belgium weather radars at Jabbeke and Wideumont, operated by KMI. The scan elevations of the 11 scans are the same for both radars, while the range and resolution vary.

Every scan of every radar used in this study contains data in polar coordinates (in range and azimuthal bins), with an azimuthal resolution of $d\theta = 1°$. The dimensions of the radar images thus depend on the range and resolutions which can be found in tables 3.1 and 3.2.

3.1.2. Clutter filter

Raw radar acquisitions can contain signals from any object in the atmosphere that is large enough to cause a detectable reflection of a radar pulse. In most cases that is some form of precipitation or hydrometeor present in the atmosphere. However also other objects can cause a reflection, like airplanes, swarms of insects or even objects that are aimed to produce erroneous reflections, like military chaff. An overview of possible sources affecting radar measurements is provided in figure 3.1.

Any reflection that is not of meteorological interest should be filtered. In the vicinity of the radar even really small objects can scatter back a detectable amount of power, since so close to the radar the incident

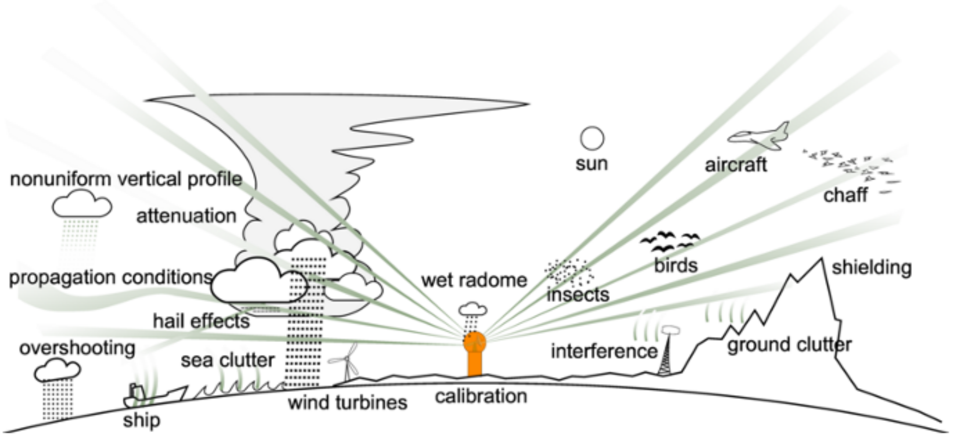


Figure 3.1: Schematic of possible error sources affecting radar measurements. Illustration by Peura et al. (2006).

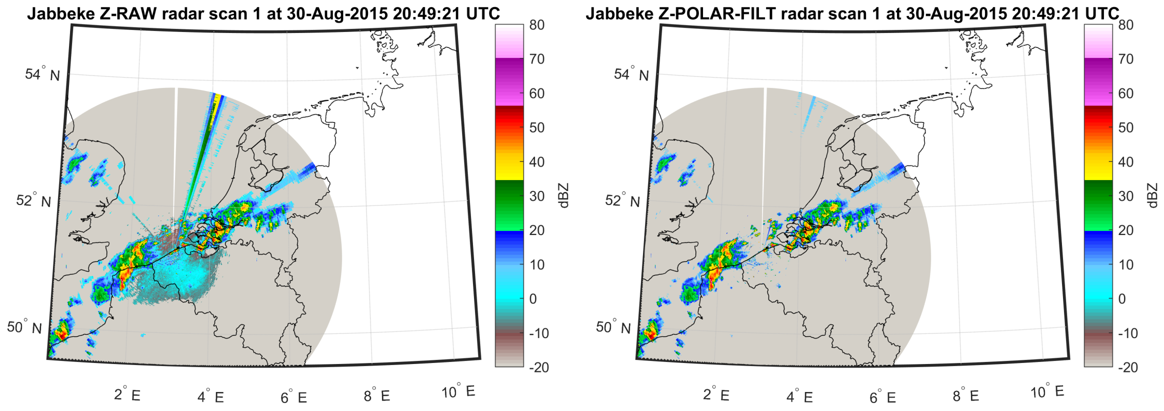
power is still really big. At distances further away objects have to be larger in order to cause a backscatter above a certain signal to noise ratio threshold. The backscatter by really small objects close to the radar means that a lot of noisy, unwanted reflections appear close to the radar in raw radar images at low elevation angles, as illustrated by the radar image of the lowest elevation scan of the JAB radar in figure 3.2a. Usually the first part of the range measurements $r < r_{\mathrm{min}}$ are discarded. In this case the first 20 kilometers are removed:

$$r < 20\,km \tag{3.2}$$

Also, some clutter is present at some distance from the radar. Since these noisy reflections are relatively low in magnitude compared to reflections of hydrometeors, radar reflections $Z < Z_{min}$ are also discarded:

$$Z < 7\,dBZ \tag{3.3}$$

Besides these basic filtering steps that are applied to the raw radar images, an additional step attempts to remove unwanted line reflections. An example of such a line reflection is given in figure 3.2a, which shows a raw radar image of the second elevation scan of the radar at Jabbeke (JAB). In essence these kinds of line reflections are detected by diagnosing a certain positive trend in reflectivity along a constant azimuth combined with observing that reflections in adjacent azimuth angles do not occur. Figure 3.2a shows the result of applying the mentioned filtering steps to the raw image in figure 3.2b. Evidently, the area of clutter in the neighbourhood of the radar is removed, as well as the majority of the unrealistic line reflections. A smaller apparent line reflection persists and is not removed because it is situated behind an area of precipitation, so that the linear trend estimation that is used to detect lines fails.



(a) Raw radar image

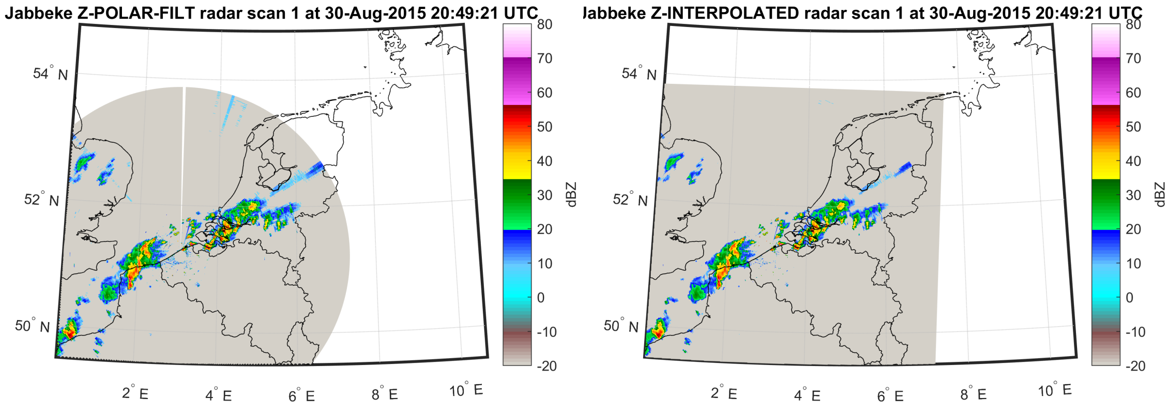
(b) Filtered radar image

Figure 3.2: Illustration of basic filtering step with radar images of scan 1 of the radar at JAB at 20:45 UTC on 30-08-2015.

3.1.3. Transformation from polar to geographic coordinates

The original and filtered radar images from the previous step both have polar coordinates. To be able to perform some more advanced filtering steps, a gridded image with pixels of equal area is required. The first step is to transform from polar (r, θ) to Cartesian (x, y) coordinates, so that these new coordinates can be used to interpolate to a grid in Cartesian coordinates. In this transformation it is assumed that the raw polar radar measurement locations are contained on a two dimensional plane. Technically this is not true since the radar scans actually form a cone when $\alpha > 0$, but especially for low elevation angles this is a valid approximation.

In order to obtain a reflectivity grid, the reflectivity observations from the coordinate transformed locations are interpolated to the grid locations. The rectangular grid is defined with the same resolution as the original range measurements of the corresponding scan. Logically, in a polar coordinate system the bin area increases laterally with range, resulting in increasingly larger spacing between azimuth angles with range. This is also visible in the raw radar images (as in figure 3.3a) where the pixels (and in reality also the volume of air that causes a backscatter) become more stretched. Hence the interpolation close to the radar where observations are denser is more precise. A linear, two dimensional interpolation is used to estimate reflectivity at the grid locations. The interpolation of a polar image (figure 3.3a) to a rectangular grid results in the square grid as shown in figure 3.3b. The interpolated square image is really similar to the polar image.



(a) Filtered polar radar image

(b) Interpolated square radar image

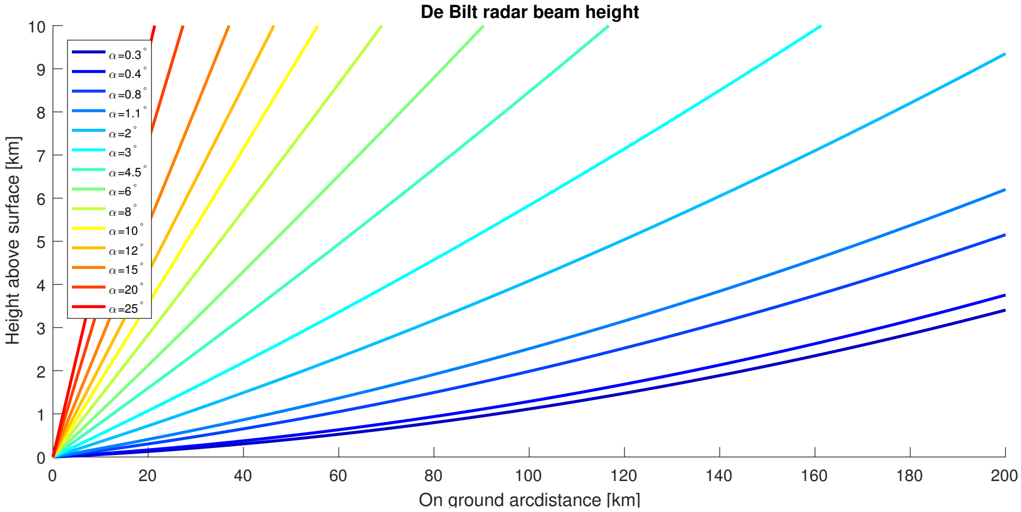
Figure 3.3: Illustration of grid interpolation step with radar images of scan 1 of the radar at JAB at 20:45 UTC on 30-08-2015.

However, before a radar observation is to be assimilated, as for any observation, it should be assigned a geodetic coordinate in terms of the latitude, longitude and (ellipsoidal) height. Being a three dimensional coordinate system, it requires estimating the height of the radar beam at a given location. To that end, the widely used approximation from Doviak and Zrnic (1993) is used:

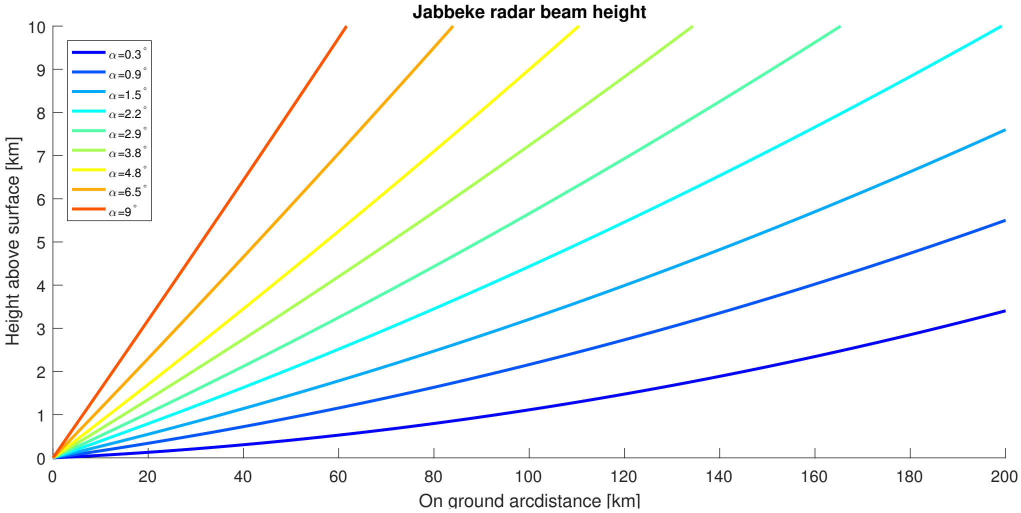
$$h = \sqrt{r^2 + (k_e a)^2 + 2 r k_e a \sin\alpha} - k_e a + h_{radar} \tag{3.4}$$

Where h represents the height above the (spherical) Earth's surface, r the range of the radar beam, a the radius of the Earth (taken to be $a = 6370\ km$) and α the scan elevation. Because the radar is usually not at the surface, the height of the radar h_{radar} above the surface has to be added to the approximation. The coefficient k_e relates to the bending of a beam with height and depends on the vertical gradient of the refractive index n with height. The refractive index of air is a function of pressure, temperature and humidity. For the Standard Atmosphere (USAF (1976)), it can be shown that $k_e = \frac{4}{3}$ (Doviak and Zrnic (1993)). Thus it is commonly referred to as the *"four-thirds Earth model"*. Of course the actual ray path of each scan is different from the modelled one because the atmosphere is different from the Standard Atmosphere. This non standard behaviour of a radar beam is called **anomalous propagation**. For example, when a strong temperature inversion exists, radar beams can be reflected back to the surface by the sudden change in density between layers. The radar beams in turn reflect back from the surface into the atmosphere. That way radar beams can get trapped in a shallow layer, knows as **ducting**. These are extreme cases however which would require a specific analysis with atmospheric profiles. In this study the approximation of the beam height in a Standard Atmosphere from Doviak and Zrnic (1993) is considered sufficient.

Using the scan elevations and ranges mentioned in table 3.1, a plot of each scan height for the radar at DB is shown in figure 3.4a. The radar at DH uses the same elevation angles and ranges, making the beam



(a) Radar beam height of the scans of the radar at Den Helder (DH).



(b) Radar beam height of the scans of the radar at DB.

Figure 3.4: Radar beam height using the approximation in equation 3.4.

height per scan identical to the one from DB. Additionally for the radar at JAB a height plot per scan using the metadata in table 3.2 is included in figure 3.4b. The radar at Wideumont (WID) uses the same scan strategy but with smaller ranges so a height plot would look the same, only extending less far. Besides the range-height plots in figure 3.4, the points of the lowest scan of 0.3 ° having the same height above the surface are plot for heights ranging from 0.5 to 8 km in figure 3.5. These points mark the lowest heights at which the radar is able to detect a reflection. Any precipitation at lower heights at this location is thus overshot (see figure 3.1).

3.1.4. Neighbourhood filtering

After the basic filtering steps described in 3.1.2, some undesired areas of reflection may still exist in the image. In the next step of filtering radar images, reflectivity areas beneath a certain threshold area are discarded, because these areas of reflection are considered to be clutter, too small and incoherent to resemble an area of precipitation. For each reflective area on the image, the group size is calculated by counting the number of connected pixels that contain a reflection. Areas with a group size smaller than a defined threshold are removed. The threshold of number of pixels that is appropriate naturally depends on the resolution. Therefore a minimal area is defined instead, that can be combined with resolution to determine a minimal group size in terms of pixels.

Besides this so called small neighbourhood filtering, also an attempt is made to suppress noisy reflections by comparing a pixel value to the average of a 3x3 block of surrounding pixels. To that end, reflections are

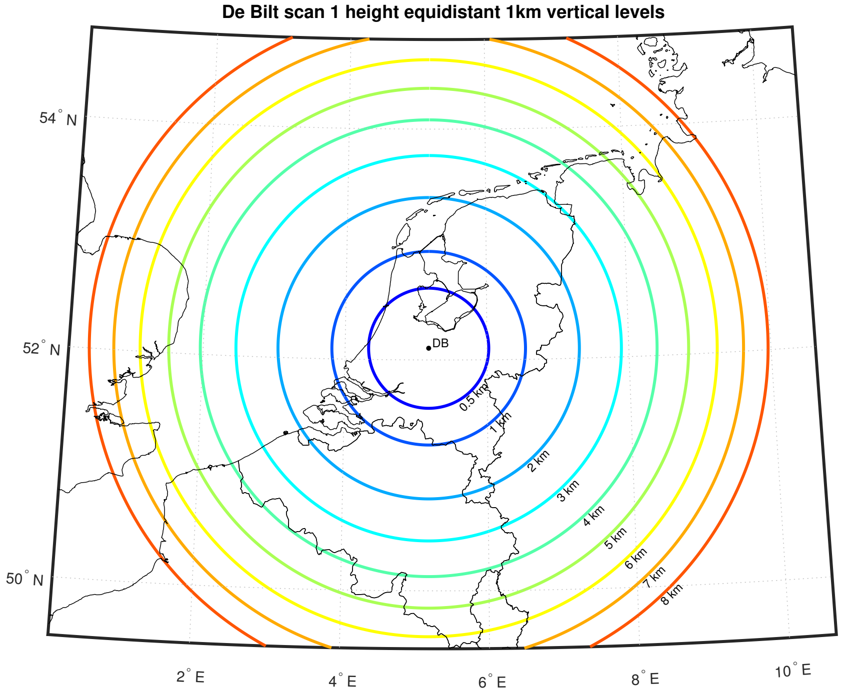
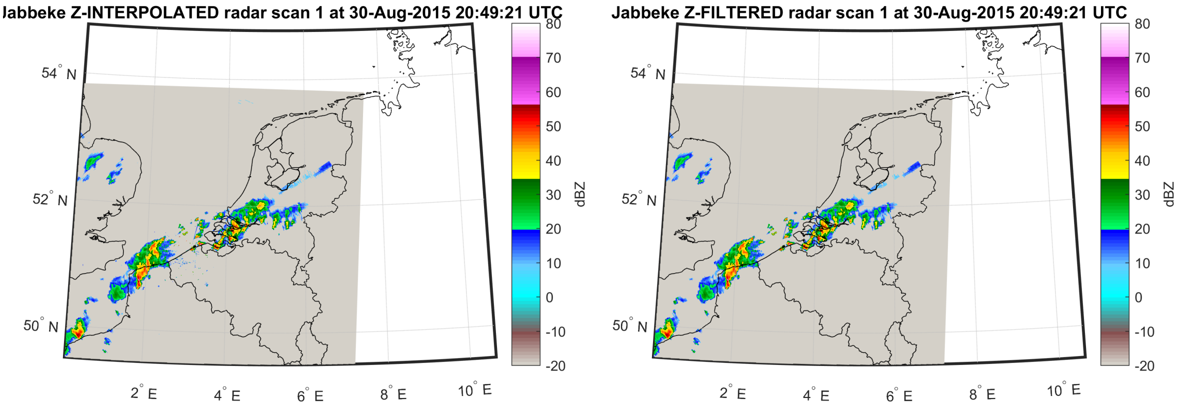


Figure 3.5: Equidistant points from the DB radar at heights ranging from 0.5 to 8 km.

considered to be unrealistic when they deviate 15 dBZ from the 3x3 block average. Applying these filtering steps to the interpolated image of 3.6a results in a filtered image as shown in figure 3.6b. Although hard to recognize, some small areas of reflection especially near the Belgian coast are removed. Other than that no significant changes are made to the image.



(a) Interpolated square radar image

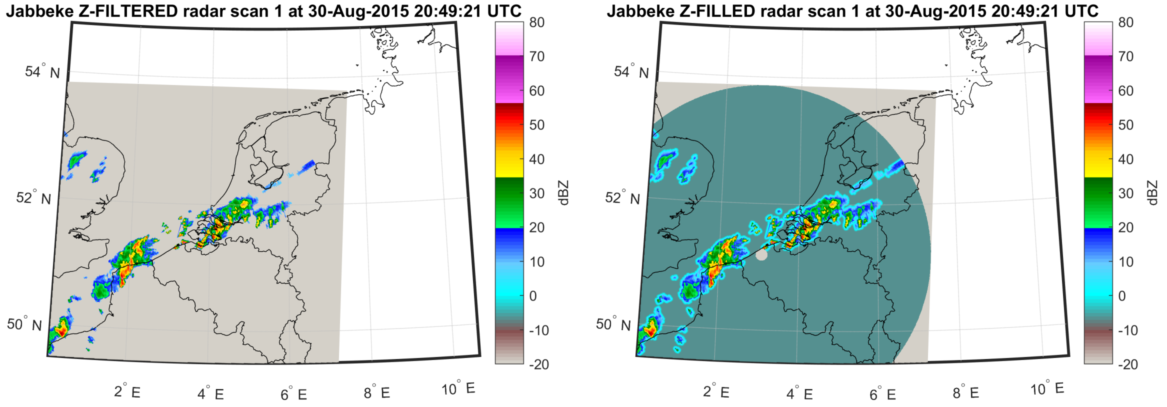
(b) Filtered square radar image

Figure 3.6: Illustration of neighbourhood filtering step with radar images of scan 1 of the radar at JAB at 20:45 UTC on 30-08-2015.

3.1.5. Null-echo and image filling

At this point an image is obtained which is thought to accurately represent precipitation. However, for locations/pixels where no significant reflections are measured a pixel is classified as containing no measurement. At the same time though, the fact that no reflection is obtained does provide important information that there is no precipitation. This is referred to as a **null-echo**, which, when assimilated as described in section 2.5.5, helps to suppress erroneous model precipitation. Without assimilation of a null-echo, it is impossible to penalize model precipitation that does not exist in reality.

Every location with no data within the original radar range per scan is flagged as a null-echo point. Except for the edges of the areas of precipitation, where radar reflectivity is first gradually relaxed to a value of



(a) Filtered square radar image

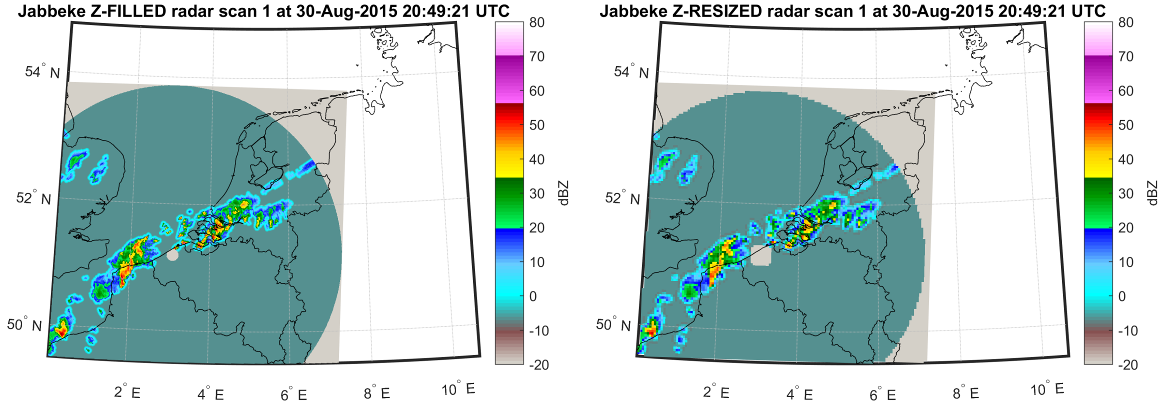
(b) Null-echo filled radar image

Figure 3.7: Illustration of null-echo filling step with radar images of scan 1 of the radar at JAB at 20:45 UTC on 30-08-2015.

$-20\ dBZ$ with a slope of $-3\ dBZ\ km^{-1}$. The relaxing is applied to allow for a more natural transition from cores of precipitation to the surrounding areas of no reflection. Also, if null-echoes are right next to the precipitation cores without any transition zone, a slightly inaccurately positioned null echo could unnecessarily suppress model precipitation. Figure 3.7b shows how the filtered image in 3.7a looks like when the no data locations are filled with null-echoes, which have been highlighted with a dark green/blue tone in a circle that contain the points within the original radar range.

3.1.6. Resampling

Finally, before the radar data is ingested, the locations are resampled to a coarser grid. The first reason being that the model resolution of 4 km is much coarser than the radar range resolution, varying from 250 m to 1 km per radar and scan (see tables 3.1 and 3.2). Such a high density network of observations is unnecessary for the current model resolution. An additional benefit is the reduced data volume and processing time required in the data assimilation minimization algorithm later. Moreover, the data thinning helps to make the observations less correlated, as in the definition of the observation error covariance matrix **R** it is assumed that the observations are uncorrelated. Evidently they are not truly uncorrelated, since the filter that is applied in 3.1.4 is based on the assumption that there is a certain degree of smoothness (and thus correlation) between adjacent pixels.



(a) Null-echo filled radar image

(b) Resampled radar image

Figure 3.8: Illustration of resampling step with radar images of scan 1 of the radar at JAB at 20:45 UTC on 30-08-2015.

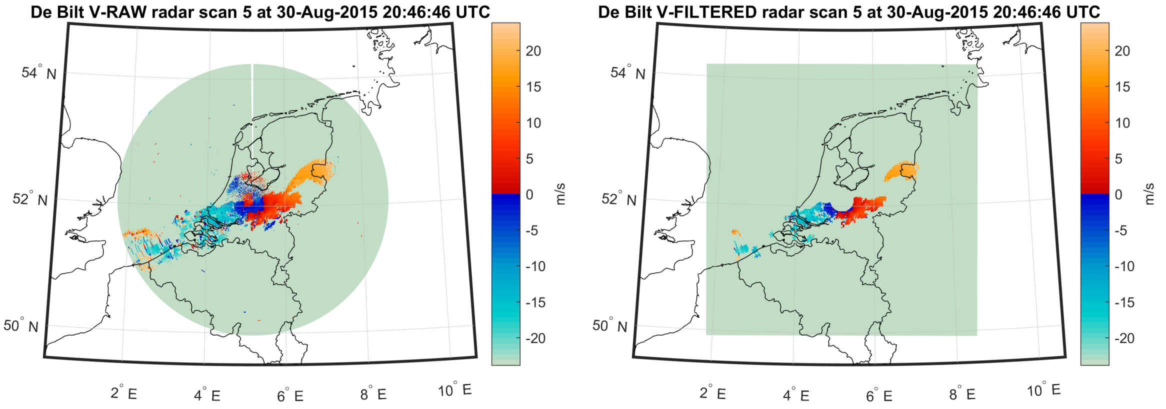
So, each radar image is resampled to a grid with a resolution being equal to the model resolution. For the resampling a bicubic interpolation method is used. An example of the resampling to a 4 km resolution is shown in figure 3.8, where 3.8b is the resampled version of 3.8a.

3.2. Radar radial velocity

Besides reflectivity radars are also capable of measuring Doppler velocity by looking at the Doppler frequency shift of moving targets, as explained in section 2.6. Just as was the case for reflectivity, the raw Doppler velocity measurements are not yet suitable to assimilate directly. Raw images might contain clutter from non meteorological objects as well. Radial velocities might even be aliased, a phenomenon unique to Doppler radial velocity as discussed in section 2.6. An example of aliased velocity can be observed in figure 3.9a, where near the Belgian coast the positive (orange) radial velocity measurements indicate a wind opposite to other observations in the image. Some efforts were made to develop a de-aliasing method, but conventional methods like one based on the velocity azimuth display (VAD) are infeasible because the lack of continuity in the velocity observations (depending on the size and shape of precipitation areas) inhibits the creation of a VAD. The method by de Haan (2013) suggests to use the upper radar scans with bigger elevation angles and higher unambiguous velocities to determine the general direction of the wind speed, which can then be used to map any aliased measurements to real velocity again. However this method did not prove to be robust enough, especially when little to no radial velocity measurements at higher elevation angles are available. Combined with the observation that often the area of aliased observations is limited and aliasing mostly occurs for scans 2-5 which have a relatively low unambiguous radial velocity (see table 3.1), the choice is made not to apply any de-aliasing algorithms and rely on the maximum innovation check as defined by equation 3.1 to reject these observations. The observation standard deviation multiplier c in equation 3.1 is set to 3 instead of 5 for radial velocity, mainly as a conservative safeguard to any observation that might just pass the maximum innovation test but significantly degrade the model wind forecast. This does imply that aliased observations unfortunately can not be used in the assimilation. In a more elaborate de-aliasing scheme one could include model wind estimates in the de-aliasing by adding $k \cdot 2v_{\max}$ to the aliased observations until they pass a tight maximum innovation test.

3.2.1. Reflectivity filtering

Only radial velocity measurements from meteorological phenomena are of interest for assimilation. As was the case for reflectivity, we would like a radial velocity image to contain only measurements of a meteorological source like precipitation. The first step aims to filter the observation locations for which also a reflectivity of $Z > 7\ dBZ$ exists in the filtered polar reflectivity image, see figure 3.2b. An example of a raw radial velocity image is given in figure 3.9a, showing the raw measurements of the 5th scan of the DB radar. If the basic minimum reflection filter is applied, a filtered image as shown in figure 3.9b is acquired. Also, the polar image is interpolated to a square grid analogous to reflectivity as described in section 3.1.3. The resulting interpolated filtered image is shown in figure 3.9b.



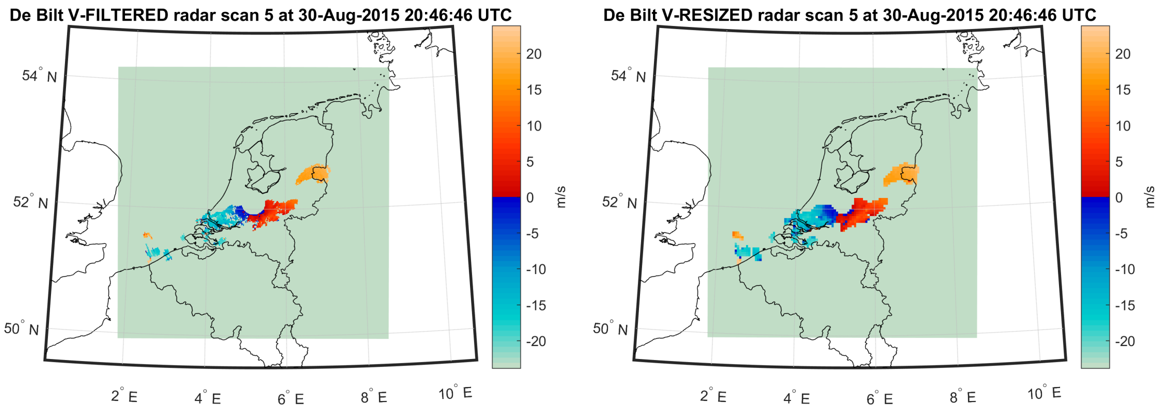
(a) Raw radar image

(b) Interpolated and filtered radar image

Figure 3.9: Illustration of filtering step with radar radial velocity images of scan 5 of the radar at DB at 20:45 UTC on 30-08-2015.

3.2.2. Resampling

The original resolution of the radar radial velocity measurements is unnecessarily high compared to the resolution of the model domain. Hence, similar as for radar reflectivity, the radial velocity measurements are resampled to the 4 km model resolution using the same procedure. The result of this resampling applied to the filtered image of figure 3.10a is shown in figure 3.10b. Approximately the same area of radial velocity measurements is now described by a coarser grid of measurements.



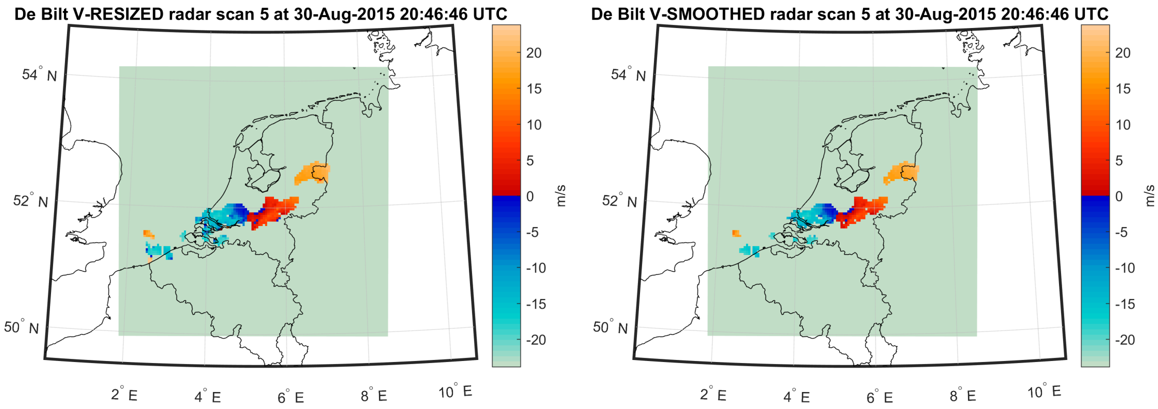
(a) Interpolated and filtered radar image

(b) Resampled radar image

Figure 3.10: Illustration of resampling step with radar images of scan 1 of the radar at JAB at 20:45 UTC on 30-08-2015.

3.2.3. Outlier removal

Several experiments with the resized radial velocities images (not shown here) indicated unrealistic changes to model wind velocity because several noisy observations degraded the DA analysis. For example, the assimilation of these observations locally resulted in a stop and reversal of jet stream winds at high altitude. To limit the amount of these suspected wrong observations, the pixels that do not conform to the surrounding observations are filtered out. That is, observations for which the measured velocity deviates more than twice the sample standard deviation from the mean of the surrounding observations are discarded. A minimum of 4 neighbouring pixels is required for the a valid mean and standard deviation, otherwise the observation is discarded as well. So at the same time this removes single observations and really small regions of observations for which the consistency of observations cannot be checked.



(a) Resampled radar image

(b) Smoothed radar image

Figure 3.11: Illustration of smoothing step with radar radial velocity images of scan 5 of the radar at DB at 20:45 UTC on 30-08-2015.

By applying these steps to the resized radial velocity image in figure 3.11a, the filtered image in figure 3.11b is obtained. Comparing the two it is evident that some individual pixels have been removed. Especially in the area of observations near the Belgian coast where the edges of the region of radial velocity measurements are not really consistent with the majority of the area, unlikely observations are removed. In the other two main areas of measurements only locally a measurement is rejected.

3.3. GPS Zenith Total Delay

3.3.1. Data source

The GPS ZTD measurements used in this study have been supplied by the EUMETNET GPS Water Vapour Programme (E-GVAP). The E-GVAP is a joint venture of several European national meteorological services and geodetic institutions aimed at providing near real-time measurements of atmospheric water vapour available for the numerical weather prediction community. Each participating institution employs a network of GPS stations, who also process the raw GPS measurements to estimate ZTD as a required step in accurately determining position. Originally the atmospheric delays that GPS signals experience are slant path delays, which are converted to ZTDs by means of a mapping function in the processing, see e.g. Niell (1996). In this study GPS ZTD measurements from Geo Forschungs Zentrum/German Research Centre for Geosciences (GFZ), KNMI, Royal Observatory of Belgium (ROB), Institut Géographique National (SGN) and United Kingdom Meteorological Office (UKMO) are used for assimilation. The locations of the GPS stations for each of these institutions have been plot in figure 3.12. As can be seen from this map, some stations are part of multiple networks. In the assimilation pre-processing as a part of the observation thinning any duplicate observations will be ignored, only the one with the best precision is kept.

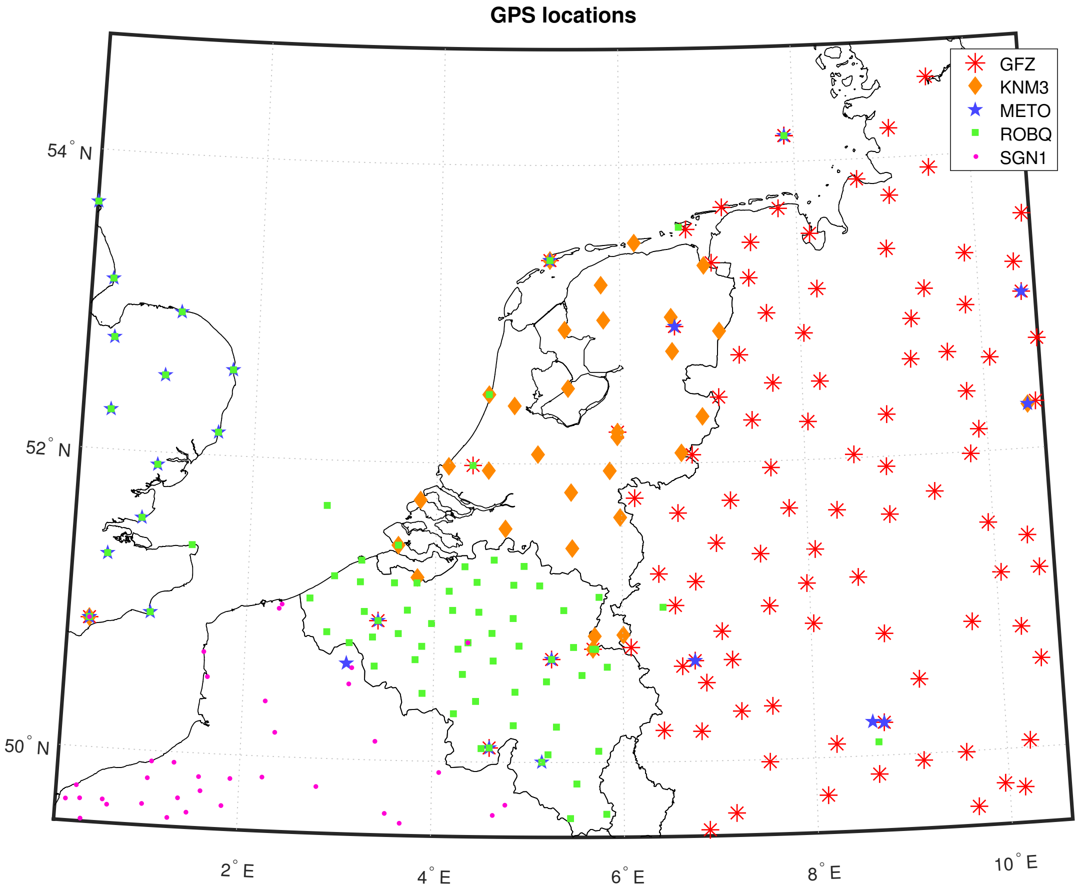


Figure 3.12: Locations of GPS ZTD measurements per institution.

3.3.2. Height correction

Last but not least, observations of zenith total delay (ZTD) from GPS receivers also require an additional preprocessing step before observations can be assimilated. Recall from section 2.7 that besides the ZWD that is proportional to atmospheric water vapour which is of interest, the major part of the ZTD consists of the signal delay caused by the interaction with air itself, referred to as the ZHD. As indicated by equation 2.102, the hydrostatic delay is estimated using pressure, which is a measure of the air mass through which a signal has to travel. If differences between model and observed ZTD are to be attributed only to differences in ZWD caused by water vapour, it requires ZHD and thus pressure to be equal for both ZTD estimates. As pressure decreases with altitude, this implies that model height should be equal to height of the GPS receiver. In practice however these heights are never the same. For example, GPS receivers are often mounted on top of a building to obtain a better signal strength and less possible interaction of the signal with the surroundings. Even more important is the possible offset between the model terrain height and the actual height of the surface. While the WRF model uses a 30 arcsecond topography dataset that is quite detailed for a mesoscale weather model, locally differences between modelled and real topography may exist, as illustrated by figure 3.13. While this height discrepancy is the largest in mountainous areas where the model elevation model is not able to describe complex local topographic features, in fact for every point with a height difference a **height correction** should be applied. The correction is the estimated additional atmospheric delay that would be caused if the GPS receiver would be moved from receiver height to the surface:

$$\Delta ZTD = 10^{-6} \int_{z_{sfc}}^{z_{GPS}} N(z)\, dz = 10^{-6} \sum_{k=1}^{k_{GPS}} N_k z_k \tag{3.5}$$

In fact this is the same approximation as used in the observation operator for ZTD in 2.104, only now integrated from the model surface height z_{sfc} to the height of the GPS receiver z_{GPS}. N_k represents the modelled refractivity for model layer k. Logically, the summation is from the surface until layer k_{GPS} at the height of the GPS receiver. In most cases the GPS receiver is above the model surface, so that the observed ZTD is smaller than the ZTD that a GPS signal would experience when the receiver would be positioned at the model surface. Thus the ZTD height discrepancy correction ΔZTD in equation 3.5 is positive. In case the GPS receiver is below model surface topography, ΔZTD is negative. Instead of a thickness weighted summation of refractivity as in 3.5, ΔZTD then is estimated from the single refractivity of the lowest model layer combined with the height difference between model surface and GPS receiver.

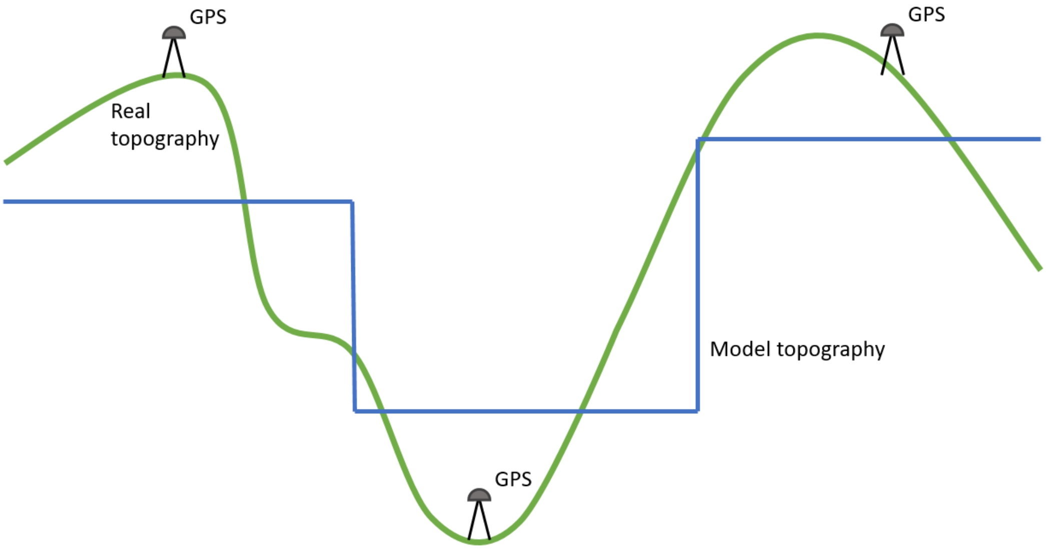


Figure 3.13: Schematic illustration of difference in model topography and GPS receiver height.

GPS measurements use an ellipsoid as a geodetic reference system. That is, the Earth is approximated with an ellipsoid and any heights measured with a GPS are with respect to the ellipsoid surface. Consequently, the heights measured determined by GPS positioning are referred to as *ellipsoidal height*. In contrast, the heights of the model topographic surface use mean sea level as a reference. Globally, the mean sea level is approximated by the *geoid*, the shape Earth's sea level would attain solely due to gravity and rotation. The height with respect to mean sea level is called *orthometric height*. For a valid computation of the ZTD height correction, GPS ellipsoidal height should be converted to orthometric height as well. This is achieved by subtracting the difference between geoid and reference ellipsoid knowns as the *geoid undulation* from the ellipsoidal height. This principle is illustrated by figure 3.14.

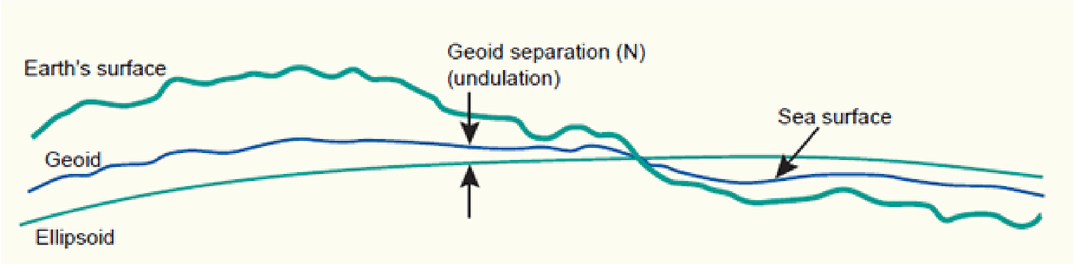


Figure 3.14: The geoid undulation is defined as the height difference between the geoid and the ellipsoid that is used in the geodetic reference system. The ellipsoidal height is the height from the surface to the ellipsoid, as measured by GPS receivers. Finally, orthometric height is the height separating the Earth's surface and mean sea level, the geoid.

3.3.3. ZTD error

In section 2.4.2 it has been discussed that an observation error used in data assimilation is more than just the instrumental error. The observation error also has to account for errors in the model observation operator and for representativeness, resulting from the inability to describe really detailed or small scale physical processes by a model. Especially for remote sensing observations which require various sets of approximations and assumptions to produce a model variable estimate. A proper estimate of the ZTD observation error for data assimilation is not well known (Vedel and Huang, 2004), other than the fact that it can be significantly larger the the instrumental error on its own, which varies between $1-3\ mm$ for the measurements from the processing institutions mentioned in section 3.3.1. Therefore the instrumental error specified by the processing is scaled with a factor 3 to account for these additional errors, similarly as done in GPS assimilation studies conducted by Vedel and Huang (2004) and de Haan (2013).

4

Experiment setup

This chapter will discuss how the data assimilation experiments are setup. First the choice of WRF model domain with its settings is explained. A variety of settings, schemes and parametrizations can be employed within the WRF model for example to describe the boundary layer, radiation, clouds and microphysics. That is followed by a description of the data assimilation strategy that involves cycling between model and assimilation. The data assimilation strategy is tied to several data assimilation experiments that are carried out, described in section 4.2. Finally, the verification using radar composite images is explained in the last section 4.3.

4.1. Model domain and setup

In section 4.1.1 a brief overview is given of the WRF model domain used for the data assimilation experiments, along with a description of how such a model is initialised. Next, in section 4.1.2, relevant model parametrizations are discussed and how they are relevant for precipitation forecasting. The section is concluded in 4.1.3 with a brief description of the data assimilation cycling strategy.

4.1.1. Domain

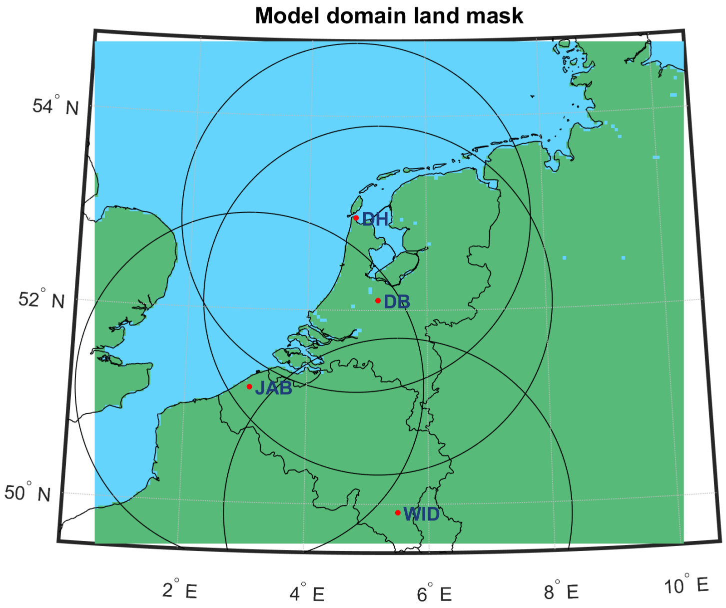


Figure 4.1: Map of the WRF model domain. The domain consists of 171 × 146 grid points at a 4 *km* resolution. The circles for each radar show the points with an arc length distance of 200 *km* to the radar, giving an indication of the range of the radars.

For the data assimilation experiments, a WRF model domain has to be defined. For the benefit of future comparisons, the model domain of the operational WRF model employed by Weerplaza has been used. This operational model consists of 171 (East-West) by 146 (North-South) grid cells of 4 km resolution. Vertically, the model is made up of 40 pressure layers from the surface to a model top of 50 hPa. A Lambert conformal projection is used to define a two dimensional grid. The Lambert projection is suited for mid latitudes because it minimizes the distance distortions (between a sphere and two dimensional surface) for these areas. Figure 4.1 shows a map of the domain that is used for all data assimilation experiments. The green colour shows the extent of the domain. It covers the Netherlands, Belgium and parts of Germany, Luxembourg, France and the United Kingdom. The domain size is limited, in the first place because computational resources are also limited. A 24 hour forecast already requires several hours of calculations. Ideally, one would like to have a large, high resolution domain that covers all observations to describe the real atmosphere as close as possible. However if the domain size for example would be doubled one could argue if the 4 km resolution is still of added value at areas where no high resolution observations are available that can provide the model with detailed information. In this case a large part of the high resolution radar observations are covered by the model, except for a fraction of the radar located at Wideumont in Belgium. The relatively small domain does imply that the effects of data assimilation possibly cannot be recognised anymore for forecasts longer than e.g. +12 hours, as the assimilated high resolution information has left the domain again by that time. Because this study focuses on short term precipitation forecasting with a primary interest for the 0-6h range, this is not considered to be a problem. However it is noted when it is desired to have a longer forecast with the same amount of detail the domain should be extended. At the same time that also requires high resolution observations in these new areas of the domain, otherwise the 4 km resolution provides apparent detail while the source data used to initialize the model is much coarser, thus not really exploiting the capabilities of the higher resolution.

The 4 km resolution is chosen for it's assumed ability to resolve large scale convection and at the same time does not require extensive computational resources. The range of resolutions varying from $1-10$ km is commonly referred to as the grey zone, because at these resolutions there is a lot of debate whether or not clouds and convection are explicitly resolved or that parametrization is required. For grid scales smaller than several hundreds of meters large eddy simulation (LES) models can capture the big eddies and clouds are described with the model equations. At resolution > 10 km cumulus parametrization is often employed to describe subgrid convection. The 4 km resolution is assumed to reasonably describe mesoscale convection and precipitation so that no parametrization for convection is employed. This assumption is discussed further in 4.1.2.

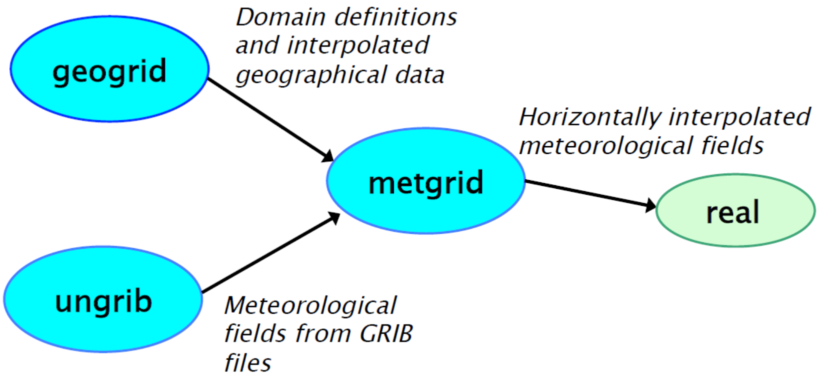


Figure 4.2: Flowchart showing the various steps in the WRF WPS.

A WRF model is initialised by invoking the WRF Preprocessing System (WPS). The flowchart in figure 4.2 shows the various components of the WPS. The first step defines a geographical grid using the specified projection, domain size and resolution and then interpolates various geographical datasets to the model grid. For example, figure 4.3 shows the model domain elevation interpolated from the 30 arc second Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010). The blue colours indicate areas below sea level. Also, a 21 class land use type dataset derived from MODIS images is used to define landuse fractions for the model grid. Other geographic datasets include soil type, vegetation fraction and surface albedo. After the static geographic data for the domain have been constructed, the WRF model requires an initial state for the atmosphere to start a forecast. The limited area WRF model for the Benelux is initialized with meteorological

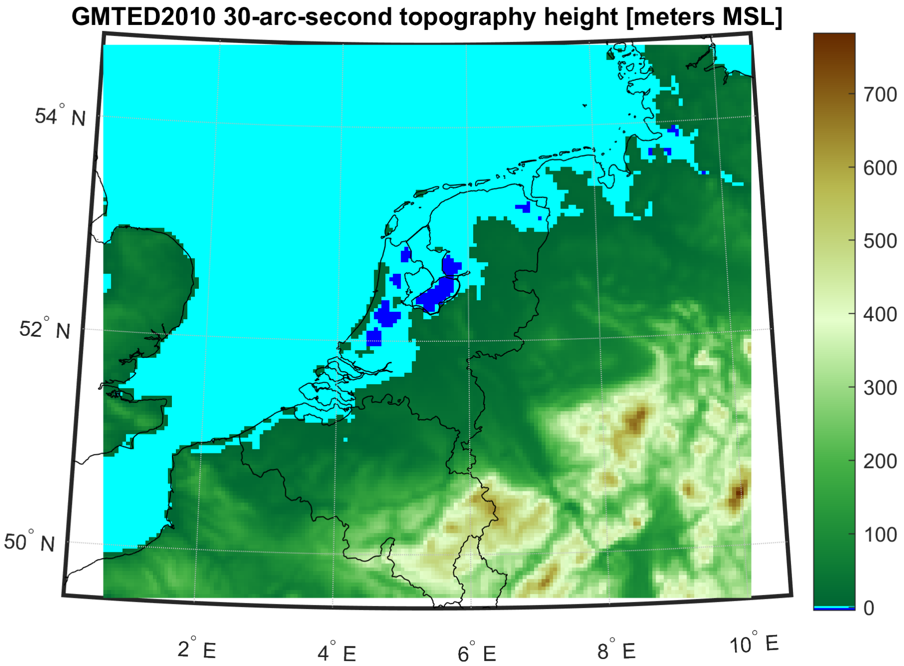


Figure 4.3: Terrain height in the WRF model domain.

fields from the Global Forecast System (GFS) model. This is a global NWP model that is maintained and developed by National Centers for Environmental Prediction (NCEP), part of NOAA, and is updated four times per day. The finest resolution of the GFS model available is 0.25°, roughly corresponding to 28 kilometer and additionally GFS has 32 vertical model levels. Meteorological variables from the GFS model are interpolated horizontally to the domain grid. E.g. horizontal wind speed components from GFS are rotated to align with the x- and y-axis of the WRF domain. Variables that are not defined in the source data are derived from related ones. For example relative humidity is computed from temperature, pressure and specific humidity. The 'real' module defines the vertical grid for the WRF model and then interpolates the meteorological data vertically to the model vertical levels. WRF uses a mass based terrain following vertical coordinate, defined as:

$$\sigma = \frac{p - p_{top}}{p_s - p_{top}} \tag{4.1}$$

Variables p, p_t and p_s represent pressure at a certain point, model top pressure and terrain surface pressure, respectively. σ ranges from 0 at the model top to 1 at the surface. Usually σ spacing is more dense near the surface, as weather is primarily driven by processes near the surface. Because the terrain surface pressure is used, the vertical coordinate follows orographic features in the model topography. Figure 4.4 shows an illustration of how the vertical coordinates behave when a mountain is encountered. For areas with flat topography the σ levels are parallel to levels of constant pressure. The meteorological variables are interpolated from pressure levels to defined σ levels. Finally, lateral boundary conditions (LBC) are calculated from GFS to provide continuity at the model boundaries throughout the forecast. Since GFS provides 3h output, the lateral boundary conditions can only be updated every 3 hours. In the mean time the lateral boundary conditions for a WRF domain are interpolated temporally.

It is common for regional weather models to nest a limited area, higher resolution domain into a big coarse resolution domain that is also solved non hydrostatically. The so called outer domain acts as a step in between the meteorological source data (in this case from GFS) and the high resolution domain that is of primary interest. The main incentive is to provide more detailed lateral boundary conditions than GFS provides and also at every time step instead of constant lateral boundary conditions with an update frequency equal to the GFS source data. Despite the fact that it provides better lateral boundary conditions, nesting is not employed since it requires additional computational effort to solve the outer domain as well. Moreover, because the interest is in the nowcasting range of 0-6h rather than several days, it is not absolutely necessary

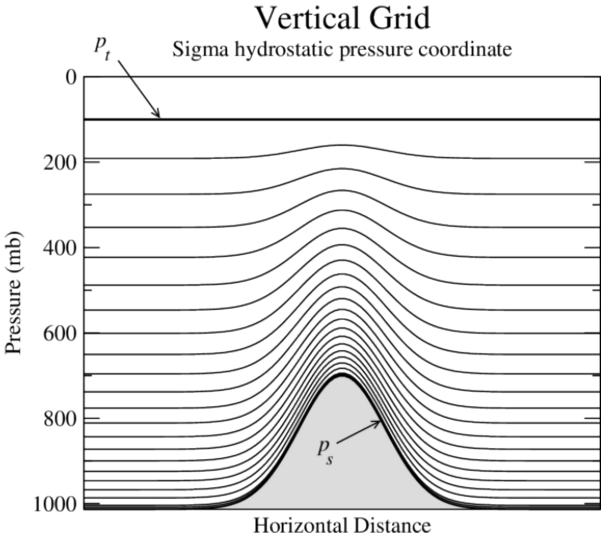


Figure 4.4: Example of σ mass based vertical coordinate. p_t and p_s represent the model top and terrain surface pressure. The vertical levels follow any features in model topography. Illustration from Boone (2017).

to have an outer domain that provides proper boundary conditions for several days.

4.1.2. Physics parametrizations

For this study the latest version of WRF and WRFDA, V3.9, has been used. In the WRF model a wide variety of parametrizations can be employed. These parametrizations can be regarded as an attempt to account for physical processes that are too small (i.e. subgrid scale) for the model to resolve explicitly using the model equations, yet are important for a realistic forecast. For example clouds might have length scales of hundreds of meters and therefore cannot be described by a model with a 4 km resolution. However, they play a major role in the atmosphere because they affect the radiative transfer and also because of precipitation formation. In general the physics parametrizations in WRF can be categorized into five classes, also illustrated in figure 4.5:

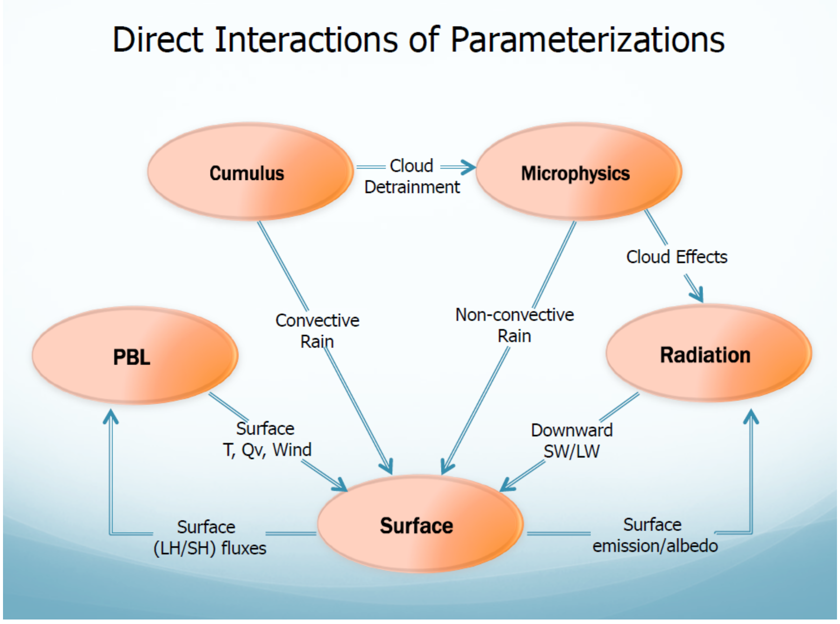


Figure 4.5: Five main categories of parametrizations within the WRF model. For each category the interaction with other parametrizations is indicated with an arrow. Illustration from WRF tutorial presentation Dudhia (2017).

1. **Microphysics**
   Cloud droplets, condensation nuclei, rain droplets, ice crystals, snow flakes and other hydrometeors are formed on scales much smaller than the model resolution (hence the term "micro"). So, for example the formation and growth of cloud droplets has to be estimated from other model variables like water vapour mixing ratio, temperature etcetera. The microphysical processes are important to model because they play a significant role in the energy and moisture budget, affecting moisture and heat tendencies in the model. For example condensation of water vapour into liquid water or freezing of liquid water causes latent heat release, which helps to drive convection in the atmosphere. The microphysics also control how much and and at which rate hydrometeors form, influencing the transportation of moisture. Moreover, microphysical parametrization enables estimation of precipitation. Rain, snow and graupel that form are too heavy to be suspended by the air and subsequently they will fall to the ground with a certain fall speed, often a function of hydrometeor concentration.

   The so called *microphysics schemes* differ in the complexity to which they describe the microphysical processes. For example, the simplest 'warm rain' schemes just consider the gas and liquid phase, so no snow or ice can form. More sophisticated schemes also accommodate modelling of ice phase particles like ice, snow, graupel and hail. Each of these hydrometeor types interact with each other in different ways. For the ice phase this can get significantly more complicated, since a wide range of solid ice particles exists that influence the growth and decay of each other. Figure 4.6 gives an impression of the variety of particles and processes that can be parametrized by a microphysics scheme. The addition of ice phase particles and how the partitioning of ice into different ice species (cloud ice, snow, graupel, hail) is implemented can significantly impact a forecast, because their key properties (fall speed, density, size distribution) are different.

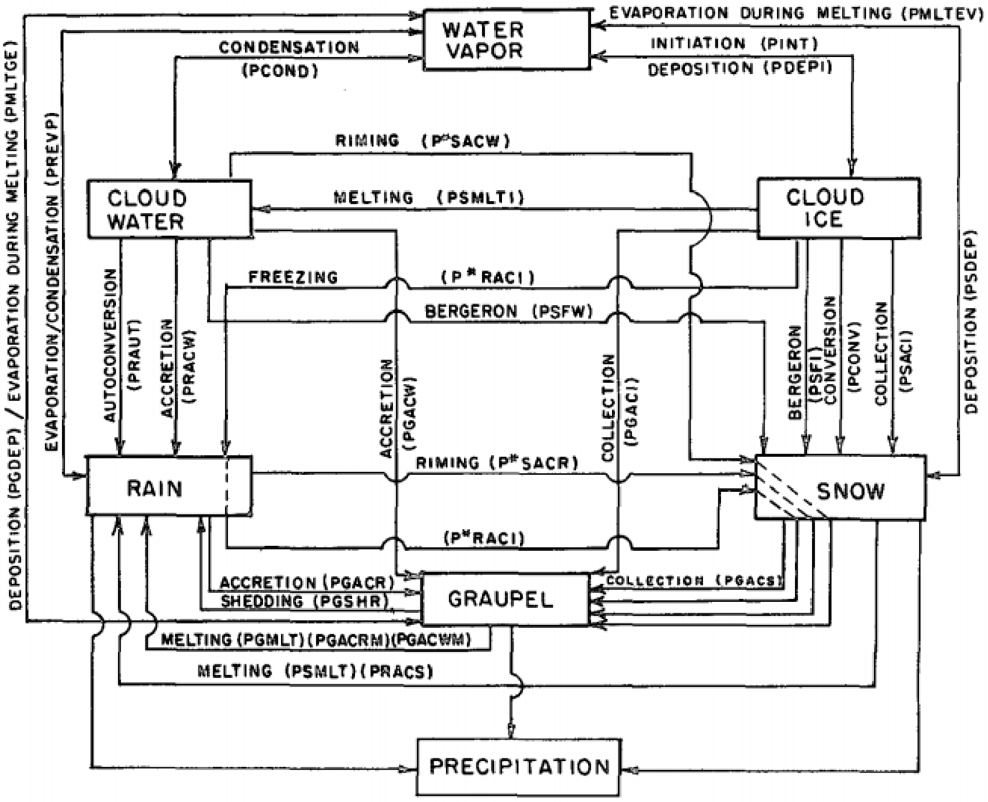


Figure 4.6: Flowchart illustrating the wide range of microphysical processes that may be parametrized. Original figure from Rutledge and Hobbs (1984).

   Microphysics schemes can be roughly classified into two categories: **spectral/bin** and **bulk** schemes. The first approach discretizes the size distribution into a fixed number of bins and is generally computationally demanding, while bulk parametrization schemes assume that the size distribution follows a certain function described by several parameters, for example the gamma distribution:

$$N(D) = N_0 D^{\mu} e^{-\lambda D} \tag{4.2}$$

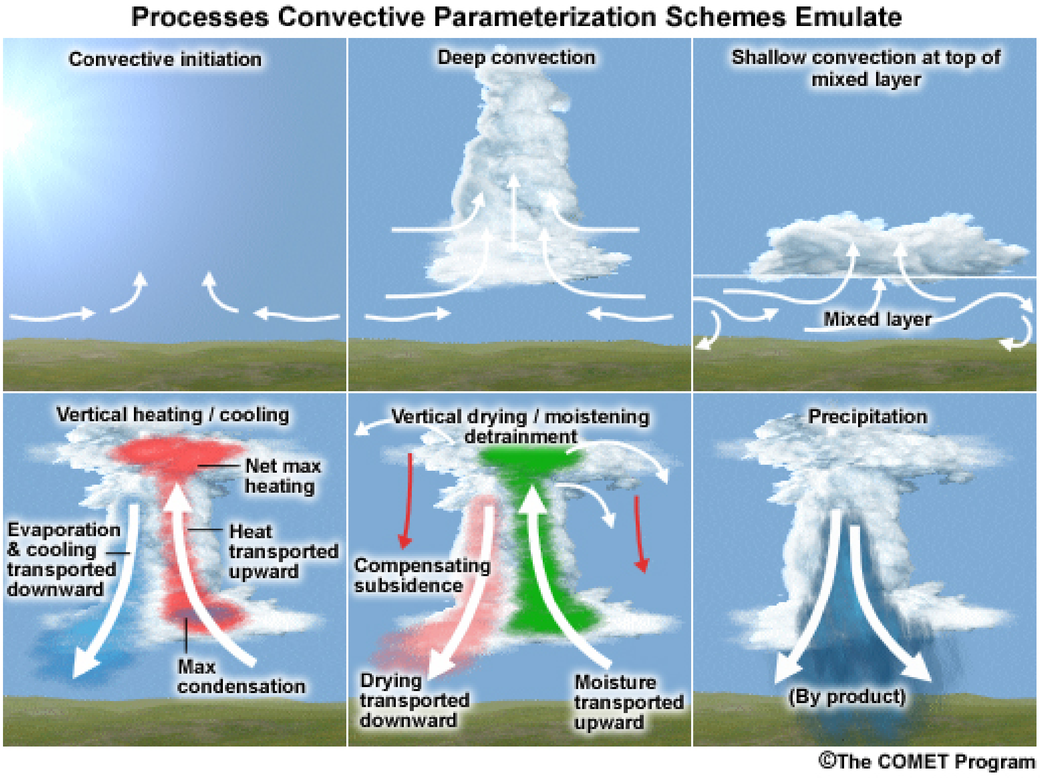
With $N(D)$ the number of droplets with diameter D, N_0 the intercept parameter and constants μ and λ that control the shape of the distribution. Note that for $\mu = 0$, the DSD in 4.2 is equal to the Marshall and Palmer (1948) DSD defined in 2.58. Basic bulk microphysical schemes assume a fixed N_0 and μ, so that a unique value for λ exists that can be estimated using the DSD in 4.2 and the DSD-LWC relation of 2.66 to find a DSD corresponding to an estimated mixing ratio. This implies a certain mixing ratio corresponds to a fixed DSD. The microphysical parametrization schemes that use this assumption are called *single moment* schemes, because they only estimate a mixing ratio. More elaborate schemes also include an estimation of N_0 per species (not necessarily for every single specie), so that the size distribution of a hydrometeor can differ for a single mixing ratio. These are referred to as *double moment* schemes and allow for greater flexibility in modelling the DSD which in turn influences the rate at which microphysical processes (e.g. accretion of cloud water by existing rain, riming of supercooled water onto ice particles or aggregation of ice crystals to produce snow flakes) occur. However, even when a double moment scheme is used to better represent DSD, uncertainty in the parametrization remains. The number of species, assumptions about hydrometeor properties (like density, fall speed etc.) and conversion rates cause uncertainty in the hydrometeor concentration estimates. Moreover, defining an optimal set of parameters is difficult since their performance is case dependent.

In this study WRF simulations are carried out using the "aerosol aware" microphysics scheme developed by Thompson and Eidhammer (2014). This scheme includes the effect of aerosols in the formation and growth of clouds and precipitation, with a focus on how aerosols affect mixed phase clouds. The aerosol aware version is an extension of the bulk microphysical scheme by Thompson et al. (2008). The latter is a scheme that predicts mixing ratios for five liquid and ice species: cloud water, rain, cloud ice, snow and graupel. It has been widely used in quantitative precipitation forecasting (QPF) studies (e.g. Rasmussen et al. (2011)) as it compares relatively well to precipitation observations. For example Liu et al. (2011) showed improved skill over other microphysics schemes in forecasting winter time precipitation in Colorado. Rajeevan et al. (2010) observed decent simulation of up- and downdrafts and better agreement of surface rainfall distribution with observations compared to three other microphysics schemes, although only a single case study of a thunderstorm in South-East India was examined.

Where the scheme by Thompson et al. (2008) predicted number concentrations (thus being double moment) for rain and cloud ice, the updated scheme by Thompson and Eidhammer (2014) includes estimation of number concentration for cloud water and additionally the number of cloud condensation nuclei (CCN) and ice nuclei (IN). The aim of including CCN and IN is to better represent the actual drop size distributions, as aerosols in general for the same mixing ratio cause a bigger number of droplets, however smaller in size. This will increase the albedo of a cloud, thus affecting the radiative transfer. Secondly, because the average drop size is reduced, the formation of larger particles that precipitate is delayed. These effects are referred to as the indirect effects of aerosols. The amount of nucleating aerosols varies in time and space and are derived from a 7 year global climatological simulation (Colarco et al. (2010)) of aerosols using the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model (Ginoux et al. (2001)). From these simulations mass mixing ratios of sulfates, sea salts, organic carbon, dust and black carbon were extracted and converted to number concentrations by assuming a lognormal CCN distribution. The number of CCN and IN can then be employed in the relations with other parameters determining the DSD. The aerosol aware scheme also models interaction of ice, snow, and cloud water particles with radiation, recently implemented by Thompson et al. (2016). This way, an attempt is made to also account for the first indirect effect of aerosols. Thompson and Eidhammer (2014) found that using the aerosol aware scheme the DSD indeed featured more droplets with smaller size on average, however the impact on longwave radiation at the surface and precipitation was rather small and additionally, only observing a reduction in rainfall of several percent. Khain et al. (2016) investigated the sensitivity of WRF hurricane simulations to microphysics schemes with aerosol interaction and also concluded that while the aerosol aware microphysics scheme produced results closer to observations than other tested (some also aerosol aware) bulk microphysical schemes, the sensitivity to aerosol concentration is limited.

Besides the aim of more realistic modelling of microphysical processes by including aerosol effects, the inclusion of the various ice phases through their mixing ratios is required if hydrometeor concentrations retrieved from radar reflectivity are to be assimilated.

2. **Cumulus/Convection** Another process that is often parameterized in a NWP model is subgrid scale convection. Convection refers to narrow, saturated updrafts and nearly saturated downdrafts where precipitation occurs. These areas of up- and downdrafts can be recognised as cumulus or cumulonimbus (precipitating) clouds. The more wide spread areas surrounding the cumulus clouds are subsiding to compensate the updrafts in the clouds. Figure 4.7 provides a schematic illustration of these processes. Moisture, heat and momentum are transported upwards by the rising air parcels. While the air rises, it expands and cools down. Cloud droplets and precipitation form once saturation vapour pressure is reached. At the same time compensating flows from higher altitude transport dry air back to the surface. The dry air accommodates evaporation, thereby introducing a cooling effect. Other processes affecting formation and life cycle of clouds are entrainment and detrainment, which describe the interaction of turbulent air in a cloud with the non turbulent, surrounding air.



©The COMET Program

Figure 4.7: Illustration of convective processes in a cumulus cloud.

Updrafts can have dimensions in the order of hundreds of meters up to kilometers for deep convection. This is smaller than common resolutions of regional weather models that range from several kilometers to tenths of kilometers, thus they are unable to resolve important processes of heat and moisture transportation explicitly and require a parametrization scheme to account for their effects. These CP schemes redistribute temperature and moisture within a column, reducing atmospheric instability that would otherwise produce unrealistic precipitation in a later stage when convection is finally triggered.

The 4 km resolution that is employed in the simulations here however lies in what is referred to as the 'grey zone'; the range of model resolutions ($1-10$ km) where the model resolution Δ is close to the length scale l of the turbulent convective clouds. Models at these resolution are able to resolve large scale convective structures like MCSs. There is no consensus in the NWP community whether or not to use convective parametrization at these resolutions. However, most convective parametrization schemes were designed with the assumption that all convection is subgrid scale, i.e. $l \ll \Delta$. Moreover, 'convection allowing' studies (e.g. Weisman et al. (1997), Speer and Leslie (2002)) showed that models with high resolutions of $3-4$ km were able to reproduce large scale features of convective systems without convective parametrization. Weisman et al. (2008) recognised an improved ability to describe structure and evolution of large scale convection, although more advanced characteristics like location, timing and intensity of the convective system were not as accurate. Similarly, Done et al. (2004) and **?** found that forecasts with 4 km resolution were more skilful than coarser > 10 km resolution simulations with CP. Schwartz et al. (2009) also found that a 4 km resolution provides a more skilful forecast over coarser CP employing models, although noting that further decreasing the resolution to 2 km apparently provided more detail but did not result in an improvement of quantified forecast skill. Contrary, studies carried out by Deng and Stauffer (2006) and Lean et al. (2008) do show cases where CP improves precipitation forecasting skill. The latter and also Weisman et al. (1997) additionally suggest leaving out a CP possibly delays convection because of the lack of vertical transportation of moisture and heat, leading to unrealistically large mass fluxes and intense precipitation once the convection is actually triggered.

In this study is chosen not to use a CP for the ability of 4 km to resolve large scale convection and because the CP schemes are designed for grid resolutions that fully cover deep convective cells. For that reason, the WRF model technical notes Skamarock et al. (2008) also do not recommend CP schemes $< 5\ km$.

3. **Planetary boundary layer**

The planetary boundary layer (PBL) is the layer just above the Earth's surface that is influenced directly by the surface. The height of the PBL varies from several tenths of meters at night up to several kilometers on a hot day. The air above the PBL that is not affected anymore by the surface is referred to as the free troposphere. These layers are separated by an what is called a *capping inversion*. The inversion is characterised by a sudden increase of potential temperature with height, blocking any buoyant, rising air parcels from below.

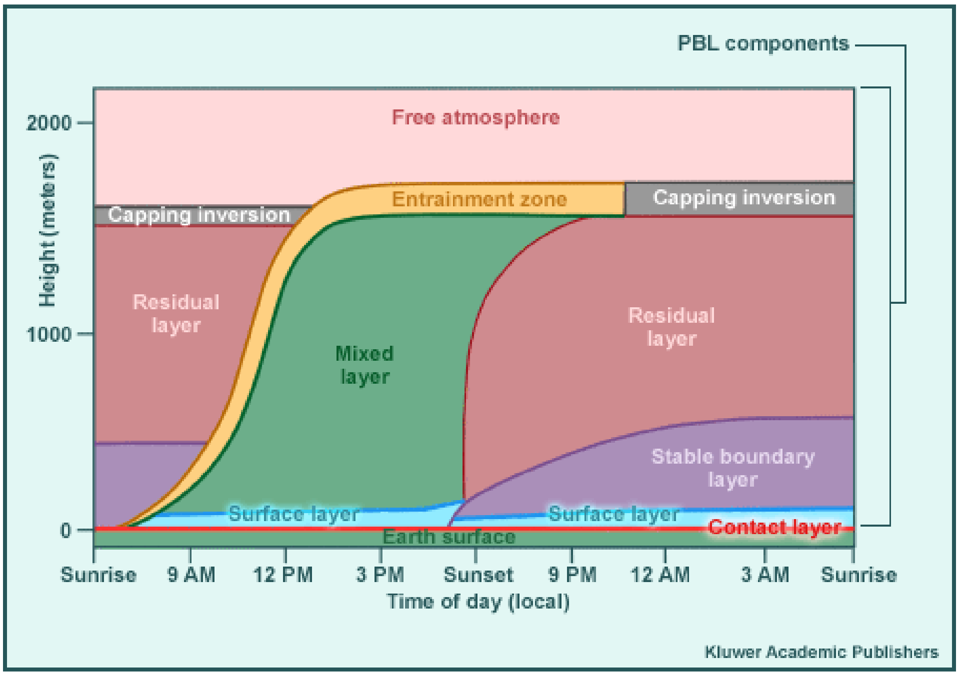


Figure 4.8: Schematic illustration of the evolution of the PBL over the course of a day. Original figure from Meted (2009).

The PBL grows during daytime when incoming radiation from the Sun heats the Earth's surface. The radiation causes sensible (temperature) and latent heat (moisture) fluxes from the surface to the air above. Local instability causes the formation of buoyant air parcels or *thermals*. These rising thermals are responsible for vertical transportation of heat, moisture and momentum. Because of their momentum they can overshoot the capping inversion, thereby entraining some air from the free troposphere when it sinks down again. Figure 4.8 illustrates an idealized evolution of the PBL during a day. In the morning the surface that is warmed by incoming radiation produces thermals in the air just above. The thermals and associated turbulence mix the air, creating the *mixed* (or *convective* boundary) layer as depicted in figure 4.8. The PBL grows in height by entrainment of upper air, as long as the Sun is heating the surface and thus maintaining the forcing from the surface. Once the Sun sets, the forcing disappears and longwave radiation that is still present starts cooling the surface, making the temperature lower than the overlying air. The gradually deepening layer of cool air is called the *stable, nocturnal* boundary layer. In the nocturnal boundary layer, most of the turbulence is *mechanical*, induced by vertical wind shear. That is contrary to daytime when turbulence is dominantly *thermal*. What remains of the mixed layer is called the *residual* layer, containing the residual heat, moisture and pollutants. The *surface layer*, usually several tenths of meters thick, constitutes the lowest layer of the PBL that directly interacts with the surface and where turbulent transportation varies little with height compared to the variability in the mixed layer above.

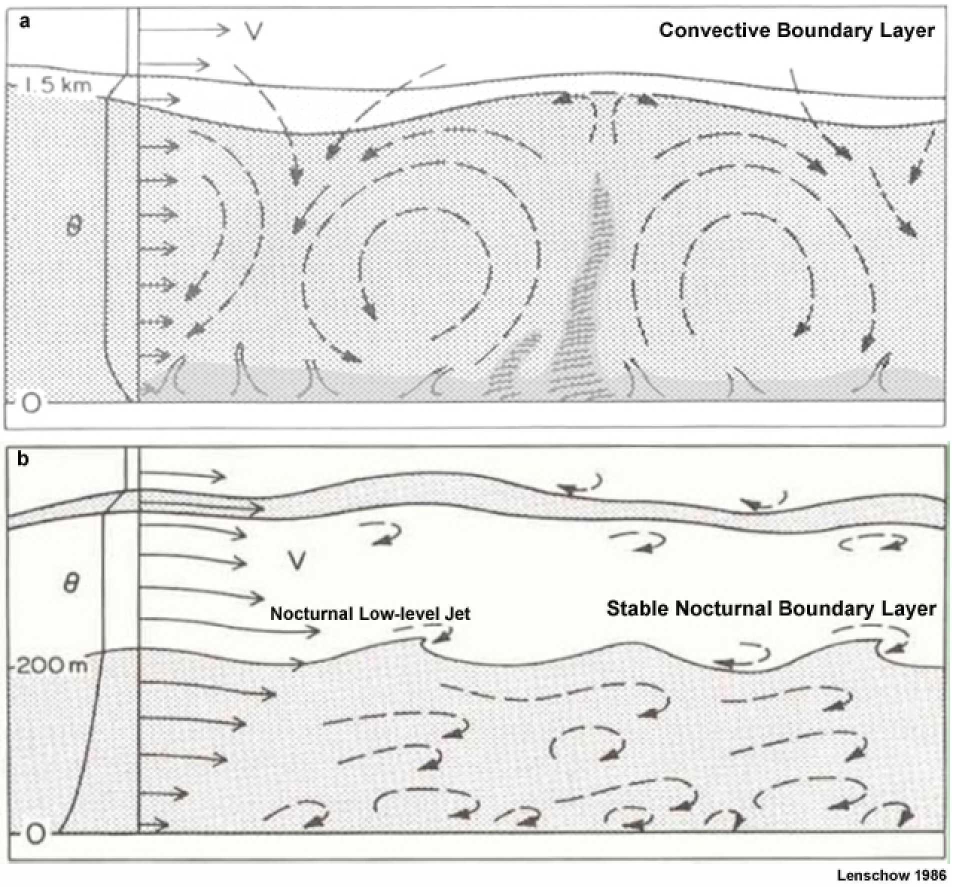


Figure 4.9: Impression of turbulent motions (a) in the convective boundary layer during daytime that are dominantly thermally driven and (b) in the shallower, nocturnal boundary layer that are mechanically driven by wind shear. Illustration from Lenschow (1986).

The PBL is characterised by turbulent eddies that redistribute heat, moisture and momentum, creating a vertical, well mixed layer. Figure 4.10 shows how the turbulence in a typical daytime PBL results in a layer of well mixed layer with constant potential temperature, specific humidity and wind speed. The turbulent motions range from kilometer length scales down to micro scales. These processes are subgrid scale for NWP model, so they have to be accounted for using parametrization. Since turbulent flow is a highly non linear and small scale phenomenon, it is not possible to explicitly predict turbulence of single eddies in the model equations. Rather, for example a statistical approach is used to describe their collective effect. Model variables like potential temperature, water vapour mixing ratio and vertical velocity are decomposed into a mean and perturbation component, e.g. $\overline{\theta}$ and θ' for potential temperature. The mean component describes the background state of the atmosphere, while the perturbed part represents the contribution of turbulence. This decomposition is applied to the time tendency (e.g. $\frac{\partial\theta}{\partial t}$, $\frac{\partial q}{\partial t}$, $\frac{\partial u}{\partial t}$, $\frac{\partial v}{\partial t}$, $\frac{\partial w}{\partial t}$) equations for model variables. For a complete introduction on this approach to describe turbulence, the reader is referred to e.g. Stensrud (2009). When the decomposed tendencies are averaged, formally known as *Reynolds averaging*, the mean of the fluctuating terms (e.g.

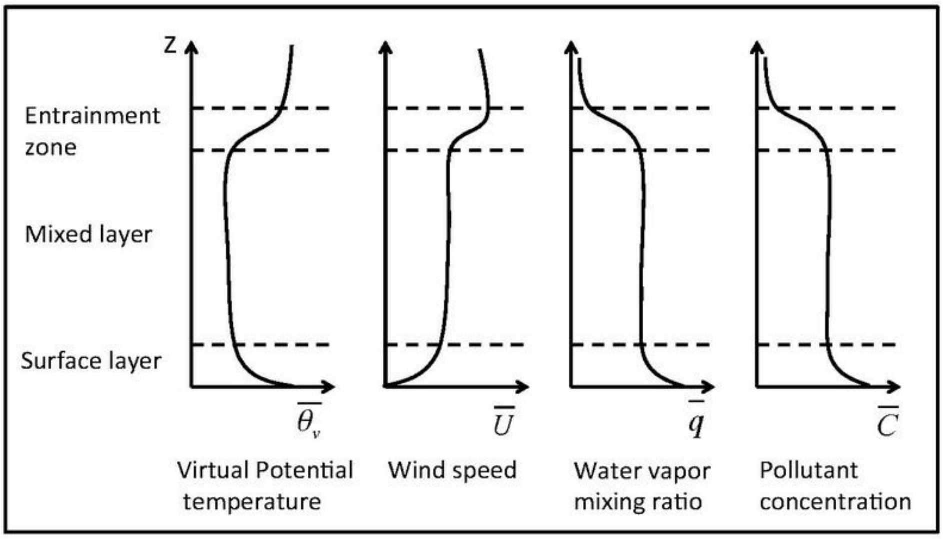


Figure 4.10: Idealized profiles of virtual potential temperature, wind speed, water vapour mixing ratio and a pollutant in a typical daytime PBL.

$\overline{\theta'}$) is assumed zero but the mean of the product of fluctuating terms (e.g. potential temperature and vertical velocity $\overline{\theta' w'}$) is not and represents the temperature flux as a result of turbulence. These product terms are the only ones in the averaged tendency equations that are not expressed in terms of the model grid scale variables and thus have to be parameterized. How they are nevertheless being tried to estimate using model variables is referred to as **closure**. Generally, two approaches exist to estimate these products:

(a) Direct parametrization using empirical relationships between the product terms and grid scale variables. A common method is known as the *eddy diffusivity theory*, *K-theory* or *gradient transfer theory*. This method assumes that the averaged product term or flux is proportional to the vertical gradient, e.g. for potential temperature:

$$F = \overline{\theta' w'} = -K \frac{\partial \overline{\theta}}{\partial z} \tag{4.3}$$

Where F represents the heat flux and K the eddy diffusivity. K is said to increase with turbulence intensity and also is a function of height, wind shear and surface heating, representing the main ways that turbulence is generated: inertially, mechanically and thermally.

Because $\overline{\theta' w'}$ is parameterized directly, these methods are said to have a *first order* closure, because that is the lowest order of covariance that is parameterized. Zeroth order closure would imply that for example $\overline{\theta}$ is parameterized directly, without decomposition into mean and fluctuating part. Sometimes also non integer values are used to characterize the closure order. For example a 1.5th order closure implies that only for a selection of variables (commonly moisture and temperature) second order covariances are estimated.

(b) Use predictive equations for the product terms. These predictive equations in turn however are a function of triple products of perturbations, like $\overline{\theta'\theta' w'}$ for example. Parameterizing these terms results in a second order closure, as the second order of covariance is parameterized. This principle can be extended to obtain higher order closures. However, predicting an n^{th} order product always includes an unknown product of order $n+1$ that has to be parameterized in terms of lower order terms. In general, a higher order closure is expected to provide better representation of the boundary layer since only higher order terms are parameterized.

Besides the order of closure PBL schemes are also characterised by how far they distribute heat, moisture and momentum vertically. Generally, two classes exist: *local* and *non local* schemes. In local

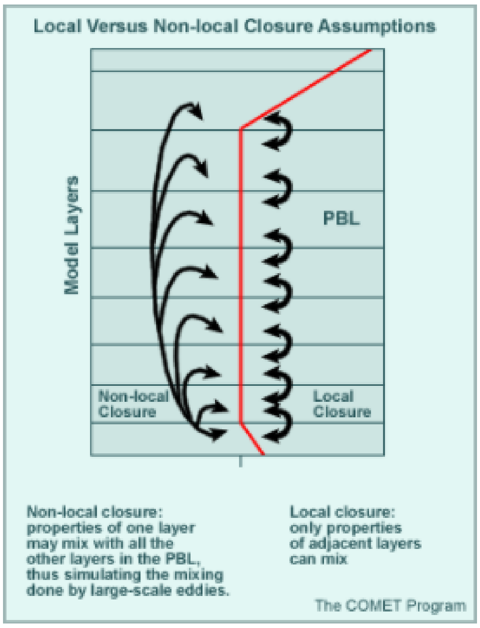


Figure 4.11: Graphic illustration of a local (right hand side) and non local closure (left). In a local closure only adjacent vertical levels are taken into account for transportation, while in non local schemes transportation beween non ajdacent vertical levels might occur. Illustration from Meted (2009).

schemes only quantities in adjacent vertical grid points are used to estimate unknown quantities, see figure 4.11. This approach implicitly assumes vertical transportation is carried out by small turbulent eddies. A known problem of local schemes is the reduced growth of the PBL when local stability (i.e. a small inversion) exists. The local gradient is negative there, so that down-gradient flow as defined by 4.3 is now from high to low altitude. See figure 4.12 for an illustration how a local scheme determines the flux direction. In reality, large eddies transport heat upwards through these locally stable parts of the boundary layer, so that the fluxes are actually locally counter-gradient because the gradient is negative there. Higher order local closure parametrizations can allow counter-gradient transport but does not model large eddies since it is still local. This does require more computational power since more, higher order covariance terms have to be estimated. On the contrary, in non local schemes, an attempt is made to include the effect of transportation by large eddies by using multiple vertical grid points to estimate fluxes. Generally, this is done by including a correction term to 4.3, e.g. for potential temperature again:

$$\overline{\theta' w'} = -K\left(\frac{\partial\overline{\theta}}{\partial z} - \gamma_\theta\right) \tag{4.4}$$

Various implementations of the correction term γ_θ exist in non local schemes but in general it is a function of the values of model variables at other grid boxes, most often the surface, so that a non local scheme as illustrated in figure 4.11c is employed. In general evaluations of PBL schemes have shown that when compared to observations vertical mixing tends to be overestimated by non local closure schemes, while contrastingly local schemes are likely to underestimate vertical mixing. Consequently, PBL height estimated by non local closure schemes is usually larger than local closure schemes. Examples of studies illustrating these findings are from Bright and Mullen (2002) investigating daytime summer PBLs and Stensrud and Weiss (2002), who validated daytime PBL in a case of the Oklahoma tornado outbreak in May 1999. Also Xie et al. (2012) found that when comparing local and non local schemes, non local schemes could more accurately represent the depth of the PBL. Additionally, Hu et al. (2010) and Gibbs et al. (2011) observed that simulations with local schemes yield relatively cool and moist conditions, while non local schemes produce a more warm and dry PBL. The differences in the predicted characteristics of the PBL between local and non local schemes are attributed to their (in)ability to represent transport by large eddies.

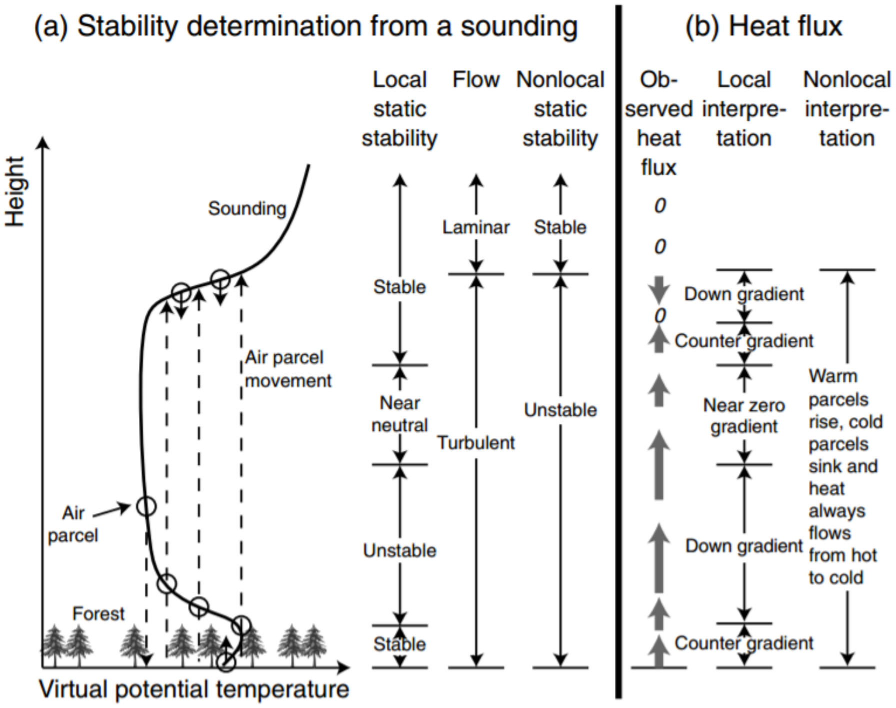


Figure 4.12: Schematic illustration of (a) an idealized sounding of the PBL with a mixed, turbulent layer. Stability is a function of the sign of the potential temperature gradient. Figure (b) shows the associated fluxes determined by local and non local schemes, respectively. Illustration from Warner (2011).

In this study the PBL parametrization by Mellor–Yamada–Nakanishi–Niino (MYNN) is employed, described in Nakanishi and Niino (2004) and Nakanishi and Niino (2006). This is a local scheme with a 1.5th order of closure predicting subgrid turbulent kinetic energy (TKE) terms, with mixing length scales (affecting eddy diffusivity K in 4.3) based on the results of LES models instead of observations to tackle common errors in earlier Mellor-Yamada schemes, like the inadequate development of the PBL and underestimation of TKE. To further compensate the deficiencies associated with a local scheme, a mass flux component was added to this scheme by Angevine et al. (2016) that is scale aware; the mass flux is reduced when grid size approaches LES scale. The concept of a mass flux approach is to separate turbulent fluxes into a component of non local updrafts and the surrounding turbulent field of local turbulence. A fraction of the total area a_u (assumed to be a couple percent) is said to be responsible for the strong, non local vertical transport by thermals. Combining with the eddy diffusivity approach leads to what is called the eddy diffusivity mass flux (EDMF) approach, for potential temperature again:

$$\overline{\theta' w'} = -K\frac{\partial\overline{\theta}}{\partial z} + M\left(\theta_u - \overline{\theta}\right) \tag{4.5}$$

Where $M = a_u\left(w_u - \overline{w}\right)$ represents the so called mass flux, the product of the fractional updraft area a_u and the difference in vertical velocity within the updraft compared to the average of the whole area. Similarly, $\theta_u - \overline{\theta}$ is the difference of the potential temperature within the fractional area containing the updrafts with respect to the average. The first term on the right hand side in 4.5 represents the local turbulent transport, while the second term aims to model the non local transport due to strong thermals. This way, the important non local, vertical transportation by large eddies and the earlier discussed countergradient fluxes that are not properly described by local schemes can be accounted for, enabling more realistic growth of the PBL height. For a more elaborate description of the EDMF approach and validation with LES experiments, the reader is referred to Siebesma et al. (2007) who pioneered this approach to describe turbulence in the PBL.

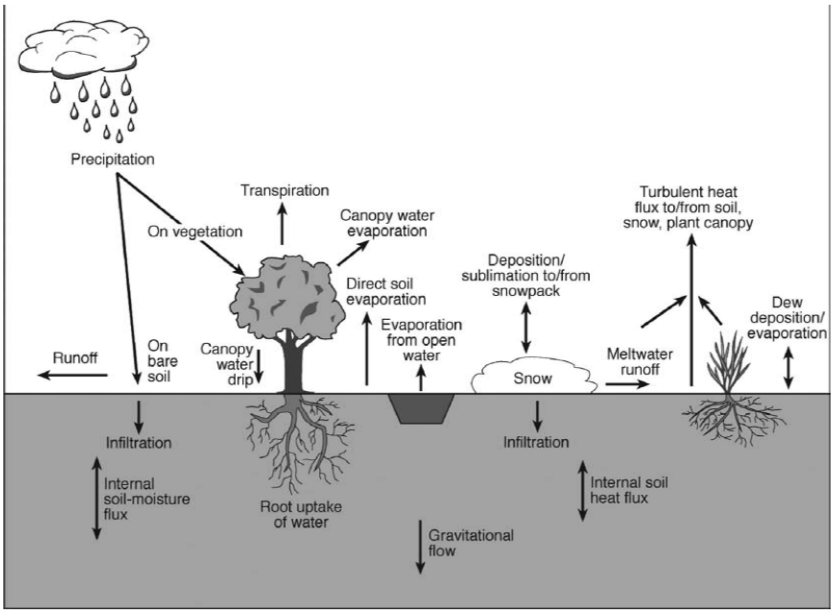


Figure 4.13: Illustration of the physical processes at the surface that a LSM can account for. Illustration by Warner (2011).

4. **Land-surface model**

   Inherently coupled to the PBL parametrization is the LSM, as it provides the surfaces fluxes of momentum, sensible heat and latent heat. Recall that the lowest layer of the PBL that is directly affected by the surface is called the *surface layer*, see figure 4.8. The surface fluxes are important because they control how much heat and moisture goes in and out of the atmosphere. The moisture flux supplies the fuel for any clouds and precipitation that might form, while the heat from the surface flux generates thermals that can carry the (moist) air upwards. Figure 4.13 gives an overview of the physical processes that a LSM accounts for. Water from trees, crops and other types of vegetation evaporates into the atmosphere. Also, water evaporates directly from the soil or from open water bodes like lakes. Water returns to the surface in the form of precipitation. Besides evaporating, the water can also infiltrate the soil and possibly be taken up by roots of vegetation, retaining the moisture in the soil for a longer time or run off to larger water bodies. Usually a LSM divides soil in several layers in between which moisture and heat can flow. Similarly to the atmosphere, each layer has its own temperature, moisture content etcetera. Depending on the soil type (e.g. sand, clay, rock), layer properties like porosity, thermal and hydraulic conductivity are defined that affect the transportation of moisture and energy between the soil layers. At the surface a land use category is defined for every grid cell. Each land use category has its characteristic properties like roughness, albedo, thermal capacity and conductivity, vegetation fraction, emissivity etc.

   The zonal τ_x and meridional τ_y momentum fluxes, sensible heat flux H and latent heat (moisture) flux E are defined as follows:

   $$|\tau| = -\rho \frac{|\mathbf{v}(z)| - |\mathbf{v_s}|}{r_m} \tag{4.6}$$

   $$H = -\rho c_p \frac{\theta(z) - \theta_s}{r_h} \tag{4.7}$$

   $$E = -\rho \frac{q(z) - q_s}{r_w} \tag{4.8}$$

   Where u and v are the familiar zonal and meridional wind velocities, ρ the atmospheric density, c_p the specific heat at constant pressure, θ the potential temperature and q the specific humidity. The subscript s denotes the surface value, as compared to the value of a quantity at height z. For momentum, heat and moisture, the *aerodynamic resistance* r is a measure of the frictional drag of the surface. The above fluxes are a result of applying the Monin and Obukhov (1954) similarity theory for the surface layer. This theory dictates that the scaled, dimensionless mean horizontal wind speed, mean potential temperature and specific humidity vertical gradients are described by a unique similarity function. That is to say, for example all vertical wind profiles for any speed or roughness, when scaled with the corresponding friction velocity u_*, are in fact described by the same curve. The similarity functions ($\varphi_m(\zeta)$, $\varphi_h(\zeta)$, $\varphi_m(\zeta)$) for momentum, heat and moisture are defined as follows:

$$\frac{k(z-d)}{u_*}\frac{\partial |v|}{\partial z} = \varphi_m(\zeta) \tag{4.9}$$

$$\frac{k(z-d)}{\theta_*}\frac{\partial \theta}{\partial z} = \varphi_h(\zeta) \tag{4.10}$$

$$\frac{k(z-d)}{q_*}\frac{\partial q}{\partial z} = \varphi_w(\zeta) \tag{4.11}$$

The dimensionless constant $k \approx 0.40$ is the von Kármán constant. z is the height within the surface layer, d a certain displacement height (defining the surface height where the wind speed approaches zero), $|v|$ the horizontal wind magnitude. The constants u_*, θ_* and q_* represent characteristic scales for wind speed, potential temperature and specific humidity, used for scaling the gradients:

$$u_*^2 = \sqrt{\left(\overline{u'w'}\right)^2 + \left(\overline{v'w'}\right)^2} = \frac{|\tau|}{\rho} \tag{4.12}$$

$$\theta_* u_* = -\overline{\theta' w'} = -\frac{H}{\rho c_p} \tag{4.13}$$

$$q_* u_* = -\overline{q' w'} = -\frac{E}{\rho} \tag{4.14}$$

The overbar terms denote the kinematic fluxes or covariance terms discussed in the previous section on the parametrization of the PBL. $|\tau|$ is the absolute value of the horizontal friction stress, hence u_* is often called the friction velocity. Recall H and E are the heat and moisture surface fluxes. Note that the heat and moisture fluxes are often called **sensible** and **latent** heat fluxes in meteorology, referring to the measurable heat flux and the indirect transportation of energy by evaporation.

$\zeta = \frac{z-d}{L}$ in the similarity functions 4.9 to 4.11 represents a parameter relating the height above the surface to the Obukhov length scale L:

$$L = -\frac{u_*^3 T_v}{kg\overline{\theta_v{}' w'}} \tag{4.15}$$

With T_v the virtual temperature, g the gravitational acceleration and $\overline{\theta_v{}' w'}$ the turbulent kinematic temperature flux. The absolute value of the Obukhov length scale can be interpreted as the height below which mechanically generated turbulence dominates over buoyant/thermally generated turbulence and usually ranges from meters to several tens of meters. The Obukhov length is infinitely large for neutral conditions, while positive for stable and negative for unstable conditions. ζ can be regarded as the dimensionless Obukhov height and acts as an indicator for stability: $\zeta < 0$ implies instability while $\zeta > 0$ implies stable conditions. For a neutrally stable atmosphere (i.e. no potential temperature gradient) $\zeta = 0$ and $\varphi_m = \varphi_h = \varphi_w = 1$. It can be shown (e.g. Monin and Obukhov (1954), Oleson et al. (2010)) by integration that wind velocity, potential temperature and humidity follow an exponential profile in the surface layer for this case. E.g. the vertical profile of wind is given by integration of 4.9, and assuming the surface velocity $V_0 = 0$:

$$V = \frac{u_*}{k}\ln\left(\frac{z}{z_0}\right) \tag{4.16}$$

The *aerodynamic roughness length* z_0 is defined as the height above the surface at which the wind speed is zero, when the air is thermally neutrally stable (i.e. no buoyant forces). Because $\zeta = 0$ and $\varphi_m = 1$, there is no stability correction and the vertical profile of horizontal wind speed follows a logarithmic profile, as plot in figure 4.14. For stable and unstable empirical corrections are applied so that the profile slightly deviates from the logarithmic profile of a neutrally stable surface layer. In general the correction takes the following form, e.g. for wind speed:

$$V = \frac{u_*}{k}\left[\ln\left(\frac{z}{z_0}\right) - \psi_m(\zeta)\right] \tag{4.17}$$

With $\psi_m(\zeta)$ a correction term resulting from integration of integration of similarity function φ_m when $\varphi_m \neq 1$. $\psi_m(\zeta)$ and similarly for ψ_h temperature and ψ_w for moisture are empirical functions based on fit to observations, see e.g. Zeng et al. (1998).

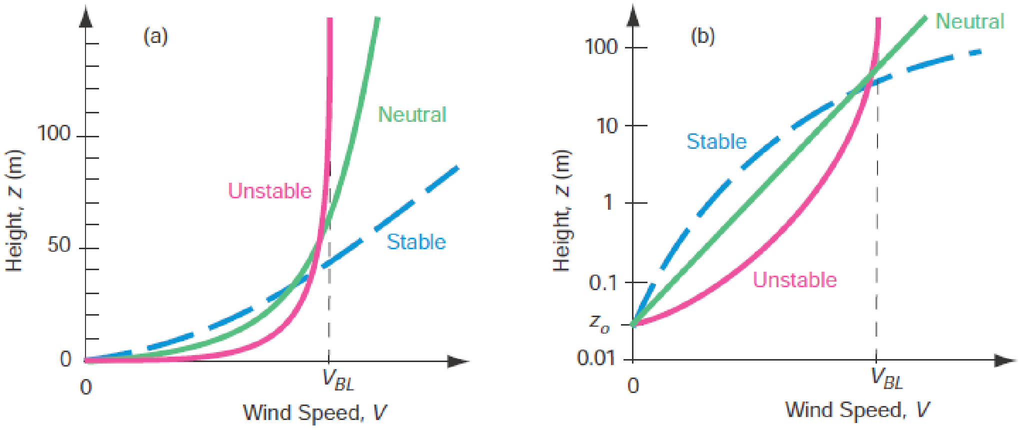


Figure 4.14: Illustration of the vertical wind profile (a) on a linear scale and (b) on a semilog scale. Stability affects the vertical gradient of the wind speed: instability causes larger gradient while stable air causes a less steep wind profile. Figure from Wallace and Hobbs (2006).

Figure 4.14 shows how stability affects the vertical profile of the horizontal wind speed. In an unstable surface layer ($L < 0$, $\zeta < 0$, $\phi_m < 1$, $\psi_m > 0$) characterised by thermals, the wind speed increases faster with height as the stronger turbulence better mixes momentum downward. The friction stress and thus friction velocity are higher in this case than for neutral stability, leading to a steeper initial slope near the surface. At height $z > 50$ the $-\psi_m$ term dominates the wind speed contribution over the logarithmic term, flattening the profile. On the other hand, in a stable layer ($L > 0$, $\zeta > 0$, $\phi > 1$, $\psi_m < 0$), the turbulence is not strong enough to mix the surface layer well. The friction velocity is smaller with respect to the neutral conditions, as indicated in figure 4.14 by the smaller gradient for stable conditions. At an altitude above $35m$, the positive ψ_m correction term ensures the vertical gradient of the wind speed does not decay as much for the neutral condition, leading to higher wind speeds at these altitudes. Besides wind speed, profiles for potential temperature and water vapour mixing ratio similar to 4.17 are also logarithmic, or close to logarithmic when the surface layer is not neutrally stable. However, typically during daytime, contrary to the wind speed, the surface boundary condition is not zero but actually higher than the boundary layer, i.e. $\overline{\theta_{sfc}} > \overline{\theta_{BL}}$ and $\overline{q_{sfc}} > \overline{q_{BL}}$. The vertical profile of (virtual) potential temperature and water vapour mixing ratio is still logarithmic, but is therefore decreasing with height, as shown in figure 4.10.

Coming back to the surface fluxes defined in 4.6, 4.7 and 4.8, the aerodynamic resistance r is one of the factors affecting the flux magnitude. Combining 4.6, 4.12 and 4.16, the aerodynamic resistance is defined as:

$$r_m = \frac{V_{BL}}{u_*^2} = \frac{1}{k^2 V_{BL}}\left[\ln\left(\frac{z_{BL,m} - d}{z_{0,m}}\right) - \psi_m(\zeta)\right]^2 \tag{4.18}$$

This is a measure of the resistance for momentum between a height $z_{BL,m}$ in the boundary layer and the surface at height $z_{0,m} + d$. Similarly, aerodynamic roughness for heat and humidity can be defined:

$$r_h = \frac{\theta_{BL} - \theta_{sfc}}{\theta_* u_*} = \frac{1}{k^2 V_{BL}} \left[\ln\left(\frac{z_{BL,m} - d}{z_{0,m}} \right) - \psi_m(\zeta) \right] \left[\ln\left(\frac{z_{BL,h} - d}{z_{0,h}} \right) - \psi_h(\zeta) \right] \tag{4.19}$$

$$r_w = \frac{q_{BL} - q_{sfc}}{q_* u_*} = \frac{1}{k^2 V_{BL}} \left[\ln\left(\frac{z_{BL,m} - d}{z_{0,m}} \right) - \psi_m(\zeta) \right] \left[\ln\left(\frac{z_{BL,w} - d}{z_{0,w}} \right) - \psi_w(\zeta) \right] \tag{4.20}$$

Evidently, aerodynamic roughness is affected by the aerodynamic roughness length z_0 of the surface. The order of magnitude is quite small, ranging from millimeters for smooth surfaces to decimeters for more rougher surfaces. The roughness lengths are a measure of the initial length scales of turbulent eddies originating at the surface and logically they depend on the type of land-use. For example, trees and shrubs in a forest are taller, resulting in a bigger roughness length and consequently larger eddies near the surface. For water the roughness length is a function of wind speed, as that controls the wave height. From the expressions 4.18, 4.19 and 4.20 it is obvious that a larger roughness length will lead to a smaller aerodynamic resistance and therefore a larger surface flux. Also, a larger difference e.g. in temperature $\theta_{BL} - \theta_{sfc}$ between the boundary layer and the surface logically causes a larger surface flux. The surface fluxes are important as they are the drivers of convection and possible precipitation. The moisture flux supplies the fuel for formation of clouds and precipitation, while the heat flux supplies the energy to generate buoyancy.

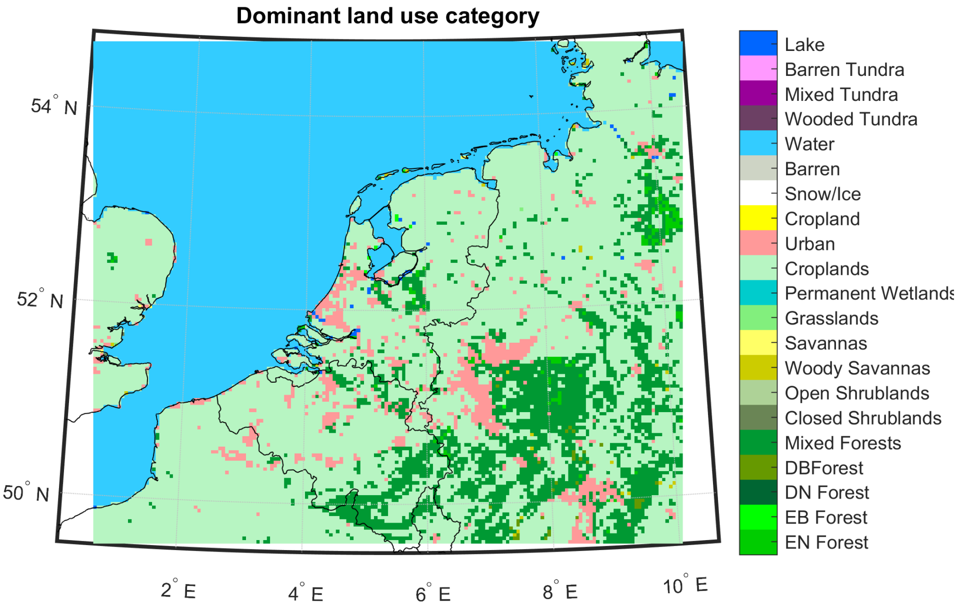


Figure 4.15: Land use in the WRF domain used in this study. The land use classification uses 21 categories. DB, DN, EB and EN denote deciduous broadleaf, deciduous needleleaf, evergreen broadleaf and evergreen needleleaf types of forest. The most abundant land use types in the domain are croplands, water and urban areas.

In this study the Community Land Model version 4 (CLM4) is used as land-surface model and is fully described by Oleson et al. (2010) and Lawrence et al. (2011). CLM4 has been developed at the NCAR with many collaborators from other institutions and is a scientific land surface process model under active development. It features sophisticated treatment of biogeophysics, hydrology, biogeochemistry, and dynamic vegetation. In the CLM4, the land surface in each model grid cell is divided into five primary sub-grid land cover types (glacier, lake, wetland, urban, and vegetated). The definition of sub-grid land cover types allows to account for land cover features smaller than the model resolution. For example, small lakes or ponds that cause significant evaporation but are not captured by the interpolated model land use grid can then be accounted for. The sub-grid vegetation consists of up to 15 plant functional types (PFT's) that differ in leaf optical properties concerning reflection, transmittance, absorption of radiation, parameters controlling the uptake of water by vegetation, aerodynamic parameters (e.g. resistance) and additionally some photosynthetic parameters. The PFT is determined from the input land use classification using a lookup table. The land use classification is based on MODIS satellite data and features 21 classes. Figure 4.15 shows the dominant land use category for the WRF domain interpolated to the 4 km grid interpolated from a global 30 arc second (1 km) land use classification map based on MODIS data. Finally, the CLM4 models the soil with 10 layers, can employ a 5 layer snow pack and accounts for vegetation canopy with a single vertical layer.

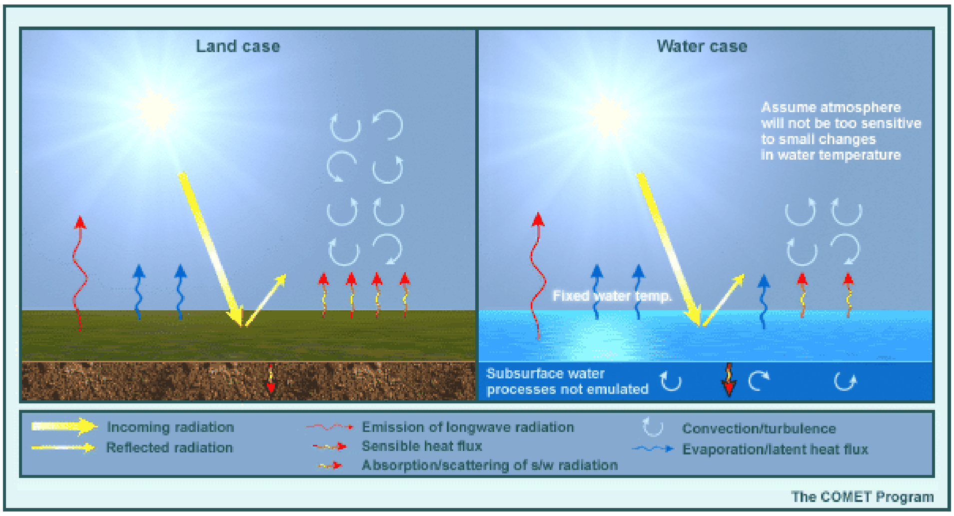


Figure 4.16: Schematic illustration of surface temperature (sensible heat) and moisture (latent heat) fluxes. The left panel shows the fluxes on a dry land surface, the right panel for a water surface. Illustration by Meted (2009).

Land cover type significantly affects the surface fluxes. For example, consider a dry, bare surface as in the left side of figure 4.16 with a limited amount of moisture that is heated by incident solar radiation during the day. Because the surface is much warmer than the air above, according to 4.7 a large sensible heat flux is present. The air above is heated, creating positive buoyancy and giving rise to convection and turbulence. On the other hand, on a water surface (see the right side of figure 4.16), more energy of the incoming radiation is used to evaporate water. Compared to a dry surface, there is a larger latent heat flux and a smaller sensible heat flux and therefore less turbulence and convection, resulting in a lower PBL height. Note that water surfaces are assumed to maintain a constant temperature throughout a short range forecast (i.e. no significant heating by radiation, because of the high heat capacity of water), thus acting as a heat sink. Comparing a dry land surface with vegetated surfaces, the plants on a vegetated surface contain moisture, resulting in a larger moisture and smaller sensible heat flux as compared to dry land. Similarly a sandy soil will also produce larger moisture fluxes than clay, because sand is more porous and therefore more moisture is readily available for evaporation. In general surfaces with relatively more moisture will cause less vertical instability in the air above because convection is not as strong as for a dry surface. However, if deep convection is triggered, the high moisture content at the surface can cause significant precipitation when the moist air rises and clouds form.

5. **Radiative transfer**

   Last but not least, the radiation balance plays an important role in the atmosphere. Incoming shortwave radiation ($\lambda < 4\mu m$) from the sun heats the surface and is the driver of atmospheric processes like the Hadley circulation, tropical cyclones, monsoons, trade winds but also on a smaller scale, convection and coastal circulation. The effects of radiative processes include temperature inversions above the surface because of longwave ($\lambda > 4\mu m$) emission by the surface, evaporation of water at the surface, damping of diurnal surface temperature variations by clouds and development of buoyant thermals giving rise to convective weather phenomena. To be able to describe these processes, a NWP needs to model the interaction of radiation with the various components of the atmosphere and the surface, see figure 4.17. The bulk of these interactions occur at molecular and micro scales and are complex functions of the wavelength spectrum, thus these processes cannot be described directly by the model equations of a NWP and need to be parameterized.

   It is essential for a good forecast to accurately represent the radiative flux at the Earth's surface, because differential heating at the surface causes most of the weather that we experience. As discussed in the

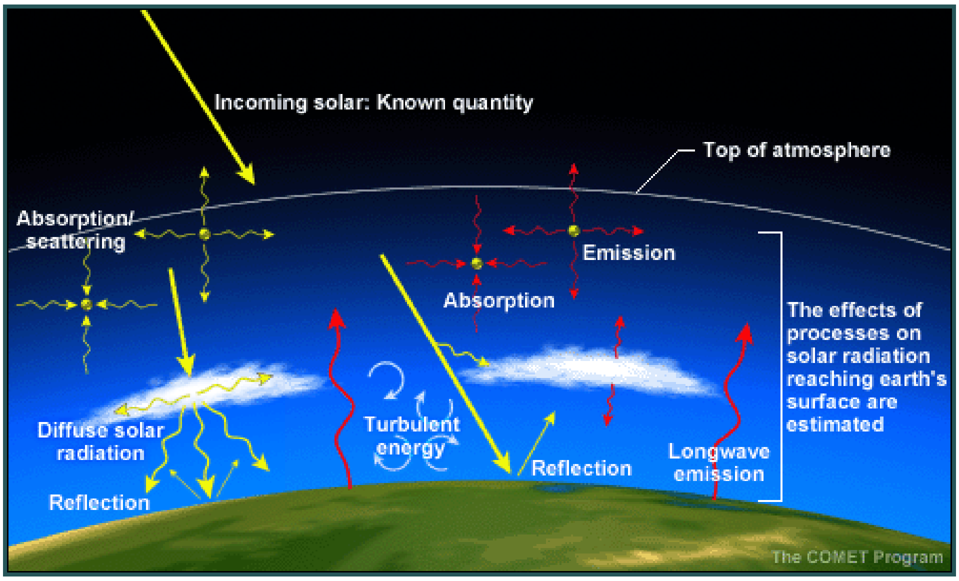


Figure 4.17: Schematic illustration of radiative processes in the atmosphere. Illustration by Meted (2009).

previous point, the land surface model calculates the amount of evaporation, as well as the amount of longwave (infrared) radiation emitted by the surface, which is a function of temperature according to the Stefan-Boltzmann law of black body radiation. Also, the albedo that determines the amount of incident shortwave radiation on the surface that is reflected back to space is dependent on the land cover type and is thus given by the land surface model. The PBL parametrization scheme is responsible for estimation of surface flux of sensible and latent heat. Latent heat is also released at higher altitudes by condensation, which is modelled by the microphysics parametrization scheme. The other radiative processes shown in figure 4.17 have to be accounted for by the radiative transfer model. Before radiation reaches the surface, it undergoes various interactions with the atmosphere. For example, solar radiation is absorbed by various gases and aerosols in the atmosphere. The wavelength band of the solar radiance spectrum that is absorbed is unique for every gas specie, as shown in figure 4.18. For example UV radiation is absorbed by ozone, while most absorption for $C0_2$ takes places in the longwave spectrum. The amount of absorption is proportional to the total amount of the gas that the solar beam encounters. These gases also scatter shortwave radiation into other directions, causing diffuse solar radiation. Longwave radiation is not scattered, but emitted back to the surface and to space. Another important process is reflection, in which clouds play a major role. Clouds reflect solar radiation back to space, as well as absorb and emit infrared radiation back to the surface. During the day less solar radiation can reach the surface so that there is less heating and thus lower temperatures, while at night the clouds trap infrared radiation in the atmosphere, so that there is less radiative cooling than compared to a clear sky. The modelling of clouds remains one of the biggest uncertainties in NWP models. Clouds bigger than model resolution can be explicitly described by the microphysics scheme, although the microphysics parametrization itself remains an approximation of the real physical processes. For example the assumed DSD for a certain mixing ratio might assume relatively bigger droplets than in reality, therefore underestimating the albedo of a cloud. Information on hydrometeor (cloud water, ice, snow, etcetera) concentration can then be passed on to the radiative transfer model that can subsequently estimate reflection, absorption, scattering and emission of radiation by clouds. However, clouds can also have scales significantly smaller than model grid resolution, so that while a model grid cell average RH might not indicate saturation, within a grid cell locally saturation is reached, subgrid clouds exist and thus affect the radiative transfer. Generally speaking, two popular methods exist to estimate subgrid cloud cover:

(a) **Probabilistic** The first approach assumes a certain PDF for the humidity in a grid cell. Then, the cloud fraction is inferred from the area of the PDF that exceeds saturation. This requires specification of an assumed distribution with the associated parameters describing its shape.

(b) **Relative humidity** The second approach uses the relative humidity to diagnose subgrid cloud cover. If the relative humidity exceeds a critical value, a certain (horizontal) cloud fraction is estimated assuming a certain RH-cloud cover relation. An example is the commonly used relationship that has been proposed by Sundqvist et al. (1989):

$$C = 1 - \sqrt{\frac{1 - RH}{1 - RH_{crit}}} \tag{4.21}$$

With C the cloud fraction in a grid cell, RH the relative humidity and RH_{crit} the critical relative humidity above which subgrid scale clouds are assumed to exist. Evidently, $0 \leqslant C \leqslant 1$ and $RH_{crit} \leqslant RH \leqslant 1$. The cloud fraction monotonically increases from 0 to 1 when RH from RH_{crit} approaches 1. This approach to modelling cloud fraction has been implemented in WRF. Because of the simplistic nature of the approximation in 4.21 other $RH - C$ relations have been proposed e.g. to account for different cloud types. Another method by Xu and Randall (1996) that has also been implemented in WRF uses cloud water and ice mixing ratio combined with RH to estimate cloud cover. However this method tends to produce black and white values for cloud cover (Thompson (2016)). Thompson (2016) modified the Sundqvist et al. (1989) approach to let RH_{crit} vary depending on horizontal model resolution, with larger critical RH thresholds for smaller grid size resolution. A distinction between land and ocean is made for the critical RH value and no subgrid clouds can exist below LCL. For the subgrid clouds to affect radiation, estimates of liquid water content (LWC) an ice water content (IWC) have to be made based on humidity and temperature. Note that the estimated water contents only affect radiative transfer calculations and has no coupling with other model components. It is not a part of the LWC nor IWC, which are controlled only by the microphysics scheme. Experiments by Thompson (2016) showed that this fractional cloud cover approach reduced the downward shortwave radiation bias at the surface compared to simulations with the Xu and Randall (1996) method.

Radiative transfer models attempt to describe the effects of radiation that passes through a medium. For the atmosphere, that equates to calculating the up and downward fluxes for every vertical model layer of radiation over a wavelength spectrum that covers the Earth's and Sun's emission spectrum. The radiative transfer models in WRF are one dimensional, treating every vertical column independently without interaction. This is a valid assumption as long as the horizontal spacing remains much larger than in the vertical. They are split up into separate modules for short- and longwave radiation. The shortwave scheme assumes a certain amount of incoming solar radiation at the top of the atmosphere (TOA) based on the time of the day and season. Then, for every vertical layer the effects of absorption, scattering and reflection on the solar irradiance are estimated so that downward flux through each layer can be calculated. The irradiance at a vertical level z is defined as the frequency integrated transmission of incident solar radiation:

$$F_{SW,D}\left(z,\mu_0\right) = \mu_0 \int_0^{\infty} S_{TOA}(v)\,\tau\left(v,z,z',\mu_0\right) dv \tag{4.22}$$

Where $F_{SW,D}\left(z,\mu_0\right)$ is the downward shortwave radiation flux through a level z of a beam of solar radiation with a fraction $\mu_0 = \cos\theta_0$ of the solar zenith angle θ_0. $S_{TOA}(v)$ is the solar irradiance at frequency v at the TOA and $\tau\left(v,z,z',\mu\right)$ the transmittance from a level z to z', defined by Beer's law:

$$\tau\left(v,z,z',\mu\right) = \exp\left(-\frac{1}{\mu}\int_{u(z)}^{u(z')} k\left(v,p,T\right) du\right) \tag{4.23}$$

In this case $z' = z_{TOA}$, because the incident solar radiation is at the top of the atmosphere. $k\left(v,p,T\right)$ denotes an absorption coefficient, u the concentration of gas that attenuates the solar radiation. The amount of absorption is estimated by radiative transfer schemes using approximations to 4.23 and depends on the frequency or wavelength (see figure 4.18) of the radiation and concentration of absorbing gases in that layer. The amount of trace gases relevant for absorption like $CO2$, O_2, O_3, CH_4, N_2O is not modelled explicitly but assumed well mixed and constant for the whole atmosphere (e.g. for $C0_2$, N_2O and CH_4) or monthly varying profiles based on climatology, as for ozone (O_3). Only water vapour is modelled explicitly, which is convenient since it is possibly the most important greenhouse gas in our atmosphere, regarding the large absorption spectrum in figure 4.18. As mentioned in the previous

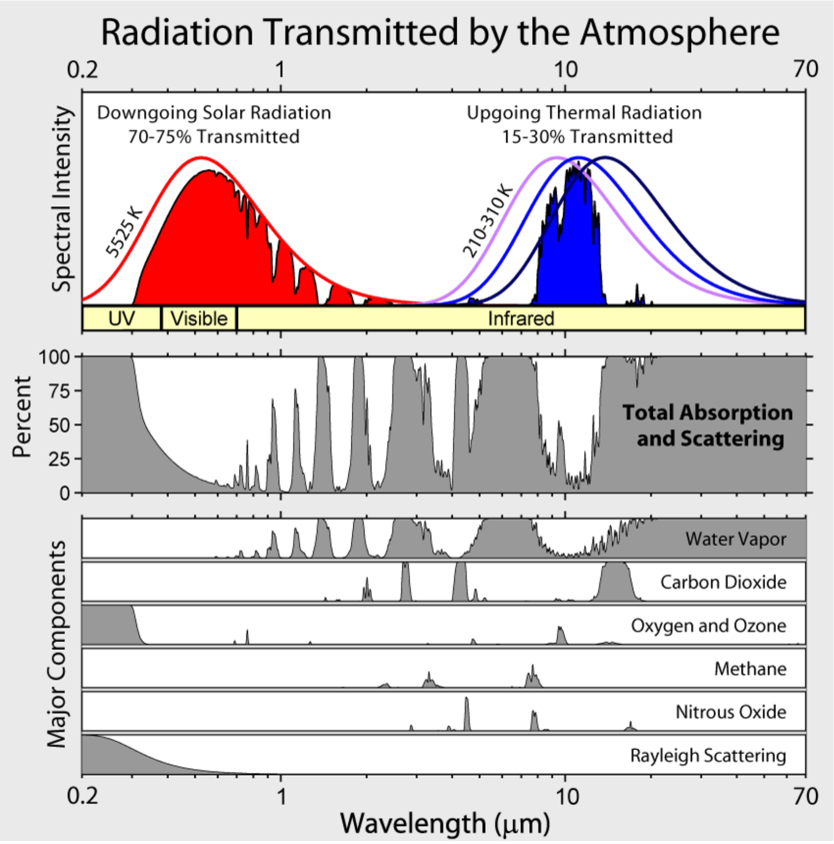


Figure 4.18: The lower graph shows the atmospheric absorption bands of the most abundant absorbing gases. For each gas (methane, oxygen, ozone, carbon dioxide, water vapour and nitrous oxide) the amount of radiation absorbed is shown as a function of wavelength. The middle graph shows the total combined absorption of gases the atmosphere. The top graph shows the solar (red) and terrestrial black body radiation spectrum and the intensity of radiation that is transmitted after absorption. Image from Wikipedia (2017).

point regarding microphysics parametrization, the scheme Thompson and Eidhammer (2014) enables spatial and temporal evolution of aerosols for cloud condensation and ice nuclei. These aerosol concentrations are used directly in the radiative transfer model instead of assuming constant values.

The longwave component of a radiative transfer model estimates the amount of both up and downwelling fluxes at each layer. The upward longwave radiative flux at a height z is defined as:

$$F_{LW,U}(z) = \int_0^\infty \pi B(v, z=0)\, \tau_f(v, z, z_{TOA})\, dv + \int_0^\infty \int_0^z \pi B(v, z') \frac{d\tau_f}{dz'}(v, z, z')\, dz' dv \tag{4.24}$$

and similarly, the downward longwave flux:

$$F_{LW,D}(z) = \int_0^\infty \int_z^\infty \pi B(v, z') \frac{d\tau_f}{dz'}(v, z, z')\, dz' dv \tag{4.25}$$

Where $B(v, z)$ denotes the Planck's law that gives the radiation intensity at a certain frequency, for a temperature at height z and $\tau_f(v, z, z')$ is the diffuse transmission function defined by the hemispheric integral over all cosines of zenith angles μ:

$$\tau_f(v, z, z') = 2\int_0^1 \tau(v, z, z', \mu)\mu d\mu \tag{4.26}$$

The upward longwave flux consists of a contribution from the emission of the Earth, as represented by the first term in 4.24, and a part by emission of longwave radiation by the atmosphere, represented by the double integral. The upward longwave radiative flux at the surface ($B(v, z=0)$) is determined by the land surface model, as the black body radiance spectrum is based on temperature. Again, radiative transfer models use approximations to these integrals to determine up- and downward longwave fluxes.

The radiative transfer model used in the simulations is called the rapid radiative transfer model for general circulation models (GCM's) (RRTMG) by Iacono et al. (2008), for both shortwave and longwave radiation. It is an adaptation of the earlier developed rapid radiative transfer model (RRTM) from Mlawer et al. (1997) to make it applicable in a NWP model. RRTMG employs the so called *correlated-k* method, which was developed to significantly reduce the number of computations compared to a line-by-line radiative transfer models, while maintaining a comparable accuracy. For a full introduction to this method the reader is referred to e.g. Iacono et al. (2008). In essence, the correlated k method rearranges the irregularly wavelength dependent absorption coefficient k of a certain wavelength interval to a smooth cumulative density function. The cumulative density function is then split into wavelength subintervals, for which a certain characteristic absorption coefficient is determined. For each interval the transmitted radiation is approximated using the characteristic absorption coefficient and Beer's law for extinction. The total radiance is obtained by taking the sum of radiances of each wavelength subinterval weighted by their size. The term correlated refers to the assumption that the mapping of wavelength absorption spectrum to cumulative distribution function is the same for every vertical layer or equivalently that k does not depend on pressure or temperature, which is not exactly true. Nevertheless it remains a sufficiently accurate and relatively inexpensive way to approximate the radiative transfer.

4.1.3. Data assimilation cycling

As discussed previously, data assimilation is employed to alter a NWP model state to one that corresponds better to actual observations. The most basic way to perform variational data assimilation is to assimilate observations at a discrete time, which has been introduced earlier as three dimensional variational assimilation (3DVar). Three dimensional refers to the fact that information from observation can spread in 3D model space space. More advanced methods like 4DVar use the NWP as an additional constraint to fit the observations to. That is, the model state of which the subsequent forecast trajectory best fits the observations at later times is found rather than finding a model state that is optimal only at a discrete time. However, similarly to ensemble methods, 4DVar is computationally very demanding because it requires computing the linearised model trajectories at every iteration in the minimization of the cost function. For that reason, operational forecasting favours 3DVar as it already significantly improves model forecast without requiring a lot of extra computational resources. To somewhat mitigate the inability to properly incorporate a dynamic constraint as with 4DVar, it is custom to apply 3DVar several times during a forecast. This strategy of alternating forecast and data assimilation is known as data assimilation *cycling*. For example, during the first hours of a forecast after initialisation, 3DVar can be applied every hour to better force a model towards the observations. In this study it is chosen to assimilate observations every 15 minutes. In the first place that is because the cases that will be studied feature strong convection which is really dynamic and evolving quickly. Modelling these features to at least some extent requires rapid update cycling between model and data assimilation. Although radar scans are available with an even higher update frequency, 15 minutes is believed to be frequently enough to have a decent impact on the forecast. Also, GPS ZTD measurements are updated with a 15 minute frequency so that the indirect radar humidity measurements are always constrained by another observation source where available.

Figure 4.19: Principle of cycling between data assimilation and forecast

Generally in all the experiments that are described in the following section, observations will be assimilated during a 3 hour window with a 15 minute interval between the assimilation times. The timeline in figure 4.24 provides a graphical illustration of the timewise setup of cycling between data assimilation and forecast. Before the data assimilation window starts, the model is given a spinup time of 3 hours. Spinup time refers to the time a NWP model needs to reach equilibrium after it is initialised from another source, in this case GFS model output. Before the large scale forcing can control the forecast, a thermal-dynamic balance has to be established. Generally, the initial conditions of a NWP model are not physically and dynamically consistent in the NWP model itself and model fields will have to adjust for some time to reach a stable state. Sometimes these adjustments are visible in the beginning of the forecast as gravity waves. The so called spinup time logically depends on the quality of the initial conditions but also on the model physics and the corresponding preferred stable state. Spinup time for NWP model usually ranges from 1-2 to 12 hours, also depending on the weather regime. For example regarding situations with significant convection, a model initialized from GFS will only contain horizontal winds and thus needs time to initiate vertical flow and become dynamically consistent with other model variables. Initialising a NWP model with coarser resolution model output or even just observations is referred to as a *cold start*. Contrarily, initialising a NWP model with an own model state in which data has been assimilated is known as a *warm start* and helps to reduce the spinup issue. In the following experiments the effects of assimilating observations on the quality of a forecast will be examined, in essence comparing cold and warm started forecasts. After the 3 hour spinup time and 3 hour data assimilation cycling, a 6 hour simulation is carried out which will be used in the verification described in section 4.3. In the experiments where no data assimilation is carried out, the 6 hour verification forecast is basically initialised with a cold started forecast that has spun up for 6 hours, thereby assuming that the model has reached equilibrium significantly enough for a valid comparison.

4.2. Case studies

As indicated before, the goal of data assimilation in a NWP is to find an optimal model state that minimizes the difference with a background forecast and observations and subsequently improve the forecast accuracy, primarily in forecasting precipitation in this study. How that accuracy is quantified will be addressed in section 4.3. To test the ability of data assimilation to improve the forecast, various case studies are examined. The setup of the experiments will be discussed in the next sections. First, it is interesting to investigate the effect of each observation type or source to see how much it improves or possibly degrades the forecast accuracy. To that end, so called *data denial* experiments are carried out where only a single observation type is assimilated. Secondly, we would like to know how data assimilation performs in several weather regimes producing precipitation. Therefore, three cases with different precipitation mechanisms are discussed.

4.2.1. Experiment 1: Observation type

The first experiment consists of repeating the data assimilation-forecast cycling only assimilating a certain type of observation. That way, the effect that assimilation of each observation has on the subsequent forecast can be isolated. This provides guidance whether or not to include these types of observations in further experiments or even an at operational stage later. Several types of observations are available for assimilation:

- **Conventional**
  Automated weather observation stations e.g. from KNMI provide frequent measurements of temperature, pressure, humidity, wind speed and direction, precipitation, et cetera. These observations from dedicated weather stations are classified as surface synoptic observation (SYNOP) measurements by the World Meteorological Organisation (WMO). Other surface measurements include observations from airports, known as meteorological aerodrom report (METAR). At sea, measurements can be obtained from ships or buoys. Although these observation types come from various sources with slightly different precision, they are grouped together because they provide direct measurements of basic model variables near the surface. Additionally, aircraft reports (AIREP's) are included because they are quite scarce and are also direct measurements of atmospheric variables.

- **Radar reflectivity - Hydrometeor content**
  In section 2.5.3 it has been discussed how radar reflectivity observations can be used to estimate hydrometeor contents. Radar reflectivity is linked to model mixing ratios for rainwater, snow and graupel based on temperature (equation 2.78).

- **Radar reflectivity - Humidity**
  Besides directly retrieving hydrometeor contents, radar reflectivity can act as indirect observations of humidity assuming that at a point of radar reflection water vapour pressure is close to saturation. How to deduce a humidity observation from a radar reflectivity observation has been shown in section 2.5.4. The humidity and hydrometeor observations are mentioned separately because they concern different quantities but moreover because their retrieval methods are significantly different.

- **Radar radial velocity**
  Next to measuring the reflectivity of objects in the atmosphere, the radar can also measure their radial velocity. The radial velocities can be assimilated as three dimensional wind speeds with components (u, v, w).

- **Global Positioning System (GPS) zenith total delay (ZTD)**
  Finally, GPS ZTD observations are available. Recall from section 2.7 that ZTD or in fact the inferred ZWD observations are a measure of the total, integrated water vapour in a vertical column.

With these observations, several experiments can be setup that use a combination of these observations. Table 4.1 provides an overview of all the experiments that are defined. The first experiment is a control (CTRL) run, where no observations of any kind are assimilated. In essence this is just a regular, unaltered WRF model run. The CTRL experiment serves as a benchmark for the other data assimilation experiments. Secondly, the CONV experiment assimilates only conventional, mostly surface observations that can be regarded as a robust baseline. The conventional observations will be used in any subsequent experiment, because they provide a reliable set of observations besides the generally more uncertain remotely sensed observations. Experiments with radar observations are split up into three cases, of which the first one is labelled CZQR, where besides the conventional observations also hydrometeors are assimilated. The next experiment CZQV assimilates radar derived humidity estimates and the CZRV experiment uses a set of radial velocity measurements. Additionally, an experiment with only GPS is designed to isolate the effect of GPS ZTD measurements. Finally, an experiment featuring all observations (ALL) is carried out to find out if the combined effect leads to the best forecasting of precipitation.

Scenario	Name	Conventional	Radar hydrometeor	Radar humidity	Radar radial velocity	GPS
1	CTRL					
2	CONV	x				
3	CZQR	x	x			
4	CZQV	x		x		
5	CZRV	x			x	
6	CGPS	x				x
7	ALL	x	x	x	x	x

Table 4.1: Data denial experiment scenarios. For every experiment number the name and the included observations are shown.

Case 1: Mesoscale convective system (MCS)

The above experiments will be performed using a WRF simulation of 26-08-2015, when a squall line passed over the Netherlands and Belgium. The squall line resulted from a sharp cold front extending from the North Sea down to the north of Spain. The surface reanalysis by the UKMO in figure 4.20 provides a synoptic overview of the atmosphere over Europe. In the Netherlands the cold front approaches with a south westerly wind, resulting from the cyclonic flow around a large scale low pressure area centered west of Scotland. During the late afternoon the cold air collides with the warmer and moist air that was advected from the Atlantic the day before by the same low pressure area. The cold front causes positive buoyancy for the warm, moist air that is present above the Netherlands, causing a narrow line of deep convection along the cold front and compensating downdrafts with heavy precipitation just behind, as indicated by the high $> 40\ dBZ$ reflectivity measured by the radar (see figure 4.21). An idealized cross section of this process is shown in figure 4.23. The cold front preceded by the gust front lifts the warm and moist air through the inversion to its level of free convection (LFC). Besides the upward forcing by the cold front, a significant amount of low level vertical wind shear as indicated by a $15\ m/s$ increase in wind speed between the surface and the $850\ hPa$ level

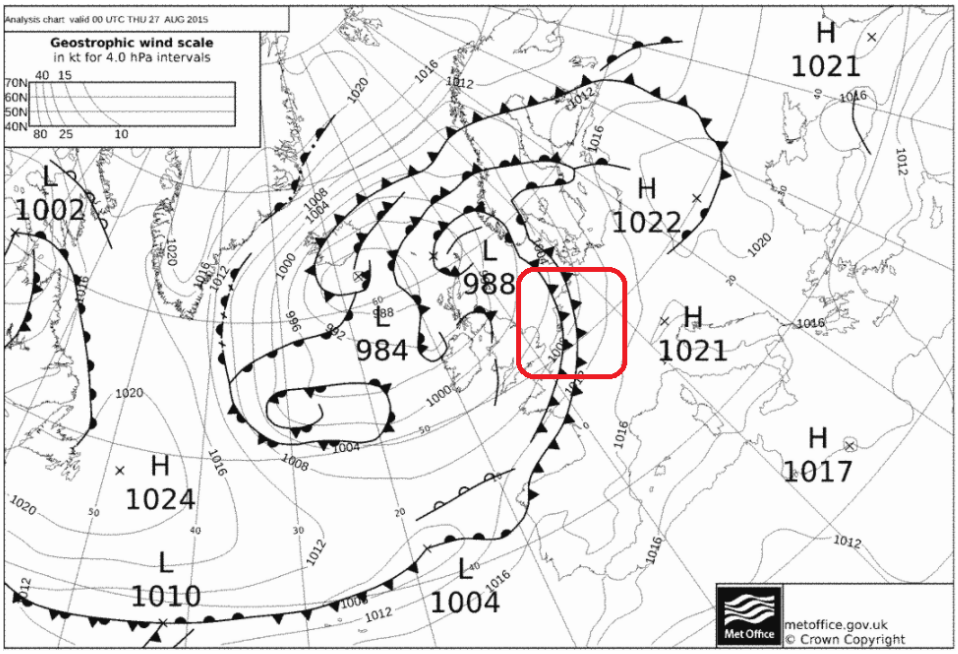
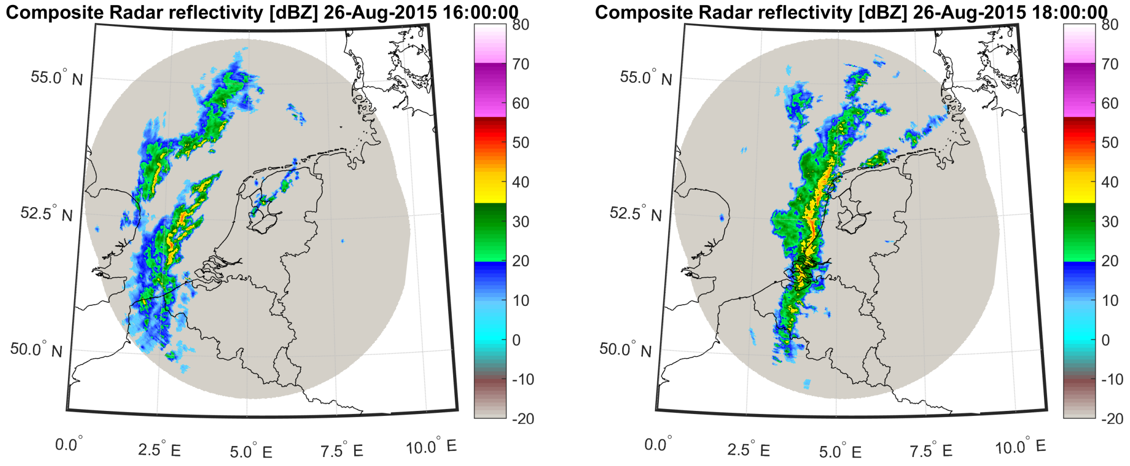


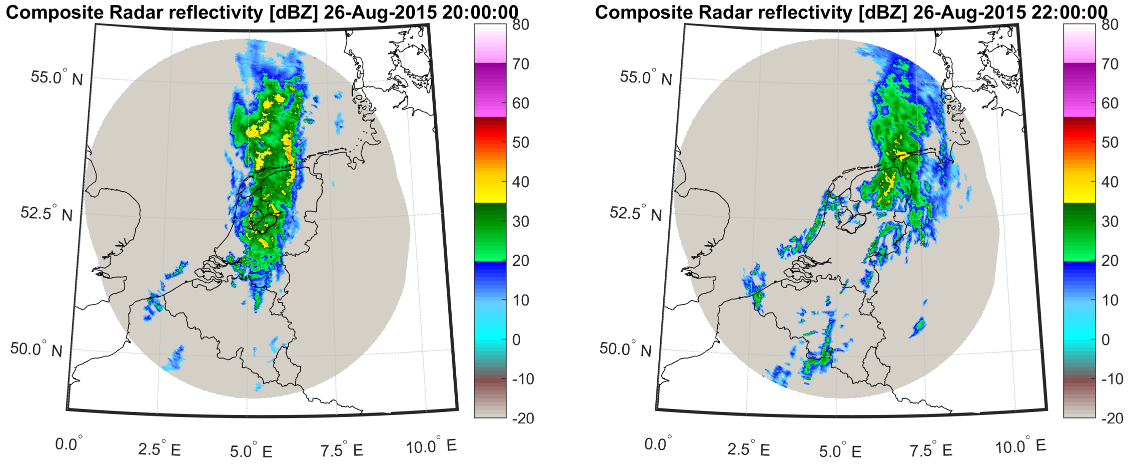
Figure 4.20: UKMO synoptic scale surface reanalysis at 00 UTC on 27-08-2015. The red box marks the region of interest above the Benelux. Conventional surface weather map symbols are used to indicate surface pressure, low (L) and high pressure (H) cores and front lines.

contributed to intensity of and longevity of the squall line. Vertical wind shear aids in the generation of vertical convection at the cold front since the vorticity generated by the vertical shear is opposite in sign to the vorticity generated by the cold front that lifts up the warm air that it encounters. Figure 4.22 illustrates this principle where positive vorticity resulting from the vertical wind shear balances the negative vorticity originating at the downstream (right) side of the cold front, strengthening the vertical motions. This is opposed to the upstream part of the cold pool, where the vorticity generated by the cold pool and the vertical shear add up, bending the air flow over the cold pool rather than forcing it upwards. For maximum longevity of the squall line the wind shear vector should be oriented perpendicular to the squall line or equivalently parallel to the direction of propagation, so that a maximum amount of warm and moist air can be convected. In this case the normal is oriented towards the east while the direction of propagation is northeast (i.e. a SW wind), so that the squall line orientation is suboptimal and thus contracts and starts to dissipate later in the evening. Radar images in figure 4.21 show the squall line forming, making landfall at 18 UTC and further contracting and losing intensity later during the evening as it moves northeast ward.



(a) Radar reflectivity at 16 UTC

(b) Radar reflectivity at 18 UTC



(c) Radar reflectivity at 16 UTC

(d) Radar reflectivity at 18 UTC

Figure 4.21: Composite radar reflectivity on 26-08-2015. A squall line makes landfall over the Netherlands. Each image shows radar acquisitions at different times.

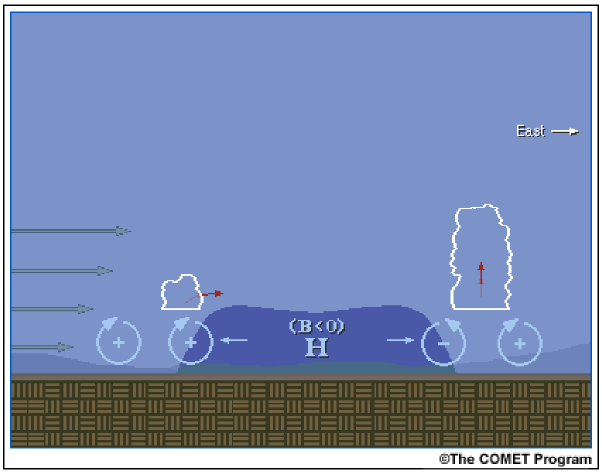


Figure 4.22: Schematic illustration of the interaction of a cold pool and environmental vertical wind shear. The cold pool generates positive vorticity at the upstream (left) side and negative vorticity at the downstream (right) side when the cold air forces the warm air upwards. This results in stronger downstream curvature of the air going over the cold pool on the upstream side, while on the downstream side the opposite signs cause a decrease in vorticity, thus generating more vertical motions. Illustration credits to Program (2016).

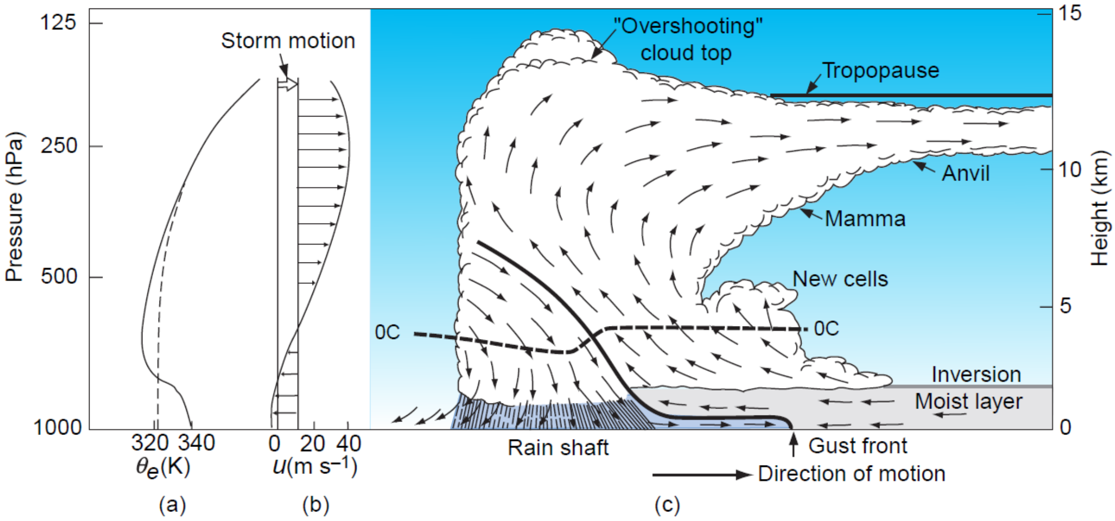


Figure 4.23: Schematic illustration of a vertical cross section of a multicell storm. Figure (a) and (b) show the equivalent potential temperature and wind speed profile with height. Figure (c) depicts an cross section showing the development of an idealized storm. Illustration by Wallace and Hobbs (2006).

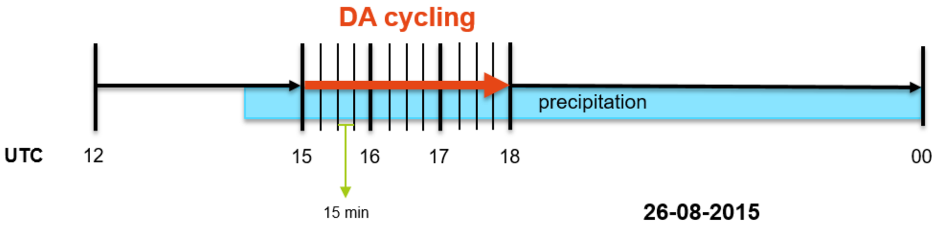


Figure 4.24: Time setup of the data denial experiment. The forecast is initialized at 12 UTC and observations are assimilated during three hours from 15 to 18 UTC with a 15 minute interval. A subsequent 6 hour forecast is run using the data assimilation output. The blue bar indicates the time significant precipitation was measured by the radar.

The timeline in figure 4.24 shows how the data assimilation cycling experiment is setup timewise. The simulation is initialised at 12 UTC from GFS output. Observations are assimilated during a 3 hour interval between 15 and 18 UTC, as shown in figure 4.24, with a 15 minute update interval. At 18 UTC, the data assimilation cycle ends and a 6h subsequent forecast is started, of which the estimated precipitation will be used to compare with actual radar measurements at the valid times to assess the quality of the forecast. The blue bar indicates the time window in which a decent amount of precipitation is within the domain and in range of the radar.

4.2.2. Experiment 2: Precipitation mechanism

In the second case study the performance of data assimilation will be investigated for different types of precipitation mechanisms. The several experiments that are conducted are summarized in table 4.2. The first type that is considered is the MCS of 26-08-2015 that featured a cold front with a squall line as discussed in the previous section. Along the cold front boundary convection is forced on a large scale and thus this is labelled as a MCS. Since this is a relatively large scale forcing, the WRF model should be able to describe this convective system quite well, even without assimilation of observations. For this mechanism two experiments are conducted: a control run with no assimilation (MCS-CTRL) and a data assimilation scenario in which all observations are assimilated (MCS-DA).

Case 2: Local convective system (LCS)

In contrast with case 1 where data assimilation during a MCS is examined, case 2 features a local convective system with relatively small yet intense convective cells that occurred a couple days later on 30-08-2015. A synoptic scale overview of the surface weather by UKMO valid at 00 UTC 31-08-2015 is shown in figure 4.25.

Case	**Name**	**Date**	**Cycling**	**Forecast**	**Description**	**DA obs**
1a	MCS-CTRL	26-08-2015	15-18 UTC	18-00 UTC	Mesoscale convective system	-
1b	MCS-DA	26-08-2015	15-18 UTC	18-00 UTC	Mesoscale convective system	ALL
2a	LCS-CTRL	30-08-2015	19-22 UTC	22-04 UTC	Local convective system	-
2b	LCS-DA	30-08-2015	19-22 UTC	22-04 UTC	Local convective system	ALL
3a	STRAT-CTRL	20-06-2016	10-13 UTC	13-19 UTC	Stratiform	-
3b	STRAT-DA	20-06-2016	10-13 UTC	13-19 UTC	Stratiform	ALL

Table 4.2: Description of the six precipitation data assimilation experiments. Three dates are examined that each feature a different type that drives precipitation. For each date a control (CTRL) and a DA experiment is performed

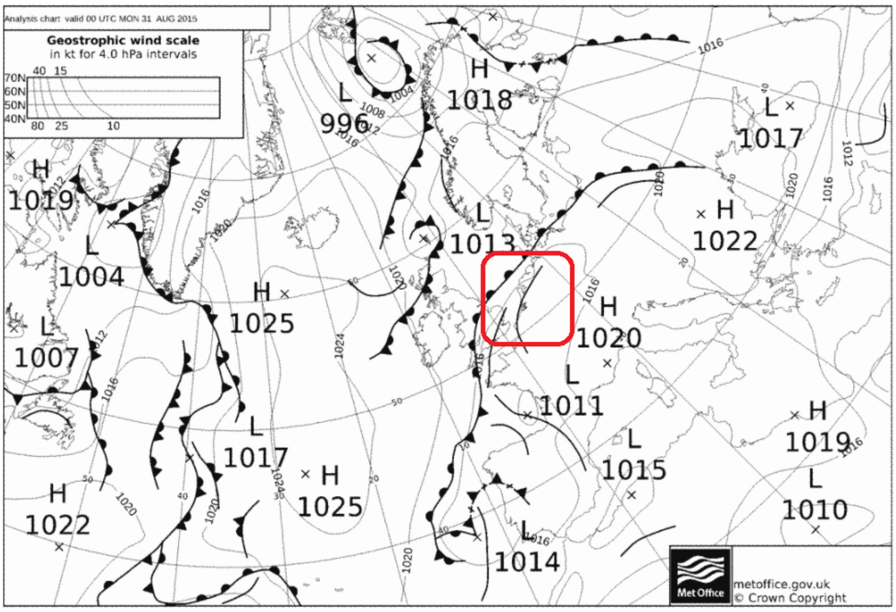
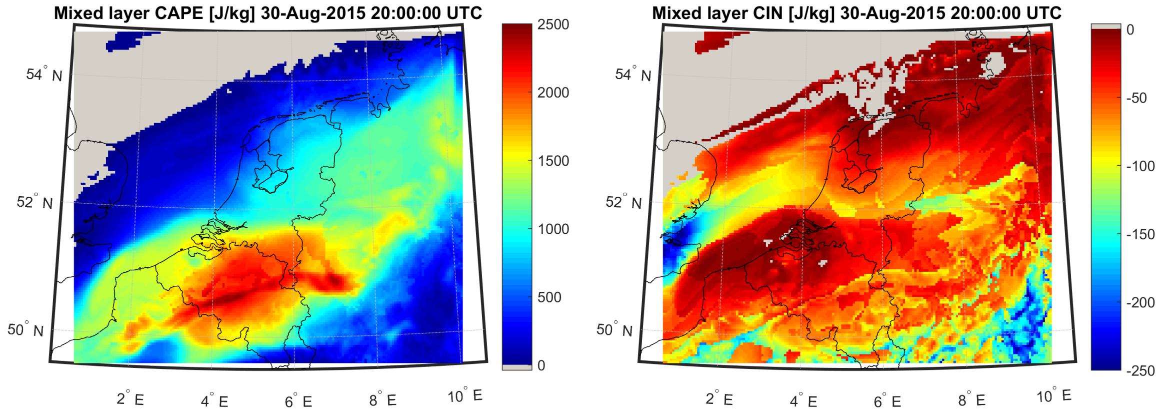


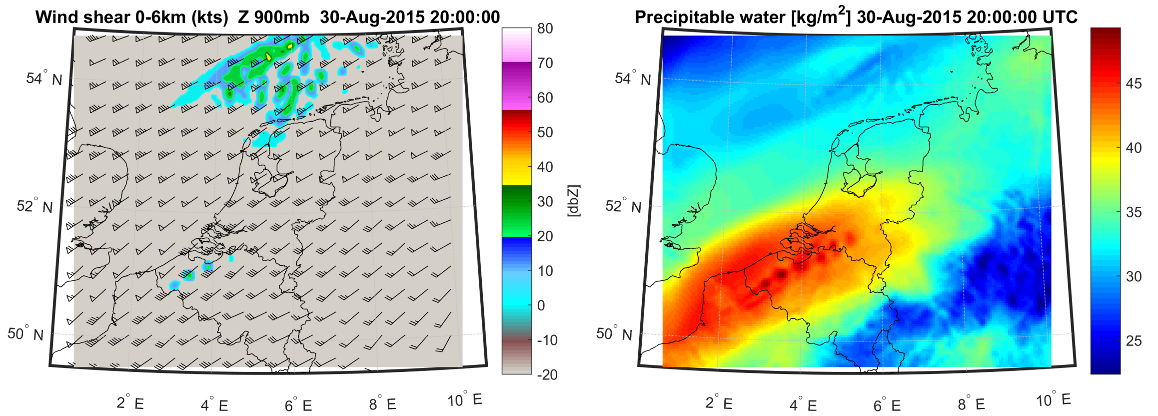
Figure 4.25: UKMO synoptic scale surface reanalysis at 00 UTC on 31-08-2015. The red box marks the region of interest above the Benelux. Conventional surface weather map symbols are used to indicate surface pressure, low (L) and high pressure (H) cores and front lines.

The short convergence line over the Netherlands just after the warm front marks a stretched area of relatively low pressure, which is associated with updrafts and an corresponding increased chance of precipitation. The local convective cells form around 19 UTC on this line of convergence along the Belgian coast and south-west part of the Netherlands. Along this line very unstable, moist and warm air provided a favourable environment for deep convection. The instability is indicated by significant ($> 1500\,J\,kg^{-1}$) amounts of convective available potential energy (CAPE), as shown in figure 4.26a where the CAPE is determined using temperature and dewpoints of the mixed layer. CAPE as the name suggests is an indicator of the potential intensity of convection. However to trigger convection, the energy preventing convection, known as convective inhibition (CIN), has to be overcome. On the evening of 30-08-2015, CIN values in the boundary layer approached zero along this line (as shown in figure 4.26b) and gave rise to the initiation of intense convection. The IWV as plot in figure 4.26d shows the band of moist air approaching from France. A strong SW jet stream also caused additional dynamical forcing from above which helped to develop strong convection. In figure 4.26c the wind shear between the surface and 6 km height is plotted, showing a wind shear of approximately 45 knots. Additionally the model simulated radar reflectivity is shown with contours in the background. Notice that in the southwest part of Belgium the first convective cells forms as indicated by the small area of high reflectivity. In that same area a dip in CAPE in figure 4.26a can be witnessed as the convection converts the potential energy to kinetic energy. Similarly, a small peak amid an area of near zero CIN is present in figure 4.26b, resulting from the vertical transportation of warm, moist air leading to warmer air at higher elevations. The cells in the model form slightly later and more south than compared to the radar image at 20 UTC in figure 4.27a.

Figure 4.27 shows other radar images acquired between 20 and 02 UTC. The dark red $> 50\ dBZ$ areas highlight the cores of the local convective cells that start to develop around 19 UTC. At 22 UTC the cells have



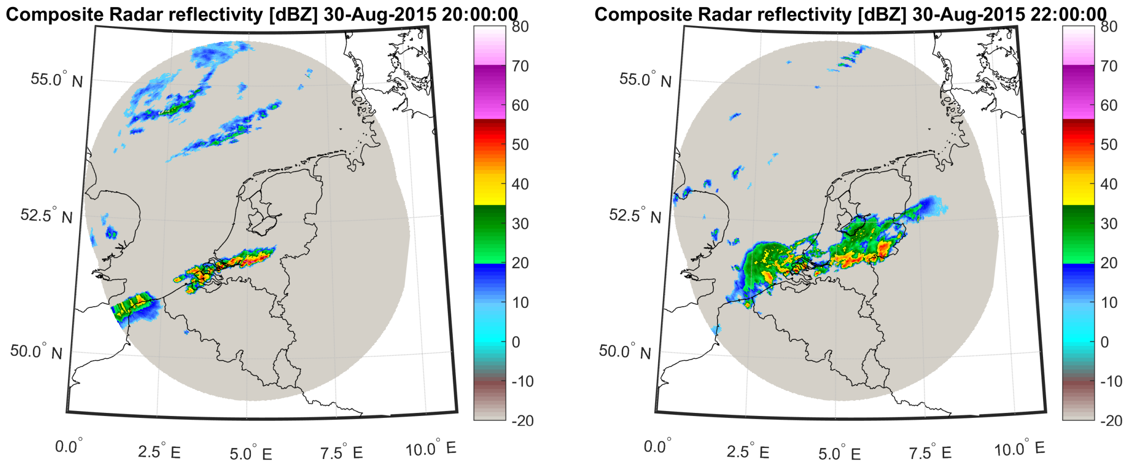
(a) CAPE at 20 UTC (b) CIN at 20 UTC



(c) Wind shear 0-6km, 900mb reflectivity at 20 UTC (d) IWV at 20 UTC

Figure 4.26: WRF model state at 30-08-2015 20:00 UTC.

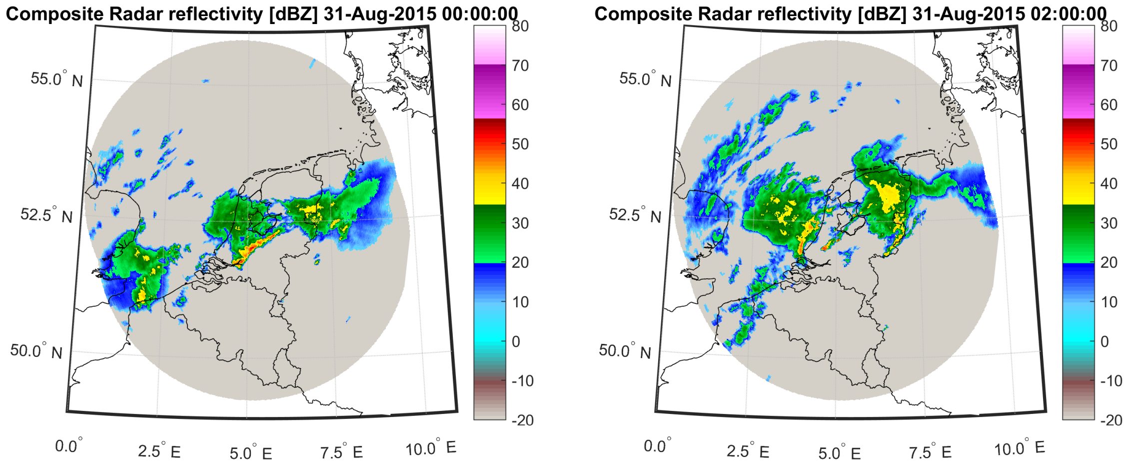
grown in size and a east-west band of precipitation covers the central part of the Netherlands. Observations are assimilated during an interval from 19 UTC when the cells form to a developed situation at 22 UTC as shown in figure 4.27b. In the following 6 hours two larger, separated areas of precipitation form, as seen in figure 4.27d. The really local, intense cells develop into a bigger system with a slightly smaller intensity, yet still producing significant precipitation. The left area of precipitation in figure 4.27d has only entered the domain after 22 UTC and therefore cannot be assimilated directly using radar data assimilation within the assimilation window. Consequently, it is expected that the added value of radar data assimilation and thus the forecasting skill will degrade quite rapidly after 00 UTC, as the areas of precipitation that were assimilated leave the domain. Similarly as for the MCS case, the local convection case features a control run (LCS-CTRL) and a simulation with assimilation of all observations (LCS-DA), refer to table 4.2 again for details. Out of all cases, the model will likely have the hardest time accurately representing this LCS case as it is a really local phenomenon with scales comparable or even smaller than model resolution.


Composite Radar reflectivity [dBZ] 30-Aug-2015 20:00:00

Composite Radar reflectivity [dBZ] 30-Aug-2015 22:00:00

(a) Radar reflectivity at 20 UTC

(b) Radar reflectivity at 22 UTC


Composite Radar reflectivity [dBZ] 31-Aug-2015 00:00:00

Composite Radar reflectivity [dBZ] 31-Aug-2015 02:00:00

(c) Radar reflectivity at 00 UTC

(d) Radar reflectivity at 02 UTC

Figure 4.27: Composite radar reflectivity on the night of 30-08-2015 to 31-08-2015. Intense, localized convection cells form over Belgium and the Netherlands. Each image shows radar acquisitions at different times.

Case 3: Stratiform (STRAT)

The third case concerns a day with stratiform precipitation: 20-06-2016. The core of an extratropical cyclone is positioned just south of Iceland, as shown in the surface reanalysis from UKMO in figure 4.28. In the Northern hemisphere air rotates in an anticlockwise manner around such a low pressure core, resulting in dominantly western winds in the Netherlands. An occluded front stretches from the low pressure center over the North Sea to just north of the Netherlands, where a triple point marks the junction of a large scale cold front approaching from the Atlantic ocean, overtaking a smaller scale warm front that extends into France. This warm front crosses the Netherlands during the afternoon between 12 and 18 UTC. Radar images in figure 4.30 show the advance and development of the rain associated with the warm front. Warm, moist air from the warm front is less dense than the cool air it encounters to the East and thus it is gently forced upwards over the cold air. As the air rises, it expands and cools down, forming clouds and precipitation. A warm front usually has quite a gradual slope, first leading to formation of high cirrus clouds followed by cirrostratus when the warm front is approaching. If the air contains sufficient moisture, as is the case here, precipitating or nimbostratus clouds develop along the warm front. Since the ascent of an air parcel is not that steep, water vapour condenses gradually and only light to moderate precipitation is formed. The radar images in figure

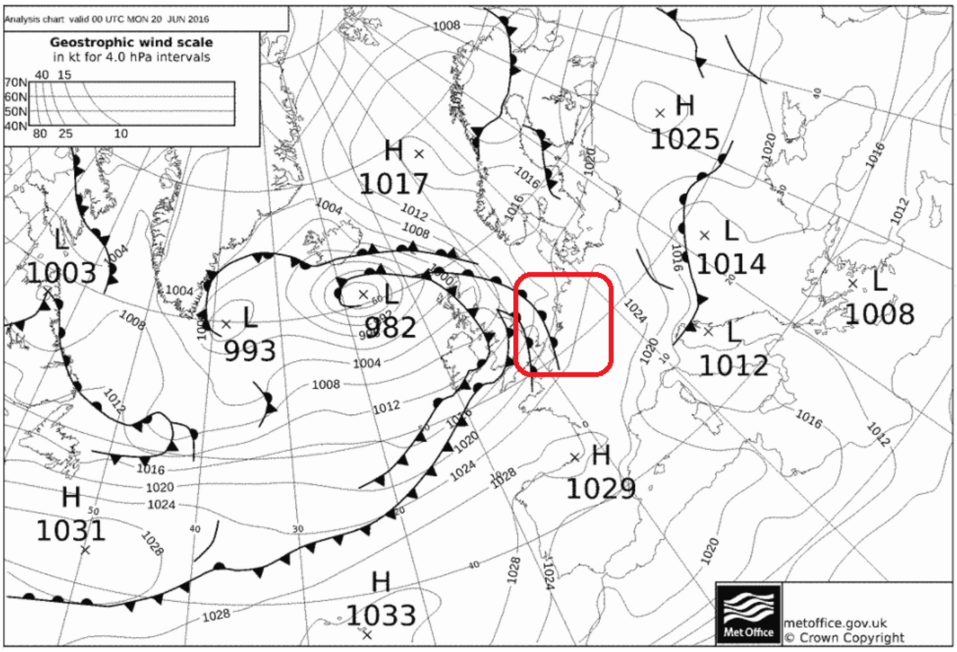


Figure 4.28: UKMO synoptic scale surface reanalysis at 00 UTC on 20-06-2016. The red box marks the region of interest above the Benelux. Conventional surface weather map symbols are used to indicate surface pressure, low (L) and high pressure (H) cores and front lines.

4.30 show only moderately intense precipitation, with maximum reflectivity observations of 35 dBZ (refer to table 2.2). Around 18 UTC, most of the precipitation has passed, leaving only some trailing edges. As the warm front is a mesoscale weather phenomenon, it is likely to assume that the process causing the stratiform precipitation is captured by a NWP quite well already without any data assimilation. Because of this large scale nature, it is expected that assimilation of high resolution observations will not provide as much added value for the STRAT case compared to the more dynamic and local events in the MCS and LCS cases.

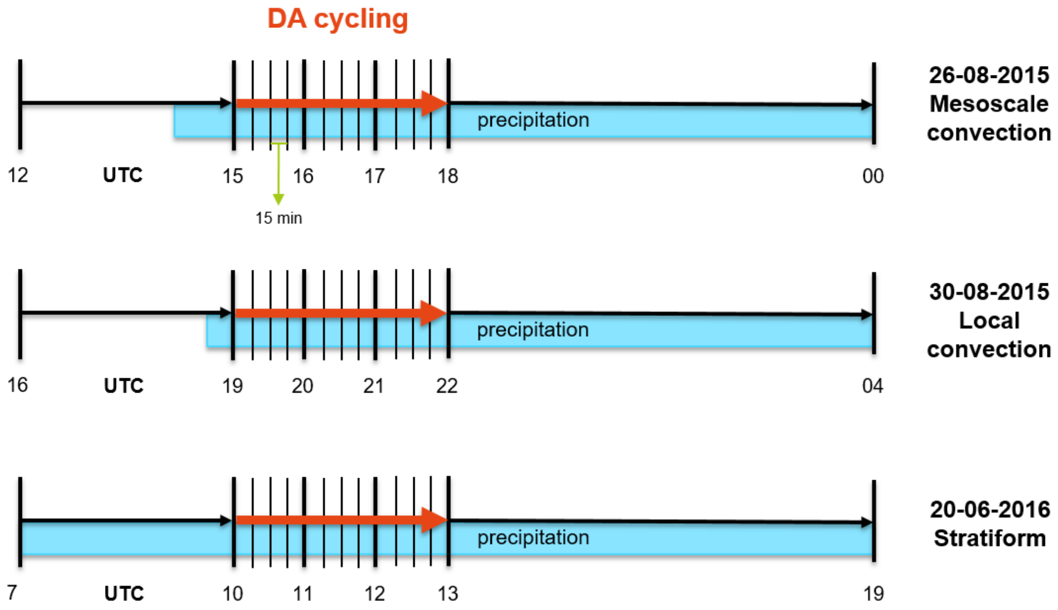
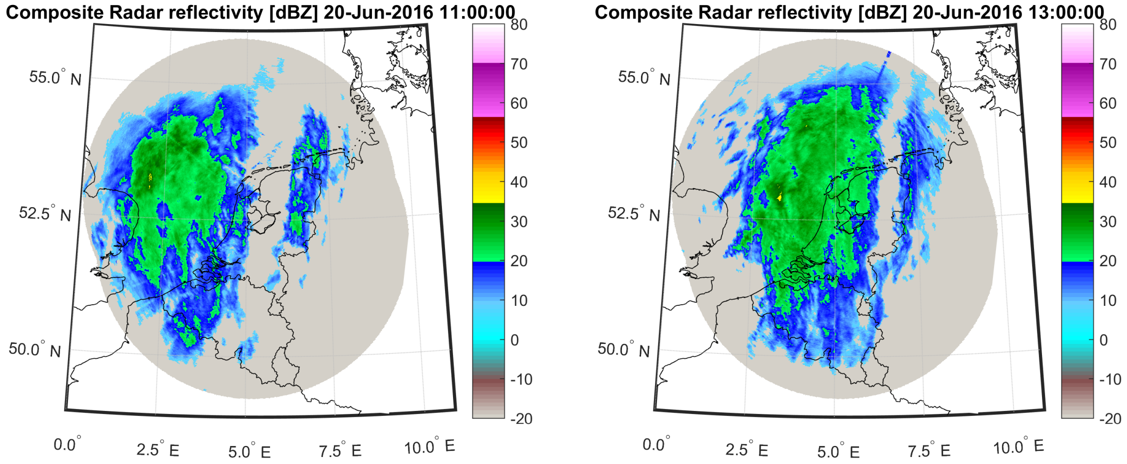


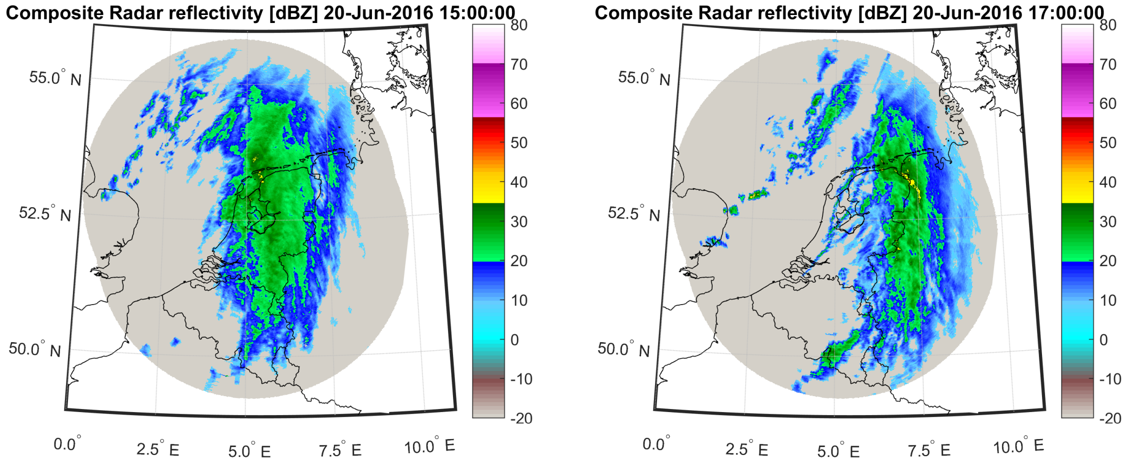
Figure 4.29: Timeline of the MCS, LCS and STRAT cases. Left time is the initialisation time of the simulation. The orange arrow indicates the cycling window where observations are assimilated with a 15 minute interval. A subsequent 6 hour forecast is made for each case. The blue bar indicates the time in which significant amounts of precipitation detected by the radar are within the domain.

A timeline for each of the three dates used for case studies is given in figure 4.29. The leftmost time



(a) Radar reflectivity at 11 UTC

(b) Radar reflectivity at 13 UTC



(c) Radar reflectivity at 15 UTC

(d) Radar reflectivity at 17 UTC

Figure 4.30: Composite radar reflectivity on the 20-06-2016. Stratiform precipitation covers the Netherlands throughout the most part of the day.

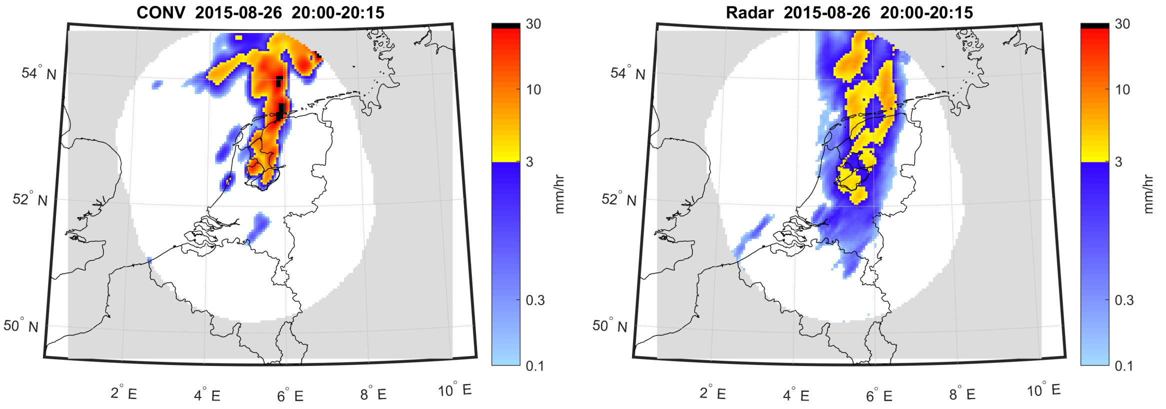
marks the time of initialization from the GFS model. A three hour forecast is generated using the initial state, allowing the model some to spin up. After the spinup time observations are assimilated every 15 minutes for a duration of 3 hours. when the cycling window is over, a 6 hour WRF simulation is made which will be used to verify the precipitation forecast. The blue bar in the timeline highlights the time during which significant amounts of precipitation are present within the domain.

4.3. Verification

When assessing the forecasting skill of precipitation, several objective measures of verification have to be defined. Primarily, forecasting skill is associated with the ability to predict whether it is actually raining or not. This qualitative method of verification is discussed in the next section, 4.3.1. Additionally however one would like to have a good estimate on the precipitation quantity. Therefore, some basic QPF verification scores are defined in section 4.3.2.

4.3.1. Qualitative verification: Rain/no rain

In the previous sections several experimental setups to produce precipitation forecasts have been discussed. These precipitation forecasts from the WRF model should be verified in order to quantify the skill of correctly forecasting precipitation. A qualitative way to do this is to classify areas (or in this case grid cells) whether they rain or not, for both model and observations. In this study the 15 minute averaged rainfall rate is used to determine whether it rains or not. In practice this requires defining a rainfall rate threshold R_{min} that is associated with perceivable rain, usually between 0.1 $mm\,h^{-1}$ and 1 $mm\,h^{-1}$. Refer to table 2.2 for an overview of rainfall rates and the corresponding classification of perceived intensity. A commonly used threshold is 0.3 $mm\,h^{-1}$, which will also be employed in this study unless indicated otherwise. Naturally, defining a different threshold will yield slightly different results, especially if the rainfall intensity distributions of model and observation are not similar.



(a) WRF model

(b) KNMI radar composite

Figure 4.31: Example of 15 minute average rainfall rate

Radar composites from the KNMI are used as 'truth' in the verification. These composites have been constructed by linearly interpolating reflectivity observations from all scans of the radars at De Bilt and Den Helder to a height of 1.5 km. Refer to table 3.1 again for an overview of the scans that are used to construct the composites from. Although the radar composites have a range of 320 km according to the metadata specification, at this range the altitude of the lowest elevation scan at which a reflection is measured is approximately 7.7 km when using the approximation from Doviak and Zrnic (1993) for the radar beam height as a function of range for a standard atmosphere. Often at this height the radar overshoots any clouds that contain precipitation. Consequently, the radar might not measure any precipitation, while an underlying precipitating cloud produces rainfall at the surface which is also correctly forecast by the model. In that case the verification would suggest a false positive while in truth the model accurately predicted rain. For this reason the radar range used in the verification is brought down to 220 km, with a corresponding beam height of 4.0 km. While the radar beam still might overshoot some precipitation, it is assumed that up until this range the rainfall intensity estimates are representative for surface rainfall intensity. There are some caveats to this assumption, like the possibility of rain evaporating before it reaches the surface or strong winds that can significantly advect hydrometeors. Additionally, to compensate for errors introduced by the Z-R relation radar retrieved rainfall rates would have to be corrected with surface rain gauge measurements to obtain an even more accurate estimate, however that is beyond the scope of this work.

Having defined a threshold R_{min} above which pixels are precipitating, each pixel can be classified for both the radar and model images. With all the pixels classified, a **confusion matrix** is constructed, which shows in a tabular form the skill of forecasting an event. In some other literature it is better known as the **contingency**

		Observed		
		Yes	**No**	**Total**
Forecast	**Yes**	a	b	a+b
	No	c	d	c+d
	Total	a+c	b+d	a+b+c+d=n

Table 4.3: General form of the confusion matrix. a represents the number of times an event was correctly forecast when it actually happened. Numbers b and c represent occurrences of false positives and negatives, respectively. Finally, d is the number of times an event was correctly forecast as non occurring. These classification numbers can be used to derive subsequent verification scores characterizing the skill of verification.

table, although a contingency table need not necessarily be a confusion matrix as a contingency table can be multivariate and thus have dimensions larger than 2x2. The number a represents the amount of times an event (in this case whether it rains) occurs and is forecast correctly. Similarly, d is defined as being the number of non precipitating grid cells measured by radar which were correctly forecast by the model. Finally, b represents the occurrences of precipitation forecast when it actually did not occur (generally known as *false positives*) and vice versa for c, known as *false negatives*. These statistics can be used to derive useful verification scores that describe several aspects of the forecast skill:

1. **Percentage correct (PC)**

$$PC = \frac{a+d}{n} \tag{4.27}$$

The percentage correct (PC) measures how much events in total are correctly forecast. Ranges from 0 (poor) to 1 (good). The PC can be misleading when a low frequency event is considered: $a \ll d$. Then the correctly forecast non occuring events d will cause a high PC although there is no skill in correctly forecasting occurring events a.

2. **Hit rate (H)**

$$H = \frac{a}{a+c} \tag{4.28}$$

The H is the fraction of observed events that is forecast correctly. Ranges from 0 (poor) to 1 (good).

3. **Probability of false detection (POFD)**

$$POFD = \frac{b}{b+d} \tag{4.29}$$

probability of false detection (POFD) is the fraction of observed non occurring events that is falsely forecast. Ranges from 0 (good) to 1 (poor).

4. **False alarm ratio (FAR)**

$$FAR = \frac{b}{a+b} \tag{4.30}$$

false alarm ratio (FAR) measures the fraction of false forecasts of the total positive forecasts. Ranges from 0 (good) to 1 (poor).

5. **Critical succes index (CSI)**

$$CSI = \frac{a}{a+b+c} \tag{4.31}$$

critical succes index (CSI) combines H and FAR to be a measure of how well low frequency events are forecast, contrary to PC. Note that d is not used in this score. CSI range from 0 (poor) to 1 (good).

6. **Classification bias (CB)**

$$CB = \frac{2b}{b+c} - 1 \tag{4.32}$$

Finally, classification bias (CB) is defined to describe the biasedness in forecasting, by taking the ratio of false positives b and false negatives c according to 4.32. The CB is centered around zero when $b = c$. For the CB, $CB < 0$ indicates underprediction, while $CB > 0$ signals overprediction.

7. **Gilbert skill score (GSS)**
   The GSS is similar to the CSI, with the difference that the correct positives a are corrected with the correct positives that would be expected by chance:

$$GSS = \frac{a - a_{random}}{a - a_{random} + b + c} \tag{4.33}$$

   The hits expected by random probability are:

$$a_{random} = P_c \cdot n = \frac{(a+b)(a+c)}{n} \tag{4.34}$$

   In general, a_{random} is relatively small so that the GSS will be close to the CSI. Similarly to the Heidke skill score (HSS), a $GSS = 1$ mean a perfect forecast. At $GSS = 0$, the number of correct positives are equal to the number expected by chance, showing no skill over a CSI from a random forecast. A $GSS < 0$ signals that even less correct positive forecasts were made than a random forecast.

4.3.2. Quantitative verification: rainfall intensity

The previous section 4.3.1 provides useful tools to asses the models ability to accurately predict whether it is raining or not. However this qualitative description does not provide any information on the amount of precipitation. Although accurate quantitative precipitation forecasting (QPF) remains a challenge for mesoscale models with resolutions in the order of kilometers, it is certainly interesting to examine how well the model forecasts precipitation quantity. In order to do so two additional quantities are employed, adapted from Brasjen (2014): the relative bias and the absolute ratio of the average 15 minute rainfall rates of model R_{mod} and radar R_{radar}. The **relative bias** δ_R is defined as follows:

$$\log_{10}\left(\delta_{R|R>0}\right) = \frac{1}{n}\sum_{n}\log_{10}\left(\frac{R_{\text{mod}\,|R>0}}{R_{radar|R>0}}\right) \tag{4.35}$$

Because the rainfall rate ranges several orders of magnitude, a logarithmic scale is preferred so that a difference in rainfall rate between 0.1 $mm\,h^{-1}$ and 1 $mm\,h^{-1}$ is half as bad as a difference between 0.1 $mm\,h^{-1}$ and 10 $mm\,h^{-1}$, as they vary one and two orders of magnitude, respectively. δ_R ranges from nearly 0 to infinity and a value $0 < \delta_R < 1$ indicates that on average the model underestimates the intensity of precipitation, while for $\delta_R > 1$ rainfall rate in the model is too high compared to the radar retrieved rainfall rate. Notice that δ_R can only be calculated for pixels that are precipitating in both the model and radar images, hence the symbol $R > 0$. So dry conditions or even displacement errors of estimated precipitation areas can result in only a small amount of points available for determining δ_R. This should be kept in mind when interpreting δ_R. The threshold for determining when $R > 0$ remains 0.3 $mm\,h^{-1}$ as before.

To give an indication of the average magnitude difference, additionally the **absolute ratio** ρ_R is defined:

$$\log_{10}\left(\rho_{R|R>0}\right) = \frac{1}{n}\sum_{n}\left|\log_{10}\left(\frac{R_{\text{mod}\,|R>0}}{R_{radar|R>0}}\right)\right| \tag{4.36}$$

The absolute ratio ρ_R ranges from 1 to infinity. A ρ_R close to 1 implies that the model and observations agree really well on the rainfall intensity, whereas higher values indicate on average large scale differences between forecast and observed precipitation.

5

Results

This chapter will discuss the results of the data assimilation assimilation experiments described in the previous section 4 on the experiment setup. The first experiment concerns the effect that assimilating several types of observation has on the quality of the precipitation forecast. The second experiment has been set up to investigate the performance of data assimilation in three cases with different mechanisms that drive precipitation formation. As a last experiment the cycling window in which observations are assimilated is shifted to see how crucial it is for the precipitation forecast that observations of which primarily radar that cover the first stages of precipitation are assimilated.

5.1. Experiment 1: Observation type

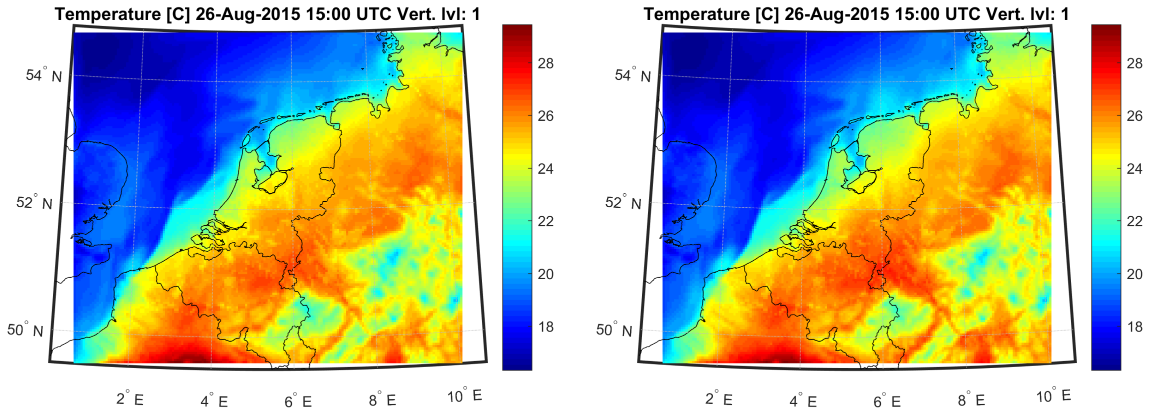
This experiment will investigate the influence that assimilating each observation has on the model state and consequently how that affects the quality of the precipitation forecast. Five types of observations are being investigated, namely: conventional, radar retrieved hydrometeor mixing ratios, radar derived humidity, radar radial velocity and GPS ZTD. Out of these, seven scenarios were formulated that have been described in section 4.2.1. Refer to table 4.1 for an overview of the observation types that are assimilated for every experiment.

5.1.1. Single data assimilation

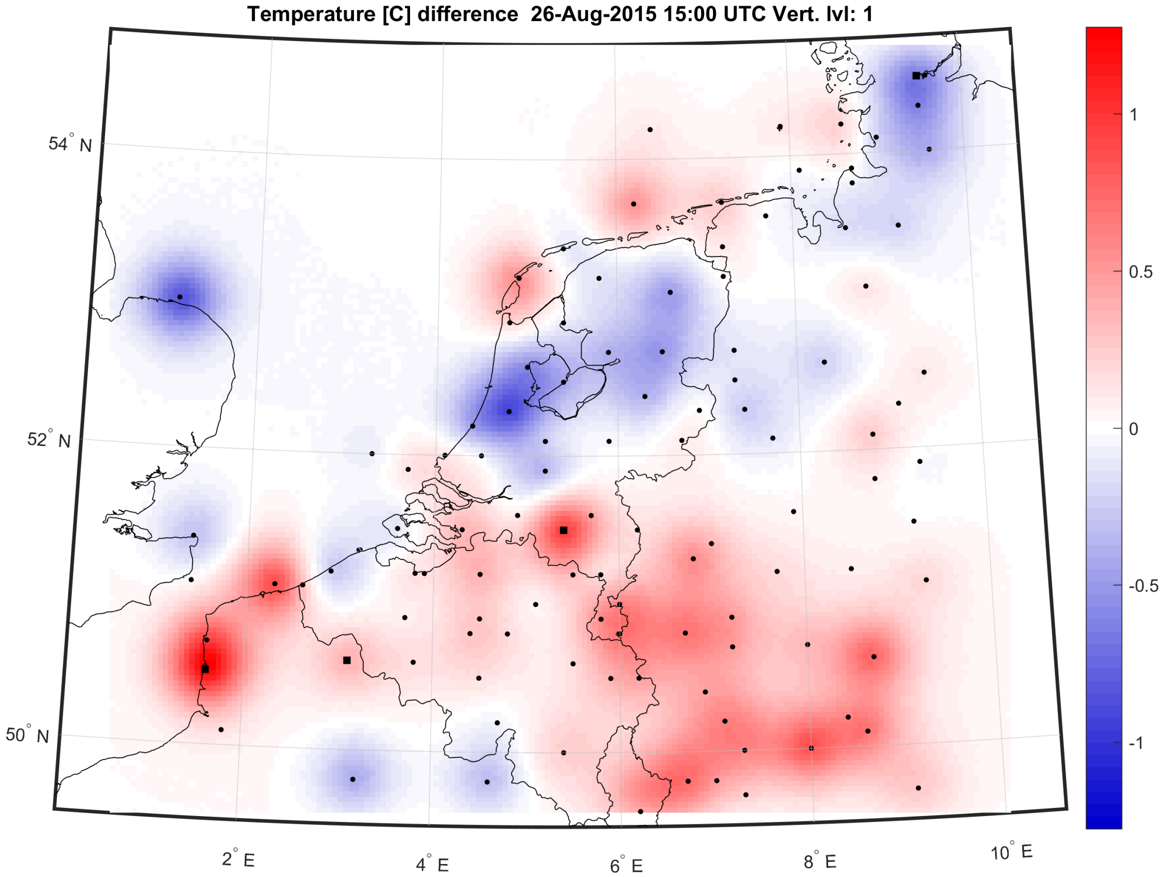
Before the verification is carried out with the subsequent forecast after data assimilation cycling, it is worthwhile to examine the adjustments that assimilation of observations make on the model state at a single assimilation time. Especially the start of the assimilation window at 15 UTC is interesting to look into as it is the first time in the forecast that observations are assimilated into an unaltered model state.

2. CONV

The first scenario where observations are assimilated is the CONV scenario, in which conventional, mostly surface observations from e.g. weather stations and airports are assimilated into the model. As a result, most of the data assimilation analysis increments are near the surface. The near surface model temperatures at the lowest model level are shown in figure 5.1. Figure 5.1a shows the model state before observations are assimilated, previously defined as the first guess (FG). The surface cold front as mentioned in section 4.2.1 is clearly visible, extending from the northern tip of France until north of the Netherlands. Especially on the North Sea a sharp temperature gradient can be observed. The optimal model state (analysis) resulting from the assimilation of the conventional observations is shown in figure 5.1b and the corresponding differences with the first guess (FG) in figure 5.1c. Additionally, the locations of temperature observations that have been assimilated are shown, where the dot symbols indicate SYNOP and the squares METAR locations respectively. The biggest change to the model temperature field is made near the coast of north France, where apparently the cold front is positioned slightly too far east in the FG according to the observations. Moreover, throughout the southern part of the model domain (latitudes $< 52°N$), with the exception of a few stations near the Belgian coast and France, temperatures are underestimated by the model for this time. For the Netherlands, temperatures are mostly overestimated in this case regarding the decrease of model temperature in the data assimilation analysis.



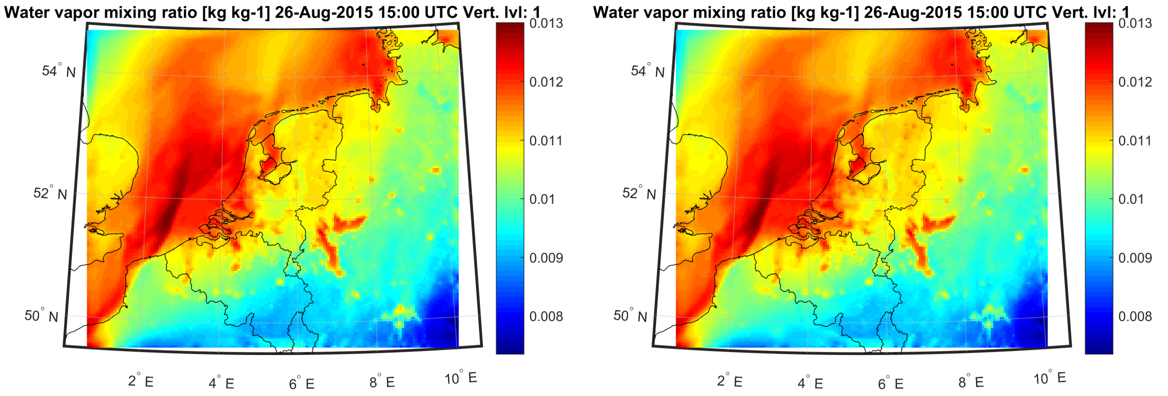
(a) First guess (b) Data assimilation analysis



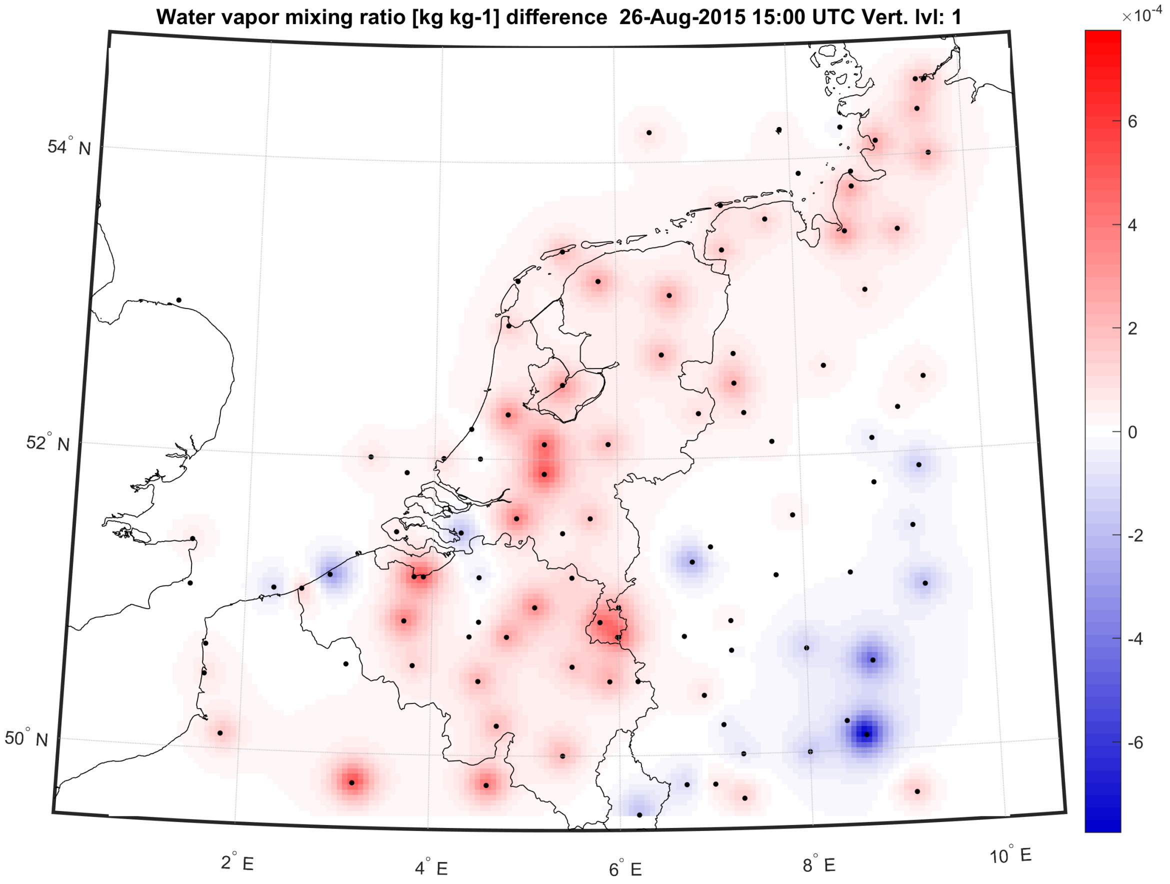
(c) Difference analysis and first guess. Dots denote SYNOP and squares METAR observation locations.

Figure 5.1: Horizontal model level 1 temperatures in the CONV scenario at 26-08-2015 15 UTC

Similar as for temperature, the assimilation makes changes to the water vapour mixing ratio where observations are acquired. The FG and DA analysis of the lowest horizontal model level are shown in figure 5.2a and 5.2b, respectively. The highest concentrations of water vapour are evidently just behind the cold front, where rain forms at higher altitudes. When the precipitation falls down and reaches the surface again, a part evaporates, adding to the near surface moisture on the North Sea where it is already relatively moist. The water vapour innovations introduced by DA have been plot in figure 5.2c. The dot symbols denote SYNOP stations again. As indicated by figure 5.2c, SYNOP measurements are the only source of moisture observations at the surface for this time. Judging from the increments, generally the humidity is underestimated as the major part of the innovations are positive, except for some areas in Germany and some single stations near the Belgian coast. The surface humidity increments provide more moisture that could precipitate later, possibly intensifying the rainfall rate. However, the moisture increments are not that big in magnitude, given that the maximum increments of near 6 $kg\,kg^{-1}$ are roughly 5% of the initial field. It is also noticed that the so called radius of influence of an observation is not extensive and thus observation increments will not affect large areas. The primary reason for this is the downscaling of the length scale of the moisture control variable (refer to table 2.3 for an overview of the applied scaling). However although the scaling factor for temperature and humidity are the same, moisture increments are still significantly smaller in extent, which is explained by the smaller natural autocorrelation length of moisture derived from the climatology on which the model background covariance matrix is based. Keeping these notions in mind no really significant changes are expected in the precipitation forecast, other than slight tweaks in intensity or timing.

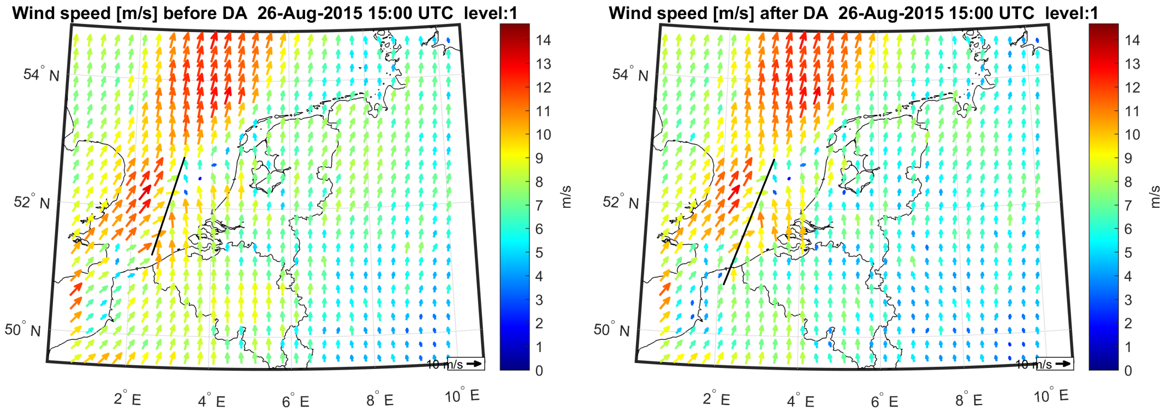


(a) First guess (b) Data assimilation analysis



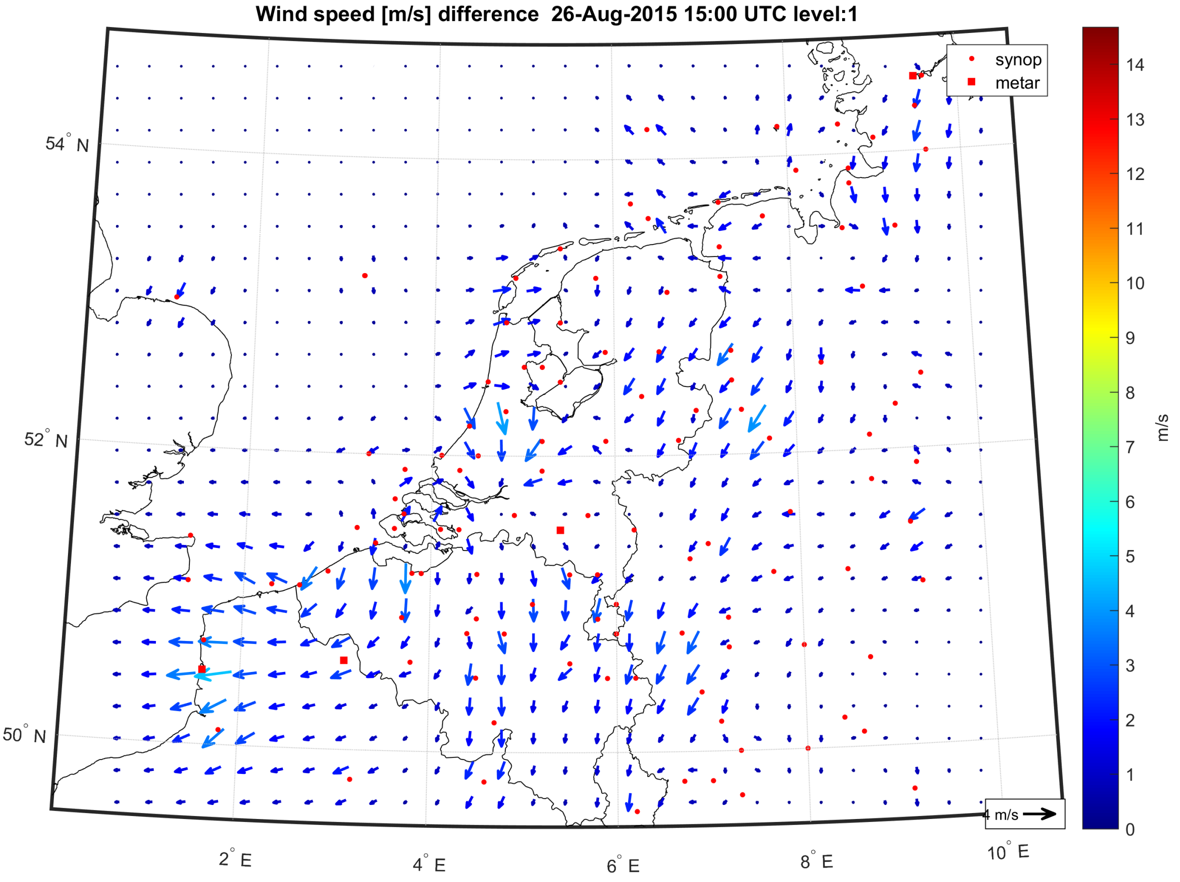
(c) Difference analysis and first guess. Dots denote SYNOP locations.

Figure 5.2: Horizontal model level 1 horizontal water vapour mixing ratio in the CONV scenario at 26-08-2015 15 UTC



(a) First guess

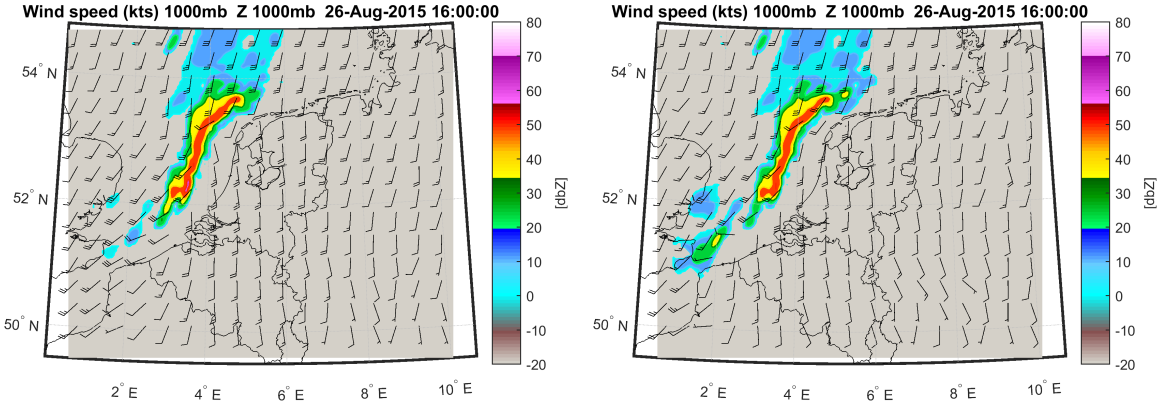
(b) Data assimilation analysis



(c) Difference analysis and first guess

Figure 5.3: Horizontal model level 1 horizontal UV wind speed in the CONV scenario at 26-08-2015 15 UTC

Besides temperature and humidity the assimilation of surface observations also results in changes to the wind speed and direction, expressed in terms of the zonal and meridional wind components U and V. The near surface model wind speeds of the lowest model level have been plot in figure 5.3a. The area on the North Sea just north of the Belgian coast where the squall line is forming (as indicated by the high moisture areas in 5.2a) can be recognised by a line of converging winds as depicted in figure 5.3a. The assimilation of surface wind observations from SYNOP and METAR results in a model wind field as shown in figure 5.3b. The analysis increments in figure 5.3c show how winds are adjusted locally to the observations. The maximum wind speed increment is about 5 ms^{-1} in the northern part of France. The westerly increments here result in a more southern wind in the DA analysis as compared to the SW winds in FG in figure. This extends the line of convergence, as shown in figure 5.3a. The larger area of surface wind convergence results in an elongated squall line. Comparing the model state of the CTRL and CONV scenarios one hour later in the assimilation cycling window at 16 UTC, which are shown in figure 5.4, a new trailing area at the end of the squall line is formed in the forecast where conventional observations are assimilated. The area is not elongated as much as the radar composite at 16 UTC in figure 4.21a observes, but it definitely improves the accuracy of the overall precipitation forecast and shows the potential effect of surface observations on precipitation formation.



(a) Scenario 1: CTRL

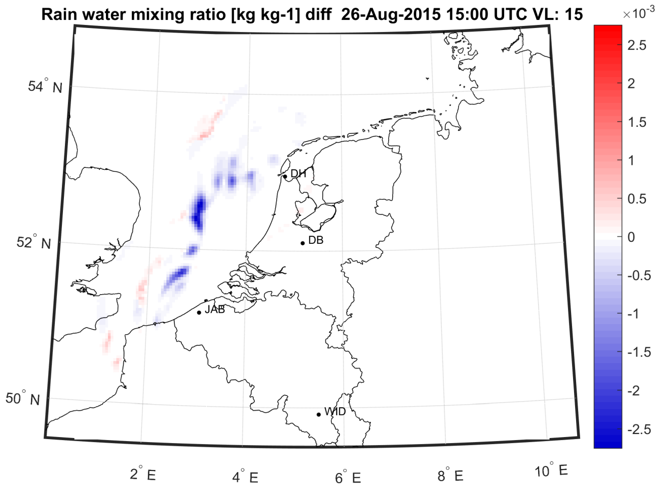
(b) Scenario 2: CONV

Figure 5.4: Model wind fields with reflectivity overlay at 26-08-2015 16 UTC.

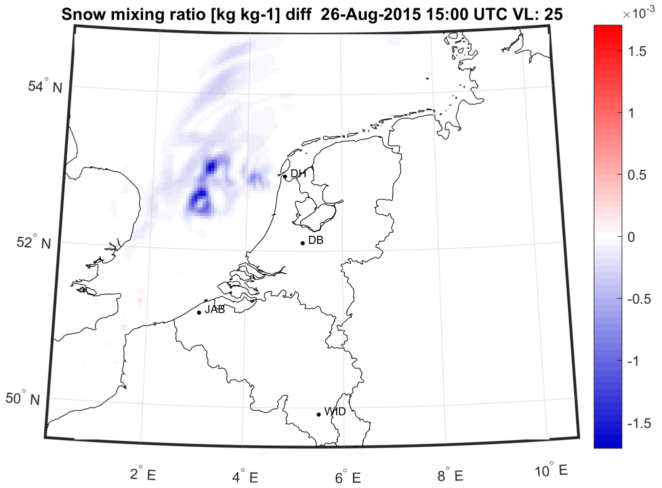
3. CZQR

In the third scenario CZQR of table 4.1 radar retrieved mixing ratios of hydrometeors like rain, snow and graupel are assimilated as described in section 2.5.3. Figure 5.5 shows DA analysis analysis increments for rain water, snow and graupel mixing ratio at 15 UTC, resulting from the first assimilation. The FG and DA analysis are omitted as the areas in figure 5.5 already highlight the areas of interest with significant nonzero mixing ratios in the FG. For rain water the 15th model level ($1.8\ km$ geopotential) is plot, while for snow and graupel the 25th model level ($4.8\ km$ geopotential) is shown, as only rain water is assimilated at temperatures above five degrees, while only snow and graupel are assimilated at $T < -5°C$ (see equation 2.78) with a temperature weighted mixture in between. Also, at lower model levels like 1-10 ($10 - 900\ m$ geopotential), no significant changes can be made by the assimilation of radar observations since this precipitation is too far away and too low for the radar to be detected. The minimal height at which increments can be made is logically equal to the beam height of the lowest scan angle, which is a function of distance (recall). In this case the middle of the squall line is reached by the lowest scans of the radars at Den Helder and Jabbeke at model level 15, meaning that no precipitation below this level at this location can be added nor removed. In figure 5.5a the middle of the squall line can be recognized by the small area of decreased mixing ratio on 52° latitude. Actually, a large part of the rainwater mixing ratio along the forming squall line is removed, since the model formation of precipitation is too far east. Also a lot of the model snow that is present at higher altitude mainly at the leading area of the squall line around $53°N\ 3°E$, in this case for example the 25th model level, is suppressed by the radar data assimilation. Similar as for rain, the graupel that exists in higher levels of the squall line is also mitigated. Some small positive increments are made above the UK and along the Dover Strait but are relatively small compared to the negative increments of the squall line. The main reason for this is the fact that the squall line has initiated slightly earlier in the model and thus simulated reflectivity ($50 + dBZ$) and mixing ratios are significantly higher than the reflectivity and the retrieved mixing ratios that the radar observe.

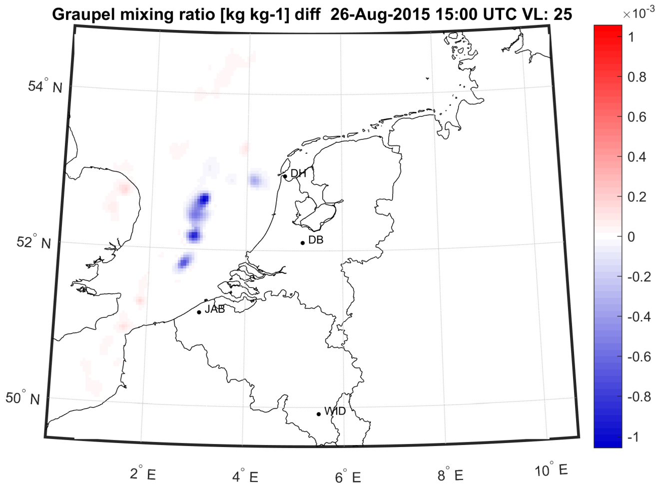
Note that the areas where increments to mixing ratios are made are the intersection of radar beams with the model. In this case with an abundance of negative increments, the areas of increment are the intersections of the radar beam with model precipitation where the radar does not measure any significant reflectivity. Logically, the increments thus appear circular in the analysis, as is perhaps best illustrated by the snow mixing ratio changes in figure 5.5b. The increments, especially further away from the radar where only one radar (e.g. Den Helder for locations far out on the North Sea) intersects the horizontal model level, are on equidistant lines from the radar, creating a wave like pattern. To homogenize the increments one would have to readjust the scaling of the correlation length for hydrometeors again. At the same time however that means degrading



(a) Difference analysis and first guess



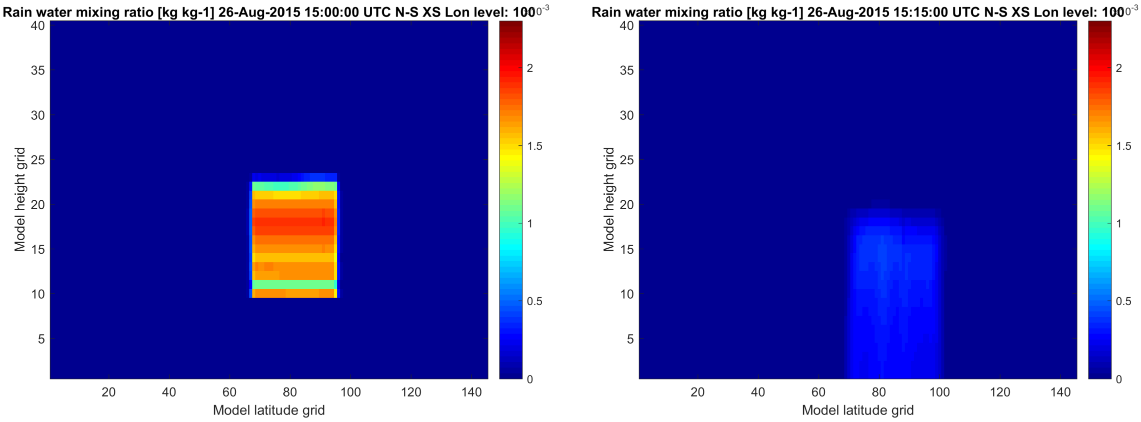
(b) Difference analysis and first guess



(c) Difference analysis and first guess

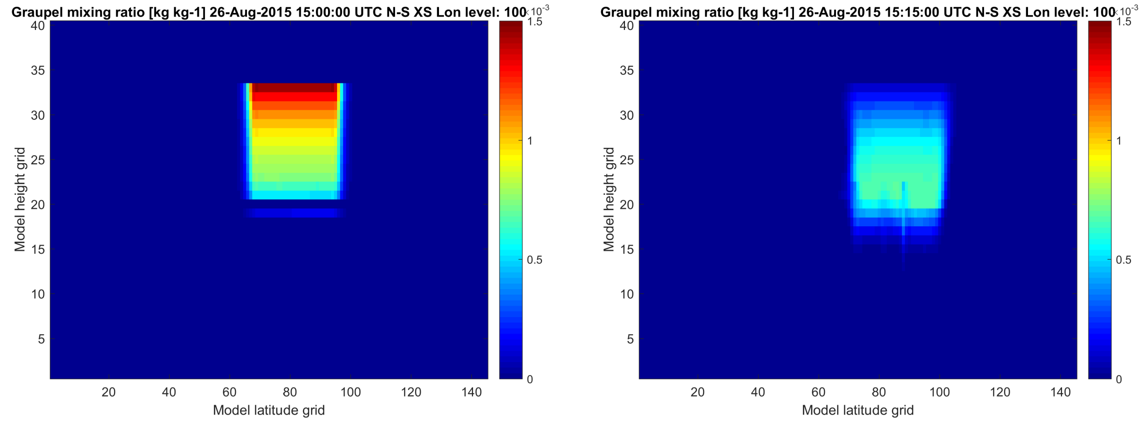
Figure 5.5: Horizontal model level 16 rain water mixing ratio in the CZQR scenario at 26-08-2015 15 UTC.

the ability to add detail to the model hydrometeor field since local, high frequency features cannot be fitted anymore.



(a) Rain water mixing ratio at 15:00 UTC.

(b) Rain water mixing ratio at 15:15 UTC.



(c) Graupel mixing ratio at 15:00 UTC.

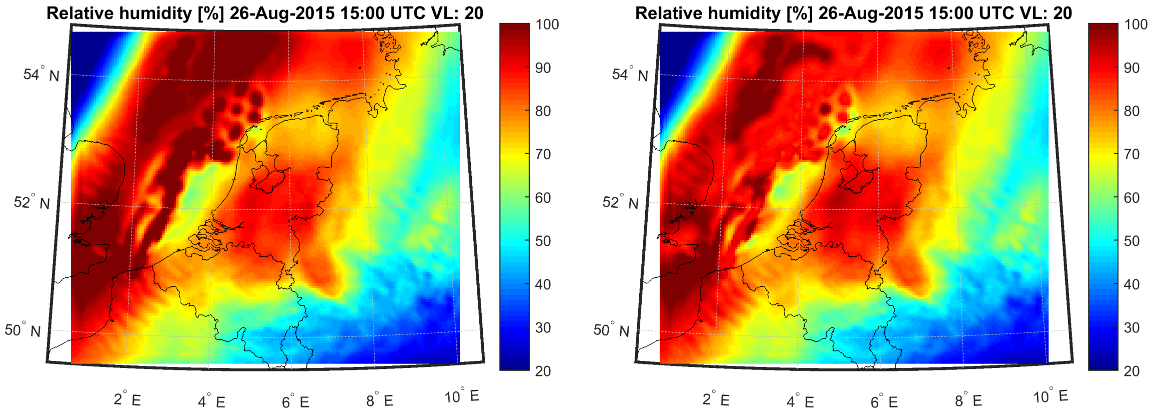
(d) Graupel mixing ratio at 15:15 UTC.

Figure 5.6: Vertical south-north cross sections of mixing ratios for rain and graupel at 15:00 and 15:15 UTC.

Throughout the remainder of the data assimilation window between 15-18 UTC, changes to model hydrometeor fields similar as to figure 5.5 can be observed. That is curious, since one would expect erroneous precipitation to be suppressed after one to a few assimilations. However, the assimilation of hydrometeors only proves to be insufficient to change the process governing the formation of precipitation. Actually, this is not surprising since precipitating hydrometeors can be regarded in a way as the final stage of the precipitation formation process with limited feedback to the what is initiating the process. Assimilation of rainwater does not directly change the temperature other than e.g. through evaporative cooling near the surface. The inclusion of extra liquid water does result in a slight decrease of the virtual (potential) temperature, which in turn slightly reduces buoyancy and thus updraft velocity. Similarly, removing rainwater would even slightly increase buoyancy and strengthen the precipitation formation process. Also, removing existing hydrometeors reduces the rate at which new hydrometeors form, as for example less snow is available for collection to form graupel and similarly less rainwater is present for accretion to graupel as well, so that the amount of graupel reduces. However more importantly, the large amounts of available moisture combined with the updrafts that drive the precipitation formation are not impacted significantly and will continue despite the 'perturbation' of the microphysical processes by radar hydrometeor data assimilation. Analogous to this, in 'dry' areas of the model the process of precipitation formation will not be initiated just by adding hydrometeors as this does not change the amount of water vapour available for precipitation nor does it significantly affect stability of the air. Moreover, in these areas where positive mixing ratio increments are made, the assimilated hydrometeors (rain, snow, graupel) are relatively large particles and thus their large fall speed assigned by the microphysical parametrization will make them fall down quite rapidly, depending on the altitude at which they are formed. A single experiment has been carried out where in a vertical column between $1-10$ km a hypothetical reflectivity of 50 dBZ is assimilated into the model state at 15 UTC. In figure 5.6 this is illustrated with a north-south vertical cross section of rain water and graupel mixing ratio, for both the assimilation time at 15 UTC and 15 minutes into the subsequent forecast. Obviously the rain water has been almost completely drained after 15 minutes (see figure 5.6b), with some residual from melting of ice at higher levels. Similarly, the graupel concentration dilutes and precipitates in the 15 minutes after data assimilation as shown in figure 5.6d. Although from figure 5.6 it might seem that rainwater precipitates much faster, bear in mind that the vertical pressure level spacing is much denser for lower levels near the surface. For example, graupel falling from level 32 to 24 has fallen about 4.4 km while rain has to descend 2.5 km from level 18 to 1 to reach the surface. So, in essence the information that is added by the hydrometeor assimilation to the model trough assimilation is lost once the assimilated hydrometeors have precipitated down. The model state is then similar to the pre assimilation state, except for the changes resulting from the remaining (surface) observations that are assimilated.

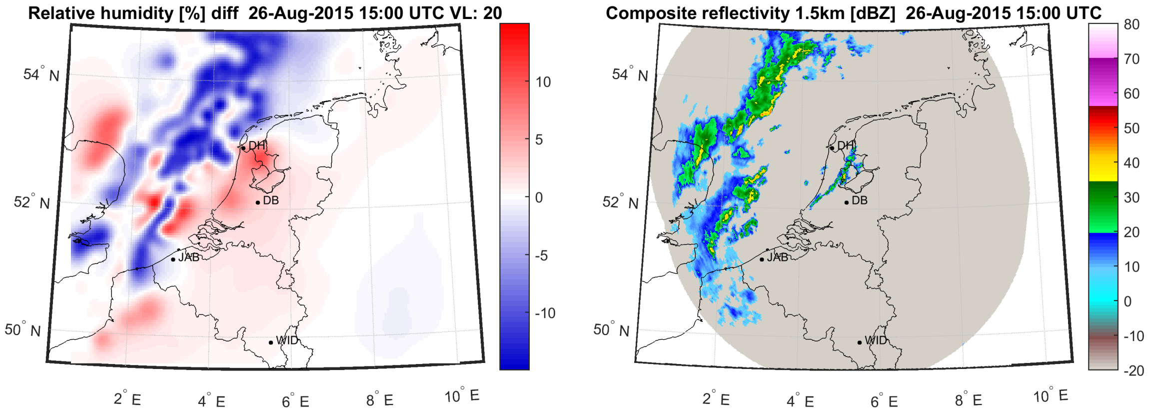
4. CZQV

Instead of assimilation of hydrometeors, in the fourth scenario CZQV only RH observations derived from radar are assimilated, besides the conventional surface observations. The procedure to derive moisture observations from radar have been explained in section 2.5.4 and a radar humidity PSOT has been carried out in 2.8. Similarly to previous scenarios, the initial model FG, DA analysis and differences are shown in figure 5.7, but now for the relative humidity field at model level 20, with geopotential heights of approximately 3.0 km at sea.



(a) Relative humidity first guess.

(b) Relative humidity DA analysis.



(c) Relative humidity DA increments

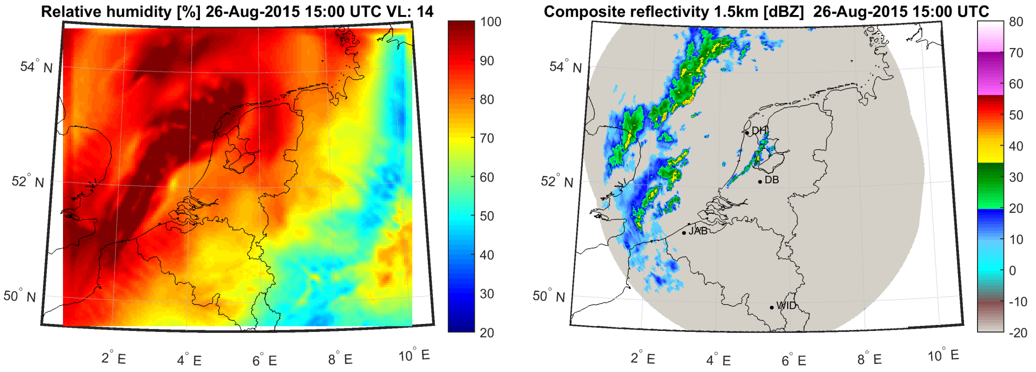
(d) KNMI Radar composite 15 UTC at 1.5km altitude.

Figure 5.7: Horizontal cross section of model relative humidity at model level 20 (3.0 km geopotential height), 15:00 UTC.

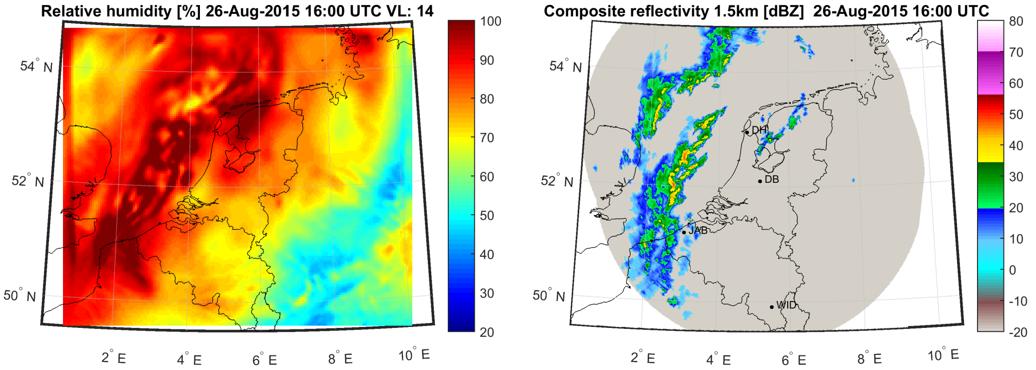
In the FG in figure 5.7a, the original model squall line is easy to distinguish over the middle North Sea, with a clear line of saturated air parallel to the Dutch coast. However, as can be seen by the KNMI radar composite image (which is valid for 1.5km altitude and does not include the Belgian radars at JAB and WID of which data is also assimilated) in figure 5.7d, in reality the squall line has not developed quite as much as in the model.

Consequently, large areas of the squall line are suppressed by assimilation of assumed sub saturated humidity 'observations'. For example in the northern leading part with some visible individual plumes humidity is reduced significantly with as much as 15% and has homogenised with the surrounding relative humidity. Only the areas where forecast precipitation is also measured with significant radar reflections (in this case $Z > 20\ dBZ$, see section 2.5.4) the RH remains near saturation of 100%, as indicated by the remaining areas of high RH in figure 5.7b in the (north)west part of the domain. Here, the assimilated RH observation (88% or 95%, depending on the reflectivity) is really close to the model RH so no significant changes can be observed. In areas where model relative humidity is not near saturation where however a reflectivity $Z > 20\ dBZ$ is measured, significant increases in relative humidity are made. For example above the middle of the Netherlands in between the DH and DB radars a small detached line of reflections is observed, resulting in a moistening of this area. Zooming in at the squall line, near $52°N$ latitude both west and east of the model squall line relatively large increments are made according to figure 5.7c. To the west because significant reflections of $Z > 35\ dBZ$ indicated by the radar result in an assimilation of 95% RH with respect to the 80% background and to the east of the squall line not because the radar measures a reflection but due to the adjacent model squall line where sub saturated RH of 90% is assimilated to suppress erroneous precipitation. In the warm area preceding the model squall line just to the east RH in the FG is so low (around 65%) that the autocorrelation of the suppressing 90% RH observation nearby actually causes a significant RH increment. This shows the importance of the variance length scale in producing realistic increments. In case of a narrow squall line it is essential to make relatively local adjustments to be able to reposition the sharp line realistically without unnecessary smoothing the environment. This does require a dense grid of observations as the radar provides, otherwise only a very limited area can be changed. Mind however that the dense radar observations are only at hand at altitudes above the first scan elevation. For example the precipitation in the west of the domain above mainland UK only results in noticeable moisture increments at model level 20 (as shown in figure 5.7c) and onwards since at this location the lowest reflection of all radars is at roughly 3 km for this level. The positive humidity increments in the north of France result from reflections obtained by the JAB radar which is not visible in the KNMI composite because it is overshot and partly out of range of the Dutch radars. In general comparing the FG 5.7a, DA analysis 5.7b and the radar composite 5.7d, the assimilation of humidity based on radar reflection better confines the areas of near saturation humidity to the locations where radar reflections are measured.

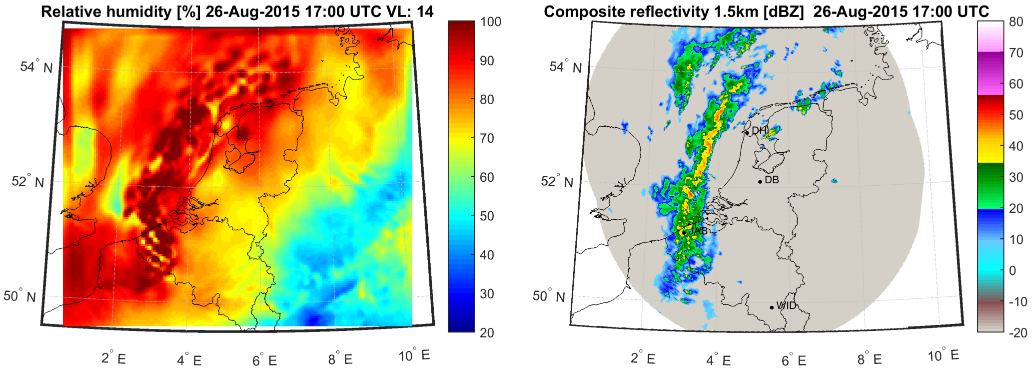
It is worthwhile examining how well data assimilation is able to change the dynamic process of precipitation formation. Figure 5.7 nicely showed how assimilation of radar observations is able to positively adjust model humidity towards observations at a certain point in time. However, equally important is how the model deals with the DA analysis in the subsequent forecast after data assimilation. A known limitation of 3DVar is its inability to really impact dynamical processes since observations are only assimilated at a single time. As mentioned in section 2.2, more advanced data assimilation methods like 4DVar use the NWP model itself as an additional constraint to fit observations to. That is, the optimal DA analysis is the model state of which the subsequent forecast trajectory best fits model observations at later times, not just at a single time as with 3DVar. Recall that the data assimilation cycling approach has been employed as a computationally less demanding alternative to 4DVar. It is interesting to see whether or not the data assimilation cycling is able create new sustained areas of precipitation, instead of the short lived precipitation from hydrometeor assimilation in the previous CZQR scenario. For the formation of rain and other types of precipitation favourable conditions have to exist, of which moist, near saturated air is essential. To examine how well the model humidity represents the actual conditions, the model FG of the relative humidity is shown alongside radar reflectivity composite images, starting from the assimilation window at 15 UTC up to 18 UTC, with a 1 hour interval. It is preferred to show the FG rather than the DA analysis since it is the 'unaltered' model state which better shows the tendency of the model than the DA analysis since there the model shortcomings are masked again by assimilation of observations. Vertical level 14 is plot since in a large part of the domain mainly at sea the geopotential height is close to the altitude for which the KNMI radar composite is valid: 1.5 km. Although model relative humidity need not be saturated at locations where a radar receives a reflection since precipitation might form at higher altitudes and fall to lower altitudes, at least relatively high RH is expected at these areas relative to the environment. At 15 UTC, high model humidity does not really coincide with areas with strong radar reflections when comparing figures 5.8a and 5.8b. As noted before, right before the first data assimilation time at 15 UTC the moist area of the squall line is positioned too far east compared to the radar images. One hour and 4 assimilations later at 16 UTC the RH in figure 5.8c shows better agreement with radar observations above north France, however still some areas of relatively high humidity are present above the north of the Netherlands which seems unlikely given the limited amount of precipitation there (figure 5.8d).



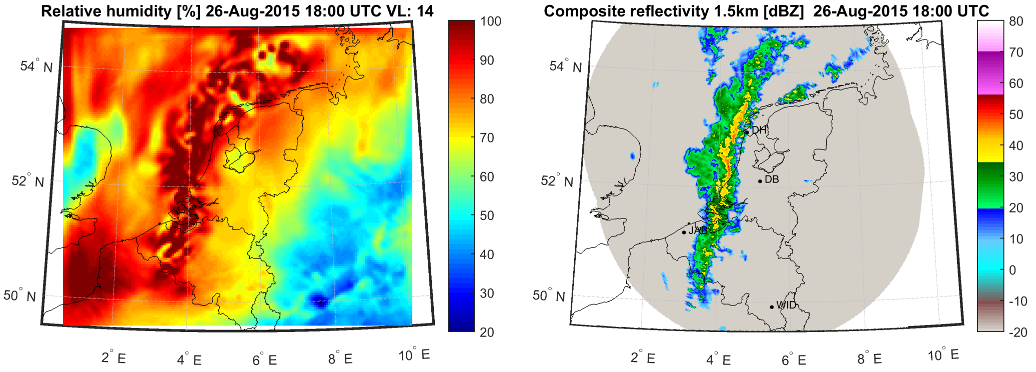
(a) FG 15 UTC (b) Radar composite 15 UTC



(c) FG 16 UTC (d) Radar composite 16 UTC



(e) FG 17 UTC (f) Radar composite 17 UTC



(g) FG 18 UTC (h) Radar composite 18 UTC

Figure 5.8: Model level 20 RH first guess together with radar composites at 15, 16, 17 and 18 UTC.

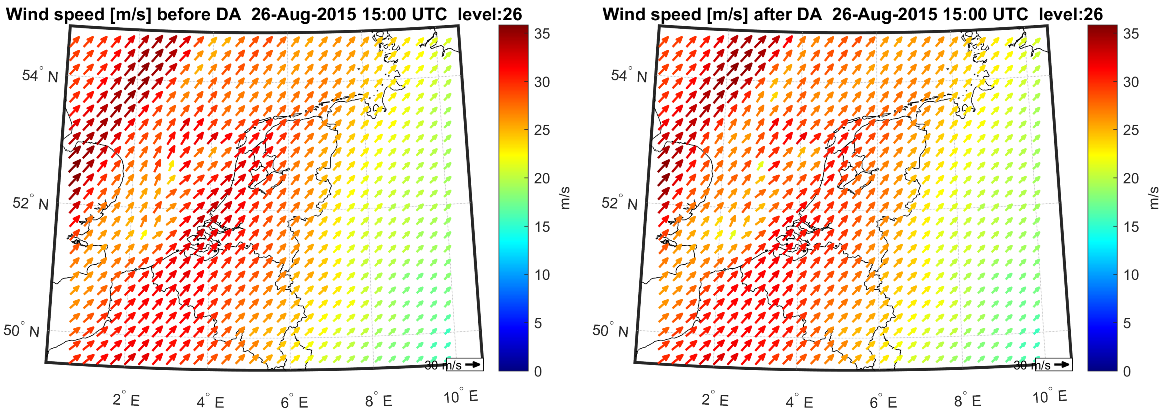
Another hour later at 17 UTC (figure 5.8e) moist areas seem to resemble the reflectivity pattern (5.8f) quite well. The positioning of the squall line is more on point and the extent of the area is now more in line with the radar. Finally, at the last assimilation time at 18 UTC, areas of relative humidity near saturation (5.8g) are even more lined up from north to south and correspond relatively well with the position and size of the squall line in the radar composite of figure 5.8h. Judging from these observations, it seems that the model is able to adjust the squall line to the observations through radar data assimilation, although it requires a decent amount of assimilation cycles. Finally it is noted that the areas of high humidity become quite scattered throughout the cycling period. The localized cells are mainly a consequence of the spatial length scale of the autocovariance, which was scaled down to allow for detailed increments, especially in the case of local convection. In this case of a squall line, the extent in the direction along the squall line, normal to the direction of propagation is relatively large which requires a longer length scale to make the line of RH smoother. But also since the autocovariance is modelled to be isotropic, accurate repositioning of a smooth squall line would require a different description of the autocovariance that better represents the case specific anisotropic variance, e.g. through ensemble/hybrid assimilation.

5. CZRV

In the fifth scenario the last type of radar observations is assimilated: radar radial velocity. In section 2.6 it is addressed how radar radial velocity is assimilated as well as some characteristic features of this type of observation, like the aliasing of radial velocity. As for the previous scenarios, it will be briefly discussed how the assimilation of radial velocity affects the model wind field.

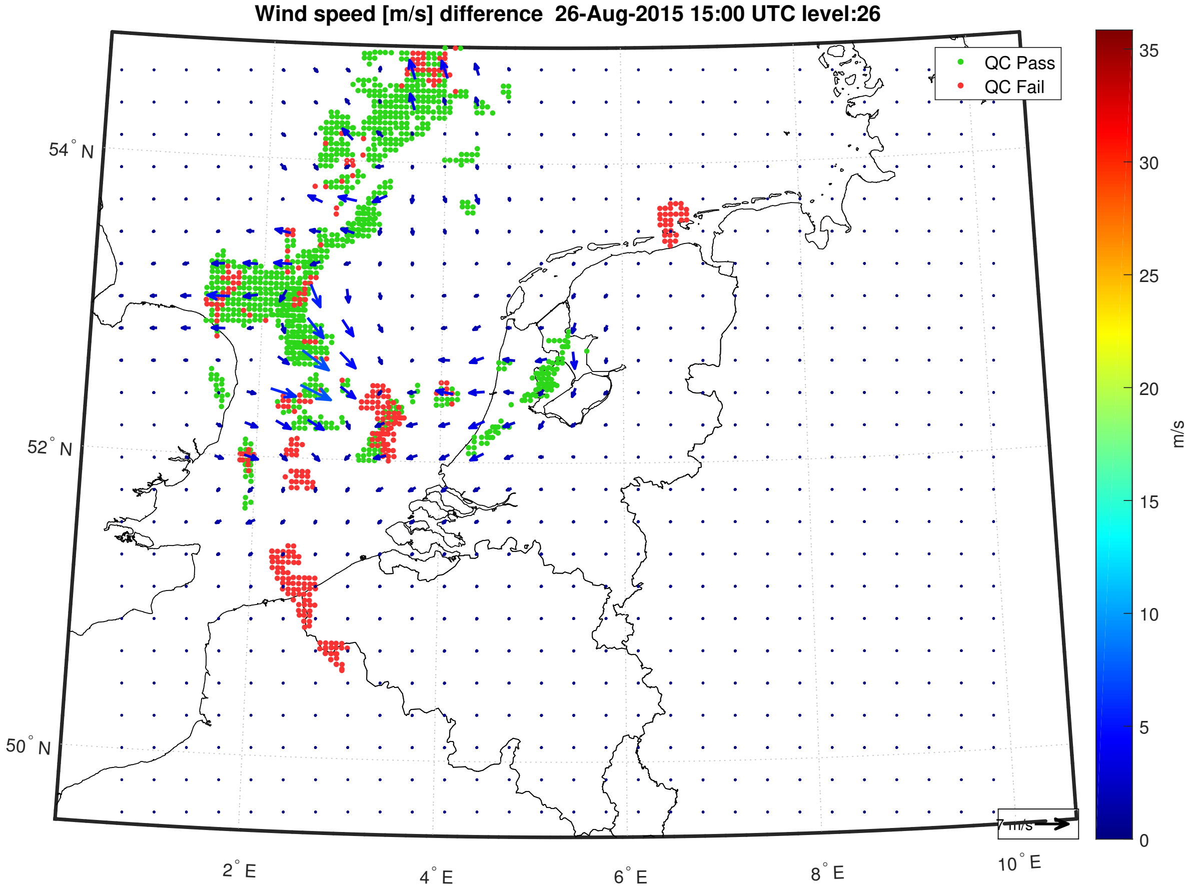
An illustration of the changes that assimilation of radial velocity results in is given figure 5.9. Similar to the wind increments resulting from assimilation of surface observations in the CONV scenario as shown in figure 5.3, the model DA increments are highlighted in figure 5.9c as the difference between the model FG and DA analysis. Vertical level 26 is chosen as it intersects with large areas of radial velocity of scans 3-5 of the DB and DH radar. These scans have been included in figure 5.10, where the left column shows the filtered scans for DB and the right for DH. For example, a large portion of the radial velocity measurements far on the North Sea acquired by the DH radar intersects the level 26 at a geopotential height of roughly 5.2 km. Similarly, velocity measurements covering the middle of the Netherlands intersecting this model level come from higher elevation scans (7-8) which are not shown here, but are really similar to the velocity observations in the lower scans in figure 5.10. The observations on the North Sea located between $53-55°N$ show a southwesterly wind, corresponding with the wind field in the FG in figure 5.9a which shows a SW oriented jet stream. These observations pass the maximum innovation check $O-B<3\sigma$ (see equation 3.1) for radial velocity and its locations are plotted with green dots on top of the analysis wind increments in figure 5.9c. The maximum innovation threshold has been set to 3σ because experiments with the default of 5σ produced unrealistic local stop and reversal of the jet stream wind at 500 hPa as a result of noisy and aliased radial velocity observations. Besides the radar reflections on the 26th model level the observation locations at slightly different heights intersecting adjacent model levels are plot as well, because they still influence the wind field increments through the vertical correlation that is described by the eigenvectors of u and v (see the eigenvectors for u and v in figure 2.9). Observations that fail the maximum innovation check are shown in red. Apart from some isolated areas in the northwest of the domain, most of the velocity observations that fail the maximum innovation or quality control (QC) check correspond with the aliased velocity measurements of the radar scans. For example, the velocity observations at the border between France and Belgium in scan 3 and 4 of the DB radar in figure 5.10a and 5.10c appear to be aliased regarding the strong SW flow in the FG of figure 5.9a. Also the positive radial velocity measurements around $52°N$ in scans 3-5 of DH erroneously indicate a strong NE wind component that is obviously aliased regarding the dominating SW wind flow in the model. These observations are discarded since they produce huge innovations with respect to the FG (logically close to $2v_{\max}$ when the dealiased observations correspond with the model wind field), while for a relatively large velocity standard deviation (also referred to as Doppler velocity spectrum width) of $\sigma_D = 4\ m\,s^{-1}$ the maximum allowed innovation is 12 $m\,s^{-1}$. Thus the aliased regions above the southern part of the North Sea fail the maximum innovation check and therefore do not affect the model wind field.

The observations shown in green in figure 5.9c that do pass the check cause some adjustments to the wind field, with a maximum wind speed magnitude increase of 7.3 $m\,s^{-1}$ on the middle of the North Sea at the southern part of the large area of radial velocity measurements acquired by scan 3-4 of both DB and DH radars. Around this area winds are adjusted from a SSW towards a more SW orientation. In general throughout the domain the wind speed magnitude is mostly slightly overestimated in the FG, since most increments have a negative radial component with respect to the FG SW wind direction.



(a) First guess

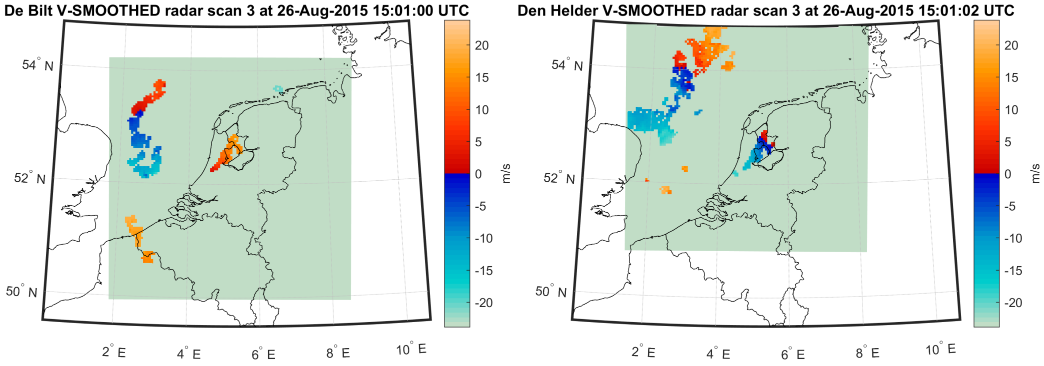
(b) Data assimilation analysis



(c) Difference analysis and first guess. Dots denote observation locations of radial velocity measurements with a height similar to this model level that pass (green) and fail (red) the maximum innovation QC check.

Figure 5.9: Horizontal model level 26 wind speed in the CZRV scenario at 26-08-2015 15 UTC.

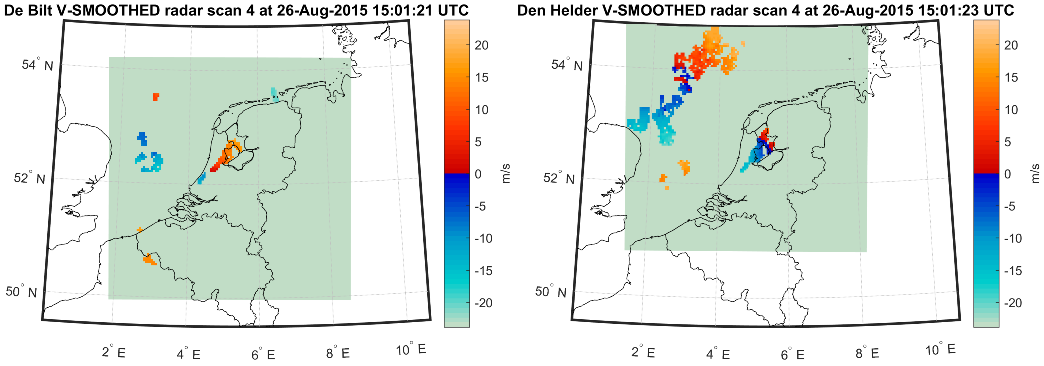
Finally, it is worthwhile comparing the model wind fields of the CZRV scenario with the CONV scenario at the end of the data assimilation cycle at 18 UTC. This is the model state that is used as a starting point for the subsequent verification forecast. In figure 5.11 the wind field at the 10th model level is shown for both the CONV and CZRV scenario. The 10th model level, at which a large part of the domain has a pressure close to 900 mb, is highlighted because in the case of a squall line the mechanical forcing of the near surface wind with which the cold front is approaching largely determines the intensity of convection. In the CONV scenario most of the areas with high $> 20\ dBZ$ reflectivity (as denoted with the black lines) in figure 5.11a can be found north of the Netherlands. The precipitation front coincides with a line of converging winds from north of the


De Bilt V-SMOOTHED radar scan 3 at 26-Aug-2015 15:01:00 UTC

Den Helder V-SMOOTHED radar scan 3 at 26-Aug-2015 15:01:02 UTC

(a) Scan 3, DB.

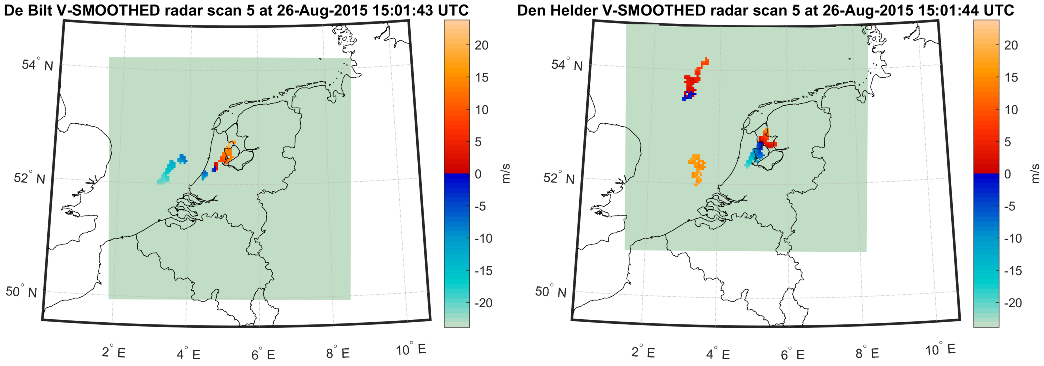
(b) Scan 3, DH.


De Bilt V-SMOOTHED radar scan 4 at 26-Aug-2015 15:01:21 UTC

Den Helder V-SMOOTHED radar scan 4 at 26-Aug-2015 15:01:23 UTC

(c) Scan 4, DB.

(d) Scan 4, DH.


De Bilt V-SMOOTHED radar scan 5 at 26-Aug-2015 15:01:43 UTC

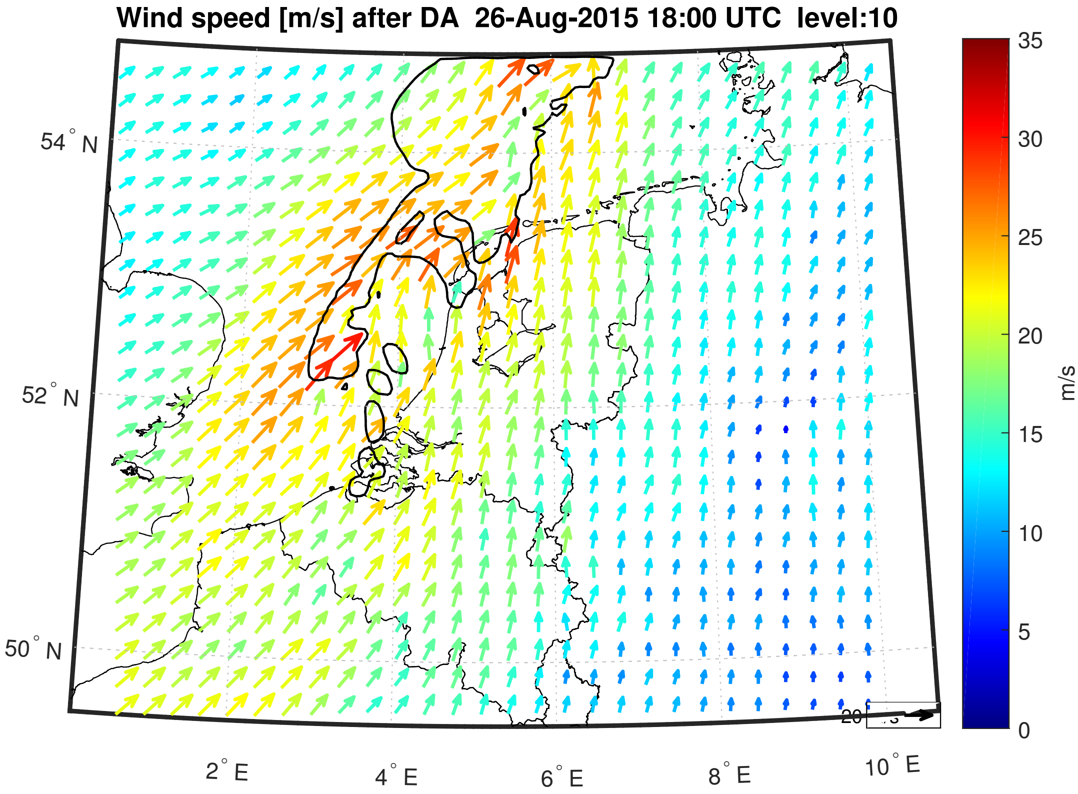
Den Helder V-SMOOTHED radar scan 5 at 26-Aug-2015 15:01:44 UTC

(e) Scan 5, DB.

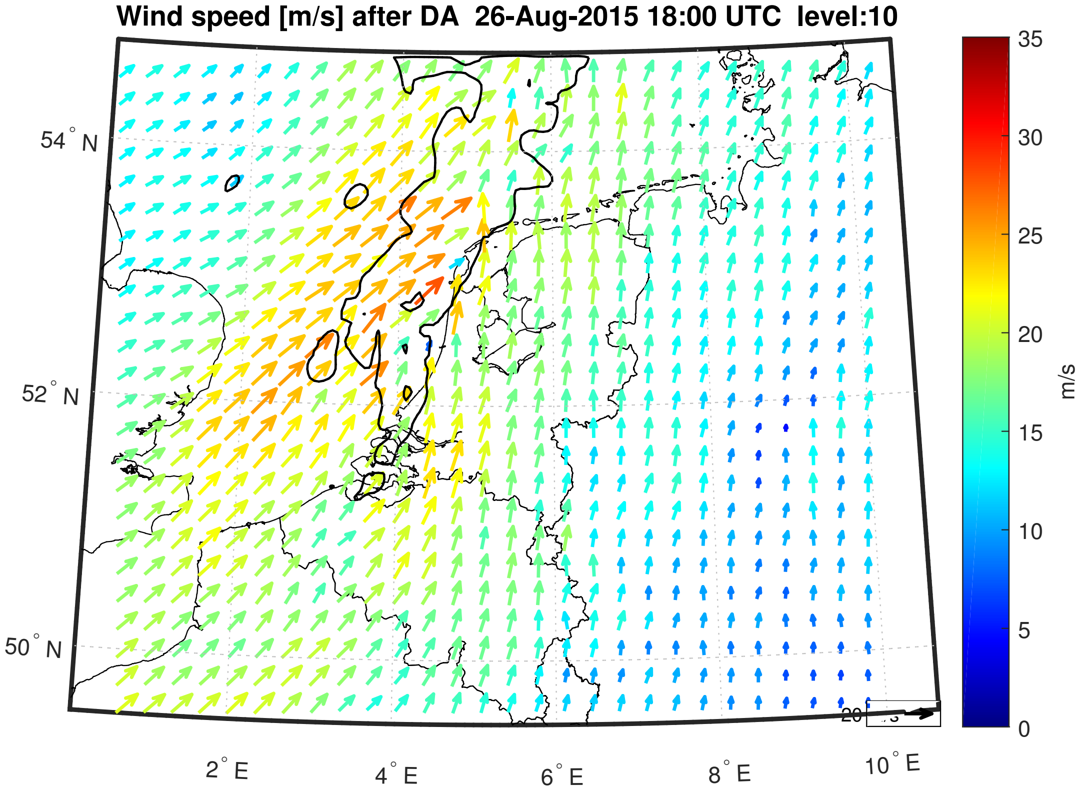
(f) Scan 5, DH.

Figure 5.10: Radar radial velocity measured at KNMI radars at De Bilt (DB) (left) and Den Helder (DH) (right).

domain down to Den Helder. Also, on the North Sea right at around $52.5^{\circ} N$ $4^{\circ} E$ locally an intense cell of convection forms because of the convergence of really strong near 30 ms^{-1} SW winds with slower southerly winds. In contrast, after the three hours of radial velocity assimilation, the wind field in the CZRV scenario as shown in figure 5.11b is adjusted so the line of convergence is moved eastward, more or less parallel to the Dutch coast line. Also, the line of convergence has been slightly extended southwards. Along the convergence line convection is forced, which leads to a more easterly positioned, elongated area of precipitation (figure 5.19b) which corresponds significantly better with radar images, see figure 5.19h.



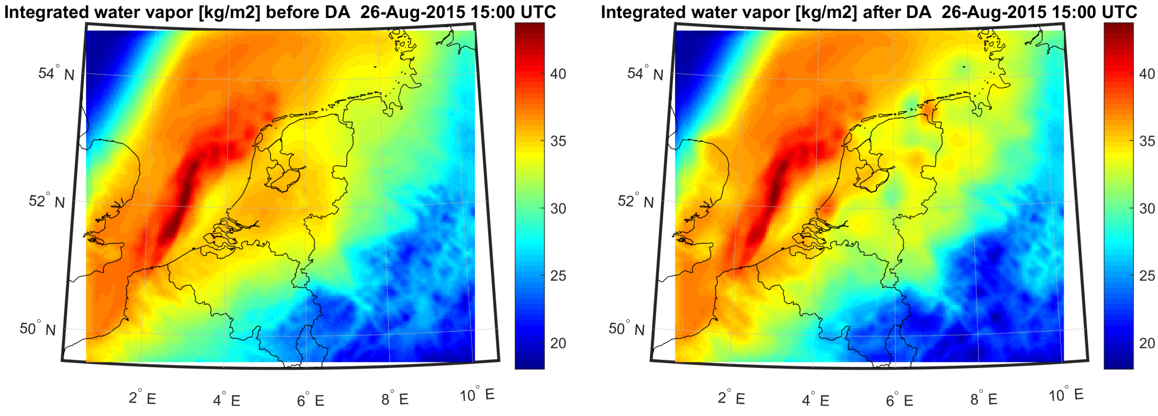
(a) CONV



(b) CZRV

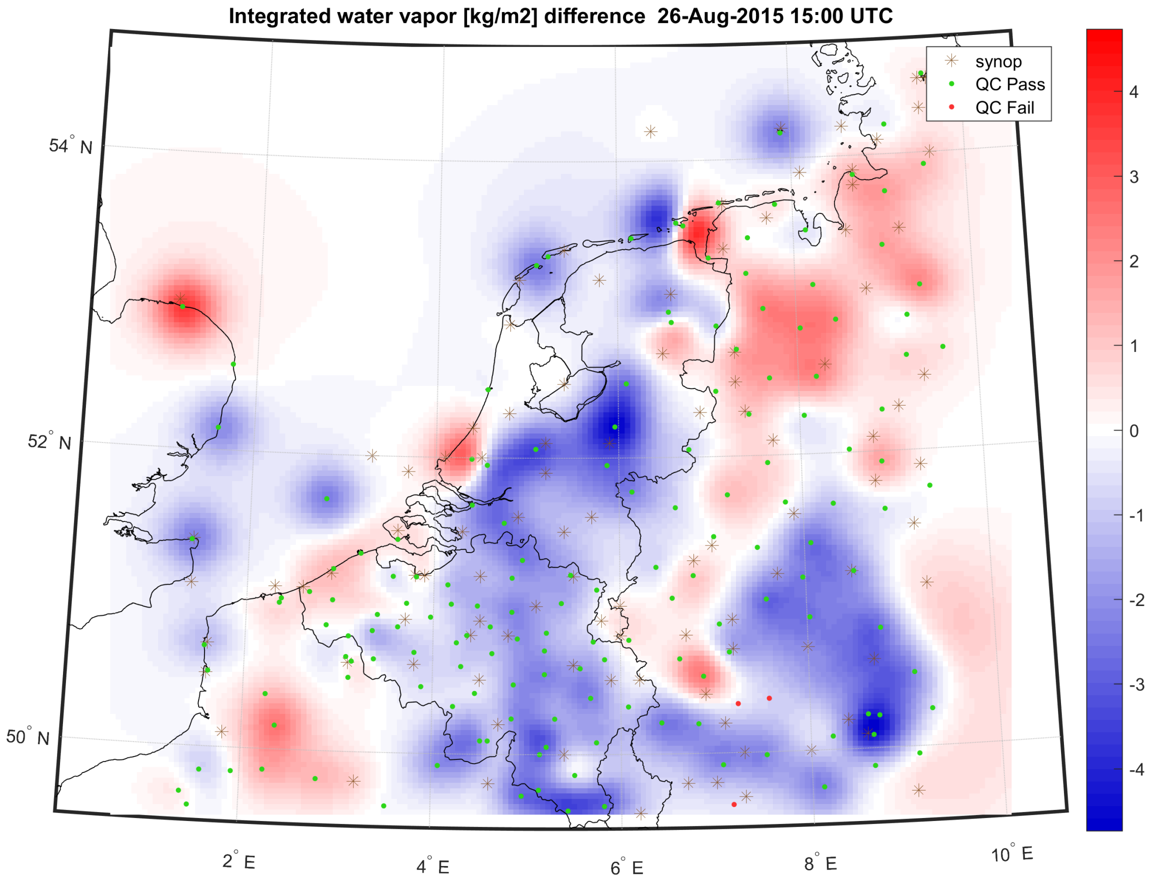
Figure 5.11: Comparison of horizontal model level 10 (900 mb) wind speed at 26-08-2015 18 UTC for the CONV and CZRV scenarios. The black line indicates the contours of 20 dBZ model reflectivity.

6. CGPS



(a) First guess

(b) Data assimilation analysis

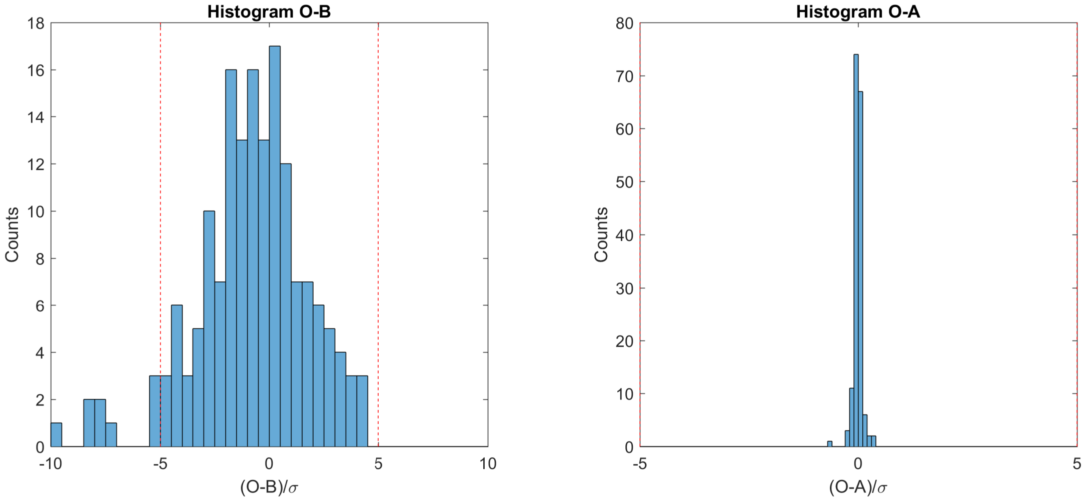


(c) Difference analysis and first guess. Dots denote observation locations of GPS ZTD measurements that pass (green) and fail (red) the maximum innovation QC check. Additionally SYNOP humidity observation locations have been plot.

Figure 5.12: IWV of the model in the CGPS scenario at 26-08-2015 15 UTC.

In the last isolated observation scenario only GPS zenith total delay (ZTD) measurements are assimilated besides the conventional surface observations. Refer back to 2.7 for a description of how ZTD observations are assimilated into the model. As described there, GPS ZTD measurements and more specifically ZWD are measures of the delay of a GPS signal caused by water vapour in the atmosphere. Therefore, it makes sense to look at the IWV of the model before and after the assimilation of ZTD measurements. The FG for the integrated water vapour at 15 UTC is shown in figure 5.12a. The squall line in the model on the North Sea is clearly visible. The DA analysis for the IWV has been included in figure 5.12b, showing a similar state as the FG. The squall line itself is mainly unaltered because there are no GPS stations at sea, except for one measurement location. Similar to the surface based conventional observations, it is a limitation that generally only on land observations are made. That way, radar humidity changes at sea cannot be constrained by GPS measurements. Nevertheless, the observation network on shore is relatively dense, allowing for local changes as illustrated by the DA increments in IWV resulting from ZTD assimilation, shown in figure 5.12c. In a large part of the Netherlands and Belgium the amount of moisture is overestimated, judging from the dominantly negative increments. Along the Dutch and Belgian coast some positive increments can be observed, creating a sharper gradient with the less moist regions of negative increments to the east. The region of moist air west of this gradient corresponds with an elongated band of precipitation visible in the composite radar image at the time of assimilation in figure 5.8b, at approximately $52.5° N\ 5° E$. Also, on the northeast coast of the UK a strong positive IWV increment is made of approximately $4\,kg\,m^{-2}$, more than 10% of the original IWV. In this area also a significant area of reflectivity is present, which makes it likely that this significant moisture increment is justified.

The observation locations of GPS ZTD measurements have been included in the DA increments plot in figure 5.12c. The green points show the locations that pass the maximum innovation test, while the red points correspond to the locations that fail the test. The maximum innovation for GPS ZTD is unaltered and remains 5σ. Histograms in figure 5.13 further illustrate this maximum innovation test. In figure 5.13a, a histogram of O-B shows the distribution of the σ normalized ZTD observation introduced innovations. The observations outside the 5σ interval denoted by the red lines fail the maximum innovation test and are thus ignored. The DA changes result in a model state in which the difference between the observations and the DA analysis (O-A) is minimized, as indicated by the histogram of normalized O-A in figure 5.13b. The histogram of O-A is so small because in this scenario only ZTD and some conventional measurements are assimilated and thus the cost function that is minimized only penalizes model differences with these types of observations. Second, because of the small autocovariance length scale of moisture observations can only be constrained by other observations if the network is dense enough, as is the case in Belgium. In less dense areas data assimilation can just fit the model to ZTD observations without any further constraints from other observations. When other observations are included the relative weight of ZTD measurements in the cost function is smaller so that O-A values will be larger for ZTD.

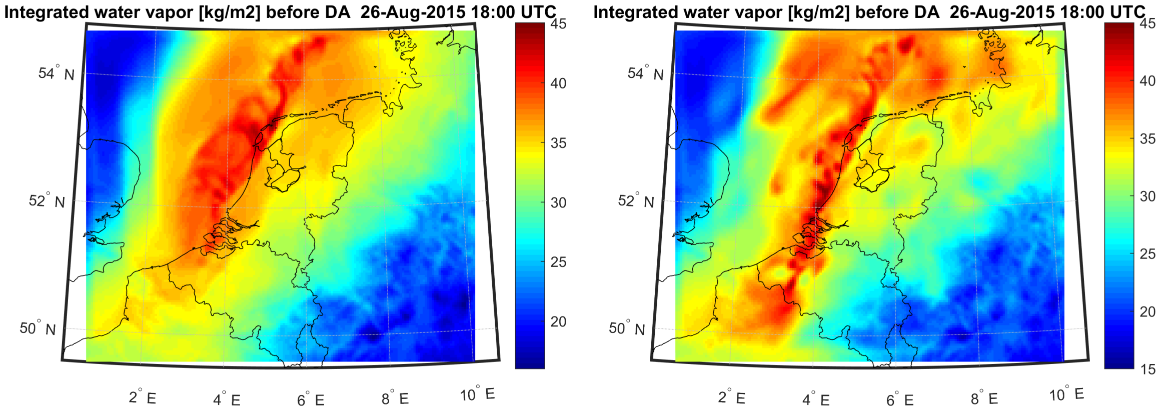


(a) O-B

(b) O-A

Figure 5.13: Histograms of ZTD O-B and O-A, normalized by the error standard deviation σ of each observation. The red lines indicate the maximum innovation (O-B) test boundaries.

Again it is interesting to see how the total humidity has changed during the assimilation cycle of ZTD measurements compared to the CONV scenario. Figure 5.14a shows the IWV for the CONV scenario and figure 5.14b for the CGPS variant. Interestingly enough the assimilation of ZTD measurements results in a IWV field much more similar to the precipitation areas on the radar composite image at 18 UTC in figure 5.8h, without the assimilation of any radar observations. The band of high precipitable water vapour has become much narrower compared to the CONV scenario. Keeping in mind the southwestern wind, the GPS stations upstream of the squall line on the UK coast consistently measure lower ZTD than the model equivalent, resulting in a drying of the moist air west of the model squall line at 15 UTC (compare figure 5.12a and 5.12b again) and onwards. At the same time the GPS stations near the Dutch and Belgian coast cause some moistening at the eastern side of the squall line. Also, stations in Belgium and the north of France indicate larger precipitable water quantities than in the CONV scenario, which is also in correspondence with the fact that the actual squall line on the radar composite at 18 UTC in figure 5.8h also stretches through Belgium into France. The two most northern GPS locations in the UK in figure 5.12c cause a plume of humid air on the North Sea around $53.4° N$ $2.7° E$, producing some significant precipitation as will be shown in the verification.



(a) CONV (b) CGPS

Figure 5.14: Comparison of model IWV at 26-08-2015 18 UTC for the CONV and CGPS scenarios.

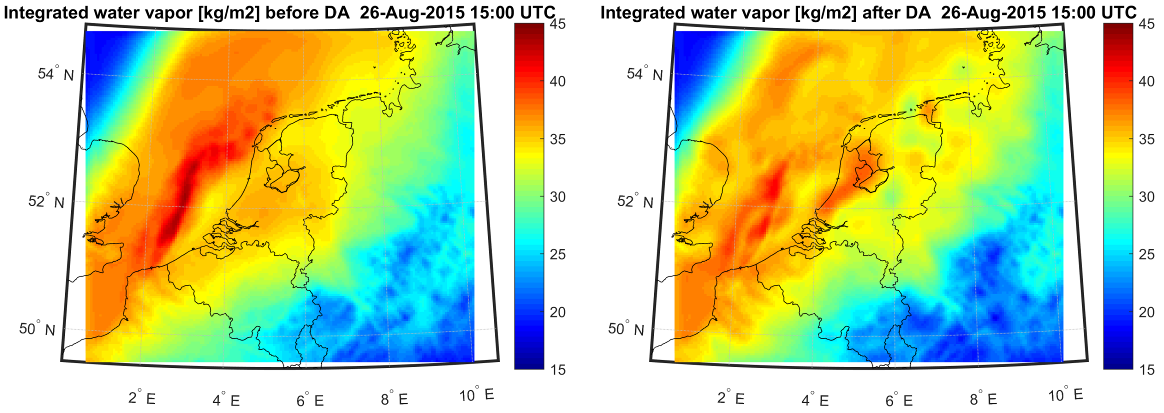
7. ALL

In the final scenario all available observations are being assimilated. Conventional observations provide measurements of temperature, pressure, humidity and wind speed at the surface. Radar reflectivity produces hydrometeor and humidity observations with high spatial resolution that should improve the positioning of precipitating areas. The assimilation of radar radial velocity improves the dynamics driving the intensity of convergence at the squall line front. At the same time, GPS ZTD measurements are supposed to constrain radar humidity observations and provide additional humidity observations in areas where no precipitation (and thus no detectable reflectivity) has formed yet.

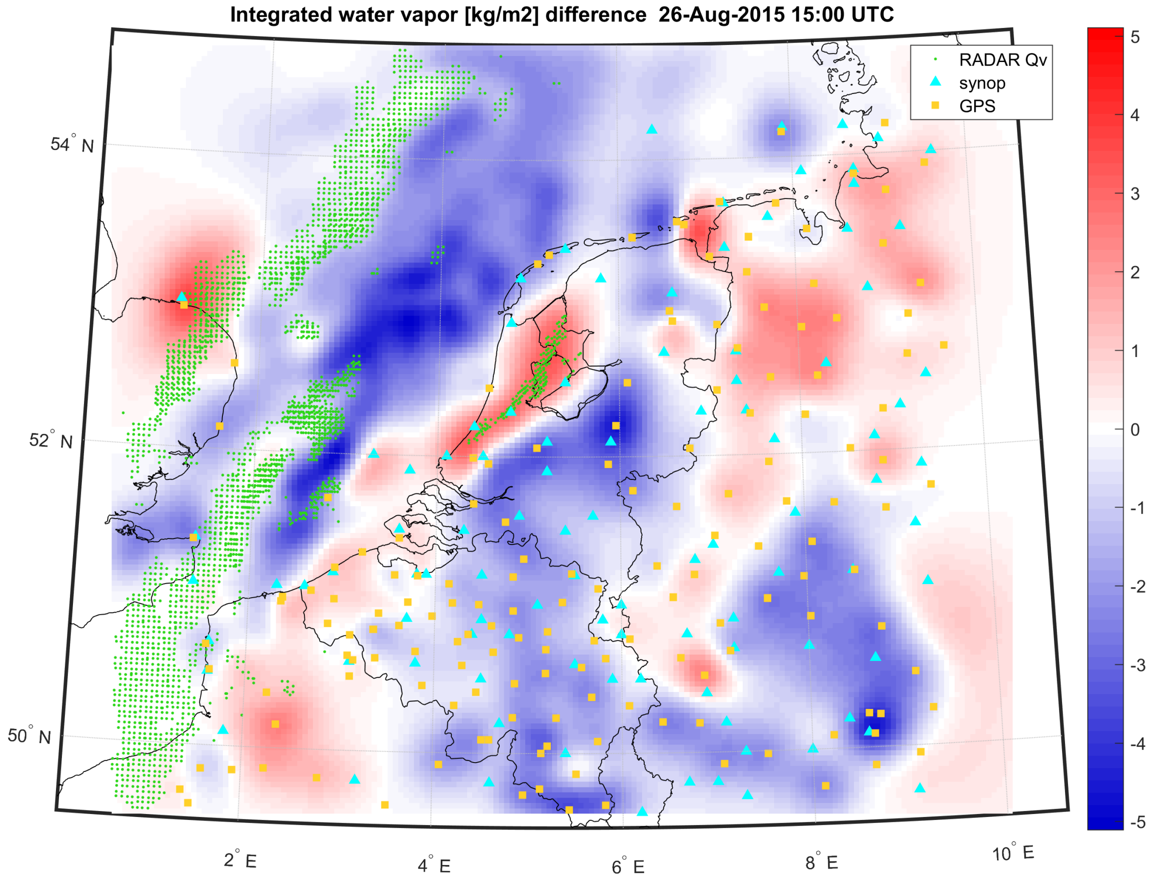
	Synop	Metar	Buoy	GPS	Radar Zqv	Radar Zqr	Radar Zqg	Radar Zqsn	Radar RV
Used	116	4	19	162	24368	25274	43598	43598	3674
Rejected	2	0	0	3	0	0	0	0	1075
Total	118	4	19	165	24368	25274	43598	43598	4749

Table 5.1: Number of assimilated observations at 15 UTC per observation type.

Table 5.1 lists all the observations that are assimilated into the model, for every observation type our source. Radar 'Zqv' refers to the humidity retrieved observations from radar and 'qr', 'qg' and 'qsn' to the other assimilated hydrometeors of rain, graupel and snow respectively. Finally, also radial velocity (RV) observations are assimilated. The number of observations failing the maximum innovation that are then rejected have also been included in the table. For observation sources supplying observations of multiple variables, the amount of available and rejected observations might deviate slightly from the values listed in table 5.1. From table 5.1 it is evident that most conventional observations pass the QC test, except for two SYNOP stations. A relatively large part of the GPS measurements are assimilated as well, considering that ZTD measurements often lack very good precision. For the radar hydrometeors no maximum innovation test exists, since most often it occurs that no hydrometeors exist before assimilation at radar observation locations, yet we would like to be able to assimilate the observations anyway. As radar is a remote sensing instrument with a high spatial resolution, a large quantity of observations is readily available for assimilation. Note that there are less radar humidity measurements than other hydrometeors, primarily because humidity observations are only assimilated at $Z \geqslant 20\ dBZ$. Also, not every reflectivity observation location has a corresponding radial velocity observation. A significant part of the radial velocity observations is rejected because some aliased observations exist in the lower to middle scans, as was demonstrated in section 5.1.1 and specifically in figure 5.10 where radial velocity of the lower scans are shown.



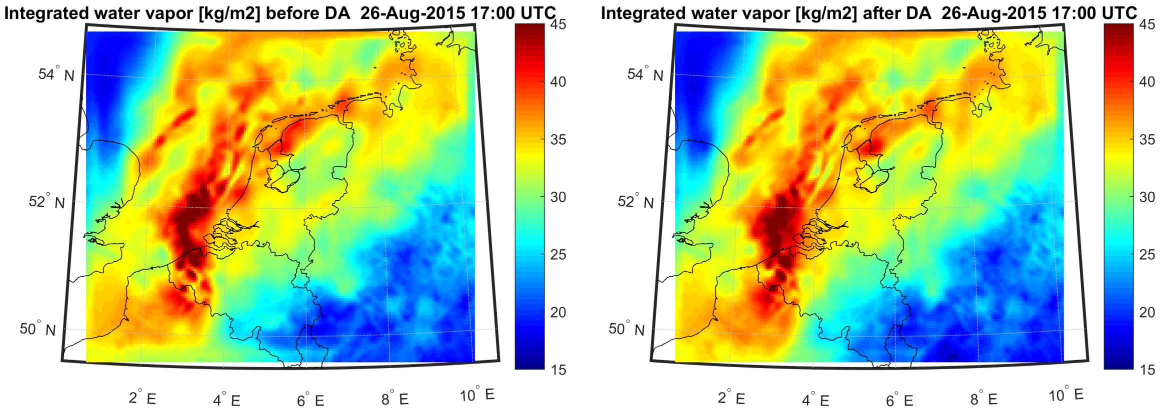
(a) First guess (b) Data assimilation analysis



(c) Difference analysis and first guess. Markers show the observation locations of SYNOP and ZTD measurements. Additionally dots indicate the locations of radar humidity observations where $Z > 20\ dBZ$.

Figure 5.15: IWV of the model in the ALL scenario at 26-08-2015 15 UTC.

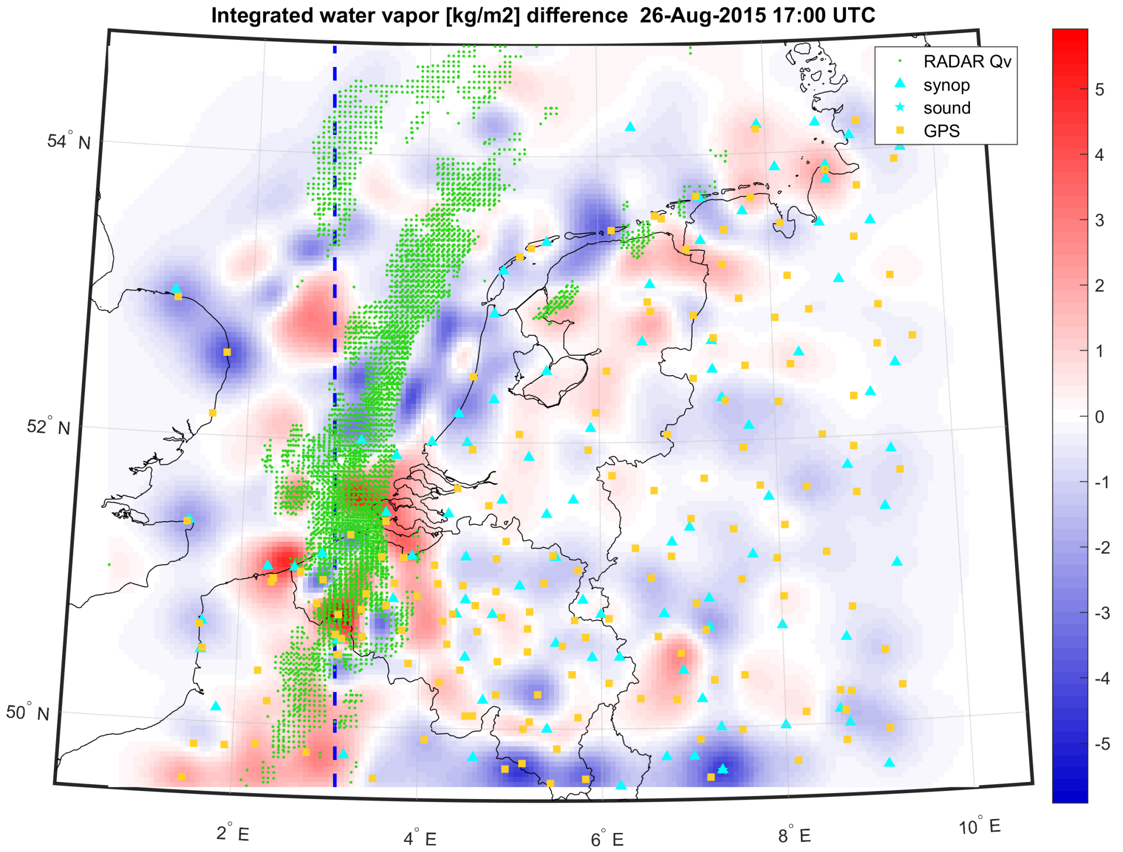
Similarly as for the other scenario's, a map of data assimilation changes together with the FG and DA analysis has been included in figure 5.15. The map shows the increments in IWV, which is influenced by radar derived humidity observations at higher altitudes, at the surface through SYNOP humidity measurements and throughout the whole column by GPS ZTD. Other observation sources listed in 5.1 do not provide humidity observations. The map of increments in figure 5.15 shows that in most areas where radar humidity


Integrated water vapor [kg/m2] before DA 26-Aug-2015 17:00 UTC

Integrated water vapor [kg/m2] after DA 26-Aug-2015 17:00 UTC

(a) First guess

(b) Data assimilation analysis


Integrated water vapor [kg/m2] difference 26-Aug-2015 17:00 UTC

(c) Difference analysis and first guess. Markers show the observation locations of SYNOP, sounding and ZTD measurements. Additionally dots indicate the locations of radar humidity observations where $Z > 20\ dBZ$. The dashed line shows where the cross section in figure 5.17 is made.

Figure 5.16: IWV of the model in the ALL scenario at 26-08-2015 17 UTC.

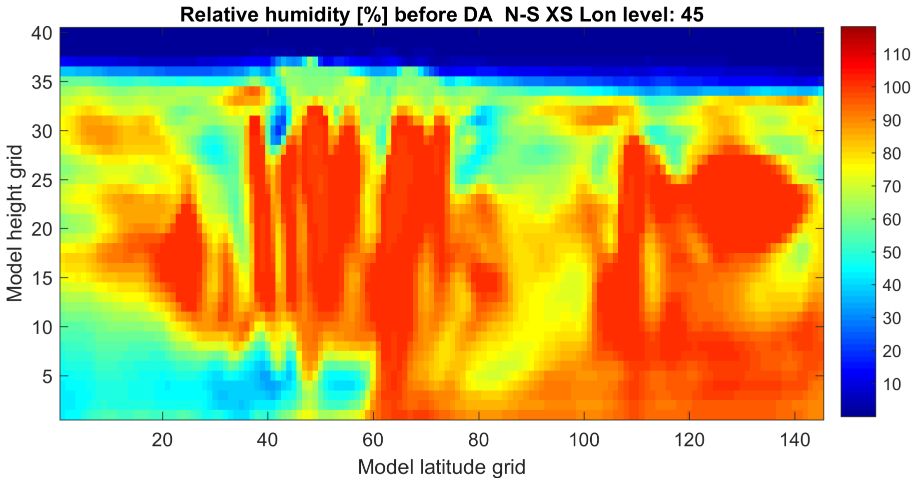
observations are assimilated, like the large strip of reflectivity stretching from the southwest corner of the domain until approximately $52.5^\circ N\ 2.5^\circ E$, no really big changes in IWV are made, the reason being that is air is already relatively moist and close to saturation here. In a large area on the North Sea between the Netherlands and the UK, the amount of water vapour is suppressed because null echoes in the radar image do not indicate any precipitation while it exists in the model. As noted in the discussion of the CZQV scenario in section 5.1.1, humidity in the model squall line that is positioned just slightly too far east is also significantly suppressed, judging from the large negative IWV increments around $52^\circ N\ 3^\circ E$ just in between the radar humidity observations. The strongest increase in humidity and IWV by radar can be observed above the middle of the Netherlands. A small line of reflectivity causes a moistening of this area that is underestimated in the FG, see 5.15a. Also above the UK the radar measurements cause an increase in moisture, although in this case the majority of the moisture increments is due to the GPS measurement nearby.

Looking at IWV gives an overview of the net effect of humidity assimilation over the depth of the atmosphere. However, this two dimensional approach does not show the vertical structure of changes to water vapour. It may happen that near the surface a strong drying is forced by a surface observation while a radar reflectivity observation indicates near saturation at higher altitudes in a column so that the net moisture increment in a column is near zero. To examine the interaction between the observation sources that produce humidity observations, the optimal situation for such a comparison is when two close SYNOP and ZTD locations are passed by a band of precipitation so that reflectivity humidity observations are also contained by the zenith humidity column to which the ZTD is proportional. That is the case two hours into the assimilation cycle, at 17 UTC. A map of the model IWV first guess and DA analysis and DA increments is included in figure 5.16, analogous to the map for 15 UTC in figure 5.15. Again, observations that change model water vapour have been plot. The green points indicate points where $Z > 20\ dBZ$ and thus near saturation RH observations are assimilated, while the other markers indicate locations of SYNOP, sounding and GPS measurements. The dashed line shows where the south-north cross section in figure 5.17 is made. This line intersects a large area of radar humidity measurements, and with a decent amount of SYNOP and GPS measurements nearby.

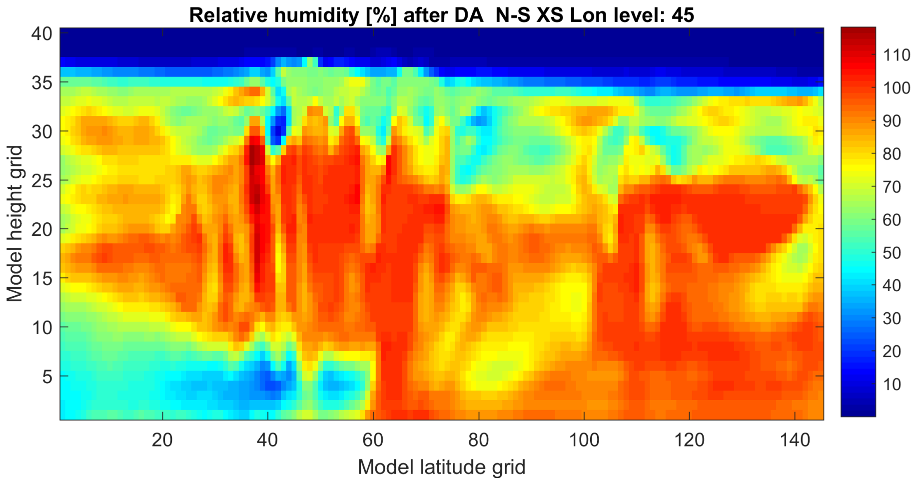
The north-south cross section of model RH is shown in figure 5.17 for the first guess (5.17a), DA analysis (5.17b) and DA changes (5.17c). In the last plot of DA increments the observations have been plot that are in proximity of the cross section so that they affect the humidity in the cross section. SYNOP and radar observations have been plot at their respective heights, while GPS measurements are indicated with the dashed vertical lines to keep in mind that a ZTD measurement is proportional to the integrated water vapour in a column. Do note that vertical model grid spacing increases upwards, as discussed in the model grid description (section 4.1), so that thick, higher model levels obtain a bigger weight in the vertical integration. The RH first guess already shows quite a lot of areas that are (nearly) saturated, corresponding to the radar measured reflections in figure 5.17c. One area in particular is underestimated though in terms of humidity, in the vertical of the 34th latitude grid point and its surroundings. A significant moisture increment is made here, judging from the near 30% RH change. Radar reflections at higher altitudes and the 3 middle GPS observations in figure 5.17c indicate that the moisture content in the air is underestimated. The left (south) most GPS observation measures ZTD slightly smaller than the model equivalent, so along this column negative increments are made, also because radar observations are not abundant to indicate otherwise. Interestingly both surface SYNOP stations measure a smaller humidity than the model, leading to a drying near the surface. Sometimes the combination of radar and GPS measurements can lead to a compensation effect when no other observations are available for constraint above or below radar observations. That is, the atmosphere is dried in the parts of the column where no observations are at hand, compensating for moistening of the radar measurements to still fit the ZTD measurements. This occurs to some extent above and below the area of largest humidity increase, e.g. in columns 35-40 of the latitude grid, although at the surface there is some influence of the drying SYNOP measurements. In general however this compensation effect occasionally might lead to surface drying this way.

The strong water vapour increment also results in a super saturation of the column just north of the area of largest increments, near the column of the 38th model latitude grid cell, because of the horizontal autocorrelation. The water vapour increments resulting from radar retrieved RH observations are continuous, in the sense that they are not suppressed when it leads to supersaturation in nearby grid cells. The excess water vapour is swiftly converted to precipitation by the microphysics scheme. This phenomenon adds to the model rainfall intensity which is already quite big.

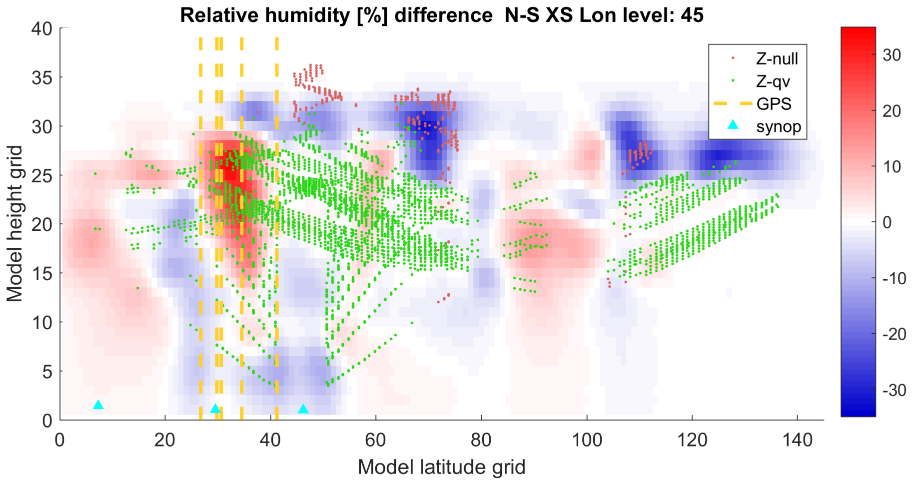
Finally, it is interesting to see how the assimilation cycling of all observations during three hours has affected the water vapour field with respect to the CONV scenario, where only conventional observations were assimilated. The IWV fields for both scenarios have been plot in figures 5.18a and 5.18b respectively. In terms of timing and placement there is a strong improvement. The moist air of the squall line has been correctly repositioned along the Dutch coast. Apparently however, the assimilation cycle has also led to quite a significant total moisture increment due to the inclusion of radar humidity measurements. The abundance of radar measurements combined with the rapid update cycling makes possibly unrealistically big increments in humidity.



(a) First guess

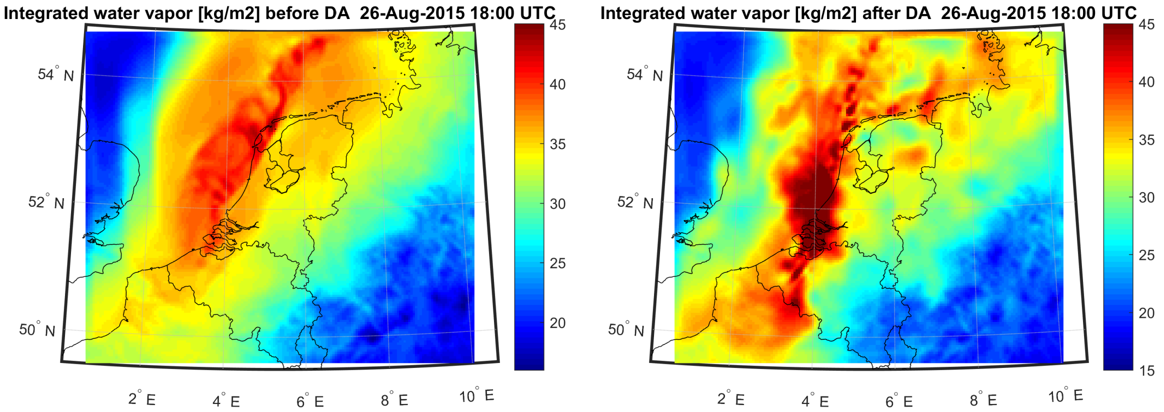


(b) Data assimilation analysis



(c) Difference analysis and first guess. The markers indicate locations of radar (null and humidity) and SYNOP measurements. GPS ZTD measurements are indicated with dashed lines to illustrate that they are proportional to the IWV.

Figure 5.17: RH of the model in the ALL scenario at 26-08-2015 17 UTC.



(a) CONV

(b) CGPS

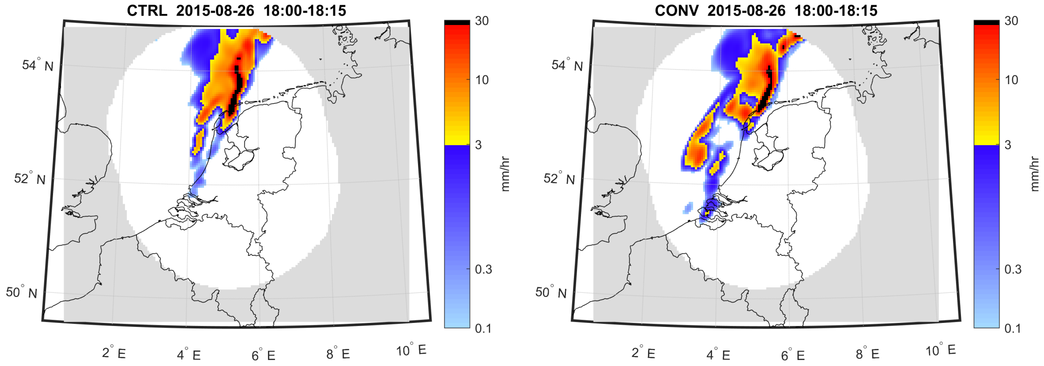
Figure 5.18: Comparison of model IWV at 26-08-2015 18 UTC for the CONV and CGPS scenarios.

5.1.2. Verification results

The primary goal of data assimilation is to obtain a model state that is closer to the real atmospheric state, which is used to initialize a forecast that should result in a more accurate forecast afterwards. As mentioned before, in this study the accuracy of the precipitation forecast is what is of primary interest. To assess the quality of the subsequent forecast after the assimilation cycle, model forecast precipitation is compared to radar images. For each observation type scenario the 15 minute accumulated rainfall is used to determine the average rainfall rate during that period. Similarly, 4 radar images are averaged to obtain the radar equivalent rainfall rate. Although the radar rainfall intensity has not been corrected with surface rain measurements, it is nevertheless assumed that the radar provides an estimate of rainfall intensity that is decently accurate enough for verification. For more details regarding the precipitation verification, refer to section 4.3. Figure 5.19 provides an overview of the average rainfall rate for every observation assimilation scenario, as well as the actual radar images.

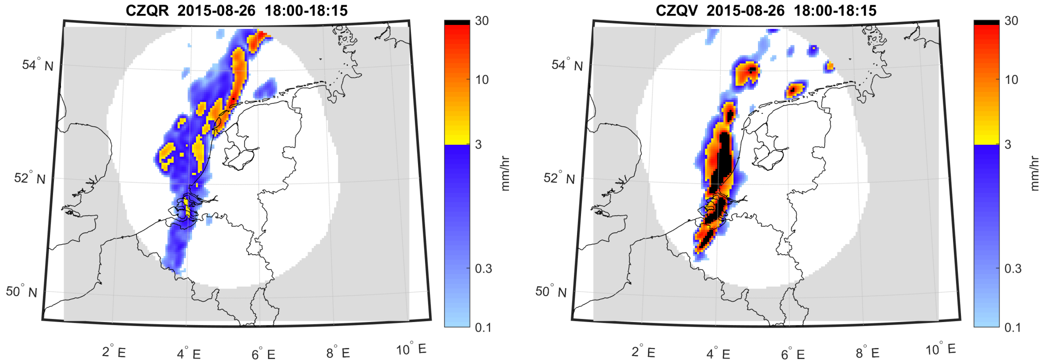
The rainfall intensity images for each scenario at later times throughout the verification forecast (20 and 22h) have been included in appendix B.1. As mentioned in the description of the verification method in section 4.3, the radar rainfall rate is used as truth in classifying whether a model grid cell produces a correct forecast. In the qualitative verification a threshold of 0.3 $mm\ hr^{-1}$ is used to classify precipitating and non precipitating cells. That way for each observation type scenario, a confusion matrix is constructed containing true positives, true negatives, false positives and false negatives. In section 4.3 some useful verification scores based on the confusion matrix to assess forecasting skill were defined. One of the most important scores is the Gilbert skill score (GSS), a skill score that attempts to quantify the accuracy of a forecast. It is similar to the critical succes index (CSI) with the addition of a correction for the chance of randomly predicting correct positives. The GSS every 15 minutes throughout the forecast is shown in figure 5.20 for each of the observation type scenarios discussed in the previous section 5.1.

When examining the GSS timeseries in figure 5.20, the first thing to notice is the added value of data assimilation on the accuracy of precipitation forecasting. The CTRL experiment evidently performs worst in forecasting precipitation, as its GSS is lowest almost throughout the whole forecast. In this scenario where no observations are assimilated the model accuracy purely relies on the quality of the initialization, i.e. the accuracy of the coarser GFS data. The major part of the squall line at 18 UTC as visible in the radar image in figure 5.19h is not forecast by the model itself, see figure 5.19a. The CONV scenario with assimilation of conventional observations already notably improves forecasting accuracy, mainly because of the extra precipitation area that formed after the assimilation of surface (wind) observation led to an extended line of surface convergence. The direct assimilation of radar reflectivity in the CZQR scenario leads to the best precipitation forecast right after assimilation, since hydrometeors are directly assimilated into the model. The areal extent (5.19c) is really similar to the radar images (5.19h), primarily because the very small autocovariance length of hydrometeors allows for really local increments. However shortly after assimilation the hydrometeors have precipitated to the surface and the added value of radar hydrometeors data assimilation is lost. This is confirmed by the GSS of the CZQR scenario which plummets between 15 and 30 minutes after assimilation. From this point when the majority of assimilated hydrometeors have precipitated to the surface, the CZQR forecast



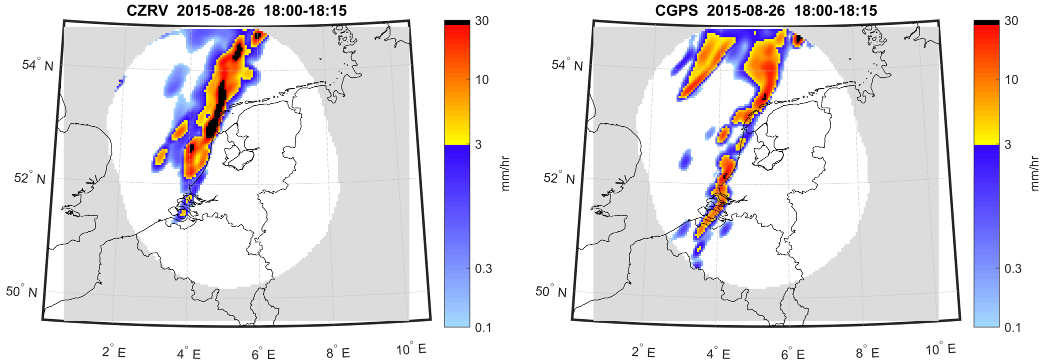
(a) CTRL

(b) CONV



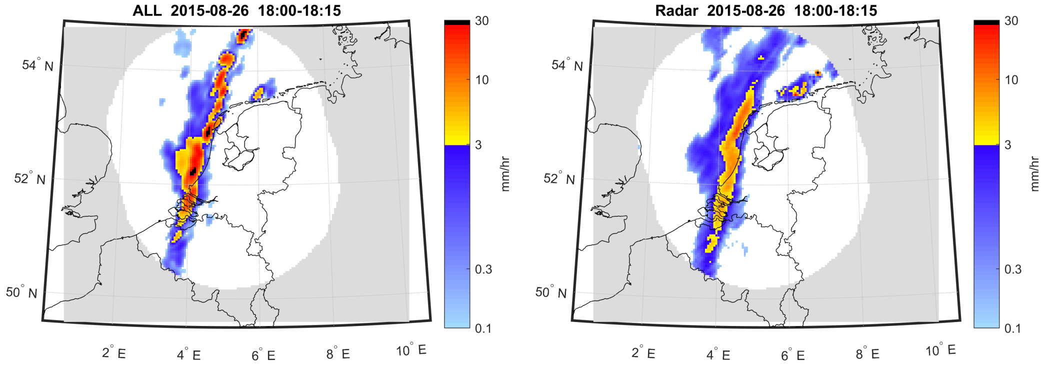
(c) CZQR

(d) CZQV



(e) CZRV

(f) CGPS



(g) ALL

(h) Radar

Figure 5.19: 15 minute average rainfall rate for each of the data assimilation scenarios and KNMI radar composite between 18:00 and 18:15 UTC.

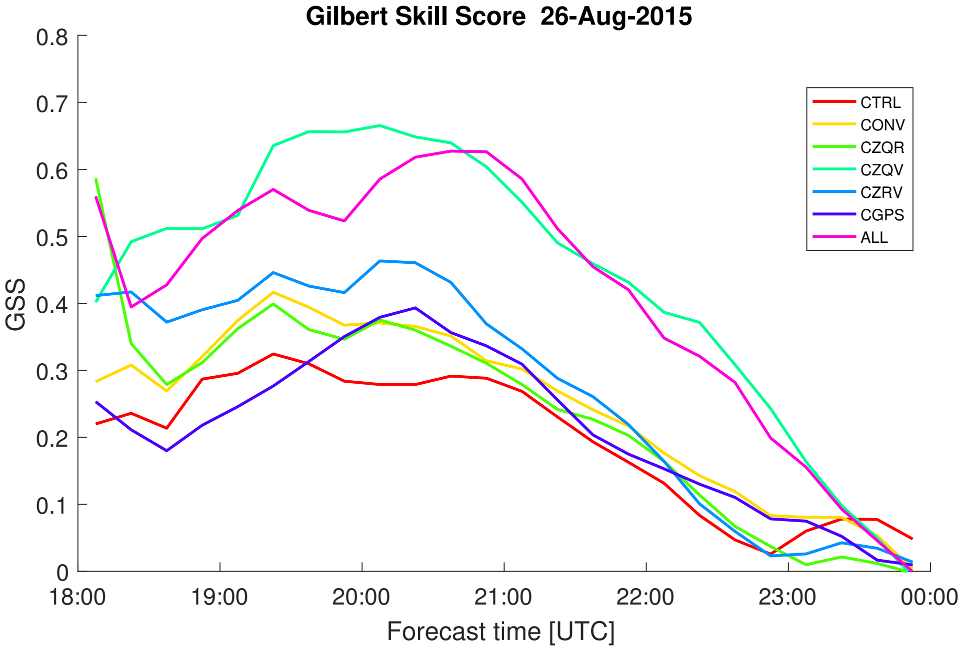
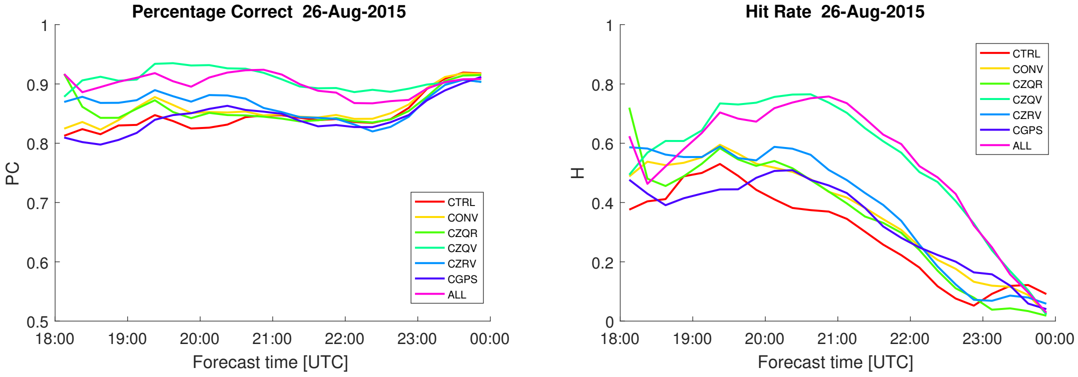


Figure 5.20: GSS of the model precipitation forecast MCS case of 2015-08-26 verified with radar composites. For each of the observation type scenarios listed in table 4.1 the GSS is shown as a function of the forecast time.

is in essence equivalent to the CONV scenario. The assimilation of near saturated water vapour derived from radar reflectivity in the CZQV scenario does a much better job in producing an improved precipitation forecasting skill that is sustained throughout the whole forecast. The model squall line was successfully repositioned and adjusted to fit observations. The near saturated areas of humidity resulting from indirect radar data assimilation provide a favourable environment for sustained precipitation. At the same time erroneous model precipitation which would appear as false positives in the verification forecast is suppressed by the assimilation of null-echoes, adding to the forecast skill. The GSS of the CZQV scenario is the best throughout the forecast except for the first 15 minutes, when the assimilation of hydrometeors in the CZQR and ALL scenarios slightly outperforms CZQV because of the temporary near copying of the radar images (compare 5.19g and 5.19h). At the very end of the forecast the CZQV scenario is surpassed by the CTRL and ALL forecast because nearly all of the model precipitation has left the domain and only some trailing precipitation is present, of which slightly more happens to be correctly positioned for the CTRL and ALL scenario. With the assimilation of radar radial velocity in the CZRV scenario the GSS also notably improves with respect to the CTRL forecast, without affecting moisture. The main reason is that in the case of a linear convective system as studied here convection is forced mechanically by an advancing cold front, so that adjusting wind speed and direction can trigger new areas of precipitation. In the CZRV scenario the assimilation of radial velocity leads to an extension of the line of surface convergence, producing a new area of precipitation on the southern flank of the model squall line. This extended squall line is what causes the improvement of the CZRV's GSS over the CTRL forecast. In the CGPS scenario where GPS ZTDs are assimilated, the precipitation forecast accuracy is really similar to the scenario CONV where only conventional observations have been assimilated. Actually the squall line at 18 UTC in figure 5.19f is quite accurately reproduced when compared to the actual radar intensity in figure 5.19h. However some erroneous areas of precipitation are forecast, primarily in the north of the domain, resulting from ZTD measurements on the east coast of the UK. These erroneous areas of precipitation also show up in the false alarm ratio (FAR) which is plot in figure 5.21c. In the first half of the verification forecast the CGPS scenario produces a significant amount of false positives. As a result the GSS is similar as for CONV although the structure of the squall line is much better from a forecaster's perspective. Once the erroneous precipitation has left the domain at around 20:00 (figure B.1f), the amount of false alarms significantly reduces and correspondingly GSS improves to scores similar to CONV again. Finally and not surprisingly, the ALL scenario is most similar to the CZQV scenario, as it is the only other scenario besides CZQV where radar based humidity is assimilated. Since in the ALL scenario also model differences with other observations like radar hydrometeor content, radial velocity and GPS have to be minimized in the cost function, the impact of the radar humidity observations is not as strong as for the CZQV scenario. The initial collapse of the forecasting skill as result of precipitation of assimilated hydrometeors can again be observed, although not as dramatic as the CZQV scenario. Just like the CZQR scenario the forecast rainfall intensity (see figure 5.19g) is most similar to the radar image, but is misleading since the hydrometeors vanish quite rapidly after assimilation. So at later times e.g. 20 UTC (appendix B.1g), the model rainfall intensity tends towards the CZQV scenario, also indicated by the GSS.

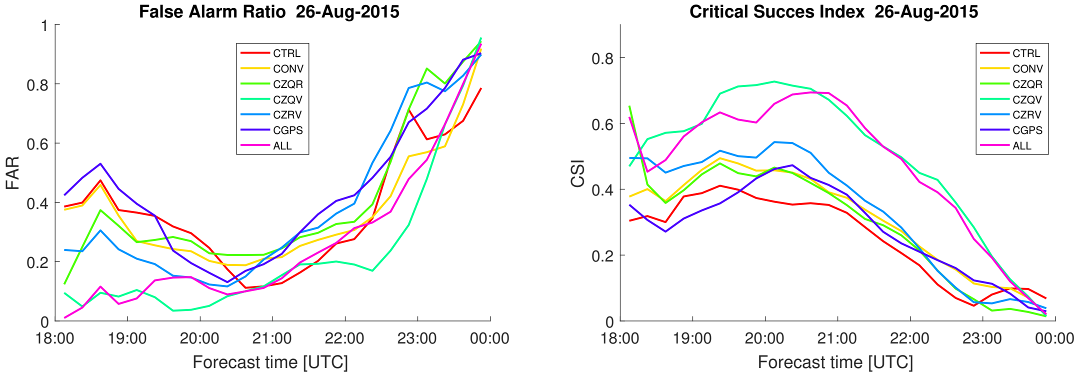
The GSS is practically zero at the end of the forecast for all scenarios. Not because the model has completely lost its forecasting skill, but because there is very few precipitation forecast by the model while the radar still shows some areas of trailing precipitation. The lack of forecast precipitation causes a small number of correct positives a, also indicated by the hit rate (H) in figure 5.21b, the fraction of correctly forecast rain grid cells. Consequently, the GSS which largely depends on a will tend towards zero, similar to H. Besides looking at the correctly forecast grid cells, the classification bias (CB) which is plot in figure 5.21e provides a measure of whether false forecasts are predominantly false negatives or false positives. In this case all scenario forecasts contain more false negatives than false positives, so generally underpredicting precipitation. Whether under- of overprediction is considered worse is subject to discussion, but the CB should always be interpreted with the fraction of false forecasts of the total, equal to $1-PC$, in mind. In this case the fraction of falsely forecast events is relatively constant for every scenario, varying per scenario roughly between $10-15\%$, see figure 5.21a. so although generally more precipitation is not forecast rather than falsely forecasting precipitation, the total amount of false classifications is limited.

The use of skill scores only provides a qualitative verification for the model forecast. This is sufficient when one is only interested in knowing whether it rains or not. However, it is also interesting to examine how well the model is capable of forecasting precipitation quantity as there is a growing demand for accurate quantitative precipitation forecasting (QPF). Just from visual comparison of the model forecast rainfall intensity at 18 UTC in figure 5.19 one can observe that generally the intensity of precipitation is substantially overestimated. Also at later times in the verification forecast this is the case, see the rainfall intensity forecasts at e.g. 20 and 22 UTC in appendix B.1. At 18 UTC, even in the CTRL forecast in figure 5.19 without any data assimilation the mispositioned core of precipitation the rainfall intensity is stronger than in the actual radar image. Generally throughout the domain more precipitation is forecast, as indicated by the domain 15 minute rainsum that has been plot in figure 5.21f. The total precipitation sum of the radar has been included for reference, as indicated by the dashed line. Two other scores to asses the ability to accurately forecast quantitatively have been introduced in section 4.3.2: the relative bias (RB) and absolute average scale ratio (ASCR), which have been plot in figures 5.21g and 5.21h respectively. These scores tell a similar story as the rainsum. Precipitation is considerably overestimated in terms of intensity, regarding the fact that $RB > 1$. Also, the precipitation is forecast on average at least three times as intense as the radar indicates for all observation scenarios, judging from the absolute average scale ratio (ASCR) in figure 5.21h. As mentioned before, even the CTRL forecast and the scenarios without assimilation of water vapour from radar significantly overestimate precipitation intensity. Apparently the microphysics parametrization struggles to produce realistic precipitation amounts in this rather extreme case of forced convection. In the scenario CZRV with radial velocity the stronger convergence leads to even more intense rainfall. The biggest overestimation of rainfall is by the scenario's that include radar humidity assimilation, CZQV and ALL. Because generally in the radar humidity assimilation scenarios positive moisture increments are made throughout the assimilation cycle, even more water is available for precipitation than for the scenarios without radar humidity assimilation. It seems that an improved accuracy in spatial terms goes at the expense of accuracy in precipitation quantity. Increasing the assimilation update interval from 15 minutes to 30 minutes or longer will reduce the amount of moisture that is added to the model, but most likely this will also reduce the accuracy of the qualitative precipitation forecast.



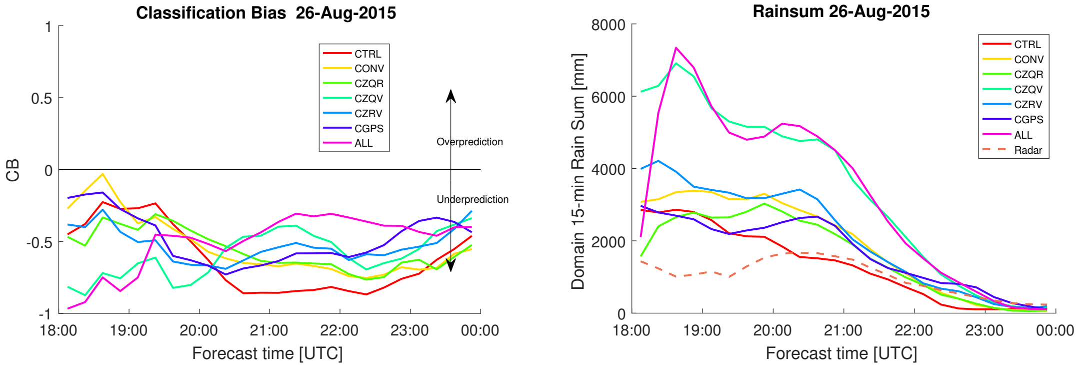
(a) percentage correct (PC)

(b) hit rate (H)



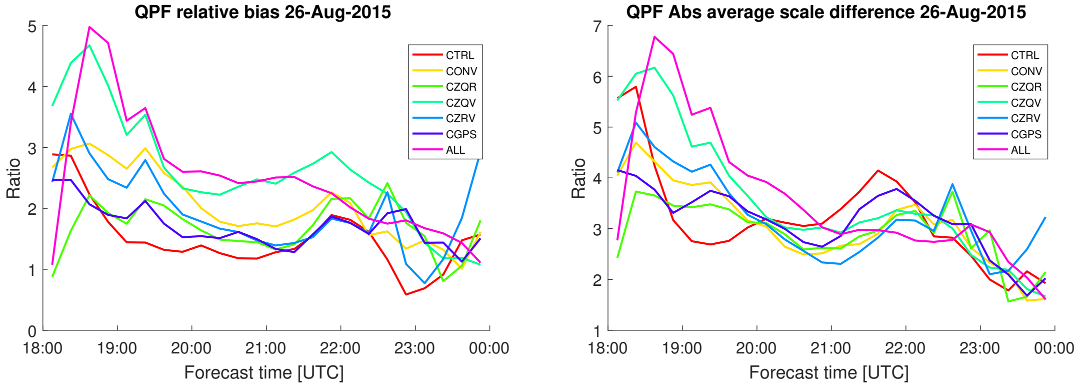
(c) false alarm ratio (FAR)

(d) critical succes index (CSI)



(e) classification bias (CB)

(f) Domain precipitation sum



(g) Relative bias

(h) Average scale ratio

Figure 5.21: Scores derived from the confusion matrix of the model precipitation forecasts the MCS case of 2015-08-26 verified with radar composites. For each of the observation type scenarios listed in table 4.1 the H is shown as a function of the forecast time.

5.2. Experiment 2: Precipitation mechanism

In the second experiment the performance of data assimilation assimilation for different kinds of precipitation mechanisms is examined. Three cases (MCS, LCS and STRAT) featuring different kinds of mechanisms causing precipitation have been selected and have been discussed in the experiment setup description in section 4.2.2. The setup for each case in terms of simulation start and end time has been summarized in table 4.2. As indicated by table 4.2, in each precipitation case the scenario where all available observations are assimilated is compared to a reference CTRL forecast without any data assimilation, hence the labels CTRL and DA. However, when examining the CZQR scenario where radar retrieved hydrometeor mixing ratios are assimilated, it was found that the added value of hydrometeor assimilation is lost once hydrometeors have precipitated. This is confirmed by the course of the GSS of the CZQR scenario in figure 5.20 which is almost equivalent to the CONV scenario from 30 minutes and onwards. Assimilation of hydrometeors however consumes unnecessary weight in the cost function of which the value is minimal during the majority of the subsequent initialized forecast. For that reason, the assimilation of hydrometeors is omitted in experiment 2 in the DA forecasts.

5.2.1. Single data assimilation

Before assessing the skill of the precipitation forecast after the data assimilation cycling window, first an impression is given of the changes to the model that assimilation of the observations results in will be given for each case. The MCS case of 26-08-2015 has already been extensively discussed in section 5.1 during the first experiment where the effect of observation type on the forecasting skill was discussed. Therefore, in this section only the remaining two cases (LCS and STRAT) are described.

LCS

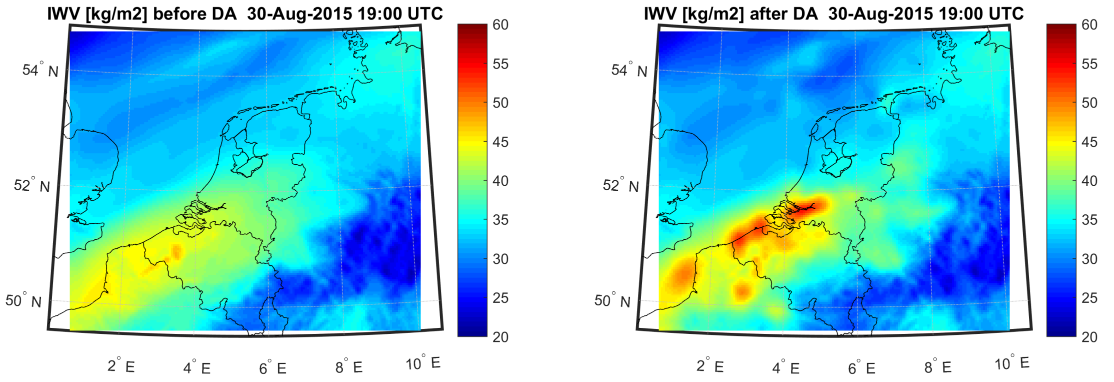
On the evening and following night of 30-08-2015, a narrow but intense local convective system (LCS) passed over the Netherlands. The first areas of strong convection were initiated just before 19 UTC near the Belgian coast and produced some heavy precipitation including some hail, see the radar images valid at 20 UTC in figure 4.27a, as shown in the description of this case in the experiment setup in section 4.2.2.

	Synop	**Metar**	**Buoy**	**GPS**	**Radar Zqv**	**Radar Zqr**	**Radar Zqg**	**Radar Zqsn**	**Radar RV**
Used	97	43	19	140	6420	0	0	0	501
Rejected	0	0	0	18	0	9292	17029	17029	285
Total	97	43	19	158	6420	9292	17029	17029	786

Table 5.2: Number of assimilated observations at 19 UTC in the LCS case per instrument.

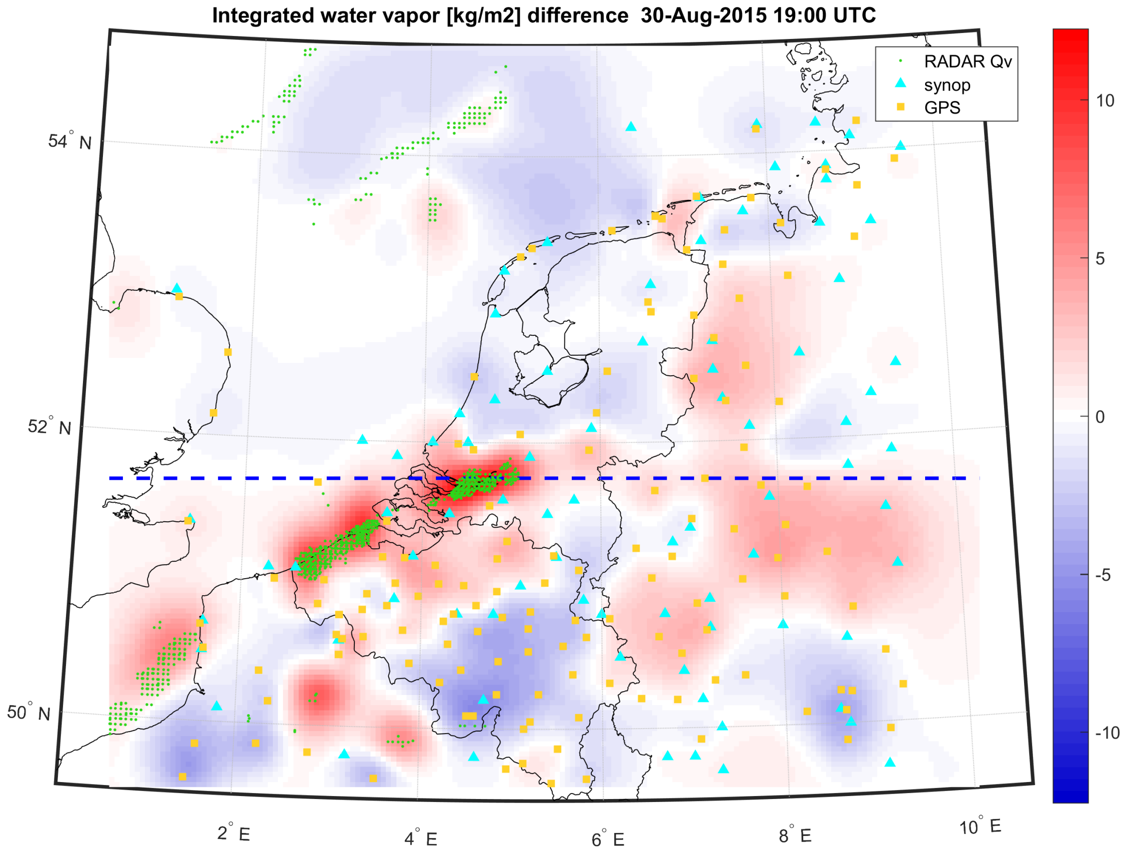
Instead of highlighting the impact of assimilating every type of observation, for this LCS case as well as for the STRAT case only the assimilation of all observations together are discussed. Similar to the previous experiment, the model changes of the first assimilation at the start of the cycling window at 19:00 UTC will be examined, as it shows the changes made by data assimilation to an unaltered forecast. An overview the number of observations that is assimilated at this time is given in table 5.2. Robust observations from SYNOP and METAR and buoys are all pass the maximum innovation test and are assimilated into the model. The GPS measurements are generally more uncertain and logically a fraction of the GPS measurements are discarded because they are too far off from the background. At 19 UTC the first convective cells have just developed along the coast of Belgium and the Netherlands, which is measured by the radars and translated into the radar hydrometeor (Zqr, Zqg and Zqsn) observations. Also a decent amount of radar based humidity observations are assimilated, although their number is limited as there are not a lot of areas with strong reflections ($Z > 20\ dBZ$) yet that lead to a humidity observation. Also, the number of radial velocity measurements is quite limited in this case because their are not many radial velocity observations that coincide with the reflectivity measurements. A significant part of the radial velocity measurements fails the stringent maximum innovation test. Compared to the MCS case (see table 5.1) the radar observations are less abundant because this case features local convection which is still in the developing stage.

First, it is interesting to examine the changes in model moisture that result from the assimilation of radar measurements, GPS ZTD and conventional observations. The unaltered model first guess for IWV at the beginning of the data assimilation cycling window is shown in figure 5.22a. Evidently, the model first guess captures the area of moist air approaching from the south-west that adds to the instability of the atmosphere.



(a) First guess

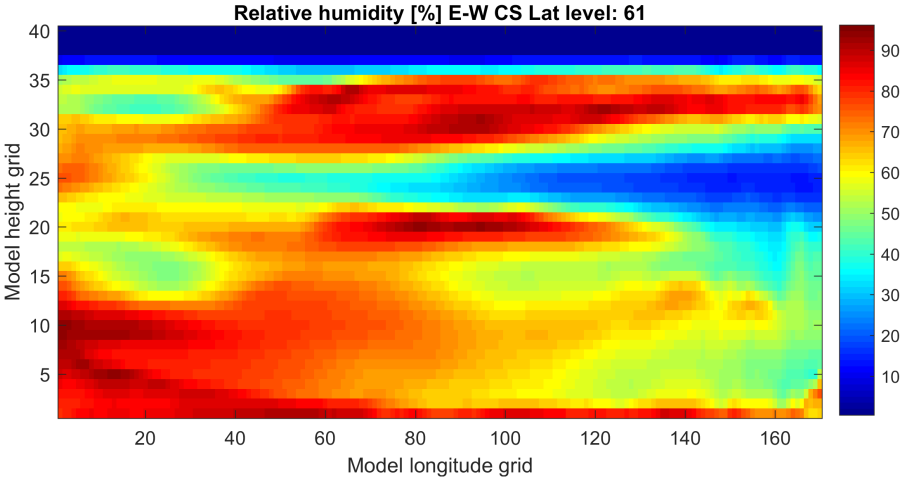
(b) Data assimilation analysis



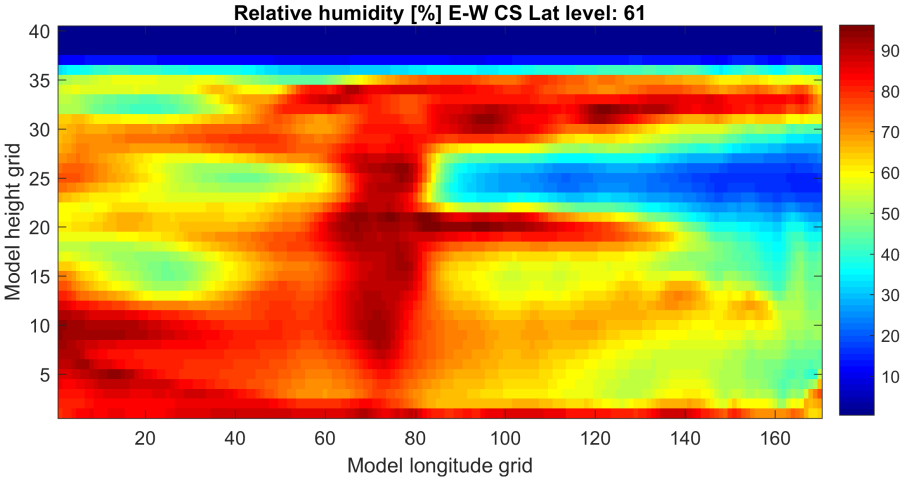
(c) Difference analysis and first guess. Markers show the observation locations of SYNOP and ZTD measurements. Additionally dots indicate the locations of radar humidity observations where $Z > 20\ dBZ$. The dashed line shows where the cross section in figure 5.23 is made.

Figure 5.22: IWV of the model in the ALL scenario at 30-08-2015 19 UTC.

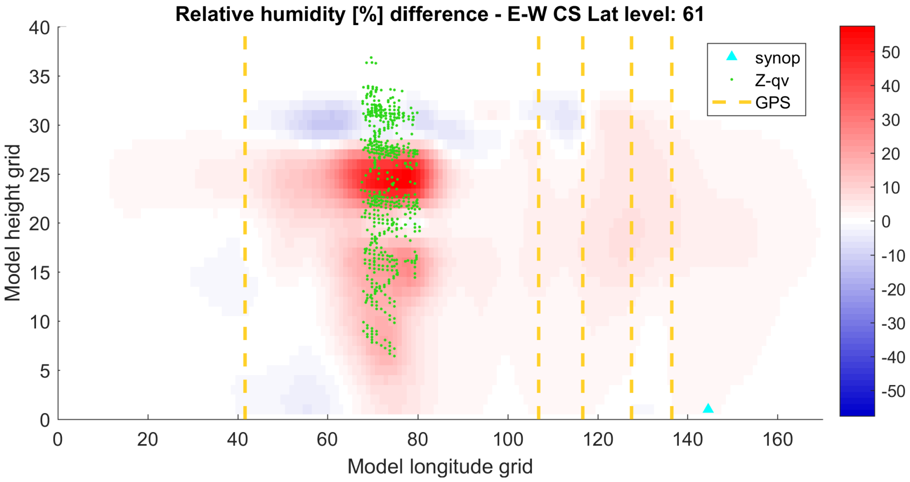
However, the detailed features that are needed to accurately initialize precipitation are not really visible. The analysis in figure 5.22b shows that after assimilation the line along which the convective cells form according to the radar is now more evident as an area containing more moisture than it's surroundings. The radar derived humidity measurements above the Belgian coast and within the Netherlands cause a significant increase in humidity locally. Another area of precipitation located even more south-west near the French coast that is measured by the JAB radar also causes a moistening, although not as big since the air is already close to saturation here and also since only radar observations at higher altitudes are available, given the range to the



(a) First guess



(b) Data assimilation analysis

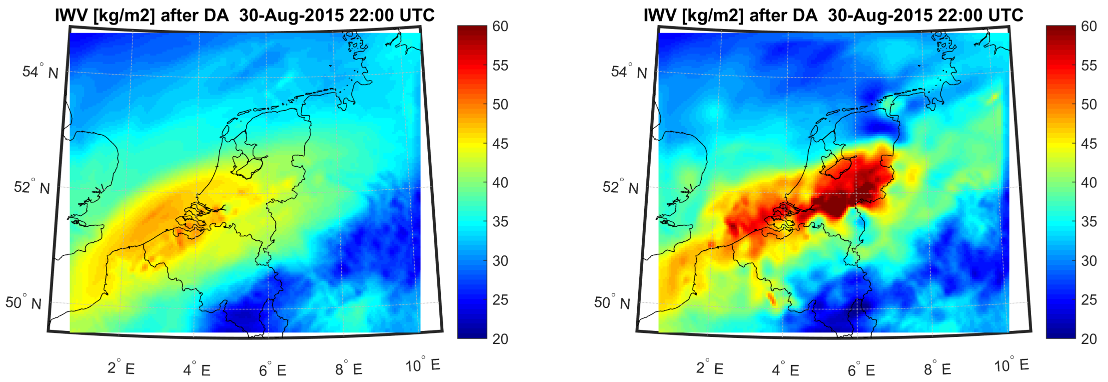


(c) Difference analysis and first guess. The markers indicate locations of radar (humidity) and SYNOP measurements. GPS ZTD measurements are indicated with dashed lines to illustrate that they are proportional to the IWV.

Figure 5.23: Cross section of RH of the model in the ALL scenario at 30-08-2015 19 UTC.

radar is bigger for this area of precipitation. In the areas with no radar reflection nearby, humidity changes result from assimilation of GPS and SYNOP observations.

Besides looking at the change in integrated water vapour, it is also interesting to examine the vertical profile of humidity after assimilation. A vertical model cross section of relative humidity is shown in figure 5.23. It intersects the area of the largest change in IWV and is shown with the blue, dashed line in figure 5.22. The model first guess of relative humidity in figure 5.23a shows that a band of moist air extends from the surface up to the 20th model level, but is not near saturation for the large part. Assimilation of radar humidity measurements results in a clear column of nearly saturated air in the analysis shown in figure 5.23b. The data assimilation increments in relative humidity shown in figure 5.23c show that the moistening is primarily a result of radar assimilation. Because the precipitation is located in between the JAB and DB radar, radar observations are available throughout almost the whole column except for the near surface levels. The strong radar reflections at high model levels (approximately 300 hPa), assuming that the estimated reflection altitudes are correct, indicate that some deep convection developed that facilitated the formation of heavy rain and graupel. Because of the abundant radar reflections moisture increments are made throughout the whole column which results in a significant near 20% change in IWV that is visible in figure 5.22. Unfortunately no GPS measurements are close enough to provide constraints on the radar humidity increments which are relatively large.



(a) CTRL (b) DA

Figure 5.24: Comparison of model IWV at 30-08-2015 18 UTC for the CTRL and DA scenarios.

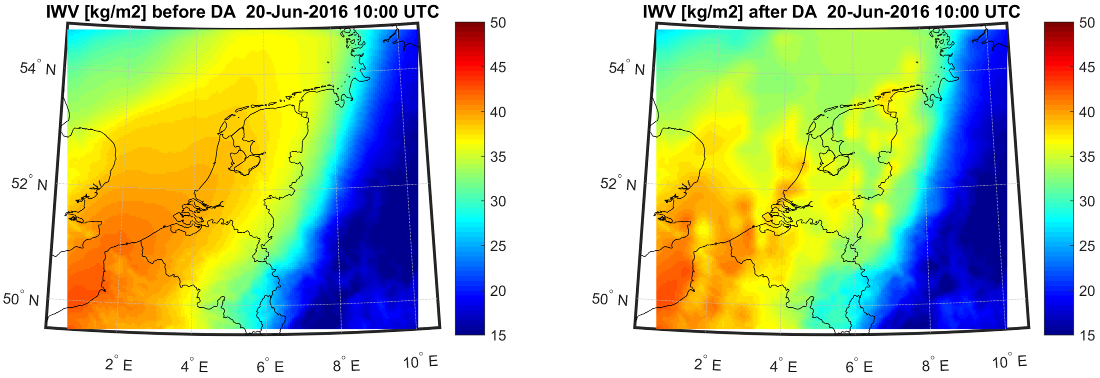
After 3 hours of 15 minute update assimilation, notable changes have been made to the moisture field in the model. In figure 5.24 the IWV of the CTRL run is compared with the IWV field from the ALL run in which every single observation type is assimilated. Evidently, a substantial amount of moisture is added as a result of the rapid updates of radar data assimilation. While the structure of the humidity field better represents the extent of the actual precipitation measured by the radar (figure 4.27b), similar as for the MCS case, it is questionable if such big increments are realistic.

STRAT

	Synop	Metar	Buoy	GPS	Radar Zqv	Radar Zqr	Radar Zqg	Radar Zqsn	Radar RV
Used	91	45	17	175	50419	0	0	0	19180
Rejected	0	0	0	7	0	80899	99420	99420	3678
Total	91	45	17	182	50419	80899	99420	99420	22858

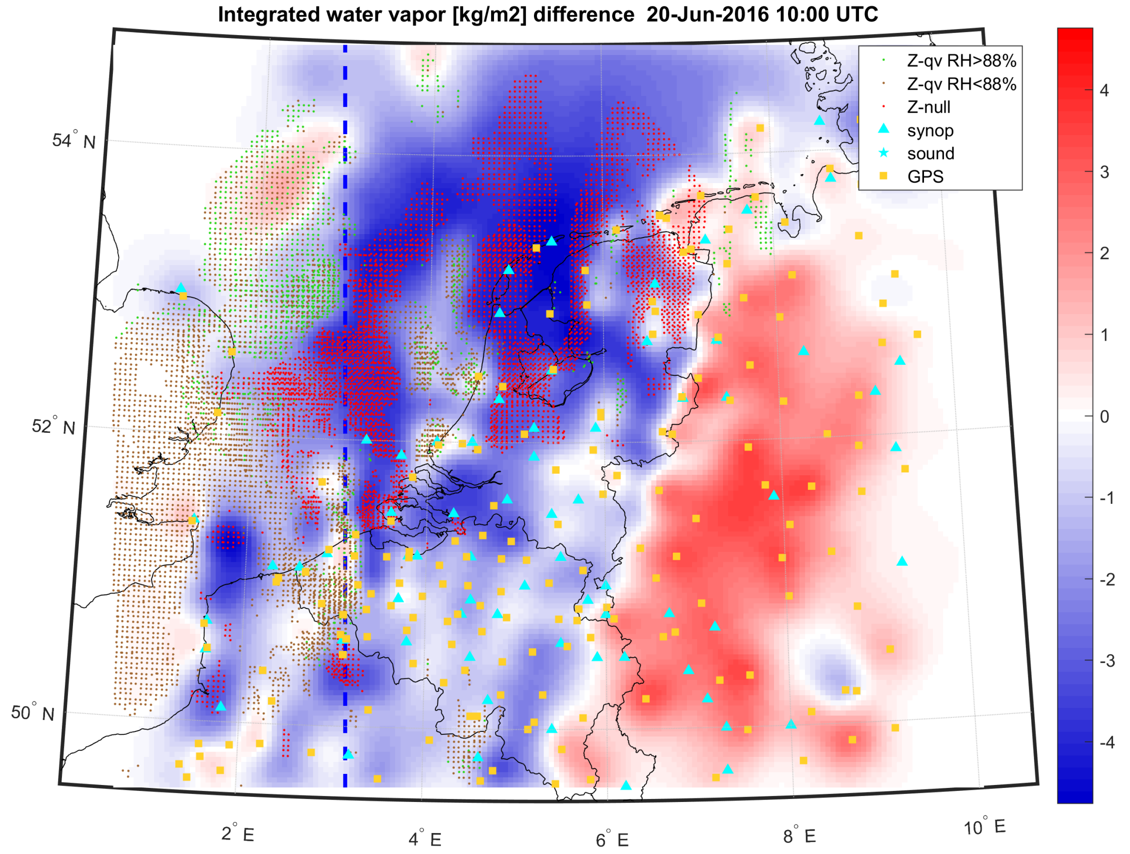
Table 5.3: Number of assimilated observations at 10 UTC in the STRAT case per instrument.

In the final case a weather situation with more gradual, stratiform precipitation is considered. During the afternoon of 20-06-2016, a warm front crossed the Netherlands which produced an elongated band of stratiform precipitation that covered the whole Netherlands and Belgium for some hours. Refer to section 4.2.2 again for a synoptic overview of this case. In table 5.3 an overview is given of all the possible observations that are assimilated and rejected per observation type. A decent amount of conventional observations are available which all pass the maximum innovation test, while a small fraction of the GPS measurements are discarded. A huge amount of radar observations is the consequence of the large extent of the precipitation that covers a large part of the domain.



(a) First guess

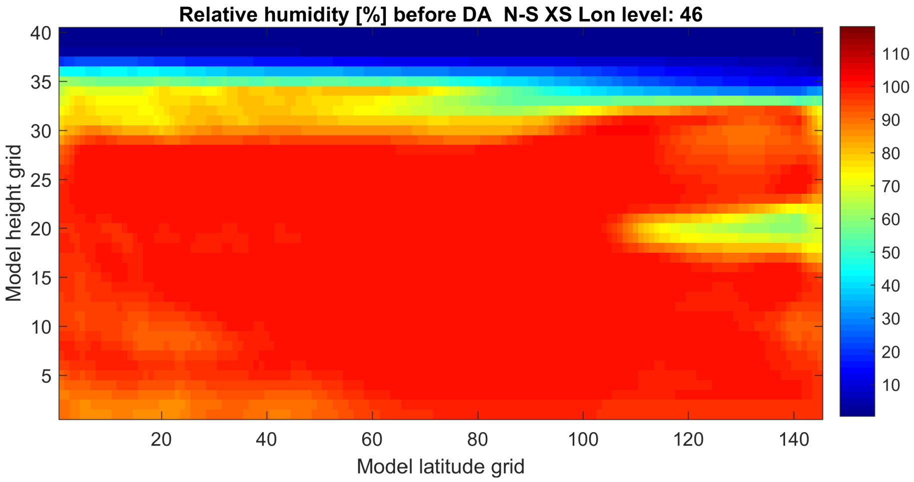
(b) Data assimilation analysis



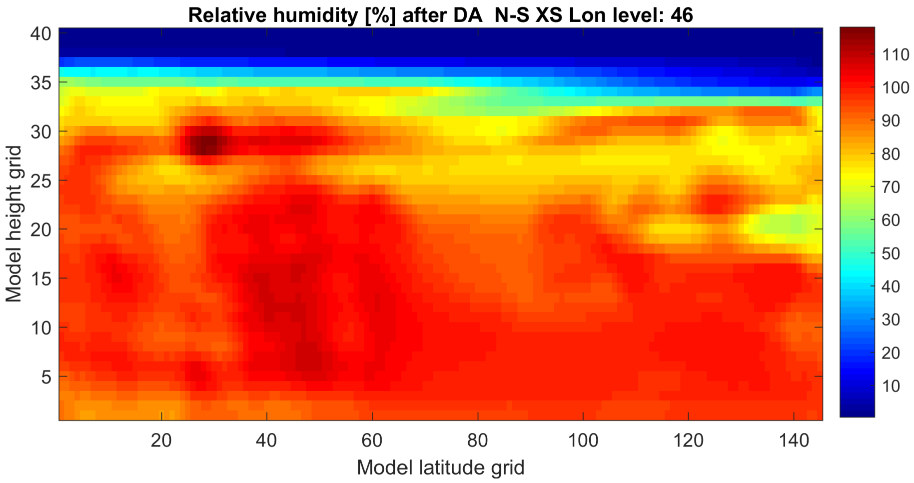
(c) Difference analysis and first guess. Markers show the observation locations of SYNOP and ZTD measurements. Additionally, green dots indicate the locations of radar humidity observations where $Z > 20\ dBZ$ and model $RH < 88\%$, while brown dots indicate $Z > 20\ dBZ$ locations that are not assimilated because the model RH is already $RH > 88\%$. The dashed line shows where the cross section in figure 5.26 is made.

Figure 5.25: IWV of the model in the ALL scenario at 20-06-2016 10 UTC.

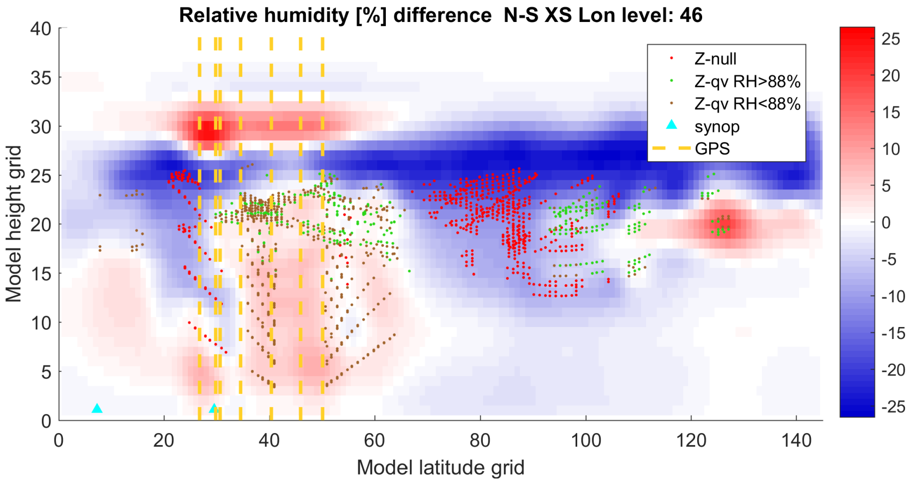
Similar as for previous cases, the changes to model moisture made by the first assimilation at the start of the cycling window are discussed. The model FG for IWV at 10 UTC is shown in figure 5.25a and the model IWV analysis after assimilation in figure 5.25b. Interestingly, changes in the IWV are dominantly negative, at



(a) First guess



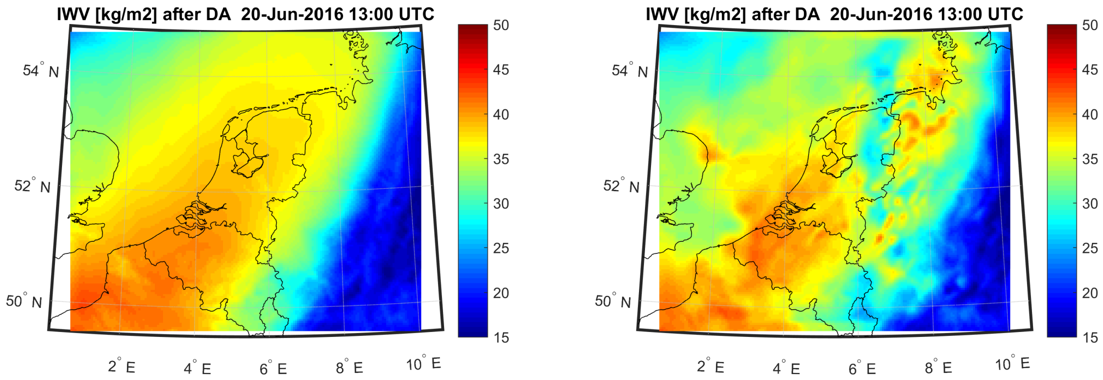
(b) Data assimilation analysis



(c) Difference analysis and first guess. The markers indicate locations of radar null, humidity, non assimilated humidity and SYNOP measurements. GPS ZTD measurements are indicated with dashed lines to illustrate that they are proportional to the IWV.

Figure 5.26: Cross section of RH of the model in the ALL scenario at 20-06-2016 10 UTC.

least in the relatively moist part of the domain where precipitation is formed. The negative increments can be seen in figure 5.25c where model changes in IWV are plot together with the various observations affecting model humidity. The majority of radar humidity observations (indicated with the brown points) where $Z > 20\ dBZ$ are not assimilated because the model RH is already bigger than 88%, while only a fraction (green points) of the radar humidity measurements are in fact assimilated because $RH < 88\%$. The fact that the DA increments are dominantly negative as a result of the numerous null echo assimilations, as shown by the red dots in figure 5.25c. In these areas model precipitation with a simulated reflectivity $Z_{model} > 20\ dBZ$, while none of the radars indicates precipitation, a $RH = 90\%$ observation is used to suppress erroneous model precipitation. This leads to a drying of the water vapour column, which also becomes evident in the vertical cross section of model relative humidity shown in figure 5.26. The location of the cross section is indicated by the dashed blue line in figure 5.25c, intersecting the large area of null echoes on the North Sea. The null echoes appear in the cross section of model increments in figure 5.26c at higher altitudes at model latitude grids 70-100. South of the null echoes, minor positive increments are found as a result of GPS assimilation combined with some valid radar humidity observations. The low altitude radar observations are made by the JAB radar that are not assimilated anyway due to the RH already being saturated, as its location is near the Belgian coast which coincides with the cross section.



(a) CTRL

(b) DA

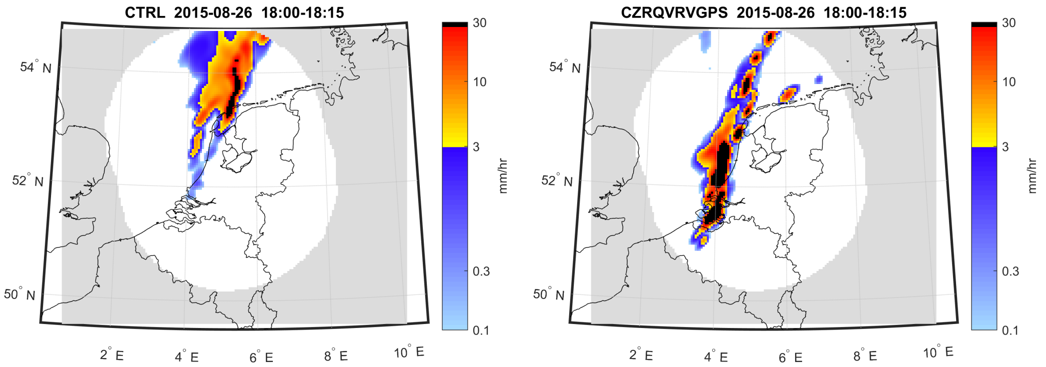
Figure 5.27: Comparison of model IWV at 20-06-2016 10 UTC for the CTRL and DA scenarios.

After the 3 hour assimilation cycling window a model IWV as shown in figure 5.27b is obtained which serves as the start of the verification forecast. Similar to the first DA increments at the start of the assimilation window as in figure 5.25c, in most areas (along the Dutch/German border at roughly $6.5°E$) changes are negative compared to the CTRL forecast (figure 5.27a), drying the column which inhibits further precipitation. Also in south Belgium and parts of the domain covering the North Sea are dried, leading to a smaller and more narrow band of precipitation, as will be shown in the verification (figure 5.30). In west Germany areas with high moisture content have become quite scattered, producing some isolated cells of with slightly more intense rainfall. The area of relatively large IWV on the coast of the UK is the result of the assimilation of GPS ZTD at this location, of which it seems that the bias correction is not really adequate since it generally produces large positive increments. This area with extra water vapour adds an area of intense precipitation that is not measured by a radar, as will become evident in the verification. In general, the small scall nature of increments combined with rapid assimilation results in a high frequent varying moisture field in the DA scenario of figure 5.27b, which is less able to produce the same kind of stratiform precipitation as the smooth IWV field of the CTRL scenario.

5.2.2. Verification results

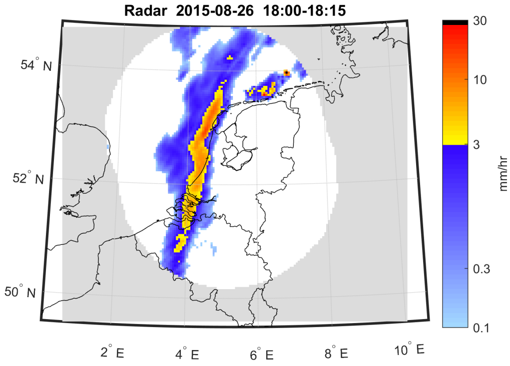
As for the verification in experiment 1, model 15 minute average precipitation rate is compared with a 15 minute average radar precipitation rate. The model precipitation rates for the CTRL and DA scenario are shown at the end of the assimilation cycle for the cases MCS (figure 5.28, LCS (figure 5.29)) and STRAT (figure 5.30), ending at 18, 22 and 13 UTC respectively. As extensively discussed in the verification of experiment 1 which investigated the effect of assimilating every observation type on the precipitation forecast, the squall line is quite successfully forecast when observations and particularly radar retrieved humidity is assimilated.

The precipitation forecast by the CTRL scenario in figure 5.28a for a large part fails to reproduce the long stretch of intense rainfall along the squall line that is present on the radar in figure 5.28c. In the DA scenario where observations are assimilated every 15 minutes during a 3 hour window however, in terms of location the forecast is significantly improved as is evident in figure 5.28b. The elongated structure is now much better present in the model and the areal extent is really similar to the radar. Nonetheless as discussed in the verification of the first experiment, the intensity of the model squall line is overestimated substantially, although the CTRL scenario also shows some sign of overestimation in the areas where precipitation is forecast.



(a) CTRL

(b) DA

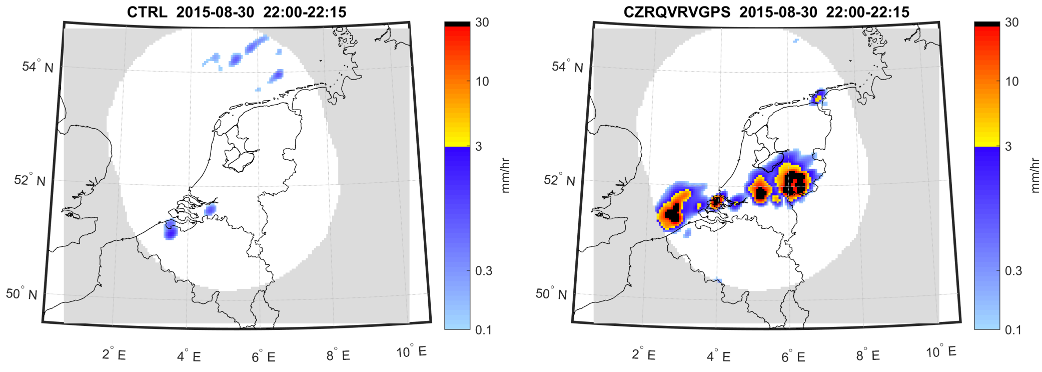


(c) RADAR

Figure 5.28: 15 minute average rainfall rate for CTRL, DA and radar at 26-08-2015 between 19:00 and 19:15 UTC in the MCS case.

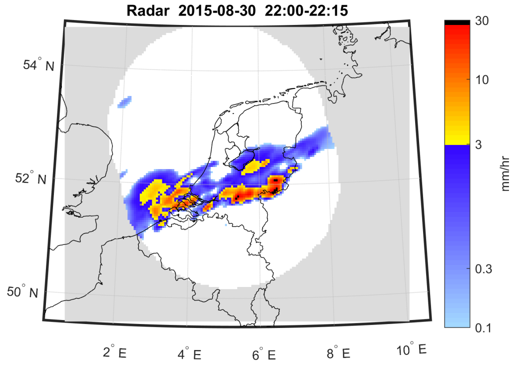
The second case LCS is a case where locally around 19 UTC really strong convection develops around the southwest part of the Netherlands. At this time the data assimilation cycling period is initiated and lasts up to 22 UTC. At 22 UTC various intense cells of convection have developed that cover the middle of the Netherlands, as shown by the radar image valid at that time in figure 5.29c. In the CTRL scenario however virtually no precipitation is forecast, see figure 5.29a. This might be related to model spinup time time which was discussed in section 4.1.3, although an experiment with an additional spinup time of 3h initiated at 13 UTC instead of 16 UTC to cover a larger part of the diurnal cycle (figure 4.8) still did not show any triggering of convection before 11 UTC just like the original CTRL scenario initialised at 16 UTC in figure 5.29a. The assimilation of all possible observations except hydrometeors again significantly improves the precipitation forecast, as shown by the DA intensity forecast at 18 UTC in figure 5.29b. The model now shows some individual cells of strong convection, as indicated by the very high $> 30\ mm\ hr^{-1}$ rainfall intensity. The extent of the area of precipitation is really similar, only the convective cells in the model have a larger structure than in the radar image. Reproducing really local and fine convective cells as the radar composite in figure 4.27a of the experiment description indicates and out of which the precipitation in the verification forecast develops is unrealistic to expect with a model resolution of 4 km. At that resolution, only large scale convective systems can be decently represented in a model. The DA model rainfall forecast in figure 5.29b is likely to be close to the best representation of the actual convective cells that the model is able to achieve with current settings regarding resolution and parametrization.

In the final STRAT case it is examined how data assimilation impacts a precipitation forecast in which precipitation forms gradually as a result of an advancing warm front. The precipitation associated with this



(a) CTRL

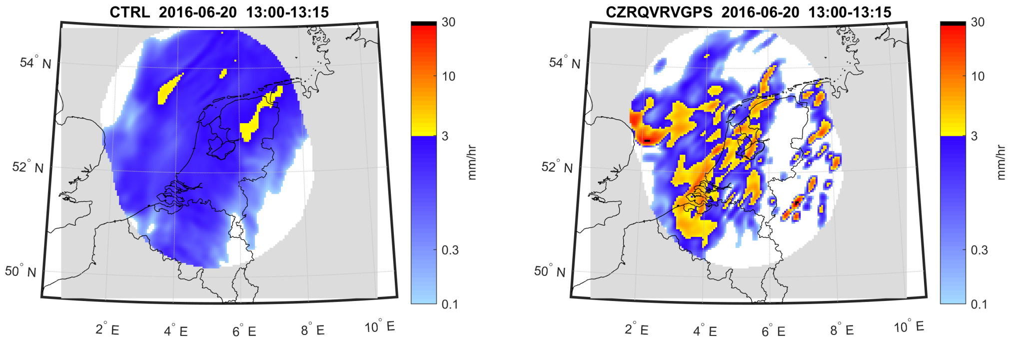
(b) DA



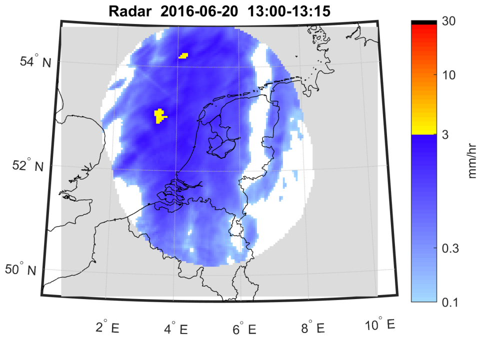
(c) RADAR

Figure 5.29: 15 minute average rainfall rate for CTRL, DA and radar at 30-08-2015 between 19:00 and 19:15 UTC in the LCS case.

warm front at the end of the data assimilation cycling window at 13 UTC as measured by the radar is shown in figure 5.30c. A large fraction of the verification domain is covered with low intensity, stratiform precipitation resulting from the gradual ascent of the warm and moist air of the front. The CTRL forecast which is shown in figure 5.30a captures this process quite well. The area of rainfall is really similar, only slightly more extensive in width. The fact that the model is able to reproduce the actual precipitation this well is mainly because of the large scale nature of the precipitation mechanism. The precipitation formation resulting from gradually rising airmasses is far less explosive and non-linear than the strong convection in the case of a squall line (MCS) or strong local updrafts (LCS). In the STRAT case the assimilation of observations actually results in a degradation of the forecast, see figure 5.30b. The rapidly successive assimilations during the cycling window combined with small covariance length scales result in more concentrated areas of heavier rainfall. In this study the small covariance length that was intended to allow for assimilation of small scale features like in the LCS case now causes unrealistic isolated areas of heavier precipitation. The moisture field in the STRAT-DA scenario contains a lot of high frequent features because of the local increments by GPS and radar observations while the actual stratiform rain has a structure that is much more homogeneous, as is evident in the radar image. Regarding the mesoscale structure of the warm front bigger covariance lengthscales and smaller model covariance (giving the model a bigger weight in the cost function) would be more appropriate for this case. Another remarkable observation is that the areal extent of precipitation has been reduced significantly with respect to the STRAT-CTRL forecast. Where the STRAT-CTRL forecast covers almost the whole of the verification domain, in the STRAT-DA forecast the extent is still substantially large but a lot less wide, not stretching from west to east of the domain anymore. In the western part and also in some areas at the northern and southern edge of the verification domain model precipitation has been suppressed by null echoes during the cycling window. Once a null echo is assimilated it is relatively difficult to initiate precipitation again in this case, for two reasons: in the first place because with this kind of stratiform precipitation only a fraction has a reflectivity $Z > 20\ dBZ$ that is required for humidity assimilation and moreover because once a reflectivity $Z > 20\ dBZ$ is actually measured assimilation of $RH = 88\%$ (as 95% corresponding to $Z > 35\ dBZ$ is unlikely) is not close enough to saturation to produce precipitation in the absence of strong vertical motions as is the case with convective precipitation.



(a) CTRL (b) DA

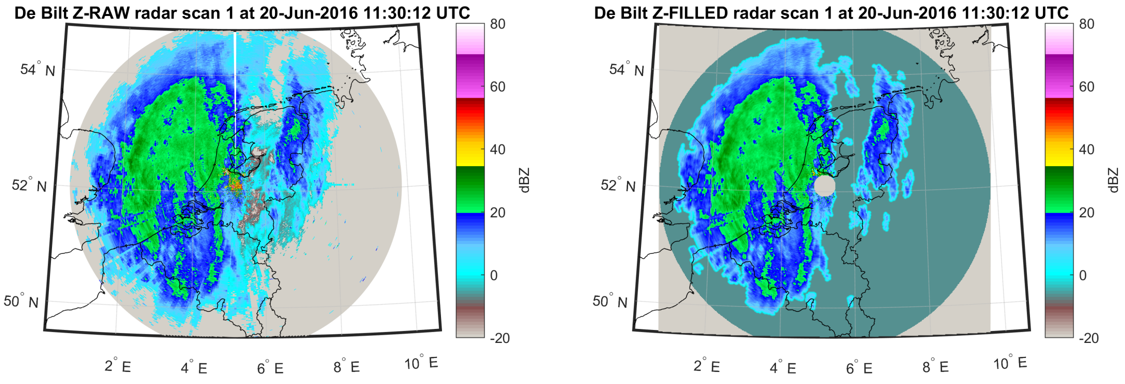


(c) RADAR

Figure 5.30: 15 minute average rainfall rate for CTRL, DA and radar at 30-08-2015 between 19:00 and 19:15 UTC in the STRAT case.

To improve the forecasting skill of the STRAT-DA case without using a different model background error covariance matrix (either by rescaling variance and lengthscales or using flow dependent methods) the most logical approach would be to reduce the extent of the null echo assimilation. In the STRAT case the assumed lapse rate of reflectivity in the filling of filtered radar reflectivity images as explained in section 3.1.5 is too steep, since precipitation intensity is really spatially uniform. This way at some points a null echo is assimilated while there is still a weak reflection from drizzle present in the raw radar image that has been filtered but nonetheless indicates that a null echo, which implies no precipitation at all is present, is not correct and the corresponding $RH = 90\%$ assimilation would be inappropriate. This point is illustrated with the radar images in figure 5.31. In the raw image in figure 5.31a a lot of weak reflections $< 20\ dBZ$ are present in between the large area of precipitation covering the west part of the domain and the small preceding band that are evident in the filtered and filled image in figure 5.31b. In the filled image the strip with $Z < 0\ dBZ$ in between these areas of more intense precipitation are flagged as null echoes, while in the raw image there are still a lot of weak reflections can be found here. One way to mitigate these false null echoes is to decrease the lapse rate or to similarly only assimilate null echoes at a minimum distance from existing precipitation. Alternatively, the filtering method will have to be improved to keep areas of low reflectivity from drizzle yet still remove the clutter that remains after the minimum reflectivity filter step. Any clutter that is not removed by the filtering might then prevent a null echo from suppressing erroneous model precipitation.

Analogous to experiment 1, the forecast skill of each case in experiment 2 is assessed by examining the GSS throughout the 6 hour verification forecast. The GSS has been plot for the MCS, LCS and STRAT case in figure 5.32. For each case the CTRL forecast is plot (with dashed lines) for reference to provide a baseline of the performance of each case without any data assimilation. Evidently in the MCS and LCS cases of strong convection there is a significant improvement when observations are assimilated, regarding the significantly higher GSS than the reference. As already mentioned before when comparing rainfall intensity images of the STRAT case in figure 5.30, in the STRAT case data assimilation actually results in a worse GSS score than the CTRL forecast. In the first because place the precipitation forecast is already quite good in the CTRL forecast and secondly, data assimilation results in unrealistic small scale features in model humidity. The frequent assimilation of radar reflections introduces local variations that are inappropriate for this case where



(a) Raw image

(b) Filtered and null echo filled radar image

Figure 5.31: Radar images at 11:30 UTC 20-06-2016, scan 1 of the DB radar. Used in the assimilation of the STRAT case.

the precipitation should be really stratiform. Also, ZTD measurements were found to produce unrealistic intensified rainfall.

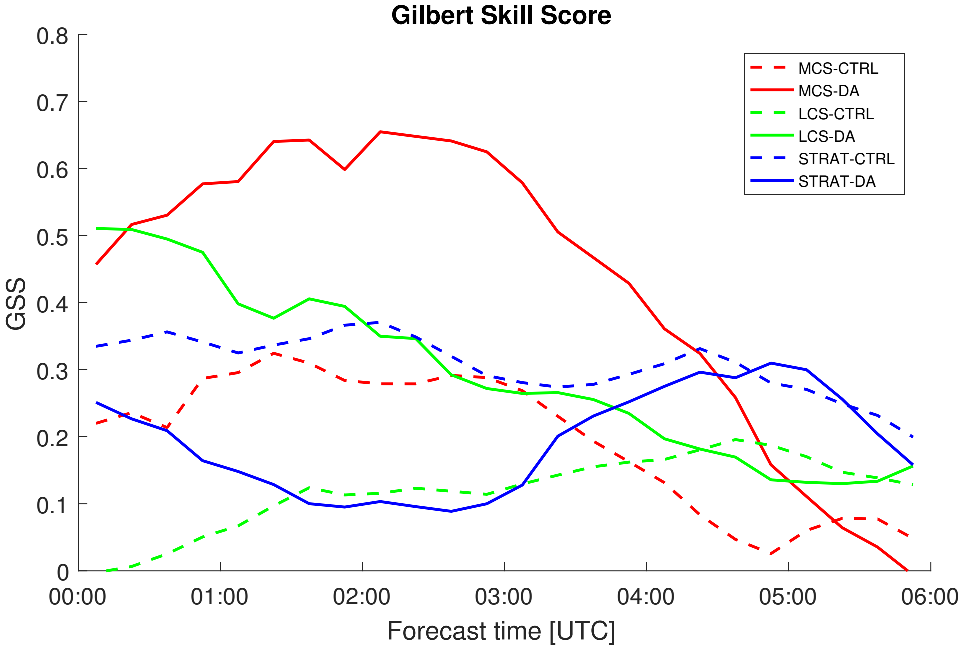
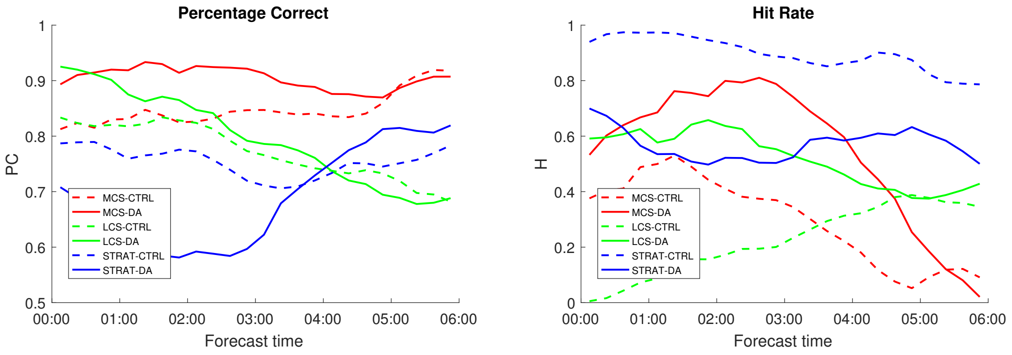


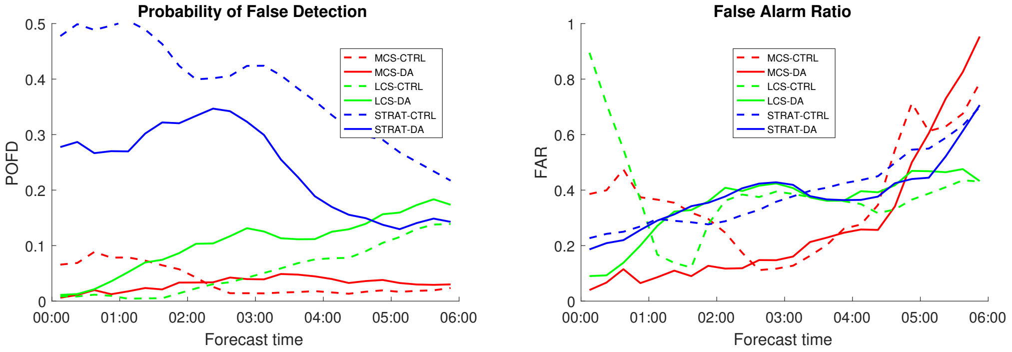
Figure 5.32: GSS of the precipitation forecast for each of the three cases MCS, LCS and STRAT. For each case the data assimilated scenario DA is plot together with the reference CTRL scenario.

Other scores than GSS that are shown in figure 5.33 provide a similar image. For both convective cases MCS and LCS the PC (figure 5.33a) is relatively high because only a fraction of the domain is covered with precipitation, so that PC is dominated by true negatives and does not really give an idea on the skill in forecasting observed precipitation. Consequently the difference between the CTRL and DA forecast is limited, although DA forecasts outperform CTRL scenarios. In the STRAT case the rainfall area is much more extensive so that the PC is more influenced by the how well observed precipitation is forecast. Judging from the H in figure 5.33b the STRAT-CTRL case forecasts pretty much all of the observed precipitation, because in the STRAT-CTRL forecast the extent of precipitation is so large that it covers all of the observed precipitation, see the rainfall intensities for STRAT-CTRL and radar at 13 UTC in figures 5.30a and 5.30c. The smaller PC for STRAT-CTRL is actually caused by the significant number of false positives, as indicated by the large POFD in figure 5.33c. Logically the CB in figure 5.33e also shows the STRAT-CTRL experiment really overestimates the extent of precipitation. In the STRAT-DA forecast the area of precipitation is much smaller, reducing false positives and thus POFD, however at the same time failing to forecast some areas of precipitation mainly at the eastern side of the front, leading to more false negatives and thus reducing the H and PC even more for the STRAT-DA scenario. Although the total sum of false positives and negatives is larger for the STRAT-DA case, they are roughly similar in quantity so that the STRAT-DA forecast does not significantly overpredict nor underpredict rainfall.



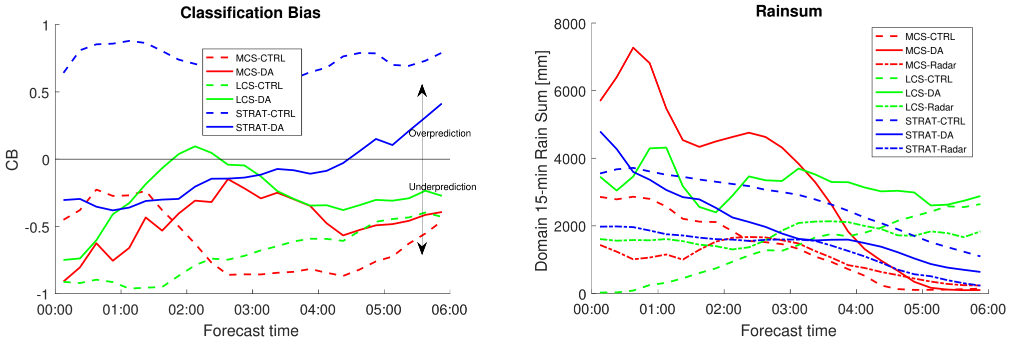
(a) percentage correct (PC)

(b) hit rate (H)



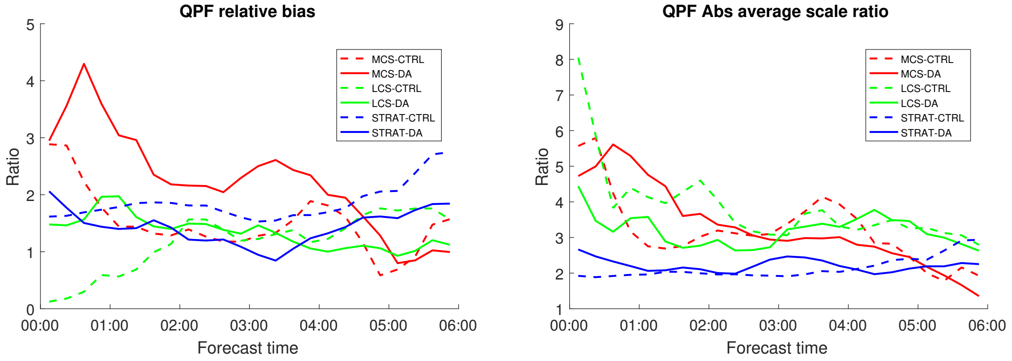
(c) probability of false detection (POFD)

(d) false alarm ratio (FAR)



(e) classification bias (CB)

(f) Domain precipitation sum



(g) Relative bias

(h) Average scale ratio

Figure 5.33: Scores derived from the confusion matrix of the model precipitation forecasts for each of the 3 precipitation cases: MCS, LCS and STRAT. For each of these cases the reference CTRL and data assimilated forecast DA are shown as a function of the forecast time.

Just like the GSS, the H is significantly improved with data assimilation for the MCS and LCS cases. In both MCS and LCS cases the number of false positives and false negatives are limited, as evident from the FAR (figure 5.33d) and POFD (figure 5.33c), so that the H, CSI and GSS are dominated by the large number of correct positives and thus follow a similar course throughout the verification forecast. The biggest improvement in qualitative forecast skill with respect to the CTRL scenario is for the LCS case of local convection, primarily because the LCS-CTRL forecast hardly produced any precipitation. The highest absolute forecasting skill is achieved with data assimilation in the MCS case with convection along the squall line. Especially in this case of mechanical forcing by the cold front non humidity related observations like SYNOP and radar radial velocity wind speeds add quality to the forecast. In the LCS case the improvement over the LCS-CTRL scenario mainly resulted from assimilation of radar retrieved humidity.

In terms of QPF, generally precipitation intensity is overestimated. The total 15 minute rainsum for every case and scenario combination is shown in figure 5.33f. As discussed in the verification of experiment 1 DA in the MCS-DA case with the squall line results in a significant overestimation of precipitation. The assimilation of radar retrieved humidity adds even more moisture to the MCS-CTRL forecast which already overestimates precipitation intensity in the first two hours of the verification forecast. The relative bias (RB) in figure 5.33g shows a strong tendency of the correct positives to overestimate rainfall intensity for the MCS-CTRL and even more so for MCS-DA. Also the ASCR in figure 5.33h indicates that the average absolute rainfall intensity ratio between model and precipitation intensity estimates is really high, specifically in the first two hours after assimilation. As noted in the verification of experiment 1, the microphysical parametrization fails to produce accurate quantitative intensity estimates in the somewhat extreme case of the squall line that is examined. The situation is less dramatic for the LCS case, however the total precipitation is still overestimated in the LCS-DA scenario compared to the total amount of rainfall regarding the total rainfall according to the radar image (LCS-Radar). The over- and underestimation of rainfall intensity however is now more balanced regarding the fact that the RB is a lot closer to 1 than was the case for MCS. In absolute terms the rainfall intensity differs with a ratio 3-4 (taking into account both under- and overestimation) from the radar measured intensity. Again it proves difficult for the model microphysics to accurately estimate precipitation quantity in a convective regime. In the STRAT case the precipitation formation is more gradual and not as explosive as in the cases of strong convection which have been discussed before, making the forecast less prone to large discrepancies with radar measured rainfall. This shows in the smaller difference in total rainsums between the STRAT-DA experiment and the radar. While the model in both STRAT-CTRL and STRAT-DA still tends to overestimate intensity, the absolute average scale ratio is smallest of all cases. The fact that this case concerns stratiform precipitation also is more forgiving if precipitation is mispositioned. Because precipitation intensity in the neighbourhood is of similar magnitude the difference in intensity does not really differ, while a perfectly modelled but misplaced convective cell could indicate large intensity differences just because their cores do not overlap.

6

Conclusion & Recommendations

6.1. Conclusion

The primary goal of the research conducted in this thesis is to improve the quality of precipitation forecasting, with a focus on the short term range from 0 to 6 hours ahead in time. In order to improve forecast quality various observations of the atmosphere are integrated into a 4km resolution Weather Research and Forecasting model (WRF) mesoscale numerical weather prediction model in a domain covering the Benelux. Observations from a variety of platforms are combined with a first guess model state using three dimensional variational assimilation (3DVar) as implemented in the Weather Research and Forecasting model's Community Variational/Ensemble Data Assimilation System (WRFDA) in order to obtain a more accurate model representation of the atmosphere. Especially data assimilation with observations from remote sensing platforms such as Doppler weather radar and GPS has been studied in detail because of the high spatial and temporal resolution that these observations provide. The assimilation of these high resolution observations is required to add the detail primarily in terms of humidity that is necessary for accurately forecasting precipitation timing and location into a limited area NWP model that is generally initialised with significantly coarser resolution data from a global NWP model.

Several experiments have been conducted to assess the influence of assimilating observations on the precipitation forecast. In every single experiment observations have been assimilated with a 15 minute interval during a 3 hour cycling window. The first experiment defines several data denial scenarios, where observations from only a single source or type are assimilated, isolating their effect on the subsequent forecast and enabling an assessment of the influence on the precipitation forecast after assimilation. Additionally the second experiment is designed to examine the performance of data assimilation in different mechanisms that drive precipitation formation. Three cases have been selected in which the first examines a mesoscale convective system (MCS) with a squall line along a cold front, the second focuses on a case of intense convection resulting from local instability (LCS) and finally the third case features stratiform precipitation formed by slowly ascending moist air along a warm front.

In the first data assimilation scenario only so called conventional observations have been assimilated. These include observations from permanent SYNOP weather stations, METAR airport measurements and a limited amount of soundings and aircraft reports that provide direct measurements of atmospheric model variables like temperature, pressure, humidity, and wind speed. Assimilation of these conventional observations alone in the MCS case of a squall line resulted in a minor improvement of forecasting skill in the verification forecast, primarily because of a better model description of wind convergence near the surface, leading to an extension of the squall line that corresponded better with radar observations. Generally though the influence of predominantly surface observations on precipitation is limited since it does not directly affect precipitation formation at higher altitude. In the case of a squall line the effect of assimilating conventional observations is notable because there is a strong forcing from the cold front near the surface. In the absence of such a strong mechanical forcing the ability of surface observations to affect precipitation is limited to for example increasing potential instability from the surface as warm and moist observations are assimilated.

Assimilation of hydrometeor mixing ratios estimated from radar reflectivity in the second observation scenario proved to be ineffective in creating sustained precipitation. The assimilated rainwater, snow or graupel do not have a noticeable affect on precipitation formation and once precipitated to the surface their added value to the forecast is lost completely. Indirect assimilation of radar reflectivity through near saturated relative humidity on the other hand produced a much better model representation of the squall line that is also expressed in the forecast skill during the verification forecast that is best out of all scenarios. The assimilation of assumed near saturated relative humidity provides a favourable environment for new precipitation to develop. The rapid data assimilation cycling however significantly overestimated rainfall intensity. While in the control forecast the rainfall intensity is already overestimated, the numerous successive radar humidity assimilations further add moisture to the model, aggravating the overestimation of precipitation intensity. The areal extent and rainfall intensity are sensitive to the relative humidity value that is assumed for the in cloud humidity. Setting a larger near saturation humidity increases the total moisture content, allowing for more extensive precipitation areas but at the same time at the expense of realistic intensity quantities. Besides reflectivity also radar radial velocity observations have been assimilated in a separate scenario. Similarly to wind observations from surface stations the assimilation of radial velocity yields an improved analysis of wind convergence at altitudes above the surface, resulting in an enhanced forecast skill on top of the conventional observations. The wind analysis might improve even more when aliased velocities from the middle radar scans would be successfully de-aliased to avoid rejection by the maximum innovation test.

In the last isolated observation scenario ZTD from GPS is assimilated as another source of water vapour observations to provide a constraint to radar humidity observations. While the assimilation of GPS ZTD measurements produced a more accurate representation of the squall line itself, also some erroneous areas of precipitation were initiated that mask the improvement in the forecast skill scores. Generally though the ZTD observations still lack decent precision because of the complex retrieval method, which in combination with the small scaled covariance length of humidity can lead to an unrealistically spatially varying moisture field, even with a reduced maximum innovation threshold. Also, ideally O-B statistics of a long range of independent ZTD assimilations should be examined to validate the bias correction for the model and receiver height discrepancy and assess whether there is still a bias after the correction is applied.

The final scenario in the first experiment combines all observations in the assimilation. There is an initial drop in the forecasting skill similar to the scenario where radar retrieved hydrometeors were assimilated, corresponding to the fallout of hydrometeors. After the drop the forecast skill is comparable to the scenario of just radar humidity assimilation, except for some time in the beginning of the forecast where radar humidity is limited by the GPS observations.

In the second experiment it is examined how data assimilation affects precipitation forecasting under different circumstances in which rain is formed. The first case featuring a cold front squall line has been studied in detail in the first experiment, where the assimilation of wind observations and mainly radar humidity significantly improved the extent and further development of the squall line. Being a mesocale system together with the fact that the main driver of precipitation is the mechanical forcing by the cold front enables nearly all assimilated observations to have a notable impact on the forecast. With a minor exception for GPS ZTD in the beginning of the forecast all observations improve the precipitation forecast for at least three hours after assimilation with respect to the control forecast. Especially humidity observation from radar reflectivity adds skill to the precipitation forecast, only with a drawback of a bigger overestimation of intensity.

The added value of assimilating each type of observations is less evident in the second case of local convection. Precipitation is produced by a limited amount of individual elevated convective cells initiated from the mixed layer, so that only observations at higher altitudes can impact the model precipitation forecast. Assimilation of GPS measurements has a minor impact, however only when radar reflections are measured and radar estimated humidity is assimilated, new and sustained cells of convection are triggered, again leading to a significant improvement in forecasting skill similar to the MCS case.

When observations are assimilated in the case of stratiform precipitation the precipitation forecasting skill is actually degraded. The rapid assimilation of moisture observations from radar and GPS results in unnecessary high frequent features in the humidity field, disturbing the homogeneous moisture field in the model that already predicts the stratiform rain quite well. Especially the relatively uncertain GPS measurements introduce local precipitation cores with strong rainfall intensity that is really unrealistic for this type of precipitation. Regarding the large scale and subtle nature of the warm front and its associated precipitation a much larger covariance scale for humidity would be appropriate for this type of forcing. Additionally the steady formation of precipitation from slowly ascending air does not require as much rapid updates as the other more turbulent, convective cases.

6.2. Recommendations

In this research the added value of variational data assimilation of weather observations to the short term precipitation forecast has been demonstrated. Especially the assimilation of radar reflectivity derived humidity observations significantly improved the precipitation forecast. Nevertheless, still several limitations for this method of radar data assimilation remain. Now the formation of new precipitation is expected to be a result of the assimilation of near saturated relative humidity, which combined with the rapid update cycling strategy generally adds quite a lot of moisture to the model, leading to a significant overestimation of rainfall intensity. Whether or not new precipitation forms is sensitive to the assumed relative humidity that is assimilated. In the current method just two different values for the near saturation humidity can be assimilated, depending on the associated reflectivity (see 2.91). Instead of these discrete intervals, a continuous function for relative humidity as a function of reflectivity could help reduce the sensitivity to the assumed relative humidity thresholds. Moreover, as shown in the pseudo single observation test with a radar reflectivity measurement, the temperature increment associated with latent heat release is minimal. By increasing the magnitude of this increment the buoyancy and the associated updrafts producing precipitation are strengthened, similar to so called latent heat nudging methods. This way, not all of the precipitation has to be initiated solely by adding moisture, reducing the total amount of moisture added to the model by the rapid update cycling.

Another way to limit the overestimation of rainfall is to add additional constraints to the humidity increments from radar. In this study that has been attempted by assimilation of water vapour estimates from GPS ZTD measurements. However these still have their limitations, such as the fact that it is a measure for the integrated water vapour and thus contains no information on the vertical distribution. Also, while the network of observations is quite dense, still a significant amount of radar humidity observations are not constrained by any nearby ZTD measurements, especially at sea. Including other sources of water vapour observations is required to provide humidity constraints in these areas, as well as a general constraint on the vertical profile of humidity. For example satellite observations can provide estimates of humidity at higher levels of the atmosphere, although assimilation of radiance is rather complex because of the highly non linear and long observation operator coupled with the radiative transfer model, the time varying bias correction that is required and the fact that usually only cloud free radiances can be assimilated. Other possible sources of humidity observations include aircraft, lidars, ceilometers and soundings. More research is needed to assess the ability of these types of observations to constrain the radar humidity increments.

Contrary to the humidity from radar observations, assimilation of hydrometeors estimated from radar reflectivity proved to have a negligible impact on the subsequent forecast, the main reason being that a feedback of hydrometeors to the model is missing. While the estimation of hydrometeor quantities itself could be improved e.g. by using dual polarization properties to better determine the partitioning, it is unlikely that this will notably improve the precipitation forecast. On the other hand, using dual polarization properties the filtering of radar reflectivity can be significantly improved, e.g. through neural networks. That way, discarding areas of low reflectivity which lead to inappropriate null echoes in areas of actual low intensity rainfall in the STRAT case is not required anymore.

While the assimilation of radar derived humidity showed to improve short term precipitation forecasting, as appeared in the verification of the case of stratiform precipitation the forecast skill can also be negatively affected. The covariance length that was downscaled to allow for detailed model variable increments in case of local precipitation is misplaced however in the case of stratiform precipitation which generally is of much larger scale. This shows the need for a flow dependent description of error covariance rather than the static error matrix used in this study, which provides an estimate of forecast specific uncertainty. That way, the increments in model variables conform to the case specific variability associated with each individual forecast. Then however it remains a challenge to acquire a background error covariance matrix that is representative of case specific uncertainty. Important problems like how to define ensembles (e.g. from different models or varying physical parametrizations), finding an adequate ensemble size and forecast length, reducing spurious long distance covariance (localization) and mitigating sampling errors should be investigated to optimally make use of ensembles in data assimilation. Nevertheless, hybrid methods that combine the robust, static description of error covariance from variational data assimilation with the flexible character of ensemble based covariance estimates have become increasingly popular and now form the basis of many operational data assimilation systems. Many NWP centres also have combined the hybrid approach with some adaptation of 4DVar, in which the evolution of the NWP model itself acts as an additional constraint for the observations. Additionally, a benefit of 4DVar is the ability to use precipitation accumulations as an observation to which the model should adhere to, which could help reduce the overestimation of intensity. There are numerous studies proving the added value of hybrid data assimilation over the purely static and flow dependent methods and of 4D over 3D methods, however the number of studies that make a comparison between these various types of data assimilation is trailing. A thorough assessment of the benefits and limitations of these methods also considering their increased computational demand would be helpful to define an optimal data assimilation setup that can be applied operationally. Additionally, most 4D hybrid methods are run for large area and global models at coarser resolution than 4 km to make them computationally feasible. More tests using these 4D hybrid methods assimilating radar observations in high resolution models only covering a limited area have to be performed in order to assess their ability to improve models at the convective scale.

Acronyms

3DVar three dimensional variational assimilation. iii, 2, 3, 12, 17, 28, 35, 81, 105, 135, *Glossary:* three dimensional variational assimilation

4DVar four dimensional variational assimilation. 2, 3, 28, 81, 105, 137, *Glossary:* four dimensional variational assimilation

AFWA Air Force Weather Agency. 7, 146

AIREP aircraft report. 82

ARPS Advanced Regional Prediction System. 35

ASCR absolute average scale ratio. 121, 133

BFGS Broyden–Fletcher–Goldfarb–Shanno. 14, *Glossary:* Broyden–Fletcher–Goldfarb–Shanno algorithm

BLUE best linear unbiased estimator. *Glossary:* best linear unbiased estimator

CAPE convective available potential energy. 87, *Glossary:* convective available potential energy

CB classification bias. 93, 119, 120, 133, 134

CCN cloud condensation nuclei. 66

CIN convective inhibition. 87, *Glossary:* convective inhibition

CLM4 Community Land Model version 4. 76, 77

CP convective parametrization. 1, 2, 67, 68

CSI critical succes index. Also known as threat score (TS). 93, 94, 118, 120, 133

DA data assimilation. xv, 45, 47, 57, 87, 96, 98, 99, 102, 103, 105, 108, 112, 113, 128, 133, *Glossary:* data assimilation

DB De Bilt. xi, xiii, 50, 52–54, 56, 57, 102, 105, 107, 108, 125, 132

DH Den Helder. xiii, 50, 52, 53, 102, 105, 107, 108

DMI Danmarks Meteorologiske Institut. v, *Glossary:* Danmarks Meteorologiske Institut

DSD drop size distribution. 28, 29, 31, 66, 78

DWD Deutsche Wetterdienst. 2, *Glossary:* Deutsche Wetterdienst

E-GVAP EUMETNET GPS Water Vapour Programme. v, 58

EC Environment Canada. 2, *Glossary:* Environment Canada

ECMWF European Centre for Medium-Range Weather Forecasts. 2, 5, 7

EDMF eddy diffusivity mass flux. 72

EOF's empirical orthogonal functions. ix, 19, 21–23, 40, *Glossary:* empirical orthogonal functions

FAR false alarm ratio. 93, 119, 120, 133, 134

FG first guess. 95, 96, 98, 99, 102–108, 112, 126, *Glossary:* first guess

GCM general circulation model. 81, 141, 143, 146

GFS Global Forecast System. 5, 63, 82, 86, 89, 118, *Glossary:* Global Forecast System

GFZ Geo Forschungs Zentrum/German Research Centre for Geosciences. 58

GMTED2010 Global Multi-resolution Terrain Elevation Data 2010 . 62

GPS Global Positioning System. iii, xi, 3, 5, 39, 40, 45, 46, 58–60, 81, 83, 95, 108, 110–116, 119, 122, 124–128, 130, 135–137

GSS Gilbert skill score. Also known as equitable threat score (ETS). xiii, 93, 94, 118, 119, 122, 131–133

H hit rate. Also known as Probaility of Detection (POD). xiii, 93, 119, 120, 132–134

HIRLAM High Resolution Limited Area Model. 7, *Glossary:* High Resolution Limited Area Model

HSS Heidke skill score. 94

IN ice nuclei. 66

IWC ice water content . 79

IWV integrated water vapour . xiii, 3, 87, 88, 108, 110–116, 122–128

JAB Jabbeke. xi, 50–55, 57, 102, 103, 123, 125, 128

KMI Koninklijk Meteorologisch Instituut . v, xv, 49, 50

KNMI Koninklijk Nederlands Meteorologisch Instituut. v, xiii, xiv, 28, 34, 37, 49, 58, 82, 92, 103, 105, 107, 117, 156–159

LBC lateral boundary conditions. 63

LCL lifting condensation level. 35, 79, *Glossary:* lifting condensation level

LCS local convective system. xiii, xv, 86, 88, 122, 130

LES large eddy simulation. 62, 72

LFC level of free convection. 83, 144

LSM land surface model. xii, 73

LWC liquid water content . 30, 31, 66, 79

MCS mesoscale convective system. xiii, 3, 67, 83, 86, 88, 122, 125, 129, 130, 135, 136

METAR meteorological aerodrom report. 82, 95, 96, 98, 122, 135

MM5-3DVAR Community fifth generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) Three Dimensional Variational Assimilation (3DVAR) system. 7, 31

MP Marshall-Palmer reflectivity-rainfall rate relation (Z-R). 30

MYNN Mellor–Yamada–Nakanishi–Niino. 72

NCAR National Centers for Atmospheric Research. 7, 76, 146, *Glossary:* National Center for Atmospheric Research

NCEP National Centers for Environmental Prediction. 2, 7, 63, *Glossary:* National Centers for Environmental Prediction

NMC National Meteorological Center. 22, 31, 33, 34

NOAA National Oceanic and Atmospheric Administration. 7, 63, 144, 146

NRL Naval Research Laboratory. 7, 146

NWP Numerical weather prediction. iii, ix, 1–3, 8, 27, 28, 40, 49, 63, 67, 69, 77, 78, 81, 82, 89, 105, 135, 137, 145, 146, *Glossary:* numerical weather prediction

O-A observaton minus analysis. xiii, 41, 44, 108, 110, 111, 143, *Glossary:* analysis residual

O-B observation minus background. xi, xiii, 41, 44, 47, 108, 111, 136, 145, *Glossary:* innovation

OI Optimal interpolation. *Glossary:* Optimal Interpolation

OU Oklahoma University. 7, 146

PBL planetary boundary layer. xi, xii, 68–74, 77, 78, *Glossary:* planetary boundary layer

PC percentage correct. 93, 120, 132–134

PDF probability density function. 15, 78, 147, *Glossary:* probability density function

PFT plant functional type. 76

POFD probability of false detection. 93, 133, 134

PRF pulse repetition frequency. 38

PSOT pseudo single observation test. x, xi, 3, 5, 26, 40–47, 102, *Glossary:* pseudo single observation test

QPF quantitative precipitation forecasting. 66, 91, 94, 121, 133

RB relative bias. 121, 133

RH relative humidity. xiii, 36, 45–47, 78, 79, 102, 104, 105, 112–116, 124, 126–128, *Glossary:* relative humidity

ROB Royal Observatory of Belgium. 58

RRTM rapid radiative transfer model. 81

RRTMG rapid radiative transfer model for GCMs. 81

SGN Institut Géographique National. 58

STRAT Stratiform precipitation case that is investigated. xiii, xv, 122, 125, 130, 131

SYNOP surface synoptic observation. 82, 95–98, 110, 112–116, 122–127, 133, 135

TITAN Thunderstorm Identification, Tracking, Analysis and Nowcasting. Nowcasting radar extrapolation algorithm developed by Dixon and Wiener (1993). 1

TKE turbulent kinetic energy. 72

TOA top of the atmosphere. 79

UKMO United Kingdom Meteorological Office. v, xii, 2, 7, 58, 83, 84, 86–88, 90, *Glossary:* UK Met Office

VAD velocity azimuth display. 56

WID Wideumont. 50, 53, 102

WMO World Meteorological Organisation. 82

WPS WRF Preprocessing System. 62, *Glossary:* WRF Preprocessing System

WRF Weather Research and Forecasting model. iii, xi, xii, 3, 7, 34, 40, 59, 61–64, 66, 68, 76, 77, 79, 83, 86, 88, 89, 91, 135, 142, 143, 146, *Glossary:* Weather Research and Forecasting model

WRFDA Weather Research and Forecasting model's Community Variational/Ensemble Data Assimilation System. iii, 3, 7, 10, 14, 18–20, 22, 23, 26, 31–35, 40, 41, 64, 135, 152, *Glossary:* Weather Research and Forecasting (WRF) model's Community Variational/Ensemble Data Assimilation System

ZHD zenith hydrostatic delay. 39, 40, 58

ZTD zenith total delay. iii, vii, xi, xiii, 3, 39, 40, 45–47, 58–60, 81, 83, 95, 108, 110–116, 119, 122–124, 126–128, 132, 136, 137

ZWD zenith wet delay. 39, 40, 58, 83, 108

Glossary

WRF Preprocessing System Preprocessing system used to convert geographic and meteorological data from other sources to initialize a WRF model. 62, 142

absolute average scale ratio Absolute average scale ratio of two precipitation intensity estimates. 133

adjoint The transpose of a matrix or vector. The adjoint turns rows into columns and vice-versa. Some operators are effectively matrices (linear operators), but cannot, due to storage limitations, be stored explicitly. The adjoint operator must then be formulated as a separate operator either manually or by automatic adjoint software. In 4d-Var, adjoint operators (matrices) help integrate the adjoint variables (cost function gradient vectors) backwards in time. ix, 8

analysis Model state that is the result of minimising the cost function. It is the best (most likely) estimate of the real atmospheric state given a first guess and a set of observations. ix, x, 6, 7, 9–12, 14–18, 20, 21, 27, 28, 31, 32, 34, 40–42, 49, 57, 95–103, 105, 106, 108, 110, 112–115, 123–127, 143, 148, 150

analysis error Difference between the data assimilation analysis and true state. 16

analysis residual Departure of observation with respect to model analysis O-A after data assimilation. xiii, 41, 141

background Model state used as input for the cost function to give a first estimate of the real atmospheric state. iii, ix, 6, 10, 14–17, 19, 20, 23, 24, 26, 31–35, 40–45, 49, 82, 97, 122, 143

background error Difference between the model background and true state. 16–19, 27, 28

background error covariance matrix Covariance matrix describing the 3D-model background error covariance. iii, ix, x, xv, 8, 19, 20, 23–25, 40, 41, 44–46, 131

best linear unbiased estimator Maximum likelihood solution of a system of equations. It minimizes the variance of the solution. 139

Broyden–Fletcher–Goldfarb–Shanno algorithm A quasi-Newton iterative minimization algorithm. 14, 139

cold start Initialisation of a GCM from observations or coarser model data. 82

condition number The ratio of the highest to the smallest eigenvalues of the Hessian. In a perfectly preconditioned system of equations, the preconditioning number is unity. 13

confusion matrix Table that shows how predicted instances of a class occur compared to observed instances of a class. xv, 92, 93, 143

contingency table Table showing the frequency distribution of variables. Often used interchangeably with confusion matrix , although the contingency table can have more dimensions/variables. 93

control variable transform Transformation of variables that is often used for preconditioning and to avoid matrix inversion. 18, 19

control variables Variables that are actually minimized in the cost function by the minimization/descent algorithm. These are often not the meteorological variables in model space, but are transformed variables in a new vector space that are ideally uncorrelated. The transformation to control variables is part of the preconditioning process. 19–21, 23, 25, 28, 34, 35, 42, 44

convective available potential energy Potential energy available for convection. In essence the vertically integrated positive buoyancy of an air parcel. Indicator of atmospheric instability and potential severe weather. 87, 139

convective inhibition Amount of energy that prevents an air parcel from reaching its LFC. In essence the vertically integrated negative buoyancy of an air parcel, from the surface to LFC. Indicator of atmospheric stability. 87, 139

cost function A function measuring the discrepancy between a guess of the model state and the background, and the guess (via the observation operator) and the observations themselves. The cost function is minimised iteratively in variational data assimilation. ix, 5, 10–15, 17, 18, 21, 22, 24, 28, 30, 31, 33–35, 37, 40, 43, 148

Cressman analysis Data assimilation method where model is equal to observation at observation location. 6

cycling Alternating data assimilation and forecast several times. 81, 82, 89

data assimilation Statistical method to obtain the best estimate of model state variables. ix, x, xii–xv, 2, 3, 5–8, 10, 11, 13, 16, 17, 19–23, 27, 28, 32, 42, 43, 46, 49, 55, 60–62, 81–83, 86–89, 95–99, 102, 105, 106, 108, 110, 112–118, 121–127, 129–133, 135–139, 143, 144, 147, 156–159

Deutsche Wetterdienst National weather service of Germany. 2, 139

ellipsoidal height Height with respect to an ellipsoidal model of the Earth's surface, as used by GPS. xi, 59, 60

empirical orthogonal functions A set of empirically found orthogonal basis of vectors of which each vector describes most of the residual variance after using the previous eigenvectors to describe variance. ix, 19, 21, 139

ensemble A group of forecasts each having been slightly perturbed. These perturbed forecasts can be used to estimate the uncertainty of the model forecast. 7, 26, 27

Environment Canada National weather service of Canada. 2, 139

error Uncertainties of information pertaining to the system. All observations have errors as do all model states. Errors can be systematic (e.g. biases) or random. It is usual to correct for biases (if known) before the assimilation procedure starts and to assume that random errors are Gaussian in nature (this leads to the quadratic form of the cost function in 3D- and 4DVar). Errors are related to variances. x, 7, 8, 10, 11, 15–17, 19, 21–24, 26–28, 31–36, 40–45

first guess Model state used as input for data assimilation to give a first estimate of the real atmospheric state. xiii, 6, 7, 10, 12, 14, 95–98, 100, 103, 104, 106, 110, 112–115, 122–127, 135, 140

four dimensional variational assimilation Variational assimilation algorithm applied to a three dimensional model with the time evolving model as an additional constraint. ix, 2, 7–9, 139

Gaussian A widely used symmetrical model for the probability density function. Also often referred to as normal distribution. 15, 16

geoid Global sea level that the sea would attain in absence of external forcing like wind and tides. xi, 59, 60

geoid undulation Height difference between geoid and a reference ellipsoid. xi, 59, 60

geopotential Work that has to be done to raise a mass of 1 kg from sea level to a certain point in space. 144

geopotential height Geopotential normalized by global average gravitational acceleration to obtain units of height. 40

geostrophic wind Wind flowing along isobars as a result of a balance between pressure gradient force and Coriolis force. 6

Global Forecast System Operational global weather model operated and developed by NOAA. 5, 63, 140

Hessian Matrix containing all second order partial derivatives of the cost function, with respect to the control variables. 14

High Resolution Limited Area Model NWP model that is a collaboration of European meteorological intstitutes to develop. Also used to refer to the group of institutions involved. 7, 140

hybrid Combination of a static (based on climatology) and flow dependent (based on ensembles) description model background error covariance. x, 2, 7, 27, 28, 137, 138

hydrostatic Vertical balance between the downward gravitational force and the upward pressure gradient force. 6

innovation Difference of observation with respect to model equivalent (through the observation operator) background value O-B. x, xi, 20, 32, 36, 41, 44–46, 49, 141

integrated water vapour Total amount of water vapour in a vertical column of the atmosphere. 108, 125

Jacobian Matrix containing the first order partial derivatives of the cost function, with respect to the control variables. 10

lifting condensation level The height at which a parcel of air becomes saturated when it is lifted assuming a dry adiabatic lapse rate. 35, 140

localization Damping of ensemble derived covariances at large distances. 27, 28

multivariate A fully consistent system where correlations are taken into account, not just between values of a variable at different points in space, but also between different variables. For example pressure is also correlated with wind speed rather than only with pressure at other points in space. 15

National Center for Atmospheric Research U.S. national center for atmospheric research. 140, 142

National Centers for Environmental Prediction Range of U.S. national centers for forecasts regarding weather, storms, ocean, space, climate, environment, aviation. 2, 7, 22, 63, 141

nowcasting Short term (0 to 6 hours ahead) forecasting of precipitation. 1

numerical weather prediction Mathematical modelling of the atmosphere to make weather forecasts. iii, 7, 8, 58, 135, 141

observation operator Function that relates model variables to observation. It is used to estimate the model equivalent of an observation in the cost function, e.g. through physical relations, spatial interpolation, etc. Usually denoted with the symbol H. ix, 10–12, 14–17, 30

Optimal Interpolation Maximum likelihood solution of a system of equations. It minimizes the variance of the solution. 11, 12, 141, 150

orthometric height Height with respect to the geoid. xi, 59, 60

parametrization Simplified mathematical description of processes too small to resolve explicitly with model equations. iii, vii, xi, 1, 3, 27, 61, 64–74, 78, 80, 102, 121, 129, 133, 137, 139

planetary boundary layer Lowest layer of the atmosphere that directly interacts with the Earth's surface. 1

preconditioning The process of choosing new control variables which are (approximately) uncorrelated, and form a unit Hessian matrix. In a perfectly preconditioned system of equations, contours of constant cost function are circles in the new control variable space and the condition number is unity. 14, 143

probability density function Likelihood of obtaining a certain value of a continuous random variable as a function of that value when sampling. 10, 15, 141, 144, 147

pseudo single observation test Assimilation experiment using a single artificial observation. x, 40, 137, 141

reflectivity Backscattered radiation from an object as measured by a (weather) radar. iii, xii, xv, 2, 3, 17, 28, 30, 32, 34, 36, 49, 50, 52, 56, 82, 83, 85, 87–89, 91, 92, 102, 105, 108, 111, 112, 118, 122, 128, 131, 135–137

relative humidity The ratio of water vapour pressure to saturation vapour pressure. x, xi, xiii, xv, 6, 19, 24–26, 35, 36, 43–47, 63, 78, 79, 102, 103, 105, 125, 128, 136, 137, 141

resolution The length scales that are used in the mathematical discretisation of the Earth's atmosphere. 5

spinup time Time needed for a GCM to adjust and reach equilibrium after being initialised with observations or coarser model data. iii, 2, 3, 23, 82, 129

three dimensional variational assimilation Variational assimilation algorithm applied to a three dimensional model. iii, ix, 7–9, 18, 81, 135, 139

true state Best possible approximation of the real atmosphere by a model. 16

UK Met Office National weather service of the United Kingdom (UK). v, 2, 7, 58, 141

variance Expectation of the squared deviation of a random variable from its mean. Covariance of a variable with itself. Diagonal elements in the covariance matrix. x, 6, 10–12, 15, 16, 21–23, 26, 27, 33, 34, 40–44

variational data assimilation Multivariate statistical method to combine model and observations based on their uncertainty. iii, x, 2, 6, 10, 27, 28, 137

warm start Initialisation of a GCM with a data assimilated own model state. 82

Weather Research and Forecasting (WRF) model's Community Variational/Ensemble Data Assimilation System Community data assimilation system coupled with the WRF model. Released and maintained by a multiagency collaboration including NCAR, NOAA, AFWA, OU and NRL. iii, 7, 135

Weather Research and Forecasting model Community NWP model. Released and maintained by a multiagency collaboration including NCAR, NOAA, AFWA, OU and NRL. iii, 3, 7, 135, 142

A

Appendix A

A.1. Derivation of cost function using maximum likelihood

In data assimilation a model estimate of the atmospheric state is combined with observations, each having an a priori uncertainty characterised by their respective probability density function (PDF). The optimal estimate of the atmospheric state **x** taking into account these observations **y** is the state the most likely. In other words, it maximizes the so called posterior PDF of the new state estimate. In general the posterior PDF can be found by applying Bayes' theorem:

$$p\left(\mathbf{x}|\mathbf{y}\right) = \frac{p\left(\mathbf{y}|\mathbf{x}\right) p\left(\mathbf{x}\right)}{p\left(\mathbf{y}\right)} \propto p\left(\mathbf{y}|\mathbf{x}\right) p\left(\mathbf{x}\right) \tag{A.1}$$

Where $p\left(\mathbf{x}\right)$ is the probability of the model state **x**, $p\left(\mathbf{y}|\mathbf{x}\right)$ the probability of observations **y** given model state **x** and $p\left(\mathbf{y}\right)$ the probability of the observations, which does not depend on **x** and is therefore just a scaling constant and not relevant for finding the maximum likelihood. If it is assumed that the PDFs have a Gaussian distribution:

$$p\left(\mathbf{x}\right) = \frac{1}{(2\pi)^{N/2}|\mathbf{B}|^{1/2}} \exp\left\{-\frac{1}{2}\left(\mathbf{x}-\mathbf{x_b}\right)^T \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x_b}\right)\right\} \tag{A.2}$$

$$p\left(\mathbf{y}|\mathbf{x}\right) = \frac{1}{(2\pi)^{N/2}|\mathbf{B}|^{1/2}} \exp\left\{-\frac{1}{2}\left(H\left(\mathbf{x}\right)-\mathbf{y}\right)^T \mathbf{R}^{-1}\left(H\left(\mathbf{x}\right)-\mathbf{y}\right)\right\} \tag{A.3}$$

By applying Bayes' theorem and taking the logarithm, a function J can be defined for the posterior probability:

$$J\left(\mathbf{x}\right) = -\log\left[p\left(\mathbf{x}\right)\right] - \log\left[p\left(\mathbf{y}|\mathbf{x}\right)\right] + c \tag{A.4}$$

Where c is a certain constant. This constant be chosen such that the normalization term before the exponent of the Gaussian PDFs in A.2 and A.3 cancel out, leading to the definition of the cost function used in data assimilation:

$$J\left(\mathbf{x}\right) = \frac{1}{2}\left(\mathbf{x}-\mathbf{x_b}\right)^T \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x_b}\right) + \frac{1}{2}\left(H\left(\mathbf{x}\right)-\mathbf{y}\right)^T \mathbf{R}^{-1}\left(H\left(\mathbf{x}\right)-\mathbf{y}\right) \tag{A.5}$$

A.2. Linearisation of cost function

The cost function that measures the difference between analysis and background is defined as follows:

$$J(\mathbf{x}) = J_b + J_o = \frac{1}{2}(\mathbf{x}-\mathbf{x_b})^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x_b}) + \frac{1}{2}\left(H(\mathbf{x})-\mathbf{y}\right)^T \mathbf{R}^{-1}\left(H(\mathbf{x})-\mathbf{y}\right) \tag{A.6}$$

In the observation term J_o, non linear observation operator H is linearised at $\mathbf{x}_b$. This linearised operator $\mathbf{H}$ is used to approximate $H(\mathbf{x})$. The validity of this approximation relies on whether how close the background is to the analysis and the degree of non linearity in H.

$$H(\mathbf{x})-\mathbf{y} = H(\mathbf{x}_b + (\mathbf{x}-\mathbf{x}_b)) - \mathbf{y} \approx H(\mathbf{x}_b) + \mathbf{H}(\mathbf{x}-\mathbf{x}_b) - \mathbf{y} \tag{A.7}$$

If the expression from A.7 is used in A.6, the cost function is rewritten into:

$$2J(\mathbf{x}) = (\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \left(\left[H(\mathbf{x}_b)-\mathbf{y}\right] + \mathbf{H}(\mathbf{x}-\mathbf{x}_b)\right)^T \mathbf{R}^{-1}\left(\left[H(\mathbf{x}_b)-\mathbf{y}\right] + \mathbf{H}(\mathbf{x}-\mathbf{x}_b)\right) \tag{A.8}$$

The cross products in the second term in A.8 can be expanded to:

$$\begin{aligned} 2J(\mathbf{x}) = {} & (\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + (\mathbf{x}-\mathbf{x}_b)^T \mathbf{H}^T \mathbf{R}^{-1}\mathbf{H}(\mathbf{x}-\mathbf{x}_b) \\ & + \left[H(\mathbf{x}_b)-\mathbf{y}\right]^T \mathbf{R}^{-1}\left[H(\mathbf{x}_b)-\mathbf{y}\right] + \left[H(\mathbf{x}_b)-\mathbf{y}\right]^T \mathbf{R}^{-1}\mathbf{H}(\mathbf{x}-\mathbf{x}_b) + (\mathbf{x}-\mathbf{x}_b)^T \mathbf{H}^T \mathbf{R}^{-1}\left[H(\mathbf{x}_b)-\mathbf{y}\right] \end{aligned} \tag{A.9}$$

Next, using the rule for the tranpose of matrix products:

$$(AB)^T = B^T A^T \tag{A.10}$$

Together with the fact that a symmetric matrix like $\mathbf{R}^{-1}$ is equal to its transpose by definition, equation A.9 can be simplified to:

$$\begin{aligned} 2J(\mathbf{x}) = {} & (\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + (\mathbf{x}-\mathbf{x}_b)^T \mathbf{H}^T \mathbf{R}^{-1}\mathbf{H}(\mathbf{x}-\mathbf{x}_b) \\ & + \left[H(\mathbf{x}_b)-\mathbf{y}\right]^T \mathbf{R}^{-1}\left[H(\mathbf{x}_b)-\mathbf{y}\right] + 2\left[H(\mathbf{x}_b)-\mathbf{y}\right]^T \mathbf{R}^{-1}\mathbf{H}(\mathbf{x}-\mathbf{x}_b) \end{aligned} \tag{A.11}$$

The quadratic $(\mathbf{x}-\mathbf{x}_b)$ terms can be combined and after dividing by 2 the following expression for the cost function is obtained:

$$\begin{aligned} J(\mathbf{x}) = {} & \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \left(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1}\mathbf{H}\right)(\mathbf{x}-\mathbf{x}_b) \\ & + \frac{1}{2}\left[H(\mathbf{x}_b)-\mathbf{y}\right]^T \mathbf{R}^{-1}\left[H(\mathbf{x}_b)-\mathbf{y}\right] + \left[H(\mathbf{x}_b)-\mathbf{y}\right]^T \mathbf{R}^{-1}\mathbf{H}(\mathbf{x}-\mathbf{x}_b) \end{aligned} \tag{A.12}$$

A.3. Gradient of cost function

Minimizing the cost function found in A.12 is equivalent to solving for **x** for which the gradient of the cost function equals zero:

$$\nabla J(\mathbf{x}) = 0 \tag{A.13}$$

The cost function is a quadratic function, of the general form:

$$J(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c \tag{A.14}$$

Of which the gradient is equal to

$$\nabla J = \left(\mathbf{A} + \mathbf{A}^T\right)\mathbf{x} + \mathbf{b} \tag{A.15}$$

For a symmetric matrix **A** this is equal to

$$\nabla J = \left(\mathbf{A} + \mathbf{A}^T\right)\mathbf{x} + \mathbf{b} = 2\mathbf{A}\mathbf{x} + \mathbf{b} \tag{A.16}$$

The cost function equivalent of **A** is $\left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)$. It can be shown that if the background and observation error covariance matrices **B** and **R** are symmetric this also applies to **A** is $\left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)$:

$$\left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^T = \left(\mathbf{B}^{-1}\right)^T + \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^T = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}^{TT} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \tag{A.17}$$

Where also the transpose matrix product rule in A.10 is used. Recall the incremental formulation of the cost function:

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\delta\mathbf{x} + \left[H(\mathbf{x}_b) - \mathbf{y}\right]^T\mathbf{R}^{-1}\mathbf{H}\delta\mathbf{x} + \frac{1}{2}\left[H(\mathbf{x}_b) - \mathbf{y}\right]^T\mathbf{R}^{-1}\left[H(\mathbf{x}_b) - \mathbf{y}\right] \tag{A.18}$$

Applying A.16 to A.18 results in:

$$\nabla J(\delta\mathbf{x}) = \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\delta\mathbf{x} + \mathbf{H}^T\mathbf{R}^{-1}\left[H(\mathbf{x}_b) - \mathbf{y}\right] \tag{A.19}$$

Setting the gradient to zero results in the analytical solution for $\delta\mathbf{x}$

$$\delta\mathbf{x} = \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\left[H(\mathbf{x}_b) - \mathbf{y}\right] \tag{A.20}$$

A.4. Equivalence 3D-Var and Optimal Interpolation

In section 2.3.2 the exact solution for the analysis increment $\delta\mathbf{x}$ is shown to be (see 2.15):

$$\delta\mathbf{x} = \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\left[\mathbf{y} - H(\mathbf{x}_b)\right] \tag{A.21}$$

Next, it will be shown that this is equivalent to the Optimal Interpolation solution (see e.g. A.C. Lorenc (1986, p. 1180)), which is equal to:

$$\delta\mathbf{x} = \mathbf{B}\mathbf{H}^T\left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T\right)^{-1}\left[\mathbf{y} - H(\mathbf{x}_b)\right] \tag{A.22}$$

Proving the following equality:

$$\begin{aligned}
\left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1} &= \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\left[\mathbf{R}\left(\mathbf{H}^T\right)^{-1}\right]^{-1} \\
&= \left[\left(\mathbf{R}\left(\mathbf{H}^T\right)^{-1}\right)\left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\right]^{-1} \\
&= \left[\left(\mathbf{R}\left(\mathbf{H}^T\right)^{-1}\right)\mathbf{B}^{-1} + \mathbf{H}\right]^{-1} \\
&= \left[\left(\mathbf{R}\left(\mathbf{H}^T\right)^{-1} + \mathbf{H}\mathbf{B}\right)\mathbf{B}^{-1}\right]^{-1} \\
&= \left[\left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T\right)\left(\mathbf{H}^T\right)^{-1}\mathbf{B}^{-1}\right]^{-1} \\
&= \mathbf{B}\mathbf{H}^T\left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T\right)^{-1}
\end{aligned} \tag{A.23}$$

A.5. Control variable transformation

The original formulation of the cost function as in A.6 was:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x_b})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x_b}) + \frac{1}{2}\left(H(\mathbf{x})-\mathbf{y}\right)^T\mathbf{R}^{-1}\left(H(\mathbf{x})-\mathbf{y}\right) \tag{A.24}$$

By applying a linearisation (see appendix A.2) the cost function has been rewritten to:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\left(\left[H(\mathbf{x}_b)-\mathbf{y}\right]+\mathbf{H}(\mathbf{x}-\mathbf{x}_b)\right)^T\mathbf{R}^{-1}\left(\left[H(\mathbf{x}_b)-\mathbf{y}\right]+\mathbf{H}(\mathbf{x}-\mathbf{x}_b)\right) \tag{A.25}$$

By defining

$$\delta\mathbf{x} = \mathbf{x}-\mathbf{x}_b \tag{A.26}$$

The cost function takes the incremental form:

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{c}+\mathbf{H}\delta\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{c}+\mathbf{H}\delta\mathbf{x}) \tag{A.27}$$

Where **c** is a constant vector measuring the offset between the observations and model equivalents.

$$\mathbf{c} = H(\mathbf{x}_b)-\mathbf{y} \tag{A.28}$$

The control variable transform is defined as follows:

$$\delta\mathbf{x} = \mathbf{U}\mathbf{v} \tag{A.29}$$

Where **U** is the preconditioner. To make sure $\mathbf{B}^{-1}$ is eliminated from the cost function A.27, **U** should satisfy:

$$\mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{T/2} = \mathbf{U}\mathbf{U}^T \tag{A.30}$$

So **U** is the square root of the covariance matrix **B**.

$$\mathbf{U} = \mathbf{B}^{1/2} \tag{A.31}$$

Now it is shown that the control variable transform indeed removes the inverse covariance matrix from the cost function. By applying A.29 to the incremental cost function A.27 and noting that $\mathbf{B}^{-1} = \mathbf{U}^{-T}\mathbf{U}^{-1}$

$$J(\mathbf{v}) = \frac{1}{2}(\mathbf{U}\mathbf{v})^T\mathbf{U}^{-T}\mathbf{U}^{-1}\mathbf{U}\mathbf{v} + \frac{1}{2}(\mathbf{c}+\mathbf{H}\mathbf{U}\mathbf{v})^T\mathbf{R}^{-1}(\mathbf{c}+\mathbf{H}\mathbf{U}\mathbf{v}) \tag{A.32}$$

Rewriting the transpose of matrix products analogous to A.10:

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T\mathbf{U}^T\mathbf{U}^{-T}\mathbf{U}^{-1}\mathbf{U}\mathbf{v} + \frac{1}{2}(\mathbf{c}+\mathbf{H}\mathbf{U}\mathbf{v})^T\mathbf{R}^{-1}(\mathbf{c}+\mathbf{H}\mathbf{U}\mathbf{v}) \tag{A.33}$$

Of which the first term J_b reduces to

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}(\mathbf{c}+\mathbf{H}\mathbf{U}\mathbf{v})^T\mathbf{R}^{-1}(\mathbf{c}+\mathbf{H}\mathbf{U}\mathbf{v}) \tag{A.34}$$

This shows the control variable transform resulted in the elimination of $\mathbf{B}^{-1}$ from the cost function. Actually, it has been absorbed into **v**.

A.6. Linearisation water vapour mixing ratio operator

In the following it is shown how the observation operator for water vapour mixing ratio is linearised to be used in the (preconditioned) cost function A.34. It is aimed to provide more detail than the description by Wang et al. (2013), who developed this method of assimilating water vapour observations derived from radar observations in WRFDA.

The observation operator H converts model relative humidity RH to water vapour mixing ratio r_v using the water vapour saturation mixing ratio r_s:

$$r_v = RH \cdot r_s \tag{A.35}$$

In general atmospheric applications, a mixing ratio (e.g. for water vapour $r_v = \frac{m_v}{m_d}$) is approximately equal to a specific humidity ($q_v = \frac{m_v}{m_d + m_v}$ for water vapour), since the contribution of water vapour (and similarly for liquid water, snow, ice et cetera) to the total weight in a volume is negligible: $m_v \ll m_d$. Therefore, it is reasonable to assume that $r_s \approx q_s$. Consequently, the saturation water vapour mixing ratio is also assumed to be approximately equal to the saturation specific humidity q_s:

$$r_s \approx q_s = \frac{\varepsilon e_s}{p - (1-\varepsilon)\, e_s} \tag{A.36}$$

Where $\varepsilon = \frac{R_d}{R_v} \approx 0.622$, p is the total pressure and e_s the water vapour saturation pressure that depends on temperature T:

$$e_s = c_1 \exp\left(\frac{c_2 T}{T + c_3}\right) \tag{A.37}$$

With constants c_1, c_2 and c_3 equal to

$$\begin{aligned} c_1 &= 6.112 \\ c_2 &= 17.67 \\ c_3 &= 243.5 \end{aligned} \tag{A.38}$$

Linearisation of A.35 yields:

$$dr_v = dRH \cdot r_s + RH \cdot dr_s = dRH \cdot r_s + RH\left(\frac{\partial r_s}{\partial p} dp + \frac{\partial r_s}{\partial e_s}\frac{\partial e_s}{\partial T} dT\right) \tag{A.39}$$

Equation A.39 shows how an increment in water vapour mixing ratio dr_v corresponds to an increment in relative humidity, pressure and temperature. When it is assumed that a water vapour saturation mixing ratio increment dr_s from a pressure perturbation dp is much smaller than from a temperature perturbation dT ($\frac{\partial r_s}{\partial p} dp \ll \frac{\partial r_s}{\partial e_s}\frac{\partial e_s}{\partial T} dT$), the second term in A.39 can be simplified so that an approximation for dr_s becomes:

$$dr_v \approx dRH \cdot r_s + RH \frac{\partial r_s}{\partial e_s}\frac{\partial e_s}{\partial T} dT \tag{A.40}$$

By using the quotient rule, the partial derivative of r_s with respect to e_s is:

$$\frac{\partial r_s}{\partial e_s} \approx \frac{\partial q_s}{\partial e_s} = \frac{\varepsilon p}{\left(p - (1-\varepsilon)\, e_s\right)^2} \tag{A.41}$$

Similarly using the chain and quotient rules for differentiation, the partial derivative of e_s with respect to T is found to be:

$$\frac{\partial e_s}{\partial T} = \frac{c_1 c_2 c_3}{(T + c_3)^2} \exp\left(\frac{c_2 T}{T + c_3}\right) \tag{A.42}$$

Substituting the above expressions for the partial derivatives A.41 and A.42 into A.40 results in:

$$dr_v \approx dRH \cdot r_s + RH \frac{\varepsilon p}{\left(p - (1-\varepsilon)\, e_s\right)^2} \frac{c_1 c_2 c_3}{(T + c_3)^2} \exp\left(\frac{c_2 T}{T + c_3}\right) dT \tag{A.43}$$

Using the definition of A.35 and recognising the expression for e_s (A.37), A.43 is simplified further:

$$dr_v \approx dRH \cdot r_s + \frac{r_v}{r_s} \frac{\varepsilon e_s p}{\left(p - (1-\varepsilon)\, e_s\right)^2} \frac{c_2 c_3}{(T + c_3)^2} dT \tag{A.44}$$

Recognizing the definition of r_s in A.44 and rewriting results in:

$$dr_v \approx dRH \cdot r_s + r_v \frac{p}{p - (1-\varepsilon)\, e_s} \frac{c_2 c_3}{(T + c_3)^2} dT \tag{A.45}$$

Finally, it is assumed that the water vapour contributes little to the total pressure:

$$\frac{p}{p - (1-\varepsilon)\, e_s} \approx 1 \tag{A.46}$$

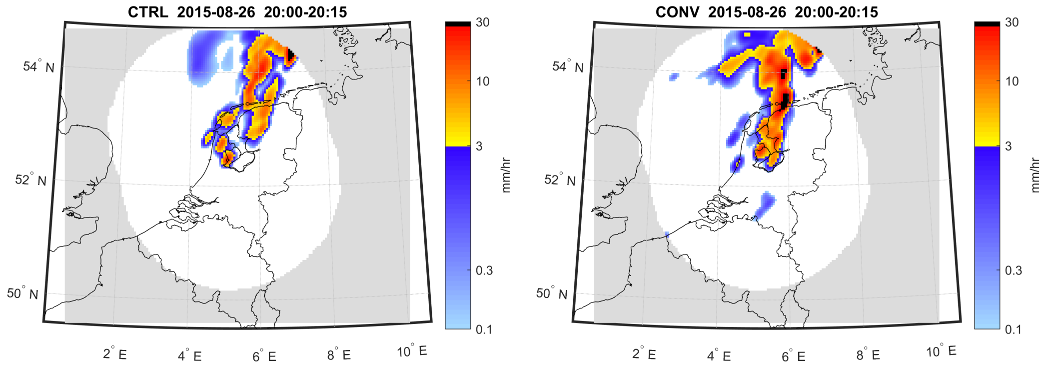
So that A.45 in the end reduces to the following approximation:

$$dr_v \approx dRH \cdot r_s + r_v \frac{c_2 c_3}{(T + c_3)^2} dT \tag{A.47}$$

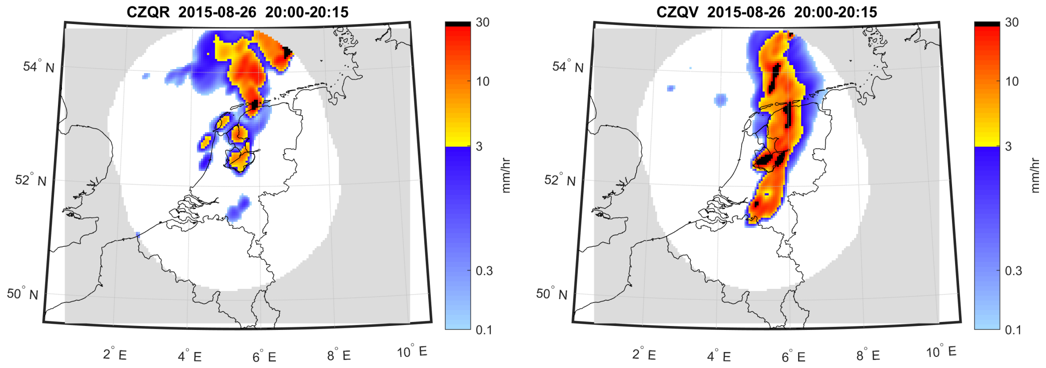
B

Appendix B: Figures

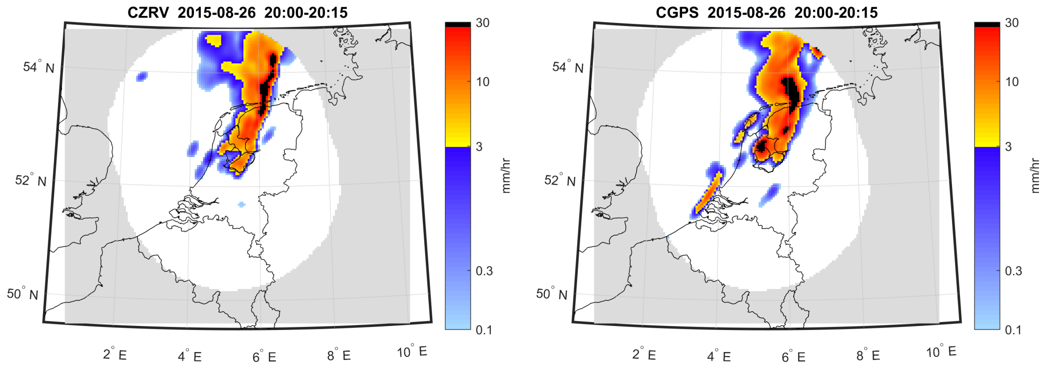
B.1. Rainfall intensity maps case 1 MCS 26-08-2015
B.2. Rainfall intensity maps case 2 LCS 30-08-2015
B.3. Rainfall intensity maps case 3 STRAT 20-06-2016



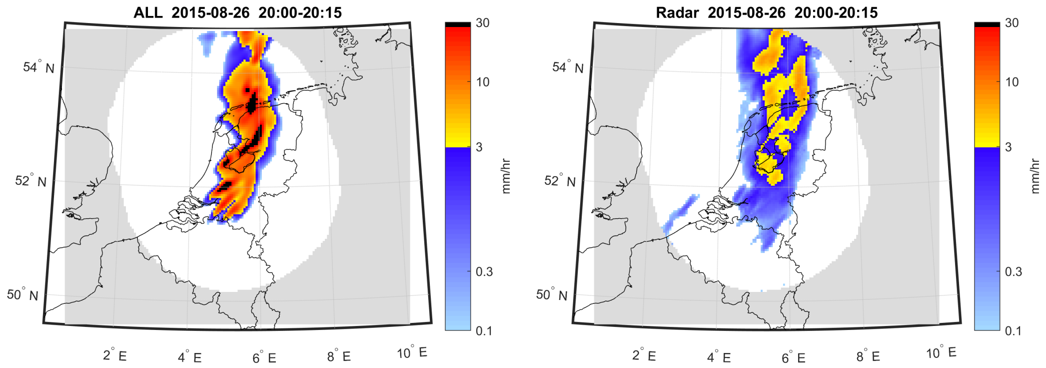
(a) CTRL (b) CONV



(c) CZQR (d) CZQV

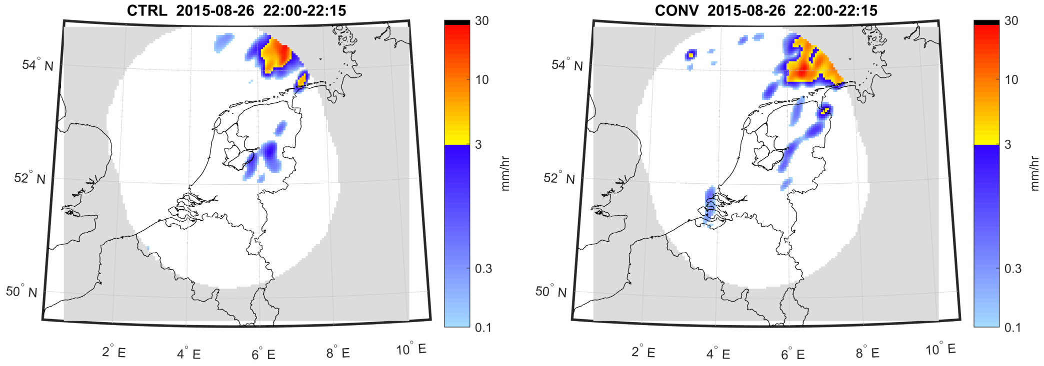


(e) CZRV (f) CGPS



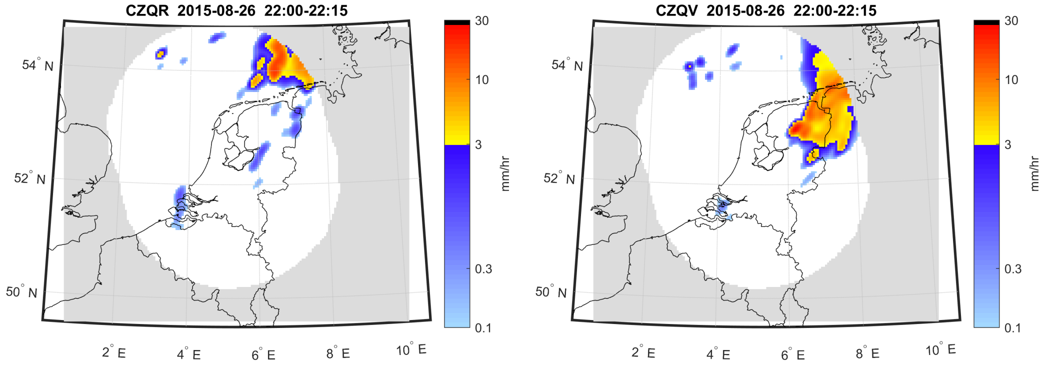
(g) ALL (h) Radar

Figure B.1: 15 minute average rainfall rate for each of the data assimilation scenarios and KNMI radar composite between 20:00 and 20:15 UTC.



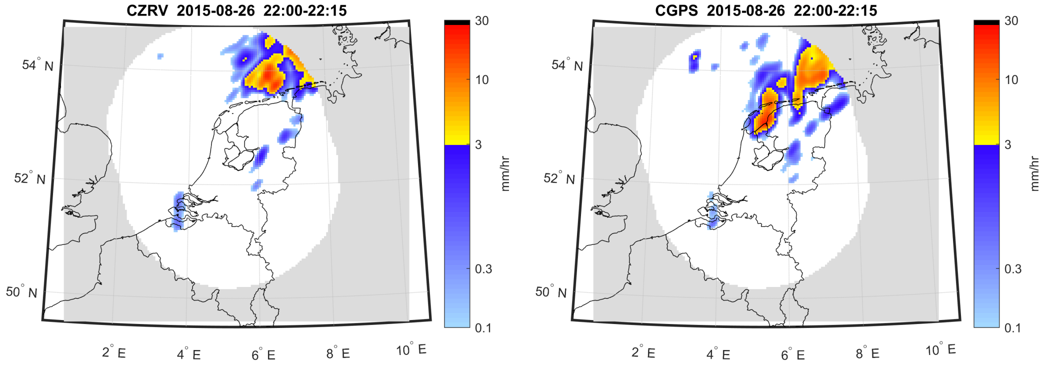
(a) CTRL

(b) CONV



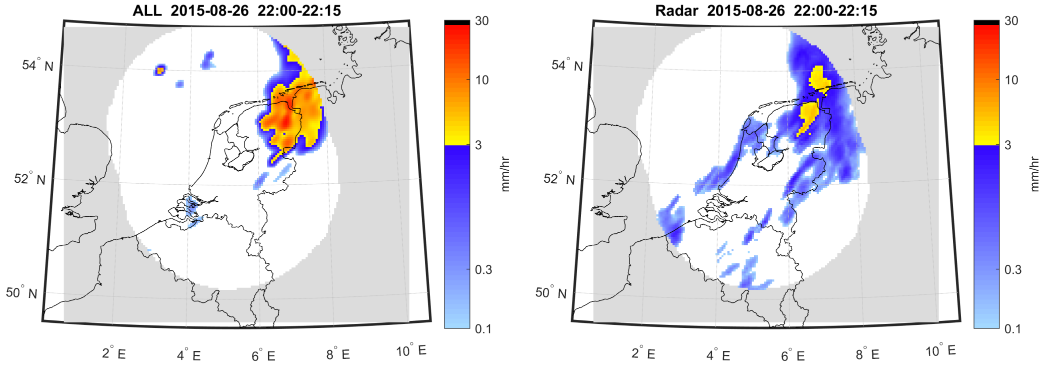
(c) CZQR

(d) CZQV



(e) CZRV

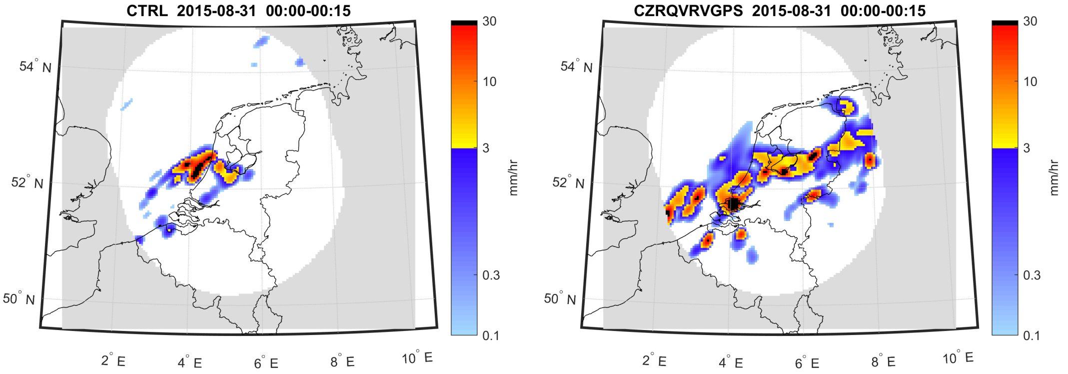
(f) CGPS



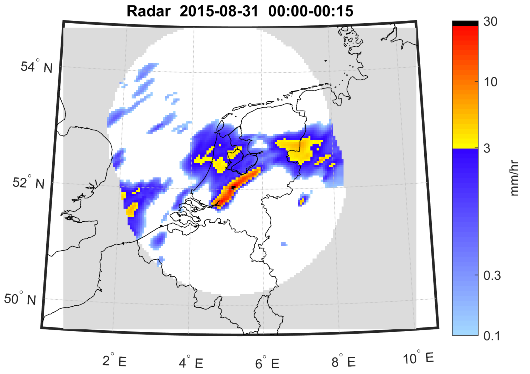
(g) ALL

(h) Radar

Figure B.2: 15 minute average rainfall rate for each of the data assimilation scenarios and KNMI radar composite between 22:00 and 22:15 UTC.

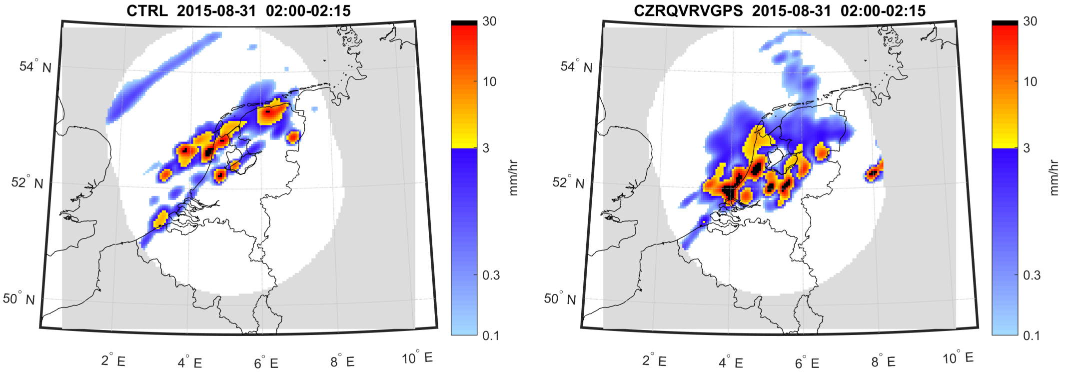


(a) CTRL (b) DA

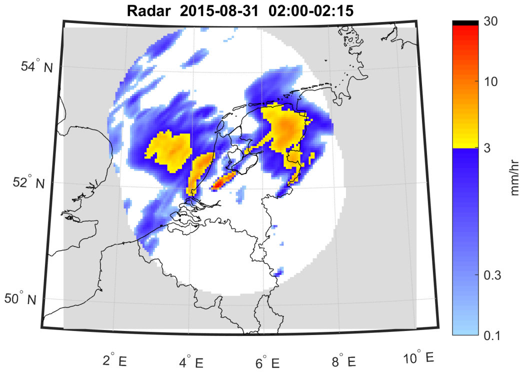


(c) Radar

Figure B.3: 15 minute average rainfall rate for each of the data assimilation scenarios and KNMI radar composite, 00:00-00:15 UTC.

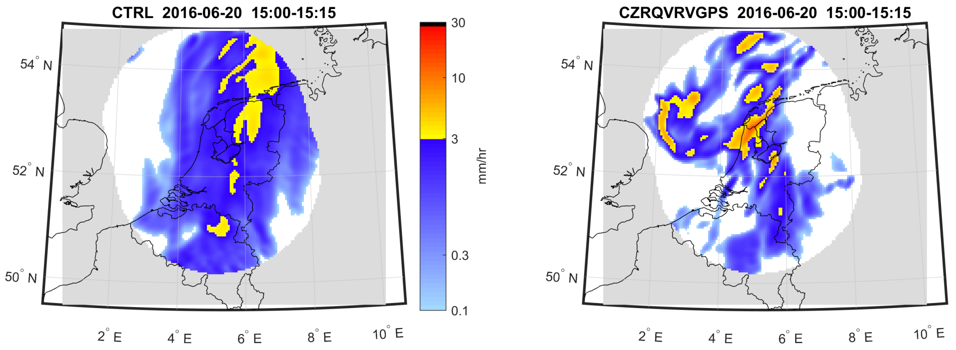


(a) CTRL (b) DA



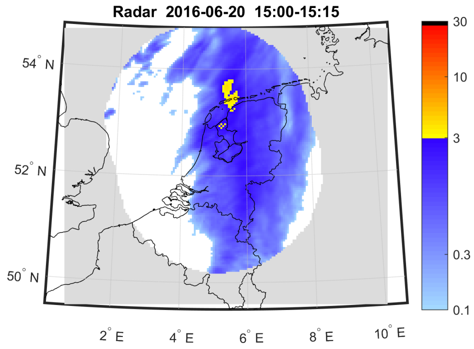
(c) Radar

Figure B.4: 15 minute average rainfall rate for each of the data assimilation scenarios and KNMI radar composite, 02:00-02:15 UTC.



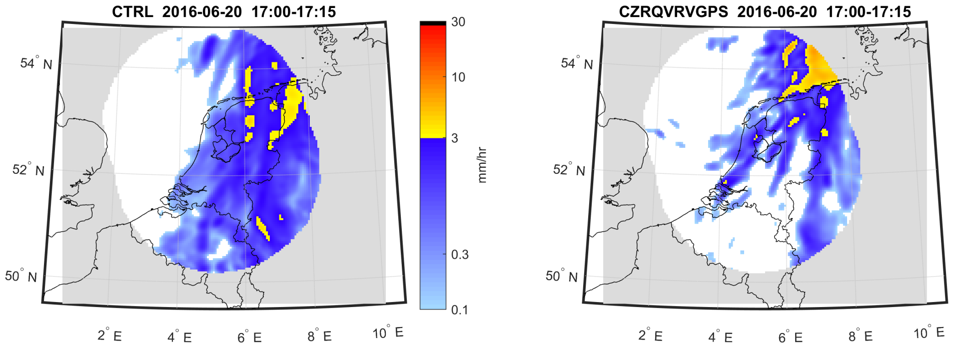
(a) CTRL

(b) DA



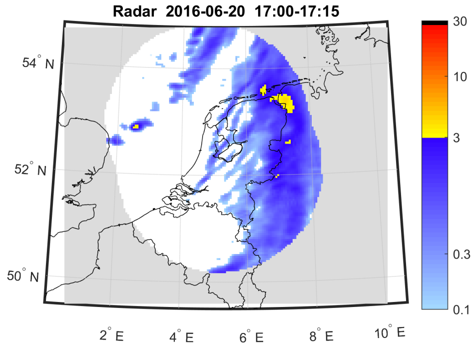
(c) Radar

Figure B.5: 15 minute average rainfall rate for each of the data assimilation scenarios and KNMI radar composite, 15:00-15:15 UTC.



(a) CTRL

(b) DA



(c) Radar

Figure B.6: 15 minute average rainfall rate for each of the data assimilation scenarios and KNMI radar composite, 17:00-17:15 UTC.

Bibliography

A.C. Lorenc. Analysis methods for numerical weather prediction. *Royal Meteorological Society*, 112(474): 1177–1194, 1986. ISSN 1477870X. doi: 10.1256/smsqj.47413.

J. L. Anderson, S. L. Anderson, J. L. Anderson, and S. L. Anderson. A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts. *Monthly Weather Review*, 127(12):2741–2758, dec 1999. ISSN 0027-0644. doi: 10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0493{%}281999{%}29127{%}3C2741{%}3AAMCIOT{%}3E2.0.CO{%}3B2`.

W. M. Angevine, J. Olson, J. Kenyon, K. Suselj, and G. Matheou. A mass flux component for the MYNN PBL scheme in WRF. WRF Workshop 2016, 2016. URL `http://www2.mmm.ucar.edu/wrf/users/workshops/WS2016/oral{_}presentations/6.7.pdf`.

R. N. Bannister. A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 29(January):1–29, 2016. ISSN 1477-870X. doi: 10.1002/QJ.2982.

D. Barker, W. Huang, Y.-R. Guo, and A. Bourgeois. A Three-Dimensional Variational (3DVAR) Data Assimilation System For Use With MM5. *NCAR Technical Note*, page 68, 2003. URL `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.453.8466{&}rep=rep1{&}type=pdf`.

D. Barker, X. Huang, Z. Liu, T. Auligné, X. Zhang, S. Rugg, R. Ajjaji, A. Bourgeois, J. Bray, Y. Chen, M. Demirtas, Y. Guo, T. Henderson, W. Huang, H. Lin, J. Michalakes, S. Rizvi, and X. Zhang. The weather research and forecasting model's community variational/ensemble data assimilation system: WRFDA. *Bulletin of the American Meteorological Society*, 93(6):831–843, 2012. ISSN 00030007. doi: 10.1175/BAMS-D-11-00167.1.

D. M. Barker, W. Huang, Y.-R. Guo, a. J. Bourgeois, and Q. N. Xiao. A Three-Dimensional Variational Data Assimilation System for MM5: Implementation and Initial Results. *Monthly Weather Review*, 132(4):897–914, 2004. ISSN 0027-0644. doi: 10.1175/1520-0493(2004)132<0897:ATVDAS>2.0.CO;2.

L. Battan. *Radar observation of the atmosphere.* University of Chicago Press, 1973.

L. J. Battan. Duration of convective radar cloud units. *Bull. Amer. Meteor. Soc*, 34:227–228, 1953.

M. Bevis, S. Businger, S. Chiswell, T. A. Herring, R. A. Anthes, C. Rocken, and R. H. Ware. GPS Meteorology: Mapping Zenith Wet Delays onto Precipitable Water, 1994. ISSN 0894-8763.

A. Boone. Vertical Hybrid-pressure Coordinate, 2017. URL `http://aaron.boone.free.fr/aspdoc/node8.html`.

F. Bouttier and P. Courtier. Data assimilation concepts and methods. *Meteorological training course lecture series. ...*, pages 1–58, 2002. URL `http://www.msi.ttu.ee/{~}elken/Assim{_}concepts.pdf`.

A. Brasjen. Precipitation estimation from infrared satellite imagery. Master's thesis, Delft University of Technology, 2014.

D. R. Bright and S. L. Mullen. The Sensitivity of the Numerical Simulation of the Southwest Monsoon Boundary Layer to the Choice of PBL Turbulence Parameterization in MM5. *Weather and Forecasting*, 17(1):pp. 99–114, 2002. ISSN 0882-8156. doi: 10.1175/1520-0434(2002)017<0099:TSOTNS>2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0434(2002)017{%}3C0099:TSOTNS{%}3E2.0.CO;2`.

R. G. Brown and P. Y. C. Hwang. Introduction to random signals and applied Kalman filtering: with MATLAB exercises and solutions. *Introduction to random signals and applied Kalman filtering: with MATLAB exercises and solutions, by Brown, Robert Grover.; Hwang, Patrick YC New York: Wiley, c1997.*, 1997. URL `http://eu.wiley.com/WileyCDA/WileyTitle/productCd-EHEP002052.html`.

C. G. Broyden. The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA Journal of Applied Mathematics*, 6(1):76–90, 1970. ISSN 0272-4960. doi: 10.1093/imamat/6.1.76. URL `https://academic.oup.com/imamat/article-lookup/doi/10.1093/imamat/6.1.76`.

J.-F. Caron, T. Milewski, M. Buehner, L. Fillion, M. Reszka, S. Macpherson, J. St-James, J.-F. Caron, T. Milewski, M. Buehner, L. Fillion, M. Reszka, S. Macpherson, and J. St-James. Implementation of Deterministic Weather Forecasting Systems Based on Ensemble–Variational Data Assimilation at Environment Canada. Part II: The Regional System. *Monthly Weather Review*, 143(7):2560–2580, jul 2015. ISSN 0027-0644. doi: 10.1175/MWR-D-14-00353.1. URL `http://journals.ametsoc.org/doi/10.1175/MWR-D-14-00353.1`.

O. Caumont, V. Ducrocq, É. Wattrelot, G. Jaubert, and S. Pradier-Vabre. 1D+3DVar assimilation of radar reflectivity data: A proof of concept. *Tellus, Series A: Dynamic Meteorology and Oceanography*, 62(2):173–187, jan 2010. ISSN 02806495. doi: 10.1111/j.1600-0870.2009.00430.x. URL `https://www.tandfonline.com/doi/full/10.1111/j.1600-0870.2009.00430.x`.

A. J. Clark, W. A. Gallus, M. Xue, and F. Kong. A Comparison of Precipitation Forecast Skill between Small Convection-Allowing and Large Convection-Parameterizing Ensembles. *Weather and Forecasting*, 24(4): 1121–1140, 2009. ISSN 0882-8156. doi: 10.1175/2009WAF2222222.1. URL `http://journals.ametsoc.org/doi/pdf/10.1175/2009WAF2222222.1`.

A. Clayton, D. Barker, N. Bowler, P. Jermey, A. Lorenc, R. Rawlins, and M. Thurlow. The Met Office hybrid data assimilation scheme. 9th Adjoint Workshop, 2011. URL `https://gmao.gsfc.nasa.gov/events/adjoint{_}workshop-9/presentations/Lorenc.pdf`.

P. Colarco, A. Da Silva, M. Chin, and T. Diehl. Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth. *Journal of Geophysical Research Atmospheres*, 115(14), 2010. ISSN 01480227. doi: 10.1029/2009JD012820.

P. Courtier, E. Andersson, W. Heckley, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, M. Fisher, and J. Pailleux. The {ECMWF} implementation of three-dimensional variational assimilation (3D-Var). {I}: Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124(550):1783–1807, 1998. ISSN 00359009. doi: 10.1002/qj.49712455003.

G. P. Cressman. An operational objective analysis system. *Monthly Weather Review*, 87(10): 367–374, 1959. URL `https://docs.lib.noaa.gov/rescue/JNWP/50th{_}Symp{_}2004{_}CD.PDF/JNWPU{_}2004{_}All/mwr-Cressman{_}1959{_}Oct.PDF`.

S. de Haan. Assimilation of GNSS ZTD and radar radial velocity for the benefit of very-short-range regional weather forecasts. *Quarterly Journal of the Royal Meteorological Society*, 139(677):2097–2107, oct 2013. ISSN 00359009. doi: 10.1002/qj.2087. URL `http://doi.wiley.com/10.1002/qj.2087`.

A. Deng and D. R. Stauffer. On improving 4-km mesoscale model simulations. *Journal of applied meteorology and climatology*, 45(3):361–381, 2006.

M. Dixon, Z. Li, H. Lean, N. Roberts, and S. Ballard. Impact of Data Assimilation on Forecasting Convection over the United Kingdom Using a High-Resolution Version of the Met Office Unified Model. *Monthly Weather Review*, 137:1562–1584, 2009. ISSN 0027-0644. doi: 10.1175/2008MWR2561.1.

M. Dixon and G. Wiener. TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting—A Radar-based Methodology. *Journal of Atmospheric and Oceanic Technology*, 10(6):785–797, 1993. ISSN 0739-0572. doi: 10.1175/1520-0426(1993)010<0785:TTITAA>2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0426{%}281993{%}29010{%}3C0785{%}3ATTITAA{%}3E2.0.CO{%}3B2`.

J. Done, C. A. Davis, and M. Weisman. The next generation of NWP: Explicit forecasts of convection using the weather research and forecasting (WRF) model. *Atmospheric Science Letters*, 5(6):110–117, 2004. ISSN 1530261X. doi: 10.1002/asl.72.

R. J. Doviak and D. Zrnic. *Doppler radar and weather observations.* Courier Corporation, 1993.

V. Ducrocq, D. Ricard, J.-P. Lafore, and F. Orain. Storm-Scale Numerical Rainfall Prediction for Five Precipitating Events over France: On the Importance of the Initial Humidity Field. *Weather and Forecasting*, 17(6):1236–1256, dec 2002. ISSN 0882-8156. doi: 10.1175/1520-0434(2002)017<1236:SSNRPF>2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0434{%}282002{%}29017{%}3C1236{%}3ASSNRPF{%}3E2.0.CO{%}3B2`.

J. Dudhia. Overview of WRF physics. WRF Winter tutorial 2017 Lecture, 2017. URL `http://aaron.boone.free.fr/aspdoc/node8.html`.

R. C. Elvander. An evaluation of the relative performance of three weather radar echo forecasting techniques. In *Conference on Radar Meteorology, 17 th, Seattle, Wash*, pages 526–532, 1977.

R. Fletcher. A new approach to variable metric algorithms. *The Computer Journal*, 13(3):317–322, mar 1970. ISSN 0010-4620. doi: 10.1093/comjnl/13.3.317. URL `https://academic.oup.com/comjnl/article-lookup/doi/10.1093/comjnl/13.3.317`.

G. B. Foote and C. G. Mohr. Results of a randomized hail suppression experiment in northeast Colorado. Part VI: Post hoc stratification by storm intensity and type. *Journal of Applied Meteorology*, 18(12):1589–1600, 1979.

J. Gao and D. J. Stensrud. Assimilation of Reflectivity Data in a Convective-Scale, Cycled 3DVAR Framework with Hydrometeor Classification. *Journal of the Atmospheric Sciences*, 69(3):1054–1065, 2012. ISSN 0022-4928. doi: 10.1175/jas-d-11-0162.1.

G. Gaspari and S. E. Cohn. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125(554):723–757, jan 1999. ISSN 00359009. doi: 10.1002/qj.49712555417. URL `http://doi.wiley.com/10.1002/qj.49712555417`.

D. Geophysics. Map Projections, Datums, GIS and GPS for Everyone, 2009. URL `https://www.slideshare.net/dr.geophysics/map-projections-datums-gis-and-gps-for-everyone`.

J. A. Gibbs, E. Fedorovich, and A. M. J. Van Eijk. Evaluating weather research and forecasting (WRF) model predictions of turbulent flow parameters in a dry convective boundary layer. *Journal of Applied Meteorology and Climatology*, 50(12):2429–2444, 2011. ISSN 15588424. doi: 10.1175/2011JAMC2661.1.

P. Ginoux, M. Chin, I. Tegen, T. Goddard, and G. In. model Brent Holben •' Oleg Dubovik •,•' and Shian-Jiann Lin e CART) model . In this model all topographic from the Goddard Earth Observing System Data Assimilation System (GEOS emission is estimated to be between. 106, 2001.

D. Goldfarb. A Family of Variable-Metric Methods Derived by Variational Means. *Mathematics of Computation*, 24(109):23, jan 1970. ISSN 00255718. doi: 10.2307/2004873. URL `http://www.jstor.org/stable/2004873?origin=crossref`.

S. Haykin. *Kalman Filtering and Neural Networks*, volume 5. Wiley Online Library, 2001. ISBN 0471369985. doi: 10.1002/0471221546. URL `http://onlinelibrary.wiley.com/doi/10.1002/0471221546.fmatter{_}indsub/pdf`.

S. G. Henry. Analysis of thunderstorm lifetime as a function of size and intensity. In *Preprints, 26th Conf. on Radar Meteorology, Norman, OK, Amer. Meteor. Soc*, pages 138–140, 1993.

F. F. Hill, K. W. Whyte, and K. A. Browning. The contribution of a weather radar network to forecasting frontal precipitation A case study. *Meteorological Magazine*, 106:69–89, 1977.

G. R. Hilst and J. A. Russo. *An objective extrapolation technique for semi-conservative fields with an application to radar patterns*. Travelers Insurance Companies, 1960.

P. L. Houtekamer, H. L. Mitchell, P. L. Houtekamer, and H. L. Mitchell. A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*, 129(1):123–137, jan 2001. ISSN 0027-0644. doi: 10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0493{%}282001{%}29129{%}3C0123{%}3AASEKFF{%}3E2.0.CO{%}3B2`.

M. Hu, M. Xue, J. Gao, and K. Brewster. 3DVAR and Cloud Analysis with WSR-88D Level-II Data for the Prediction of the Fort Worth, Texas, Tornadic Thunderstorms. Part II: Impact of Radial Velocity Analysis via 3DVAR. *Monthly Weather Review*, 134(2):699–721, 2006a. ISSN 0027-0644. doi: 10.1175/MWR3093.1.

M. Hu, M. Xue, J. Gao, and K. Brewster. 3DVAR and Cloud Analysis with WSR-88D Level-II Data for the Prediction of the Fort Worth, Texas, Tornadic Thunderstorms. Part II: Impact of Radial Velocity Analysis via 3DVAR. *Monthly Weather Review*, 134(2):699–721, feb 2006b. ISSN 0027-0644. doi: 10.1175/MWR3093.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/MWR3092.1http://journals.ametsoc.org/doi/abs/10.1175/MWR3093.1`.

X. M. Hu, J. W. Nielsen-Gammon, and F. Zhang. Evaluation of three planetary boundary layer schemes in the WRF model. *Journal of Applied Meteorology and Climatology*, 49(9):1831–1844, 2010. ISSN 15588424. doi: 10.1175/2010JAMC2432.1.

M. J. Iacono, J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins. Radiative forcing by long-lived greenhouse gases : Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, 113:2–9, 2008. doi: 10.1029/2008JD009944.

IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013. ISBN ISBN 978-1-107-66182-0. doi: 10.1017/CBO9781107415324. URL `www.climatechange2013.org`.

J. S. Kain, S. J. Weiss, J. J. Levit, M. E. Baldwin, and D. R. Bright. Examination of convection-allowing configurations of the WRF model for the prediction of severe convective weather: The SPC/NSSL Spring Program 2004. *Weather and forecasting*, 21(2):167–181, 2006.

A. Khain, B. Lynn, and J. Shpund. High resolution WRF simulations of Hurricane Irene: Sensitivity to aerosols and choice of microphysical schemes. *Atmospheric Research*, 167:129–145, 2016. ISSN 01698095. doi: 10.1016/j.atmosres.2015.07.014. URL `http://dx.doi.org/10.1016/j.atmosres.2015.07.014`.

D. M. Lawrence, K. W. Oleson, M. G. Flanner, P. E. Thornton, S. C. Swenson, P. J. Lawrence, X. Zeng, Z.-L. Yang, S. Levis, K. Sakaguchi, G. B. Bonan, and A. G. Slater. Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *Journal of Advances in Modeling Earth Systems*, 3(1):n/a–n/a, 2011. ISSN 19422466. doi: 10.1029/2011MS00045. URL `http://doi.wiley.com/10.1029/2011MS00045`.

H. W. Lean, P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell. Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Monthly Weather Review*, 136(9):3408–3424, 2008.

D. H. Lenschow. Aircraft measurements in the boundary layer. *Probing the atmospheric boundary layer*, pages 39–55, 1986.

H. Lijense. Volume gegevens radar De Bilt - archief. Datacentrum, 2016. URL `https://data.knmi.nl/datasets`.

M. Lindskog, N. Gustafsson, B. Navascues, K. S. Mogensen, X.-Y. Huang, X. Yang, U. Andrae, L. Berre, S. Thorsteinsson, and J. Rantakokko. Three-dimensional variational data assimilation for a limited area model. Part II: Observation handling and assimilation experiments. *Tellus A*, 53(4):447–468, aug 2001. ISSN 0280-6495. doi: 10.1111/j.1600-0870.2001.00447.x. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0870.2001.00447.x/fullhttp://tellusa.net/index.php/tellusa/article/view/14578`.

C. Liu, K. Ikeda, G. Thompson, R. Rasmussen, and J. Dudhia. High-Resolution Simulations of Wintertime Precipitation in the Colorado Headwaters Region: Sensitivity to Physics Parameterizations. *Monthly Weather Review*, 139(11):3533–3553, 2011. ISSN 0027-0644. doi: 10.1175/MWR-D-11-00009.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-11-00009.1`.

D. C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1-3):503–528, 1989. ISSN 00255610. doi: 10.1007/BF01589116.

A. Lorenc, S. Ballard, R. Bell, N. Ingleby, P. Andrews, D. Barker, J. Bray, A. Clayton, T. Dalby, D. Li, T. Payne, and F. Saunders. The Met. Office global three-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 126(570):2991–3012, 2000. ISSN 00359009. doi: 10.1002/qj.49712657002.

J. Marshall and W. Palmer. The size distribution of raindrops with size. *Quarterly Journal of the Royal Meteorological Society*, 76(327):16–36, 1948. ISSN 1477870X. doi: 10.1002/qj.49707632704.

J. Marshall, W. Hitschfeld, and K. Gunn. Advances in Radar Weather. *Advances in Geophysics*, 2:1, 1955. doi: 10.1016/S0065-2687(08)60310-6.

S. Massart. Assimilation Algorithms Lecture 2 3D-VAR. ECMWF training resources, 2017. URL `https://software.ecmwf.int/wiki/display/OPTR/NWP+Training+Material+2017`.

B. Y. F. R. A. Mcnally, E. Andersson, P. Courtier, P. Unden, J. Eyre, and A. Hollingsworth. The ECMWF implementation of three-dimensional variational assimilation : Structure functions. pages 1809–1829, 1998.

Meted. Influence of Model Physics on NWP forecasts. WRF Winter tutorial 2017 Lecture, 2009. URL `http://www.meted.ucar.edu/nwp/model{_}physics/print.htm{#}residuallayer`.

K.-H. Min, Y. Kim, and K. Park. Assimilation of Null-echo from Radar Observations for Short-term Quantitative Precipitation Forecasting. page 243, 2016.

E. J. Mlawer, J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. 102, 1997.

A. S. Monin and A. M. Obukhov. Basic laws of turbulent mixing in the surface layer of the atmosphere. *Contrib. Geophys. Inst. Acad. Sci. USSR*, 24(151):163–187, 1954.

M. Nakanishi and H. Niino. An improved Mellor–Yamada level-3 model with condensation physics: Its design and verification. *Boundary-layer meteorology*, 112(1):1–31, 2004.

M. Nakanishi and H. Niino. An improved Mellor-Yamada Level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Boundary-Layer Meteorology*, 119(2):397–407, 2006. ISSN 00068314. doi: 10.1007/s10546-005-9030-8.

A. E. Niell. Global mapping functions for the atmosphere delay at radio wavelengths. *Journal of Geophysical Research: Solid Earth*, 101(B2):3227–3246, feb 1996. ISSN 01480227. doi: 10.1029/95JB03048. URL `http://doi.wiley.com/10.1029/95JB03048`.

T. M. Noel and A. Fleisher. The linear predictability of weather radar signals. Technical report, MASSACHUSETTS INST OF TECH CAMBRIDGE, 1960.

K. W. Oleson, D. M. Lawrence, G. B. Bonan, M. G. Flanner, E. Kluzek, P. J. Lawrence, S. Levis, S. C. Swenson, P. E. Thornton, A. Dai, and Others. Technical description of version 4.0 of the Community Land Model, NCAR Tech. Technical report, Note NCAR/TN-478+ STR, 257, 2010.

D. F. Parrish and J. C. Derber. The National Meteorological Center's Spectral Statistical-Interpolation Analysis System, 1992. ISSN 0027-0644.

M. Peura, J. Koistinen, and H. Hohti. Quality information in processing weather radar data for varying user needs. 2006. URL `https://pdfs.semanticscholar.org/5e85/6f47b551937109b208bc8052b521c1f51ba5.pdf?{_}ga=2.39461668.624332020.1500630885-613506777.1497271681`.

C. Program. A Convective Storm Matrix: Buoyancy/Shear Dependencies. Summer Severe Weather online course, 2016. URL `http://www.geosci.sfsu.edu/geosciences/classes/m835/shear/print.htm`.

M. Rajeevan, A. Kesarkar, S. B. Thampi, T. N. Rao, B. Radhakrishna, and M. Rajasekhar. Sensitivity of WRF cloud microphysics to simulations of a severe thunderstorm event over Southeast India. *Annales Geophysicae*, 28(2):603–619, 2010. ISSN 09927689. doi: 10.5194/angeo-28-603-2010.

R. Rasmussen, C. Liu, K. Ikeda, D. Gochis, D. Yates, F. Chen, M. Tewari, M. Barlage, J. Dudhia, W. Yu, K. Miller, K. Arsenault, V. Grubišić, G. Thompson, and E. Gutmann. High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. *Journal of Climate*, 24(12):3015–3048, 2011. ISSN 08948755. doi: 10.1175/2010JCLI3985.1.

R. E. Rinehart. A pattern recognition technique for use with conventional weather radar to determine internal storm motions. *Atmos. Technol*, 13:119–134, 1981.

S. R. H. Rizvi. WRFDA Background Error Estimation, 2012. URL `http://www2.mmm.ucar.edu/wrf/users/wrfda/Tutorials/2013{_}July/docs/WRFDA{_}BE.pdf`.

S. A. Rutledge and P. V. Hobbs. The Mesoscale and Microscale Structure and Organization of Clouds and Precipitation in Midlatitude Cyclones. XII: A Diagnostic Modeling Study of Precipitation Development in Narrow Cold-Frontal Rainbands, 1984. ISSN 0022-4928. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0469{%}281984{%}29041{%}3C2949{%}3ATMAMSA{%}3E2.0.CO{%}3B2`.

J. Saastamoinen. Atmospheric correction for the troposphere and stratosphere in radio ranging satellites. *The use of artificial satellites for geodesy*, pages 247–251, 1972.

A. D. Schenkman, M. Xue, A. Shapiro, K. a. Brewster, and J. Gao. Impact of CASA Radar and Oklahoma Mesonet Data Assimilation on the Analysis and Prediction of Tornadic Mesovortices in an MCS. *Monthly Weather Review*, 139(May 2007):3422–3445, 2011. ISSN 0027-0644. doi: 10.1175/MWR-D-10-05051.1.

C. S. Schwartz, J. S. Kain, S. J. Weiss, M. Xue, D. R. Bright, F. Kong, K. W. Thomas, J. J. Levit, and M. C. Coniglio. Next-Day Convection-Allowing WRF Model Guidance: A Second Look at 2-km versus 4-km Grid Spacing. *Monthly Weather Review*, 137(10):3351–3372, 2009. ISSN 0027-0644. doi: 10.1175/2009MWR2924.1.

D. F. Shanno. Conditioning of Quasi-Newton Methods for Function Minimization. *Mathematics of Computation*, 24(111):647, jul 1970. ISSN 00255718. doi: 10.2307/2004840. URL `http://www.jstor.org/stable/2004840?origin=crossref`.

J. R. Shewchuk. An Introduction to the Conjugate Gradient Method Without the Agonizing Pain. *Science*, 49(CS-94-125):64, 1994. ISSN 14708728. doi: 10.1.1.110.418. URL `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.110.418{&}rep=rep1{&}type=pdf{%}5Cnhttp://www.cs.cmu.edu/{~}quake-papers/painless-conjugate-gradient.pdf`.

A. P. Siebesma, P. M. M. Soares, and J. Teixeira. A Combined Eddy-Diffusivity Mass-Flux Approach for the Convective Boundary Layer. *Journal of the Atmospheric Sciences*, 64(4):1230–1248, 2007. ISSN 0022-4928. doi: 10.1175/JAS3888.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JAS3888.1`.

W. Skamarock, J. Klemp, J. Dudhi, D. Gill, D. Barker, M. Duda, X.-Y. Huang, W. Wang, and J. Powers. A Description of the Advanced Research WRF Version 3. *Technical Report*, (June):113, 2008. ISSN 1477870X. doi: 10.5065/D6DZ069T.

M. S. Speer and L. M. Leslie. The prediction of two cases of severe convection: Implications for forecast guidance. *Meteorology and Atmospheric Physics*, 80(1-4):165–175, 2002. ISSN 01777971. doi: 10.1007/s007030200023.

D. J. Stensrud. *Parameterization schemes: keys to understanding numerical weather prediction models.* Cambridge University Press, 2009.

D. J. Stensrud and S. J. Weiss. Mesoscale Model Ensemble Forecasts of the 3 May 1999 Tornado Outbreak. *Weather and Forecasting*, 17(3):526–543, 2002. ISSN 0882-8156. doi: 10.1175/1520-0434(2002)017<0526:MMEFOT>2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0434{%}282002{%}29017{%}3C0526{%}3AMMEFOT{%}3E2.0.CO{%}3B2`.

J. Sun and N. A. Crook. Dynamical and Microphysical Retrieval from Doppler Radar Observations Using a Cloud Model and Its Adjoint. Part I: Model Development and Simulated Data Experiments. *Journal of the Atmospheric Sciences*, 54(12):1642–1661, 1997. URL `http://journals.ametsoc.org/doi/pdf/10.1175/1520-0469{%}281997{%}29054{%}3C1642{%}3ADAMRFD{%}3E2.0.CO{%}3B2`.

J. Sun and N. A. Crook. Dynamical and Microphysical Retrieval from Doppler Radar Observations Using a Cloud Model and Its Adjoint. Part II: Retrieval Experiments of an Observed Florida Convective Storm. *Journal of the Atmospheric Sciences*, 55(5):835—-852, 1998. URL `http://journals.ametsoc.org/doi/pdf/10.1175/1520-0469{%}281998{%}29055{%}3C0835{%}3ADAMRFD{%}3E2.0.CO{%}3B2`.

J. Sun and H. Wang. Radar Data Assimilation with WRF 4D-Var. Part II: Comparison with 3D-Var for a Squall Line over the U.S. Great Plains. *Monthly Weather Review*, 141(2009):2245–2264, 2013. ISSN 0027-0644. doi: 10.1175/MWR-D-12-00169.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-12-00169.1`.

J. Sun, H. Wang, W. Tong, Y. Zhang, C.-Y. Lin, and D. Xu. Comparison of the Impacts of Momentum Control Variables on High-Resolution Variational Data Assimilation and Precipitation Forecasting. *Monthly Weather Review*, 144(September 2016):149–169, 2016. ISSN 0027-0644. doi: 10.1175/MWR-D-14-00205.1.

H. Sundqvist, E. Berge, and J. E. Kristjánsson. Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Monthly Weather Review*, 117(8):1641–1657, 1989.

G. Thompson. Reducing WRF's high bias of shortwave radiation reaching the ground using fractional cloudiness scheme and aerosol aware direct radiative effect. WRF User's workshop 2016, 2016. URL `http://www2.mmm.ucar.edu/wrf/users/workshops/WS2016/oral{_}presentations/8.2.pdf`.

G. Thompson and T. Eidhammer. A Study of Aerosol Impacts on Clouds and Precipitation Development in a Large Winter Cyclone. *Journal of the Atmospheric Sciences*, 71(10):3636–3658, 2014. ISSN 0022-4928. doi: 10.1175/JAS-D-13-0305.1. URL `http://dx.doi.org/10.1175/JAS-D-13-0305.1`.

G. Thompson, P. R. Field, R. M. Rasmussen, and W. D. Hall. Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Monthly Weather Review*, 136(12):5095–5115, 2008. ISSN 0027-0644. doi: 10.1175/2008MWR2387.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/2008MWR2387.1`.

G. Thompson, M. Tewari, K. Ikeda, S. Tessendorf, C. Weeks, J. Otkin, and F. Kong. Explicitly-coupled cloud physics and radiation parameterizations and subsequent evaluation in WRF high-resolution convective forecasts. *Atmospheric Research*, 168:92–104, 2016. ISSN 01698095. doi: 10.1016/j.atmosres.2015.09.005. URL `http://dx.doi.org/10.1016/j.atmosres.2015.09.005`.

A. A. Tsonis and G. L. Austin. An evaluation of extrapolation techniques for the short-term prediction of rain amounts. *Atmosphere-Ocean*, 19(1):54–65, 1981.

U. S. USAF. Standard atmosphere, 1976. *NASA TMX-74335*, pages 12–15, 1976.

H. Vedel and X.-Y. Huang. Impact of Ground Based GPS Data on Numerical Weather Prediction. *Journal of the Meteorological Society of Japan*, 82(1B):459–472, 2004. ISSN 00261165. doi: 10.2151/jmsj.2004.459.

J. M. Wallace and P. V. Hobbs. *Atmospheric science: an introductory survey*, volume 92. Academic press, 2006.

H. Wang, J. Sun, S. Fan, and X. Y. Huang. Indirect assimilation of radar reflectivity with WRF 3D-var and its impact on prediction of four summertime convective events. *Journal of Applied Meteorology and Climatology*, 52(4):889–902, 2013. ISSN 15588424. doi: 10.1175/JAMC-D-12-0120.1.

T. T. Warner. *Numerical Weather and Climate Prediction*. 2011. ISBN 9780521513890. doi: 10.1017/CBO9780511763243.

M. L. Weisman, W. C. Skamarock, and J. B. Klemp. The Resolution Dependence of Explicitly Modeled Convective Systems. *Monthly Weather Review*, 125(4):527–548, 1997. ISSN 0027-0644. doi: 10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0493{%}281997{%}29125{%}3C0527{%}3ATRDOEM{%}3E2.0.CO{%}3B2`.

M. L. Weisman, C. Davis, W. Wang, K. W. Manning, and J. B. Klemp. Experiences with 0–36-h Explicit Convective Forecasts with the WRF-ARW Model. *Weather and Forecasting*, 23(3):407–437, 2008. ISSN 0882-8156. doi: 10.1175/2007WAF2007005.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/2007WAF2007005.1`.

S. S. Weygandt, S. G. Benjamin, T. G. Smirnova, and J. M. Brown. Assimilation of radar reflectivity data using a diabatic digital filter within the Rapid Update Cycle. In *Preprints, 12th Conf. on IOAS-AOLS, New Orleans, LA, Amer. Meteor. Soc*, volume 8, 2008.

Wikipedia. Absorption band. Website, 2017. URL `https://en.wikipedia.org/wiki/Absorption{_}band`.

J. W. Wilson, N. A. Crook, C. K. Mueller, J. Sun, and M. Dixon. Nowcasting thunderstorms: A status report. *Bulletin of the American Meteorological Society*, 79(10):2079–2099, 1998.

Q. Xiao, Y. H. Kuo, J. Sun, W. C. Lee, D. M. Barker, and E. Lim. An approach of radar reflectivity data assimilation and its assessment with the Inland QPE of Typhoon Rusa (2002) at landfall. *Journal of Applied Meteorology and Climatology*, 46(1):14–22, 2007. ISSN 15588424. doi: 10.1175/JAM2439.1.

B. Xie, J. C. H. Fung, A. Chan, and A. Lau. Evaluation of nonlocal and local planetary boundary layer schemes in the WRF model. *Journal of Geophysical Research Atmospheres*, 117(12):1–26, 2012. ISSN 01480227. doi: 10.1029/2011JD017080.

Y. Xie and A. E. MacDonald. Selection of momentum variables for a three-dimensional variational analysis. *Pure and Applied Geophysics*, 169(3):335–351, 2012. ISSN 00334553. doi: 10.1007/s00024-011-0374-3.

K.-M. Xu and D. A. Randall. A semiempirical cloudiness parameterization for use in climate models. *Journal of the atmospheric sciences*, 53(21):3084–3102, 1996.

X. Zeng, M. Zhao, and R. E. Dickinson. Intercomparison of bulk aerodynamic algorithms for the computation of sea surface fluxes using TOGA COARE and TAO data. *Journal of Climate*, 11(10):2628–2644, 1998. ISSN 08948755. doi: 10.1175/1520-0442(1998)011<2628:IOBAAF>2.0.CO;2.

K. Zhao and M. Xue. Assimilation of coastal Doppler radar data with the ARPS 3DVAR and cloud analysis for the prediction of Hurricane Ike (2008). *Geophysical Research Letters*, 36(12):1–6, 2009. ISSN 00948276. doi: 10.1029/2009GL038658.