BONNER METEOROLOGISCHE ABHANDLUNGEN

Heft 75 (2017) (ISSN 0006-7156)
Herausgeber: Andreas Hense

Theresa Bick

3D Radar reflectivity assimilation with an ensemble Kalman filter on the convective scale

BONNER METEOROLOGISCHE ABHANDLUNGEN

Heft 75 (2017) (ISSN 0006-7156)
Herausgeber: Andreas Hense

Theresa Bick

3D Radar reflectivity assimilation with an ensemble Kalman filter on the convective scale

3D Radar reflectivity assimilation with an ensemble Kalman filter on the convective scale

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Dipl.-Math. **Theresa Bick**
aus
Ostercappeln

Bonn, März, 2016

Diese Arbeit ist die ungekürzte Fassung einer der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn im Jahr 2016 vorgelegten Dissertation von Theresa Bick aus Ostercappeln.

This paper is the unabridged version of a dissertation thesis submitted by Theresa Bick born in Ostercappeln to the Faculty of Mathematical and Natural Sciences of the Rheinische Friedrich-Wilhelms-Universität Bonn in 2016.

Anschrift des Verfassers:

Theresa Bick
Meteorologisches Institut der
Universität Bonn
Auf dem Hügel 20
D-53121 Bonn

Address of the author:

1. Gutachter: Prof. Dr. Clemens Simmer, Rheinische Friedrich-Wilhelms-Universität Bonn
2. Gutachter: Prof. Dr. Roland Potthast, University of Reading

Tag der Promotion: 09. Dezember 2016

Abstract

Atmospheric processes are governed by strongly nonlinear dynamics, and predicting convective events is a challenging task. In numerical weather prediction (NWP), an accurate description of the current state of the atmosphere is needed to initialize NWP forecasts. In order to derive these initial conditions, data assimilation plays an essential role.

For convective-scale data assimilation, radar observations of precipitation represent an ideal database because of their unique ability to capture the spatial and temporal evolution of precipitation systems. A radar forward operator that simulates synthetic radar measurements based on the NWP model fields enables the comparison of observations and model states. Such a forward operator is integrated into a local ensemble transform Kalman filter (LETKF) for the assimilation of radar reflectivity data in areas with and without precipitation. Two experiments are examined.

In the first experiment, the data assimilation update interval is varied (ranging from 5 to 60 minutes) and its influence on the quality of analysis and precipitation forecast is studied. While the analysis quality benefits from frequent updates, forecast quality is degraded. Additional information gained from updating frequently dissipates quickly, and the best forecast quality for lead times beyond one hour can be obtained with hourly updates.

In the second experiment, radar reflectivities are assimilated for a period of seven consecutive days with significant precipitation over Germany. Precipitation forecasts initialized throughout this week are compared to those based on the current operational precipitation data assimilation scheme of Deutscher Wetterdienst (DWD), called latent heat nudging (LHN). The results show that the LETKF based assimilation of radar observations competes successfully with LHN, though LHN has been tested and tuned over several years, whereas the LETKF based system is still in an early development phase.

Furthermore, the impact of the temporal and vertical resolution of the volume radar data is examined. The results suggest that reducing the vertical or temporal resolution of radar observations does not degrade forecast quality but leads to significant speed up of runtime.

The results of this study are strongly encouraging and provide a promising step towards the operational use of radar reflectivities in an LETKF system in order to improve short-term forecasts of precipitation.

Contents

1. Introduction

Weather forecasts of precipitation play a major role in today's society and are involved in many decision-making processes of individuals and commercial organizations (Zhu *et al.*, 2002). Furthermore, there is a great demand for accurate forecasts of severe convective weather events because large hail, lightning and strong winds pose a threat to human safety and can cause severe damages.

In numerical weather prediction (NWP), the atmospheric flow is described by physical laws governed by strongly nonlinear dynamics, and solving these flow equations is strongly related to an initial value problem. Thus, in order to initialize numerical weather forecasts, an essential ingredient is the specification of an initial model state that most accurately describes the current state of the atmosphere.

Lorenz (1963) found that in models described by nonlinear equations, only slightly different initial states may evolve into substantially different solutions. The time by which this initially small uncertainty increases is one of the central questions in NWP because it determines the predictability of a model (Lorenz, 1996). Error doubling times have been found shorter the smaller the spatial scale (Ehrendorfer *et al.*, 1996). Consequently, correctly forecasting time and location of convective events is a challenging task due to the limited predictability caused by the atmosphere's chaotic and strongly nonlinear behavior. Thus, much effort is put into the derivation of accurate initial conditions for the prediction of severe weather events. This process usually involves observational information as well as the model equations describing the atmospheric flow, known as data assimilation (Talagrand, 1997).

Radar instruments are able to collect spatially dense information about precipitation by transmitting electromagnetic radiation that is backscattered by hydrometeors such as rain, graupel or snow, called radar reflectivity. By scanning on several elevations, radar measurements offer a unique description of the spatial structure of convective systems. Furthermore, radar observations are available at a high temporal resolution. Consequently, radars provide valuable information for convective-scale data assimilation. Variational methods (3D-Var, 4D-Var) have been proven successful for the assimilation of radar data at convective scales in idealized and real data frameworks (Sun and Crook, 1997, 1998; Caya *et al.*, 2005; Schwitalla and Wulfmeyer, 2014) and have been established at many operational

weather centers (Sun *et al.*, 2014).

Due to the developments in high performance computing and the increase in available computing resources, ensemble-based data assimilation methods became feasible for NWP and therefore gained popularity in recent years. The ensemble allows to estimate flow-dependent covariances that provide information about the uncertainty of the weather forecast. These estimated uncertainties are a key advantage compared to variational methods, where static covariances based on climatology have to be specified (Nichols, 2010; Freitag and Potthast, 2013). Furthermore, ensemble-based data assimilation methods are, by comparison to variational methods, easy to implement and maintain, because they neither require the formulation of the adjoint of the NWP model nor the tangent linear of the adjoint.

The ensemble Kalman filter (EnKF) as first published by Evensen (1994) extends the idea of the Kalman filter (Kalman, 1960) to an ensemble data assimilation framework. The Kalman filter calculates the model state, usually called analysis, that optimally fits observations and prior information based on least squares theory along with the associated analysis uncertainty. The EnKF of Evensen (1994) applies the Kalman filter to each member of an ensemble of model states individually, where the ensemble can be regarded as the probability distribution of the model state. Burgers *et al.* (1998) revealed an underestimation of the associated analysis error in the scheme of Evensen (1994), since the observations are not treated as the realization of a random variable. Burgers *et al.* (1998) suggest a method that generates an ensemble of observations, known as the perturbed observation method. A deterministic alternative is given by the ensemble transform Kalman filter (ETKF) (Bishop *et al.*, 2001). In the ETKF, the Kalman filter equations are commonly rearranged and the analysis of each ensemble member is expressed as a linear combination of the deviations of the prior ensemble members from the ensemble mean.

All EnKF-algorithms assume that the errors in the observations and the prior information are normally distributed. This Gaussian assumption is a major drawback, since atmospheric dynamics are strongly nonlinear. However, as the only two parameters needed to describe a Gaussian distribution are the mean and the covariance, this assumption makes ensemble filters feasible for NWP. Due to limited computing resources (memory, runtime and data storage), it is usually only feasible to run an ensemble of 20 - 100 members in NWP applications. The rank of the covariance matrices estimated by the ensemble is limited by the ensemble size, but the prognostic model state has a much higher dimension (in the order of approximately 10^7). Nonparametric estimation of probability distributions would require much larger ensembles.

The EnKF has first been applied to global NWP by Houtekamer and Mitchell (1998), and has proven successful for operational applications (Houtekamer and Mitchell, 2005).

The limited ensemble size however also poses problems for the EnKF: as the sample covariance provides only a low rank approximation of the true covariance, spurious correlations arise that hamper the benefit of the flow-dependent estimation. To alleviate this problem, localization has been introduced (Houtekamer and Mitchell, 2001; Ott *et al.*, 2004). Localization is either applied to the background covariance matrix by tapering spurious correlations (covariance localization), or applied to the observation error covariance (observation localization) by increasing the errors at larger ranges and thus restricting the influence of distant observations.

Hunt *et al.* (2007) combine the ideas of the ETKF and observation localization and propose an efficient implementation for NWP applications, called local ensemble transform Kalman filter (LETKF). Analyses are calculated independently at each model grid point which allows different linear combinations of the ensemble members in different spatial regions. Thus, the global analysis is not restricted to the low-dimensional ensemble space but can in fact originate from a higher dimensional space (Hunt *et al.*, 2007).

Assimilating radar observations with an ensemble Kalman filter approach has first been investigated by Snyder and Zhang (2003). The authors assimilated radial velocities into a cloud-scale model in an observing system simulation experiment (OSSE). In OSSE studies (Daley, 1991), synthetic observations are generated from a so-called nature run from the model and then assimilated. This approach allows easy verification of the performance of the data assimilation system because the true state, given by the nature run, is known. Snyder and Zhang (2003) revealed the potential of the EnKF to transfer information from the radar observables onto the unobserved model variables. In the OSSE studies of Caya *et al.* (2005) and Tong and Xue (2005), radial velocity and precipitation observations were assimilated. Caya *et al.* (2005) assimilated synthetic observations of the rainwater mixing ratio as a proxy for radar reflectivity. Tong and Xue (2005) assimilated synthetic reflectivity observations and found it beneficial to assimilate reflectivity also in non-precipitating areas to suppress spurious precipitation. Further OSSE studies by e.g. Xue *et al.* (2006), Gao and Xue (2008), Sobash and Stensrud (2013), and Lange and Craig (2014) corroborated the ability of the EnKF to assimilate radar data at convective scales. In most of these studies, synthetic radial velocity and/or reflectivity observations were simulated directly at the model grid points, which however differs from the true radar geometry given in polar coordinates.

In real-data studies by e.g. Dowell *et al.* (2004), Dowell *et al.* (2011), Snook *et al.* (2011), and Jung *et al.* (2012), the focus was set on the influence of radar data on the analysis quality. The studies of Zhang *et al.* (2009), Dowell and Wicker (2009), and Chang *et al.* (2014) assimilated only radial velocities and excluded radar reflectivities from the assimilation. Thus, research on the impact of the assimilation of radar reflectivity on

precipitation forecasts in an imperfect model framework is limited, and so far mostly very short-range forecasts in the nowcasting range (up to three hours lead time) have been investigated in case studies (Aksoy *et al.*, 2010; Dong and Xue, 2013; Snook *et al.*, 2015).

In this study, the km-scale ensemble data assimilation (KENDA) system (Schraff *et al.*, 2016) that provides a convective-scale LETKF framework (following Hunt *et al.*, 2007) for the NWP model COSMO-DE (Doms and Schättler, 2002; Baldauf *et al.*, 2011), is applied for radar reflectivity observations. The radar forward operator EMVORADO (efficient modular volume radar operator, Blahak, 2008a; Blahak *et al.*, 2011; Zeng, 2013; Jerger, 2014; Zeng *et al.*, 2014) is directly coupled to COSMO-DE and simulates synthetic observations based on the COSMO-DE model fields during runtime. Its integration into KENDA enables the direct assimilation of radar reflectivity observation. In two experiments, radar reflectivities are assimilated in addition to conventional observations (as done by Dong *et al.*, 2011; Snook *et al.*, 2015). The first experiment addresses the influence of the data assimilation update frequency on the quality of analysis and ensemble forecast within a mostly rainy case study. The second experiment evaluates a period of seven consecutive days with significant precipitation, where deterministic forecasts initialized from the LETKF-based radar data assimilation are compared with latent heat nudging (LHN), the radar data assimilation scheme operationally used at Deutscher Wetterdienst (DWD). Furthermore, the impact of the vertical and temporal resolution of radar observations on precipitation forecasts is studied. The results provide novel insights in convective-scale data assimilation and forecasting by assimilating radar reflectivities with an ensemble Kalman filter approach over a central European domain in a non-idealized setting, and by comparing the results to LHN.

Chapter 2 describes the NWP model COSMO and its configuration for Germany, COSMO-DE. Chapter 3 introduces the theory of the Kalman filter and the ensemble Kalman filter, gives an overview of the data assimilation systems that are currently used operationally for COSMO-DE, and describes the implementation of the ensemble Kalman filter for COSMO-DE. In chapter 4, the basic terms of radar meteorology are reviewed and the radar forward operator used in this study is presented. Chapter 5 describes the experimental design, the observation data sets and verification scores, and chapter 6 presents the results. Chapter 7 concludes with a summary.

This work is based on and extends the study of Bick *et al.* (2016). Section 3.3 (description of the implementation of the ensemble Kalman filter), Section 4.3 (description of the radar forward operator), and Chapter 5 (experimental setup, verification scores) follow closely the contents presented in Bick *et al.* (2016) but elaborated in more detail here. Experiment 1 discussed in section 6.1 is the same as in Bick *et al.* (2016), whereas experiment 2 (section 6.2) provides further results on the influence of the vertical and temporal

resolution of radar observations. The description of the NWP model COSMO (Chapter 2) and the derivation of the data assimilation theory (Section 3.1) are presented in much more detail than in Bick *et al.* (2016). Figures and tables that have been taken from Bick *et al.* (2016) are identified in the list of figures and tables, respectively.

2. Description of the numerical weather prediction model

The non-hydrostatic limited-area NWP model COSMO is used operationally at DWD since 1999 (Doms and Baldauf, 2015). Based on DWD's former NWP model Lokal-Modell, the COSMO-Model is maintained and further developed within the multi-national consortium for small scale modeling (COSMO) since 2007.

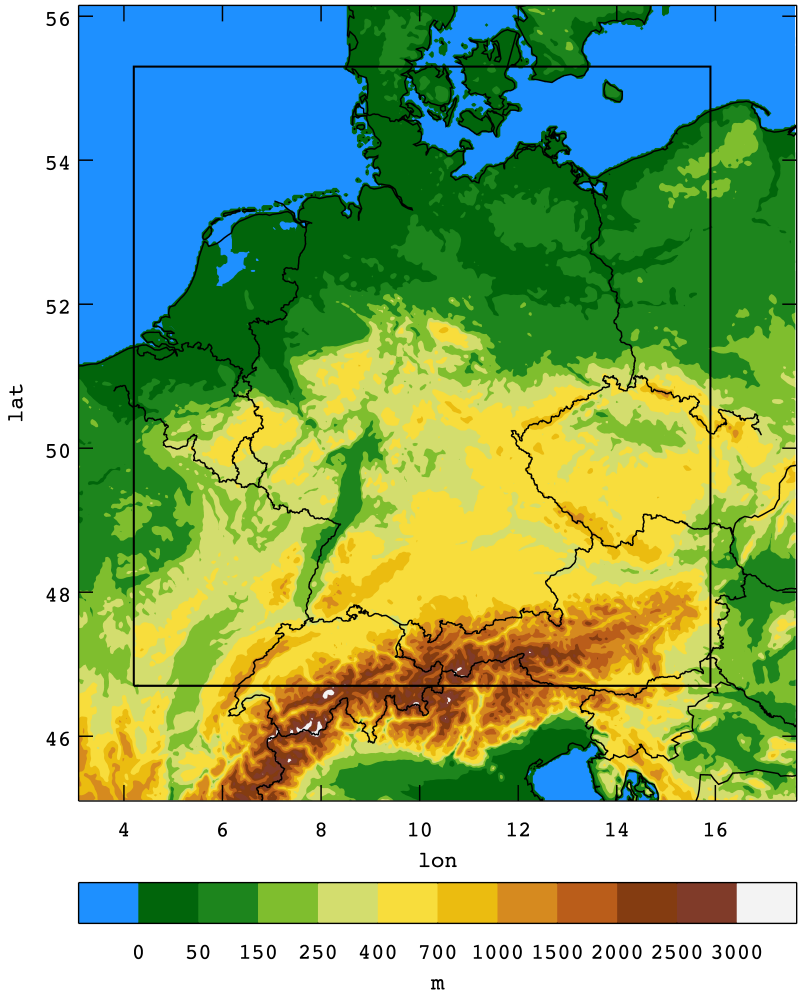


Figure 2.1.: The model domain of COSMO-DE with orography. COSMO fields shown in Chapter 6 are restricted to the subdomain identified by the inner box.

Applicable on the meso-β scale and meso-γ scale, the COSMO modeling system is based on the thermo-hydrodynamic equations characterizing compressible flow in a moist atmosphere. Various physical processes are included by parameterization schemes. At DWD, the COSMO-Model operationally uses a horizontal grid spacing of 7 km on the meso-β scale over a European domain (COSMO-EU), but is to be replaced by the European nest of the global icosahedral non-hydrostatic modeling framework (ICON) (Zängl *et al.*, 2015) soon. On the meso-γ scale, the model configuration COSMO-DE (Baldauf *et al.*, 2011) uses a horizontal grid spacing of 2.8 km. For grid spacings of the order of 1-3 km, the numerical models are able to resolve deep convection explicitly and thus to simulate severe convective events, such as supercell thunderstorms or squall-line storms, more realistically than coarse-grid models with parameterized convection.

The COSMO-DE model domain consists of 421 $\times$ 461 horizontal grid points on 50 vertical layers, thus approximately $9.7 \cdot 10^6$ grid points in total. The domain, approximately 1200 $\times$ 1300 km^2 large, covers Germany and parts of the neighboring countries, as shown in Figure 2.1.

2.1. Dynamical properties

The virtual temperature T_v[1], pressure p and density ρ form the thermodynamic state of the atmosphere at any point. Their relationship is described by

$$p = \rho R_d T_v, \tag{2.1}$$

where $R_d = 287\,\mathrm{J\,kg^{-1}\,K^{-1}}$ is the gas constant for dry air. In a static atmospheric column, pressure always decreases with height as described in the hydrostatic balance equation (e.g. Holton, 1979), that is an approximation to the Navier Stokes equation:

$$\frac{dp}{dz} = -\rho \mathbf{g}, \tag{2.2}$$

where $\mathbf{g}$ is the gravitational acceleration. However, this approximation only holds in the absence of atmospheric motion and is not justified for models at a high spatial resolution where convection plays an essential role. Pressure needs to be calculated from a three dimensional prognostic equation instead of the simple equilibrium approximation.

In general, the atmosphere is composed of dry air, water vapor, liquid water and solid water. The latter two may be further divided into subcategories such as cloud droplets, raindrops, ice crystals, graupel or hail. Being subject to several external and internal

[1] The virtual temperature T_v of a (potentially moist) air parcel is defined as the temperature of a dry air parcel with equal specific humidity and pressure (e.g. Bott, 2012).

Table 2.1.: Symbols used in Equations (2.1) to (2.7)

$\alpha = (R_v/R_d - 1)q^v - q^l - q^f$	Moisture correction term for gas constant
c_p	Specific heat at constant pressure
c_v	Specific heat of moist air at constant volume
g	Magnitude of gravity
$\mathbf{g}$	Apparent acceleration of gravity
I^x	Sources or sinks of constituent x
$\mathbf{J}^x$	Diffusion flux of constituent x
Ω	Constant angular velocity of Earth rotation
p	Pressure
$q^x = \rho^x/\rho$	mass fraction of constituent x
Q_h	Diabatic heating/cooling rate
Q_m	Impact of changes in humidity on pressure tendency
R_d, R_v	Gas constant for dry air and water vapor
ρ	Density
ρ^x	Partial density of mixture constituent x
t	Time
T	Temperature
$T_v = (1 + \alpha)T$	Virtual temperature
$\underline{\mathbf{t}}$	Stress tensor due to viscosity
$\mathbf{v}$	Barycentric velocity (3D wind vector)

processes, such as gravity, Coriolis forces, heat, mass and momentum transfers, the system is described by the following equations (following Doms and Baldauf, 2015):

$$\rho\frac{d\mathbf{v}}{dt} = -\nabla p + \rho\mathbf{g} - 2\Omega \times (\rho\mathbf{v}) - \nabla \cdot \underline{\mathbf{t}} \tag{2.3}$$

$$\frac{dp}{dt} = -(c_p/c_v)p\nabla \cdot \mathbf{v} + (c_p/c_v - 1)Q_h + (c_p/c_v)Q_m \tag{2.4}$$

$$\rho c_p\frac{dT}{dt} = \frac{dp}{dt} + Q_h \tag{2.5}$$

$$\rho\frac{dq^x}{dt} = -\nabla \cdot \mathbf{J}^x + I^x \tag{2.6}$$

$$\rho = p\,(R_d T_v)^{-1} \tag{2.7}$$

An overview of the symbols used in Equations (2.1) to (2.7) is provided in Table 2.1. The index x refers to the specific constituents of the atmosphere, described in Table 2.2. Equations (2.3) to (2.7) form the prognostic equations of the atmospheric model state variables $\mathbf{v}$, T, p, ρ, q^v, q^l and q^f. These differential equations are however mathematically only valid if time and space increments approach zero. According to Doms and Baldauf (2015), for a physically meaningful interpretation in the atmospheric context, the application of equations (2.3) to (2.7) is limited to spatial scales of $\mathcal{O}(1\,\text{cm})$ and temporal

Table 2.2.: Specific constituents of the atmosphere.

$x = d$	dry air
$x = v$	water vapour
$x = l$	liquid water
$x = f$	water in solid state

Table 2.3.: Subgrid-scale variables within simplified COSMO-Model equation set.

$\mathbf{F}^{v,l,f}$	Turbulent fluxes of water vapor, liquid water and ice
$\mathbf{H}$	Turbulent flux of sensible heat (needed in order to calculate Q_h)
$I^{l,f}$	Rates of phase changes of liquid water and ice
$\underline{\mathbf{T}}$	Reynolds stress tensor
$\mathbf{P}^{l,f}$	Precipitation fluxes of liquid water and ice
$\mathbf{R}$	Flux of solar and thermal electromagnetic radiation

scales of $\mathcal{O}(1\,\mathrm{s})$. Certainly, grid spacings of the order of 1 cm and integration time steps of the order of 1 s are computationally not feasible. Therefore it is necessary to average the basic equations over certain spatial and temporal scales.

Averaging is applied to the budget equations for momentum, total mass, mass of the water constituents and enthalpy, whereas for the thermodynamic equations of state it is assumed that thermodynamic relations between the variables are the same for their mean as on the molecular scale. Several further simplifications then yield the following final set of equations the COSMO-Model is based on:

$$\rho \frac{d\mathbf{v}}{dt} = -\nabla p + \rho \mathbf{g} - 2\Omega \times (\rho \mathbf{v}) - \nabla \cdot \underline{\mathbf{T}} \tag{2.8}$$

$$\frac{dp}{dt} = -(c_{pd}/c_{vd}) p \nabla \cdot \mathbf{v} + (c_{pd}/c_{vd} - 1) Q_h \tag{2.9}$$

$$\rho c_{pd} \frac{dT}{dt} = \frac{dp}{dt} + Q_h \tag{2.10}$$

$$\rho \frac{dq^v}{dt} = -\nabla \cdot \mathbf{F}^v - (I^l + I^f) \tag{2.11}$$

$$\rho \frac{dq^{l,f}}{dt} = -\nabla \cdot (\mathbf{P}^{l,f} + \mathbf{F}^{l,f}) + I^{l,f} \tag{2.12}$$

$$\rho = p\,(R_d T_v)^{-1} \tag{2.13}$$

Within Equations (2.8) to (2.13), it is assumed that the variables summarized in Table 2.3 can be estimated from the grid scale variables via parameterization schemes. Again, the equations for the liquid and solid forms of water $q^{l,f}$ can be divided into further precipitating categories, such as rain, snow or graupel with large sedimentation fluxes and negligible turbulent fluxes, and non-precipitating categories, such as cloud water and cloud ice, with

negligible sedimentation fluxes and larger turbulent fluxes. These particular equations can however become very complex. In the current version of COSMO-DE, the prognostic hydrometeor constituents are water vapor, cloud ice, rain drops, snow and graupel.

2.2. Numerical properties

The thermo-hydrodynamic equations derived in the previous section are valid for any coordinate system rotating with the Earth. Therefore, a spherical coordinate system is the most convenient way to represent the spherical shape of the Earth. However, due to the convergence of the meridians, numerical problems can occur. Thus, Doms and Baldauf (2015) introduce a rotated spherical coordinate system with a repositioned pole to minimize the aforementioned numerical difficulties. In the COSMO-Model equations, the appropriate equations of motion are derived from two coordinate transformations. The first transformations tilts the Z-axis, pointing through the geographical pole of the Cartesian system with the center located in the Earth's center, i.e. the transformation shifts the pole of the coordinate system. The second step transforms the Cartesian coordinates into spherical coordinates depending on longitude λ, latitude φ and the distance from the Earth's center r. The rotated meridian that crosses both the geographical and the rotated North pole is defined as the 0° meridian. Further, orography is included in the model equations by transforming the coordinate system to a surface terrain following coordinate system.

For solving the differential model equations numerically, the equations are discretized in space and in time. For this purpose, a discrete model grid is defined. In physical space, the grid is irregular but is transformed into a regular computational grid with a horizontal resolution of 2.8 km. Every point of the grid is the center point of a rectangular grid box volume. The grid point faces located between two vertical layers are usually referred to as half levels. The model state is not entirely defined at the center points but a part of the model variables is displaced onto the grid box faces halfway between two grid points (Arakawa-C/Lorenz staggering): scalars, such as temperature or hydrometeors, are defined and the grid box centers, and velocity components are defined of the box faces. For discretizing the equations in time, the time-splitting[2] third-order Runge-Kutta approach of Wicker and Skamarock (2002) is used. In the default setup, an integration time step of 25 seconds is used within COSMO-DE.

Another important component of solving differential equations numerically in a limited-area model consists of the initial and boundary conditions. Typically in NWP, initial and

[2]The prognostic equations are split into a slow part, consisting of processes such as advection and Coriolis terms, and a fast part consisting of processes such as pressure gradient terms (leading to sound expansion and buoyancy, further leading to the expansion of gravity waves), a more detailed description is given in Doms and Baldauf (2015).

boundary conditions are given by interpolating a coarser model analysis or forecast to the desired resolution. For this purpose, at DWD a nesting is set up, where the global model ICON provides initial condition and boundaries for COSMO-EU (and soon ICON-EU), which then in turn provides initial and boundary conditions for COSMO-DE. For the application in this study however, COSMO-DE is directly nested into the global model ICON, which is explained in more detail in Section 3.3.3. Interpolation is performed via the downscaling tool `int2lm` (Schättler and Blahak, 2015). The lower boundary condition is provided by the soil model TERRA (e.g. Grasselt *et al.* 2008; a more detailed description is given in Doms *et al.* 2011).

2.3. Parameterization

The model equations (2.3) - (2.7) in principle do not contain any parameters. However, as described previously, solving the equations is only feasible if several simplifications are made, such as the averaging over particular scales. By doing so, new terms arise in the equations (compare Table 2.3) that require parameterization.

Turbulence parameterization is calculated based on the prognostic TKE-equation (Raschendorfer, 2001), radiative transfer is modeled according to Ritter and Geleyn (1992). Due to the high horizontal resolution in the COSMO-DE model setup, it is assumed that deep convection is explicitly resolved. Shallow convection is parameterized following the non-precipitating part of the scheme by Tiedtke (1989). Since deep convection is calculated explicitly, complicated microphysical processes are involved. In order to meet these requirements, the microphysics parameterization scheme is a Lin-type one moment bulk scheme including the hydrometeor types cloud droplets, cloud ice, rain, snow and graupel (Lin *et al.*, 1983; Reinhardt and Seifert, 2006). The term "one moment" means that only the mass densities of the hydrometeors are predicted and a constant number concentration is assumed. A two moment scheme also predicting number concentrations has been introduced by Seifert and Beheng (2001), which is however not used in this study. For the drop size distributions of snow and graupel, an exponential distribution is assumed (Baldauf *et al.*, 2011)

$$N_{(g,s)}(D_{(g,s)}) = N_{0,(g,s)} \exp(\lambda_{(g,s)} D_{(g,s)}), \tag{2.14}$$

with the intercept parameter for graupel $N_{0,g} = 4 \cdot 10^6 \ \mathrm{m}^{-4}$, and a mass-size relation for graupel particles is set to $m_g = 169.6 D_g^{3.1}$. For snow, the intercept parameter is parameterized according to a function depending on temperature T and the snow mixing

ratio q_s via

$$N_{0,s} = \frac{27}{2} a_s(T) \left(\frac{q_s}{\alpha_s} \right)^{4-3b_s(T)}, \tag{2.15}$$

where $\alpha_s = 0.038$, and $a_s(T)$ and $b_s(T)$ are defined in Field *et al.* (2005). It has been found useful to represent the intercept parameter by this function instead of a constant value used in former versions. The new formulation allows for higher intercept parameters at cold temperatures leading to smaller snowflakes at high levels. The mass-size relation for snow particles is set to $m_s = \alpha_s D^2$. For the drop size distribution of rain drops, a Gamma distribution is assumed:

$$N_r(D_r) = N_{0,r} D_r^{\mu} \exp(-\lambda_r D_r). \tag{2.16}$$

Since deep convection is resolved explicitly, the intercept parameter $N_{0,r}$ and the shape parameter μ have to be suited for stratiform and convective rain, which involves a compromise in the choice of both parameters. Based on sensitivity studies, the parameters are set to $N_{0,r} = 3.96 \cdot 10^7\ \mathrm{m}^{-7/2}$ and $\mu = 0.5$ (Baldauf *et al.*, 2011).

3. Data assimilation

Since numerical weather prediction comes along with an initial value problem, a major challenge is the specification of the atmosphere's true state $\mathbf{x} \in \mathbb{R}^n$. This state $\mathbf{x}$ consists of all prognostic model variables on each model grid point. In meteorological applications, n is approximately of the order of $\mathcal{O}(10^7)$. The main goal of data assimilation is to infer an estimate of the true atmospheric state from all available observations $\mathbf{y} \in \mathbb{R}^m$ as accurately as possible (Talagrand, 1997). The observations are related to the atmospheric state by

$$\mathbf{y} = H(\mathbf{x}) + \eta, \tag{3.1}$$

where H is the observation operator and $\eta \in \mathbb{R}^m$ the observation error. $H : \mathbb{R}^n \to \mathbb{R}^m$ maps the atmospheric model state into observation space. In case of variables that are directly observable, such as temperature or pressure measurements from radiosondes or surface stations, H spatially interpolates the model values to the observation coordinates. In case of variables that are not directly observable, such as remote sensing observations from radar or satellites, more sophisticated observation operators are necessary. These operators usually rely on complicated relationships between the model states and the observed variables. The observation is subject to an error $\eta = (\mathbf{y} - \mathbf{y^t})$ with $\mathbf{y^t}$ the true value. Commonly, it is assumed that the observation error is unbiased, i.e. $\mathrm{E}(\eta) = 0$, and that its covariance $\mathbf{R} \in \mathbb{R}^{m \times m}$, where $\mathbf{R} = \mathrm{E}\left[(\mathbf{y} - \mathbf{y^t})(\mathbf{y} - \mathbf{y^t})^T\right] = \mathrm{E}(\eta\eta^T)$ is known. η is commonly interpreted as the sum of three error sources: instrument error, representativeness error[1] and model error of the observation operator (e.g. Janjić and Cohn, 2006).

In addition to the observations, background information obtained from a numerical model is available, denoted

$$\mathbf{x}^b = \mathbf{x^t} + \epsilon. \tag{3.2}$$

Deviations of the background state from the true state $\mathbf{x^t}$ are denoted by ϵ. Commonly, this background error is as well assumed to be unbiased, i.e. $\mathrm{E}(\epsilon) = 0$, and that its covariance $\mathbf{P}^b \in \mathbb{R}^{n \times n}$, with $\mathbf{P}^b = \mathrm{E}\left[(\mathbf{x}^b - \mathbf{x^t})(\mathbf{x}^b - \mathbf{x^t})^T\right] = \mathrm{E}(\epsilon\epsilon^T)$ is known. Usually in NWP, the

[1] In data assimilation, the error that arises if the observation resolves scales finer than the resolution of the numerical model is called representativeness error (e.g. Janjić and Cohn, 2006).

background is provided by a short range forecast initialized from a model state valid at a previous time step t

$$\mathbf{x}^b_{t+1} = M(\mathbf{x}_t), \tag{3.3}$$

where M is given by the numerical model equations. An overview of the symbols used in this section is given in Table 3.1.

Table 3.1.: Symbols used in the data assimilation framework.

a (superscript)	Analysis
b (superscript)	Background
E	Expectation
ϵ	Background error
H	Observation operator (potentially nonlinear)
$\mathbf{H}$	Linear observation operator
η	Observation error
$\mathbf{I}$	Identity matrix
J	Cost function
$\mathbf{K}$	Kalman gain matrix
m	Dimension of the observation space
M	Numerical model (potentially nonlinear)
$\mathbf{M}$	Linear numerical model
n	Dimension of the model space
N	Number of ensemble members
$\mathbf{P}$	Error covariance matrix of the model state
$\mathbf{R}$	Observation error covariance matrix
t	Time index
$\mathbf{t}$ (superscript)	True state
T (superscript)	Transpose of a vector/matrix
$\mathbf{x}$	Model state vector
$\overline{\mathbf{x}}$	Ensemble mean in physical space
$\mathbf{w}$	Weighting coefficients
$\mathbf{X}$	Matrix of ensemble perturbations
$\mathbf{y}$	Observation vector
$\overline{\mathbf{y}}$	Ensemble mean in observation space
$\mathbf{Y}$	Matrix of ensemble perturbations in observation space
$\tilde{}$	Variable in ensemble space ($\tilde{J}$, $\tilde{\mathbf{P}}$)

3.1. Algorithms based on the Kalman filter

Data assimilation is in widespread use in science and engineering and not restricted to meteorology. However, the application in meteorology (and oceanography) is unique in

regard of the high dimensionality of the numerical systems. A popular data assimilation algorithm, the Kalman filter (Kalman, 1960), was developed in electrical engineering in order to filter signals from noisy measurements. Furthermore, data assimilation is applied in the navigation of aircrafts and spaceships. Data assimilation also shows similarities to inverse problems, such as posed by plasma physics or geophysics, where information about the internal structure of a physical system or the Earth is inferred from measurements that can only be taken at the surface (Talagrand, 1997). In recent years, intense research has also been conducted in neurosciences, where data assimilation is used to estimate cortical neuronal activity (Schiff and Sauer, 2008).

In this study, the local ensemble transform Kalman filter (LETKF), a square-root filter method based on the ensemble Kalman filter (EnKF) with localization, is used. The numerical model M and the forward operator H are nonlinear in this study. In the following subsections, the algorithm applicable for this setup is derived.

3.1.1. The Kalman filter in a least-squares framework

According to Tarantola (1987) or Nichols (2010), the optimal solution, also called analysis $\mathbf{x}^a$, of the data assimilation problem stated in Equation 3.1, is given by a weighted nonlinear least-squares problem. The minimization of the cost function

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T(\mathbf{P}^b)^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y})^T\mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}) \tag{3.4}$$

yields the best estimate $\mathbf{x}^a$ given the observations $\mathbf{y}$ and the background (or prior) $\mathbf{x}^b$ and their associated uncertainties $\mathbf{R}$ and $\mathbf{P}^b$.

In general, the gradient equation

$$\nabla J(\mathbf{x}) \stackrel{!}{=} 0 \tag{3.5}$$

is not explicitly solvable. If it is assumed that $H = \mathbf{H}$ and $M = \mathbf{M}$ are linear (Equations 3.1 and 3.3) and that η and ϵ follow a Gaussian distribution, an explicit solution of Equation (3.4) exists. In this case, the gradient equation becomes

$$\nabla J(\mathbf{x}) = (\mathbf{P}^b)^{-1}(\mathbf{x} - \mathbf{x}^b) + \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{x} - \mathbf{y}) \stackrel{!}{=} 0, \tag{3.6}$$

where $\mathbf{x}^b = \mathbf{M}\mathbf{x}^a$. The solution of Equation (3.6) for $\mathbf{x}$ explicitly yields the analysis $\mathbf{x}^a$:

$$\mathbf{x}^a = \mathbf{x}^b + \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^b) \tag{3.7}$$

$$= \mathbf{x}^b + \mathbf{P}^b\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}\right)^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^b) \tag{3.8}$$

$$= \mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^b), \tag{3.9}$$

where

$$\mathbf{K} = \mathbf{P}^b\mathbf{H}^T \left(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}\right)^{-1} \tag{3.10}$$

is called the Kalman gain matrix, and $\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^b)$ is called analysis increment. If the observation errors $\mathbf{R}$ are large, the filter trusts the background $\mathbf{x}^b$ and the analysis increment is small. If the background errors $\mathbf{P}^b$ are large, the analysis is shifted closer to the observation $\mathbf{y}$. Equations (3.6) to (3.8) are derived more elaborately in Appendix A.

Analysis error

If $M = \mathbf{M}$ and $H = \mathbf{H}$, Equation (3.8) provides an exact and unique solution to the data assimilation problem. According to the Gauss Markov theorem, the solution to the linear least squares problem is the best linear unbiased estimate (BLUE) for the analysis $\mathbf{x}^a$. The BLUE is optimal in the sense that it minimizes the analysis error covariance. Under the assumption that η and ϵ are uncorrelated, the theoretical error in the analysis is given by (Evensen, 1994):

$$\begin{aligned}
\mathbf{P}^a &= \mathrm{E}\left[(\mathbf{x}^a - \mathbf{x}^\mathbf{t})(\mathbf{x}^a - \mathbf{x}^\mathbf{t})^T\right] \\
&= \mathrm{E}\left[\left[\mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^b) - \mathbf{x}^\mathbf{t}\right]\left[\mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^b) - \mathbf{x}^\mathbf{t}\right]^T\right] \\
&= \mathrm{E}\left[\left[\mathbf{x}^\mathbf{t} + \epsilon + \mathbf{K}(\mathbf{y}^\mathbf{t} + \eta - \mathbf{H}(\mathbf{x}^\mathbf{t} + \epsilon)) - \mathbf{x}^\mathbf{t}\right]\left[\mathbf{x}^\mathbf{t} + \epsilon + \mathbf{K}(\mathbf{y}^\mathbf{t} + \eta - \mathbf{H}(\mathbf{x}^\mathbf{t} + \epsilon)) - \mathbf{x}^\mathbf{t}\right]^T\right] \\
&= \mathrm{E}\left[\left[\epsilon + \mathbf{K}(\mathbf{y}^\mathbf{t} + \eta - \mathbf{H}\mathbf{x}^\mathbf{t} - \mathbf{H}\epsilon)\right]\left[\epsilon + \mathbf{K}(\mathbf{y}^\mathbf{t} + \eta - \mathbf{H}\mathbf{x}^\mathbf{t} - \mathbf{H}\epsilon)\right]^T\right] \\
&= \mathrm{E}\left[(\epsilon + \mathbf{K}\eta - \mathbf{K}\mathbf{H}\epsilon)(\epsilon + \mathbf{K}\eta - \mathbf{K}\mathbf{H}\epsilon)^T\right] \\
&= \mathrm{E}\left[(\epsilon + \mathbf{K}\eta - \mathbf{K}\mathbf{H}\epsilon)(\epsilon^T + \eta^T\mathbf{K}^T - \epsilon^T\mathbf{H}^T\mathbf{K}^T)\right] \\
&= \mathrm{E}\left(\epsilon\epsilon^T + \mathbf{K}\eta\epsilon^T - \mathbf{K}\mathbf{H}\epsilon\epsilon^T + \epsilon\eta^T\mathbf{K}^T + \mathbf{K}\eta\eta^T\mathbf{K}^T - \mathbf{K}\mathbf{H}\epsilon\eta^T\mathbf{K}^T - \epsilon\epsilon^T\mathbf{H}^T\mathbf{K}^T\right. \\
&\quad \left. -\mathbf{K}\eta\epsilon^T\mathbf{H}^T\mathbf{K}^T + \mathbf{K}\mathbf{H}\epsilon\epsilon^T\mathbf{H}^T\mathbf{K}^T\right) \\
&= \mathrm{E}\left(\epsilon\epsilon^T - \mathbf{K}\mathbf{H}\epsilon\epsilon^T + \mathbf{K}\eta\eta^T\mathbf{K}^T - \epsilon\epsilon^T\mathbf{H}^T\mathbf{K}^T + \mathbf{K}\mathbf{H}\epsilon\epsilon^T\mathbf{H}^T\mathbf{K}^T\right) \\
&= \mathbf{P}^b - \mathbf{K}\mathbf{H}\mathbf{P}^b + \mathbf{K}\mathbf{R}\mathbf{K}^T - \mathbf{P}^b\mathbf{H}^T\mathbf{K}^T + \mathbf{K}\mathbf{H}\mathbf{P}^b\mathbf{H}^T\mathbf{K}^T \\
&= \mathbf{P}^b - \mathbf{K}\mathbf{H}\mathbf{P}^b - \mathbf{P}^b\mathbf{H}^T\mathbf{K}^T + \mathbf{K}(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R})\mathbf{K}^T \\
&= \mathbf{P}^b - \mathbf{K}\mathbf{H}\mathbf{P}^b - \mathbf{P}^b\mathbf{H}^T\mathbf{K}^T + \mathbf{P}^b\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}\right)^{-1}(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R})\mathbf{K}^T \\
&= (\mathbf{I} - \mathbf{K}\mathbf{H})\,\mathbf{P}^b,
\end{aligned} \tag{3.11}$$

where $\mathbf{I}$ is the identity. According to the Sherman-Morrison-Woodbury identity[2] $\mathbf{P}^a$ can equivalently be expressed as

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\,\mathbf{P}^b = \left(\mathbf{I} - \mathbf{P}^b\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\right)\mathbf{P}^b \tag{3.12}$$

$$= \left(\mathbf{I} + \mathbf{P}^b\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{P}^b. \tag{3.13}$$

Sequential data assimilation

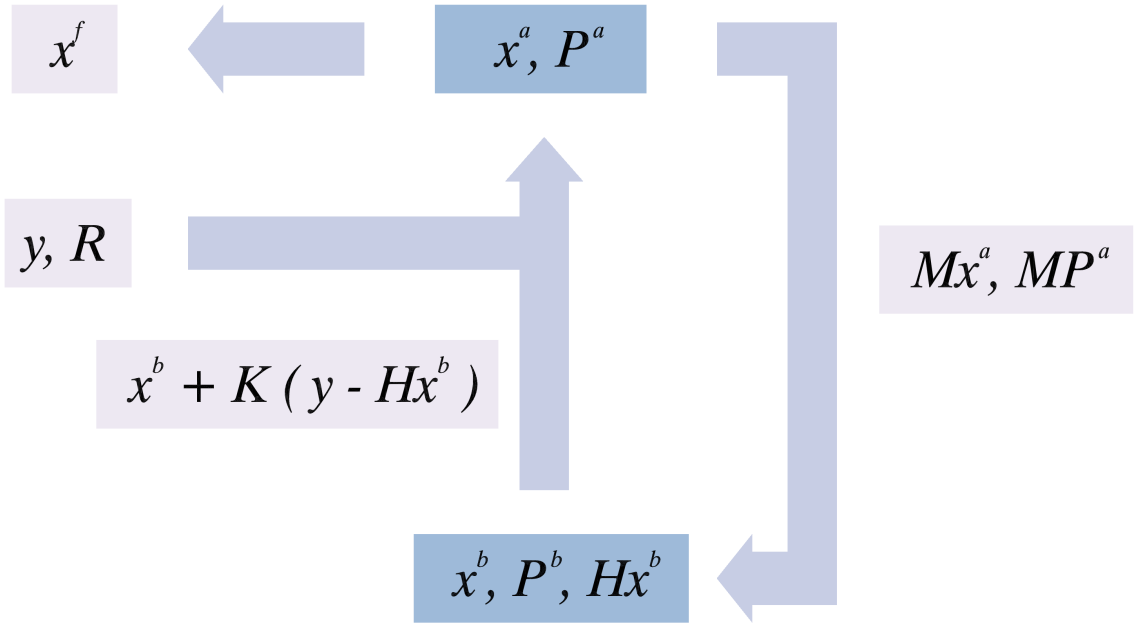


Figure 3.1.: Illustration of the data assimilation cycle. When observations $\mathbf{y}$ (with an error quantified by $\mathbf{R}$) are available, they can be compared to the background information from the model $\mathbf{x}^b$ via $\mathbf{y} - \mathbf{H}\mathbf{x}^b$, where $\mathbf{H}\mathbf{x}^b$ is the projection of the model state into the observation space. The analysis $\mathbf{x}^a$ is determined via the Kalman filter equation $\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^b)$, and the analysis error via $\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^b$. The analysis $\mathbf{x}^a$ then initializes a new model forecast ($\mathbf{M}$). At a next observation time step, the model state can again be updated to obtain a new analysis. At particular time steps (e.g. every three hours), free forecasts $\mathbf{x}^f$ can be initialized based on the analysis.

In so-called sequential data assimilation approaches, a model forward integration (for a predefined time interval) and a data assimilation update are iteratively repeated via

$$\mathbf{x}_t^a = \mathbf{x}_t^b + \mathbf{K}(\mathbf{y}_t - \mathbf{H}\mathbf{x}_t^b) \tag{3.14}$$

$$\mathbf{x}_{t+1}^b = \mathbf{M}\mathbf{x}_t^a \tag{3.15}$$

[2] Applying the Sherman-Morrison-Woodbury identity (Hager, 1989) $(\mathbf{A} + \mathbf{U}\mathbf{B}\mathbf{V})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{B} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}$ with $\mathbf{A} = \mathbf{I}$, $\mathbf{U} = \mathbf{P}^T\mathbf{H}$, $\mathbf{B} = \mathbf{R}^{-1}$ and $\mathbf{V} = \mathbf{H}$

for every time step t. In Equation (3.14), observations available at time step t are assimilated. The resulting analysis $\mathbf{x}_t^a$ then initializes a model forward step (Eq. 3.15), which is then again the background for a next data assimilation update. This sequential data assimilation cycle is illustrated in Figure 3.1. In the following, time indices are neglected.

Related estimation approaches

If $M = \mathbf{M}$ and $H = \mathbf{H}$, the Kalman filter is equivalent to several well-known estimation approaches highlighted in the following.

- The cost function (Eq. 3.4) coincides with the three-dimensional variational (3D-Var) data assimilation scheme (Nichols, 2010; Freitag and Potthast, 2013).

- The Kalman filter as introduced above is a minimum variance estimator. Using Bayes' rule, it can easily be shown that the minimum variance estimator coincides with the maximum likelihood estimator (Lorenc, 1986). In Bayes' theorem, the so-called posterior probability distribution $p(\mathbf{x} \mid \mathbf{y})$ of the model state $\mathbf{x}$ given the observation $\mathbf{y}$ can be obtained by multiplying the prior information $p(\mathbf{x})$ with an observation likelihood $p(\mathbf{y} \mid \mathbf{x})$, divided by a normalization term:

$$p(\mathbf{x} \mid \mathbf{y}) = \frac{p(\mathbf{x})p(\mathbf{y} \mid \mathbf{x})}{p(\mathbf{y})}. \tag{3.16}$$

  Assuming Gaussian distributions of both prior and observation likelihood, the distributions can be specified by incorporating the forecast error covariance $\mathbf{P}^b$ and observation error covariance $\mathbf{R}$:

$$p(\mathbf{x}) \propto \exp\left\{-\frac{1}{2}(\mathbf{x}-\mathbf{x}^b)^T(\mathbf{P}^b)^{-1}(\mathbf{x}-\mathbf{x}^b)\right\} \tag{3.17}$$

$$p(\mathbf{y} \mid \mathbf{x}) \propto \exp\left\{-\frac{1}{2}(\mathbf{y}-H\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{y}-H\mathbf{x})\right\}. \tag{3.18}$$

  The product of the prior and observation likelihood then is again a Gaussian distribution:

$$p(\mathbf{x} \mid \mathbf{y}) \propto \exp\left\{-\frac{1}{2}\left((\mathbf{x}-\mathbf{x}^b)^T(\mathbf{P}^b)^{-1}(\mathbf{x}-\mathbf{x}^b) + (\mathbf{y}-H\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{y}-H\mathbf{x})\right)\right\}. \tag{3.19}$$

  The maximum likelihood estimate of the expectation of $p(\mathbf{x} \mid \mathbf{y})$ is the value of $\mathbf{x}$ that maximizes $p(\mathbf{x} \mid \mathbf{y})$. Maximizing $p(\mathbf{x} \mid \mathbf{y})$ is equivalent to maximizing $\ln\{p(\mathbf{x} \mid \mathbf{y})\}$, since the natural logarithm is a stricly monotonic increasing function. Furthermore,

maximizing $\ln\{p(\mathbf{x} \mid \mathbf{y})\}$ is equivalent to minimizing $(-1) \cdot \ln\{p(\mathbf{x} \mid \mathbf{y})\}$. It is

$$(-1) \cdot \ln\{p(\mathbf{x} \mid \mathbf{y})\} = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T(\mathbf{P}^b)^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y} - H\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{y} - H\mathbf{x}), \tag{3.20}$$

thus, $(-1) \cdot \ln\{p(\mathbf{x} \mid \mathbf{y})\}$ equals the cost the function in Equation (3.4) and the minimum of $(-1) \cdot \ln\{p(\mathbf{x} \mid \mathbf{y})\}$ the minimum of Equation (3.4).

- In NWP, the data assimilation problem in Equation (3.1) is an ill-posed[3] problem. Since the number of available observations is usually smaller than the dimension of the model state, the data assimilation problem is underdetermined and the solution might not be unique. The background term $\mathbf{x} - \mathbf{x}^b$ serves as a regularization term and solving Equation (3.4) is equivalent to the generalized Tikhonov regularization with weighted L^2 norms (Freitag and Potthast, 2013):

$$\mathbf{x}^a = \min\left\{\|\mathbf{y} - H\mathbf{x}^b\|_{\mathbb{R}^m} + \alpha\|\mathbf{x} - \mathbf{x}^b\|_{\mathbb{R}^n}\right\} \tag{3.21}$$

$$= \min\left\{(\mathbf{y} - H\mathbf{x}^b)^T\mathbf{R}^{-1}(\mathbf{y} - H\mathbf{x}^b) + \alpha(\mathbf{x} - \mathbf{x}^b)^T(\mathbf{P}^b)^{-1}(\mathbf{x} - \mathbf{x}^b)\right\}. \tag{3.22}$$

3.1.2. Ensemble Kalman filters

The crucial part of the previously described derivation of the Kalman filter is to specify the background error covariance matrix $\mathbf{P}^b$ and to propagate $\mathbf{P}^a$ to the next time step. In order to arrive at a practical implementation of the Kalman filter, $\mathbf{P}^b$ needs to be approximated in a feasible manner. For this purpose, Evensen (1994) suggested the ensemble Kalman filter (EnKF), where an ensemble of model state vectors approximate the covariance matrices $\mathbf{P}^b$ and $\mathbf{P}^a$. Kalnay (2010) summarizes two different approaches in ensemble Kalman filtering: perturbed observation filters and square root filters.

Perturbed observation filters

Perturbed observation filters (Burgers *et al.*, 1998; Houtekamer and Mitchell, 1998) update each background ensemble member $\mathbf{x}_i^b, i = 1, \ldots, N$ according to Equation (3.9) with a randomly perturbed observation $\mathbf{y}_i$ for each ensemble member i via

$$\mathbf{x}_i^a = \mathbf{x}_i^b + \mathbf{K}(\mathbf{y}_i - \mathbf{H}\mathbf{x}_i^b), \quad i = 1, \ldots, N$$
$$\mathbf{K} = \mathbf{P}^b\mathbf{H}^T(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R})^{-1},$$

[3] An inverse problem $Ax = b$ is called well-posed if the solution x exists, if the solution is unique and if the solution continuously depends on the initial conditions, meaning small perturbations in b lead to small deviations in x.

$$\mathbf{P}^b = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}_i^b - \overline{\mathbf{x}}^b)(\mathbf{x}_i^b - \overline{\mathbf{x}}^b)^T.$$

The background error covariance matrix $\mathbf{P}^b$ is estimated by the ensemble covariance, the background ensemble mean is denoted $\overline{\mathbf{x}}^b$. No linear approximation of the model is necessary.

Burgers *et al.* (1998) show that the use of the same observation $\mathbf{y}$ for each ensemble member $\mathbf{x}_i^b$ (instead of perturbing the observation) leads to an underestimation of the real analysis error covariance $\mathbf{P}^a$. As shown in the derivation of the (theoretical) analysis error covariance matrix $\mathbf{P}^a$ in Equation (3.11), it is

$$\begin{aligned}\mathbf{P}^a &= \mathbf{P}^b - \mathbf{K}\mathbf{H}\mathbf{P}^b + \mathbf{K}\mathbf{R}\mathbf{K}^T - \mathbf{P}^b\mathbf{H}^T\mathbf{K}^T + \mathbf{K}\mathbf{H}\mathbf{P}^b\mathbf{H}^T\mathbf{K}^T \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^b(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T.\end{aligned}$$

If the same observation $\mathbf{y}$ is used to update each ensemble member, the term $\mathbf{K}\mathbf{R}\mathbf{K}^T = \mathbf{K}\overline{(\mathbf{y} - \mathbf{y}^\mathbf{t})(\mathbf{y} - \mathbf{y}^\mathbf{t})^T}\mathbf{K}^T$ vanishes and the analysis ensemble covariance matrix becomes

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^b(\mathbf{I} - \mathbf{K}\mathbf{H})^T. \tag{3.23}$$

With a simple scalar example, Burgers *et al.* (1998) illustrate the underestimation of the true analysis error: For $\mathbf{P}^b = 1$ and $\mathbf{R} = 1$, it is $\mathbf{K} = 0.5$, and the theoretical analysis error is $\mathbf{P}^a = 0.5$, whereas Equation (3.23) yields $\mathbf{P}^a = 0.25$.

Square root filters

Square root filters avoid the calculation of an ensemble of perturbed observations and perform a deterministic update of the ensemble members that, in the linear scenario, coincides with the Kalman filter equations. Various approaches have been published such as the ensemble transform Kalman filter (ETKF) by Bishop *et al.* (2001), the ensemble square root filter (EnSRF) by Whitaker and Hamill (2002), or the ensemble adjustment Kalman filter (EAKF) by Anderson (2001). A summary is provided in Tippett *et al.* (2003). The square root filter method explained here follows Hunt *et al.* (2007) which is in turn mostly based on Bishop *et al.* (2001). The ensemble background covariance matrix $\mathbf{P}^b$ can be written as

$$\mathbf{P}^b = \frac{1}{N-1}\mathbf{X}^b(\mathbf{X}^b)^T, \tag{3.24}$$

where $\mathbf{X}^b \in \mathbb{R}^{n \times N}$ is the matrix of the background ensemble perturbations. The columns of $\mathbf{X}^b$ consist of the deviations from the background ensemble mean for each ensemble member:

$$\mathbf{X}^b = \begin{bmatrix} \mathbf{x}_1^b - \overline{\mathbf{x}}^b & \dots & \mathbf{x}_N^b - \overline{\mathbf{x}}^b \end{bmatrix}. \tag{3.25}$$

Since the columns of $\mathbf{X}^b$ sum up to zero, the maximum rank of $\mathbf{X}^b$ and $\mathbf{P}^b$ is $N-1$. Therefore, $\mathbf{P}^b$ is not invertible. However, in the $(N-1)$ dimensional space S spanned by the columns of $\mathbf{P}^b$, $\mathbf{P}^b$ has full rank and is invertible. This space S is equivalent to the space spanned by the columns of $\mathbf{X}^b$, i.e. the space spanned by the ensemble perturbations. The cost function (Eq. 3.4) for the ensemble case is defined as

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \overline{\mathbf{x}}^b)^T (\mathbf{P}^b)^{-1} (\mathbf{x} - \overline{\mathbf{x}}^b) + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1} (H(\mathbf{x}) - \mathbf{y}) \tag{3.26}$$

and can be minimized in S. In this definition of the cost function, the observation operator H is no longer assumed to be linear, which will be explained in more detail during the course of this section. In order to minimize the cost function J in S, an appropriate coordinate system has to be chosen. An intuitive choice for a basis of S is given by the columns of $\mathbf{X}^b$. However, these columns are linearly dependent. Therefore, $\mathbf{X}^b$ is seen as a transformation from an N-dimensional space $\tilde{S}$ into the $N-1$-dimensional space S, i.e. for a vector $\mathbf{w} \in \tilde{S}$, it is $\mathbf{X}^b\mathbf{w} \in S$, and the corresponding model state $\mathbf{x} \in \mathbb{R}^n$ is given by $\mathbf{x} = \overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}$. Thus, $\mathbf{w}$ can be seen as a vector of weighting coefficients for $\mathbf{X}^b$. If $\mathbf{w} \in \tilde{S}$ is Gaussian distributed with mean 0 and covariance $\tilde{\mathbf{P}}^b = (N-1)^{-1}\mathbf{I}$, then $\mathbf{x} = \overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}$ is Gaussian distributed with mean $\overline{\mathbf{x}}^b$ and covariance $\mathbf{P}^b$. With the relation $\mathbf{x} = \overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}$, Equation (3.26) becomes an analog of the cost function J in S:

$$\begin{aligned} \tilde{J}(\mathbf{w}) &= \frac{1}{2}(\mathbf{X}^b\mathbf{w})^T (\mathbf{P}^b)^{-1} (\mathbf{X}^b\mathbf{w}) + \frac{1}{2}(H(\overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}) - \mathbf{y})^T \mathbf{R}^{-1} (H(\overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}) - \mathbf{y}) \\ &= \frac{N-1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}(H(\overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}) - \mathbf{y})^T \mathbf{R}^{-1} (H(\overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}) - \mathbf{y}). \end{aligned} \tag{3.27}$$

Let $\mathbf{Y}^b$, similar to $\mathbf{X}^b$, be the matrix of ensemble perturbations in observation space, i.e.

$$\mathbf{Y}^b = \begin{bmatrix} \mathbf{y}_1^b - \overline{\mathbf{y}}^b & \dots & \mathbf{y}_N^b - \overline{\mathbf{y}}^b \end{bmatrix}, \tag{3.28}$$

where $\overline{\mathbf{y}}^b$ is the mean of $\mathbf{y}_i^b = H(\mathbf{x}_i^b)$. For the observation operator H, the following linear approximation can be made:

$$H(\overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}) \approx \overline{\mathbf{y}}^b + \mathbf{Y}^b\mathbf{w}. \tag{3.29}$$

Thus, equation (3.27) can be approximated by

$$\tilde{J}(\mathbf{w}) \approx \frac{N-1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}(\overline{\mathbf{y}}^b + \mathbf{Y}^b\mathbf{w} - \mathbf{y})^T\mathbf{R}^{-1}(\overline{\mathbf{y}}^b + \mathbf{Y}^b\mathbf{w} - \mathbf{y}). \tag{3.30}$$

Similar to Section 3.1.1, calculating the minimum of the cost function yields an estimate for the analysis weighting coefficients $\overline{\mathbf{w}}^a$:

$$\overline{\mathbf{w}}^a = \overline{\mathbf{w}}^b + ((N-1)\mathbf{I} + (\mathbf{Y}^b)^T\mathbf{R}^{-1}\mathbf{Y}^b)^{-1}(\mathbf{Y}^b)^T\mathbf{R}^{-1}(\mathbf{y} - \overline{\mathbf{y}}^b), \tag{3.31}$$

and, since $\overline{\mathbf{w}}^b = 0$,

$$\overline{\mathbf{w}}^a = ((N-1)\mathbf{I} + (\mathbf{Y}^b)^T\mathbf{R}^{-1}\mathbf{Y}^b)^{-1}(\mathbf{Y}^b)^T\mathbf{R}^{-1}(\mathbf{y} - \overline{\mathbf{y}}^b). \tag{3.32}$$

According to Equation (3.13), the analysis error covariance is given by

$$\tilde{\mathbf{P}}^a = \frac{1}{N-1}\left(\mathbf{I} + \frac{1}{N-1}\mathbf{Y}^b\mathbf{R}^{-1}\mathbf{Y}^b\right)^{-1} \tag{3.33}$$

$$= \left((N-1)\mathbf{I} + (\mathbf{Y}^b)^T\mathbf{R}^{-1}\mathbf{Y}^b\right)^{-1}, \tag{3.34}$$

i.e. Equation (3.32) can be rewritten as

$$\overline{\mathbf{w}}^a = \tilde{\mathbf{P}}^a(\mathbf{Y}^b)^T\mathbf{R}^{-1}(\mathbf{y} - \overline{\mathbf{y}}^b), \tag{3.35}$$

and an equivalent to the Kalman gain in ensemble space is thus given by

$$\tilde{\mathbf{K}} = ((N-1)\mathbf{I} + (\mathbf{Y}^b)^T\mathbf{R}^{-1}\mathbf{Y}^b)^{-1}(\mathbf{Y}^b)^T\mathbf{R}^{-1} \tag{3.36}$$

$$= \tilde{\mathbf{P}}^a(\mathbf{Y}^b)^T\mathbf{R}^{-1}. \tag{3.37}$$

The minimum of the cost function corresponds to the mean of the analysis weights $\overline{\mathbf{w}}^a$, and the analysis mean $\overline{\mathbf{x}}^a$ and covariance $\mathbf{P}^a$ in physical space are obtained via

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^b + \mathbf{X}^b\overline{\mathbf{w}}^a, \tag{3.38}$$

$$\mathbf{P}^a = \mathbf{X}^b\tilde{\mathbf{P}}^a(\mathbf{X}^b)^T. \tag{3.39}$$

The analysis ensemble is sampled around $\overline{\mathbf{x}}^a$ and has to fulfill the following properties:

$$\mathbf{P}^a = \frac{1}{N-1}\mathbf{X}^a(\mathbf{X}^a)^T, \quad \text{where} \tag{3.40}$$

$$\mathbf{X}^a = \left[\mathbf{x}_1^a - \overline{\mathbf{x}}^a \quad \dots \quad \mathbf{x}_N^a - \overline{\mathbf{x}}^a\right], \tag{3.41}$$

$$\sum_{i=1}^{N}(\mathbf{x}_i^a - \overline{\mathbf{x}}^a) = 0, \tag{3.42}$$

i.e. the ensemble mean is equal to that specified by Equation (3.38), and the ensemble covariance is equal to that specified by Equation (3.39). Hunt *et al.* (2007) determine the analysis ensemble by

$$\mathbf{X}^a = \mathbf{X}^b\mathbf{W}^a, \tag{3.43}$$

where the matrix $\mathbf{W}^a$ is chosen to be

$$\mathbf{W}^a = \left[(N-1)\tilde{\mathbf{P}}^a\right]^{1/2}. \tag{3.44}$$

This decomposition of $\tilde{\mathbf{P}}^a$ uniquely exists since $\tilde{\mathbf{P}}^a$ is symmetric and positive definite. The ensemble of weighting coefficients $\mathbf{w}_1^a, \ldots, \mathbf{w}_N^a$ is calculated in $\tilde{S}$ by adding $\overline{\mathbf{w}}^a$ to the columns of $\mathbf{W}^a$. The resulting weighting vectors $\mathbf{w}_i^a$ then choose the linear combinations of the background ensemble perturbations $\mathbf{X}^b$. Then, the background mean $\overline{\mathbf{x}}^b$ is added to transform the vector back into model space:

$$\mathbf{x}_i^a = \overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}_i^a. \tag{3.45}$$

This choice of the analysis ensemble fulfills Equations (3.40) to (3.42) as shown in detail in Hunt *et al.* (2007).

Transfer of information between variables

As mentioned earlier, the particular advantage of the ensemble Kalman filter is the online estimation of the covariances, which then pass the information from the observed variables to the unobserved variables. The analysis process is illustrated with a two dimensional example[4] with one variable observed, the other variable unobserved (Fig. 3.2). In case both variables are (positively) correlated (Fig. 3.2 a), the mean of the observed component is shifted towards the observation and the variance (i.e. the spread of the dots) is decreased. The information is transferred to the unobserved variable, which is corrected accordingly, i.e. the mean is shifted as well. In case both variables are uncorrelated (Fig. 3.2 b), the observed component is corrected as in Figure 3.2 a, but the mean of the unobserved component remains unchanged. Consequently, information is passed between observed and unobserved variables if and only if these variables are correlated. This reminds of linear

[4] The example is based on an ensemble square root filter and implemented following Nakamura and Potthast (2015).

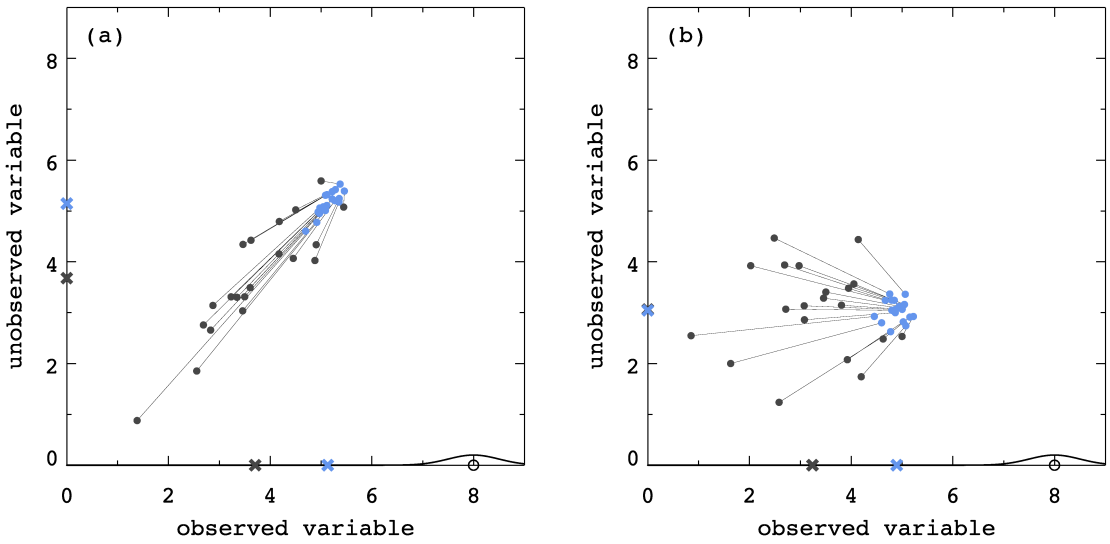


Figure 3.2.: Schematic of the ensemble Kalman filter applied to a two-dimensional example with the variable on the abscissa observed, and the variable on the ordinate not observed. The observation (circle) and its associated error are shown on the abscissa. The dots represent the ensemble members, dark gray corresponds to the first guess, blue to the analysis. Thin lines indicate how each member is shifted by the analysis. The first guess and analysis means for each variable are shown by the cross on each axis in gray and blue, respectively. In (a) the variables are correlated, in (b) the variables are uncorrelated.

regression, and in fact, it can be shown that the Kalman filter can be derived based on extended regression theory (Duncan and Horn, 1972).

3.1.3. Localization

In operational NWP applications, ensemble sizes in the order of the dimension of the model state space are not feasible. Thus, the ensemble covariance provides only a low rank approximation of the true covariance structure. This low-rank approximation can cause artificial long-distance correlations although the true covariance might be small. Localization diminishes the effect of spurious correlations and implicitly chooses analysis increments from a higher dimensional space.

There are two basic localization approaches: Firstly, in localization of $\mathbf{P}^b$, long-distance correlations are damped by multiplying $\mathbf{P}^b$ with a localization function (Houtekamer and Mitchell, 2001). Secondly, in localization of $\mathbf{R}$, the inverse of the observation error covariance $\mathbf{R}^{-1}$ is multiplied by a distance depending localization function in order to assign long distance observations a larger error (Ott *et al.* 2004; used in Hunt *et al.* 2007). For both localization approaches, a common choice for the localization function is the correlation

function introduced by Gaspari and Cohn (1999). Let G_0 be a Gaussian function

$$G_0(p_1, p_2, l_s) = \exp\left(-\frac{d(p_1, p_2)^2}{2l_s^2}\right), \tag{3.46}$$

where l_s is called length scale or Gaussian half-width, and $d(p_1, p_2)$ is the distance between two points $p_1 = (x_1, y_1, z_1)$ and $p_2 = (x_2, y_2, z_2)$ in spherical coordinates on a unit sphere. For the earth, the distance $d(p_1, p_2)$ can be approximated by

$$d(p_1, p_2) = R_e \cdot \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}, \tag{3.47}$$

with the radius of the earth $R_e \approx 6371.229\,\text{km}$. The Gaspari Cohn correlation function approximates the Gaussian function G_0 by a piecewise polynomial function defined as

$$C_0(p_1, p_2, c) = \begin{cases} -\frac{1}{4} \cdot \left(\frac{d(p_1,p_2)}{c}\right)^5 + \frac{1}{2} \cdot \left(\frac{d(p_1,p_2)}{c}\right)^4 \\ \quad + \frac{5}{8} \cdot \left(\frac{d(p_1,p_2)}{c}\right)^3 - \frac{5}{3} \cdot \left(\frac{d(p_1,p_2)}{c}\right)^2 + 1 & \text{if } d(p_1, p_2) \leq c \\ \frac{1}{12} \cdot \left(\frac{d(p_1,p_2)}{c}\right)^5 - \frac{1}{2} \cdot \left(\frac{d(p_1,p_2)}{c}\right)^4 \\ \quad + \frac{5}{8} \cdot \left(\frac{d(p_1,p_2)}{c}\right)^3 + \frac{5}{3} \cdot \left(\frac{d(p_1,p_2)}{c}\right)^2 \\ \quad - 5 \cdot \left(\frac{d(p_1,p_2)}{c}\right) + 4 - \frac{2}{3} \cdot \left(\frac{c}{d(p_1,p_2)}\right) & \text{if } d(p_1, p_2) \leq 2c \\ 0 & \text{else,} \end{cases} \tag{3.48}$$

where the parameter c is related to the Gaussian half-width l_s via

$$c = \sqrt{\frac{10}{3}}\, l_s. \tag{3.49}$$

A comparison of the Gaussian function G_0 and the Gaspari Cohn correlation function C_0 is shown in Figure 3.3 for a length scale $l_s = 16\,\text{km}$. Here, observations within a radius of $c \approx 29.2\,\text{km}$ mainly influence the analysis. The total cutoff is at approximately 58.4 km.

For observation localization, the optimal localization radius depends on the observation density and quality (Perianez *et al.*, 2014). Radar observations are dense, but prone to errors, whereas conventional observations are much sparser and of higher accuracy. Therefore, in this study, different localization radii are chosen for the different observation types. The Gaspari Cohn localization length scale l_s is set to 80 km for conventional observations, and to 16 km for radar observations.

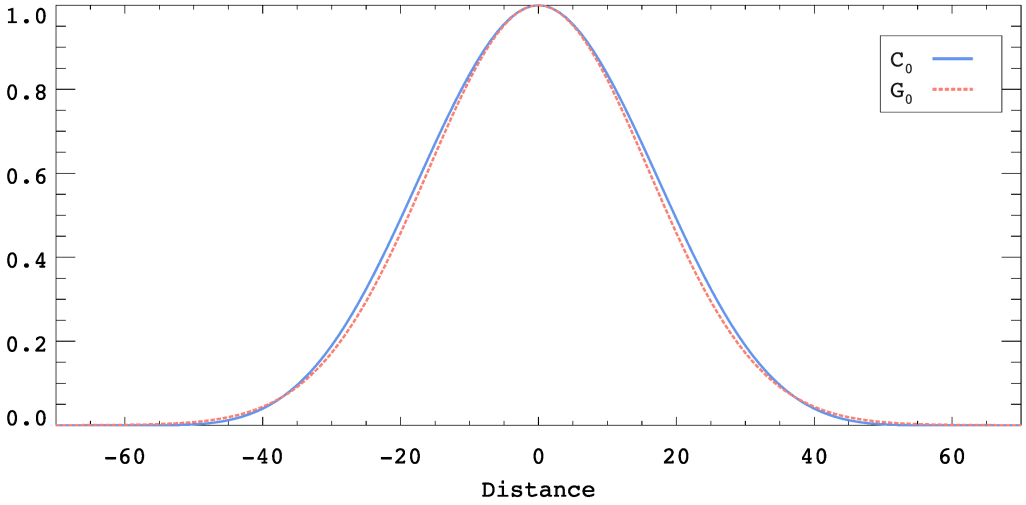


Figure 3.3.: Gaussian function G_0 (red, dashed) and the approximation given by the 5^{th} order Gaspari Cohn correlation function C_0 (blue, solid) for $l_s = 16\,\text{km}$, i.e. $c \approx 29.2\,\text{km}$, and a total cutoff at $\approx 58.4\,\text{km}$.

3.1.4. 4D-Formulation of the ensemble Kalman filter

Theoretically, observations can be assimilated whenever they are available. This approach is reasonable if observations are available at regular time intervals but not too frequently. In practice, observation can be too frequent or asynchronous in time (e.g. aircraft measurements). In the analysis step of the EnKF, the ensemble members at the analysis time are linearly combined to match the current observation. However, observations taken between two analysis time steps might contain important information about the state of the atmosphere, and these observations should neither be neglected nor should be assumed that they are valid at the analysis time.

The advantage of the 4D-Var scheme compared to the EnKF is that it finds the model trajectory that best matches all observations collected during a particular time interval. It is, however, possible to include this idea in the framework of the LETKF (Hunt *et al.*, 2007), as described in the following.

Let $\mathbf{R}_{\tau_j}$ be the observation error covariance matrix for observations $\mathbf{y}_{\tau_j}$ measured at various time steps τ_j since the last analysis. At each observation time τ_j, the observations operator is applied to the background ensemble to obtain the corresponding model equivalent and to calculate the mean $\overline{H\mathbf{x}^b_{\tau_j}} = \overline{\mathbf{y}}^b_{\tau_j}$. The matrix of the ensemble perturbations in observation space for each time step τ_j is given by

$$\mathbf{Y}^b_{\tau_j} = \left[\mathbf{y}^b_{1,\tau_j} - \overline{\mathbf{y}}^b_{\tau_j} \quad \dots \quad \mathbf{y}^b_{N,\tau_j} - \overline{\mathbf{y}}^b_{\tau_j}\right]. \tag{3.50}$$

Then, the components for each time step τ_j are simply concatenated vertically to obtain

the observation $\mathbf{y}$, the background mean $\overline{\mathbf{y}}^b$ and $\mathbf{Y}^b$. The corresponding observation error covariance matrix $\mathbf{R}$ is formed as a block diagonal matrix with blocks $\mathbf{R}_{\tau_j}$. With these notations, the minimization of the cost function can be done as described in Equations (3.27) - (3.45).

3.2. Operational data assimilation in COSMO-DE

Currently, the operational data assimilation in COSMO-DE used for conventional observations is based on nudging, also known as Newtonian relaxation (Schraff, 1997; Schraff and Hess, 2012). Nudging relaxes the model state towards given values and has a long history (Anthes, 1974; Davies and Turner, 1977; Stauffer and Seaman, 1990). In 1997, nudging replaced optimal interpolation as operational data assimilation scheme at DWD. In COSMO, the model state $\mathbf{x}$ is relaxed towards the observations and the tendency of the model state for the variable ψ at grid point k is given by

$$\frac{\partial}{\partial t}\mathbf{x}^t_{\psi,k} = M(\mathbf{x})^t_{\psi,k} + G_\psi \cdot \sum_{k_o \in K} W_{k_o} \cdot \left[\mathbf{y}^t_{k_o} - \mathbf{x}^t_{\psi,k_o}\right], \tag{3.51}$$

with the model forward integration M, the so-called nudging coefficient G_ψ, K the set of observation locations k_o that influence the model grid point location k, and W_{k_o} is an observation weight. The nudging coefficient G_ψ depends on the variable ψ[5]. The observation weight W_{k_o} varies between 0 and 1, see Schraff and Hess (2012) for details to determine W_{k_o}. Since nudging is applied during the model forward integration, it continuously relaxes the model towards the observation. The dynamics of the model should however dominate over the nudging term in order to maintain the balance of the model. The strength by which the model is corrected per time step is determined by the nudging coefficient. Nudging is used to assimilate the following observations:

- Radiosonde observations: horizontal wind, temperature, and humidity
- Aircraft observations: horizontal wind and temperature
- Wind profiler observations: horizontal wind
- Surface station observations: surface pressure, 10-meter horizontal wind and 2-meter humidity
- Drifting buoys

[5] G_ψ is set to $12 \cdot 10^{-4} s{-}1$ for surface pressure, and $6 \cdot 10^{-4} s{-}1$ for all other variables.

DWD intends to replace the nudging scheme with the KENDA based assimilation of conventional observations in due course.

Radar based precipitation observations are integrated into COSMO via latent heat nudging (LHN) (Stephan *et al.*, 2008; Milan *et al.*, 2008; Schraff *et al.*, 2016). Latent heat nudging does not use the full radar volume but only measurements from a so-called precipitation scan, a terrain-following plan position indicator (PPI) scan available every 5 minutes. In order to provide data for every time step, the observations are linearly interpolated in time between two scans. For LHN, radar reflectivity is transferred into a precipitation rate based on an empirical weather-dependent $Z - R$ relation that distinguishes between stratiform and convective cases as well as hail events. Furthermore, data from the 17 radar sites are combined in a national composite by taking the maximum of the derived precipitation rates in overlapping areas.

LHN follows the basic assumption that the latent heat release is proportional to the precipitation rate: precipitation production is linked to hydrometeor phase changes and thus also related to the release of latent heat and a change in temperature. The latent heat profile in the model is adjusted by the ratio of observed and modeled precipitation rates, R_O and R_M (Stephan *et al.*, 2008). This adjustment is accomplished by changing the temperature accordingly to enforce vertical motion, e.g. upward motion in areas where precipitation is underestimated. This vertical motion causes the air to rise, condensate and to form precipitation. The temperature increment over a path l that a hydrometeor particle follows is determined via

$$\Delta T_{LHN}(l) = (\alpha - 1) \cdot \Delta \{L_H(l)\}, \quad \alpha = \frac{R_O}{R_{ref}}, \tag{3.52}$$

where $\Delta \{L_H(l)\}$ is the amount of modeled latent heat release. For the scaling factor α, a reference precipitation rate R_{ref} is used instead of the modeled precipitation rate R_M in order to account for the delay of the initiation of precipitation. The reference precipitation rate is obtained by vertically integrating the hydrometeor fluxes of rain, snow, and graupel. Upper and lower bounds are applied to α (2 and 0.5) in order to keep the temperature increments in reasonable limits. Relative humidity is preserved by adding increments to specific humidity.

3.3. The KENDA system

The km-scale ensemble data assimilation (KENDA) system (Schraff *et al.*, 2016) implements the LETKF described above in COSMO-DE as a 4 dimensional LETKF (see Section 3.1.4). During the COSMO forward integration, observation operators are applied when-

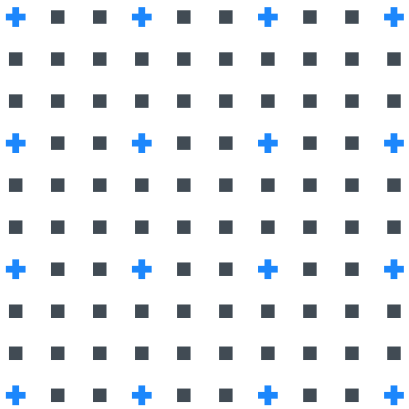


Figure 3.4.: Schematic comparing the full-resolution model grid (squares and crosses) with the sparse analysis grid (crosses). In this case, the sparse analysis grid consists of every third grid point in zonal and meridional direction of the full-resolution grid. The analysis grid uses approximately 11% as many grid points as the full-resolution grid. This figure is adapted from Yang *et al.* (2009).

ever observations are available, i.e. a model equivalent is generated for each observation within the forward integration between two analysis time steps. Thus, KENDA weights the ensemble members according to their trajectory over the first guess window. In the standard mode, KENDA uses 40 ensemble members contributing to the calculation of the weights in Equation (3.35). An additional deterministic run is updated via

$$\mathbf{w}^b_{det} = \tilde{\mathbf{P}}^a(\mathbf{Y}^b)^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{y}^b_{det}), \tag{3.53}$$

where $\mathbf{y}^b_{det} = H(\mathbf{x}^b_{det})$. KENDA implements various features relevant for convective scale data assimilation summarized in the following.

3.3.1. Sparse grid analysis

The KENDA suite allows to compute the analysis weights on a coarsened grid (Yang *et al.*, 2009). In ensemble space, the weighting coefficients are calculated from similar data sets when localization regions overlap sufficiently. In this case, the spatial variability of the weights is small and Yang *et al.* (2009) suggest to choose a limited number of grid points for the analysis, i.e. to perform the analysis on a sparse grid (Figure 3.4). After the analysis step, the information from the coarse analysis grid is transferred back to the high resolution grid by interpolation. The interpolation of the weights is performed in ensemble space before calculating the analysis ensemble in model space, i.e the background maintains the full high resolution information and the analysis increments thus may still capture small-scale features. Yang *et al.* (2009) demonstrate a significant decrease of computational

costs while accuracy is hardly affected. Furthermore, the coarse grid analysis might even be more robust against unwanted imbalances, e.g. leading to gravity waves. In this study, every third grid point in zonal and meridional direction of the full-resolution model grid is used in the analysis. Therefore, the analysis grid uses approximately 11% as many grid points as the full-resolution grid.

3.3.2. Inflation and relaxation

In practice, it is commonly found that the ensemble Kalman filter as described previously may diverge from the observations, i.e. the observations are not within the range of the ensemble. One possible reason is that the LETKF formulas do not account for model error. Even in perfect model environments, the background error tends to become too small making the filter too confident about the background (Hunt *et al.*, 2007). In case the ensemble does not provide sufficient spread, the filter fails to gain information from the observation. One possible solution is to increase the ensemble spread artificially. The most common method is multiplicative inflation (Anderson and Anderson, 1999). In the framework described above, multiplicative inflation can be included easily by replacing Equation (3.34) with

$$\tilde{\mathbf{P}}^a = \left((N-1)\mathbf{I}/\rho + (\mathbf{Y}^b)^T\mathbf{R}^{-1}\mathbf{Y}^b\right)^{-1}, \tag{3.54}$$

for an inflation factor ρ. Common choices are $\rho \in [1.03, 1.1]$.

Relaxation approaches represent an alternative to increase ensemble spread. The methods of Zhang *et al.* (2004) - relaxation to prior perturbations (RTPP), and Whitaker and Hamill (2012) - relaxation to prior spread (RTPS), are implemented in KENDA and discussed in Harnisch and Keil (2015). In RTPP, the analysis ensemble perturbations $\mathbf{x}'^a_i = \mathbf{x}^a_i - \overline{\mathbf{x}}^a$ for $i = 1, \ldots, N$ are relaxed towards the first guess perturbations $\mathbf{x}'^b_i = \mathbf{x}^b_i - \overline{\mathbf{x}}^b$ at each analysis point:

$$\mathbf{x}'^a_i \leftarrow (1-\alpha)\mathbf{x}'^a_i + \alpha\mathbf{x}'^b_i. \tag{3.55}$$

The parameter α controls the influence of the background spread. Zhang *et al.* (2004) recommend the heuristic value $\alpha = 0.75$. The second method, RTPS, relaxes the analysis ensemble spread σ^a towards the prior ensemble spread σ^b:

$$\sigma^a \leftarrow (1-\alpha)\sigma^a + \alpha\sigma^b, \quad \text{where} \tag{3.56}$$

$$\sigma^{(a,b)} = \sqrt{(n-1)^{-1}\sum\left(\mathbf{x}'^{(a,b)}_i\right)^2}. \tag{3.57}$$

Whitaker and Hamill (2012) recommend $\alpha = 0.95$ in order to give a high weight to the prior spread. Equation (3.56) can be rewritten as

$$\sigma^a \leftarrow (1-\alpha)\sigma^a + \alpha\sigma^b \qquad \big| \cdot \mathbf{x}_i'^a \tag{3.58}$$

$$\Leftrightarrow \qquad \mathbf{x}_i'^a\sigma^a \leftarrow \mathbf{x}_i'^a(\sigma^a - \alpha\sigma^a + \alpha\sigma^b) \tag{3.59}$$

$$\Leftrightarrow \qquad \mathbf{x}_i'^a \leftarrow \mathbf{x}_i'^a\left(\alpha\frac{\sigma^b - \sigma^a}{\sigma^a} + 1\right) \tag{3.60}$$

Thus, RTPS is a purely multiplicative inflation, whereas RTPP is partly multiplicative and partly additive.

3.3.3. Cycling

In the Kalman filter framework, a cycling between model runs and assimilation steps is needed as shown in Equation (3.15) and illustrated in Figure 3.1. At DWD, the NWP model COSMO-DE and the LETKF implemented by KENDA are coupled via shell scripts, called basic cycling (BACY) environment. COSMO and KENDA communicate via their output files within BACY. COSMO is reinitialized after each assimilation step. To avoid a dry start of COSMO, i.e. a start where all hydrometeor contents are set to zero, the hydrometeor contents *qc*, *qi*, *qr*, *qs*, *qg* (cloud drops, cloud ice, rain, snow and graupel) are passed between both modules.

Initial conditions and boundary data are produced within BACY via the downscaling module `int2lm` (Schättler and Blahak, 2015) applied to model fields of the global model ICON (Zängl *et al.*, 2015), which as well runs an LETKF. This means that each ensemble member of KENDA is provided with initial and boundary conditions from a different ICON ensemble member.

4. Radar measurement process and simulation

Weather radars (the name originates from Radio Detection and Ranging, an elaborate overview is given in Rinehart, 2010, which this section is based upon) became one of the most important meteorological observation instruments in recent decades. Radars probe the atmosphere by transmitting electromagnetic waves, and receiving the signal backscattered by objects such as rain drops, hail, birds, insects or trees and buildings. The transmitter generates the electromagnetic radiation traveling approximately at the speed of light. The electromagnetic radiation is transmitted and received by the antenna. Radars are able to measure the intensity of the backscattered signal, but are also able to locate the signal along the beam path (the "ranging" component) via the time delay between transmitted and received signals.

Nowadays most operational weather radars operate in a so-called volume scanning mode: Radars scan the atmosphere with a horizontal sweep at a fixed elevation angle (the so called PPI scan), then tilt the antenna to a higher elevation, and repeat the scanning procedure. So the radar is able to detect signals above the surface and determine the vertical structure of precipitation systems. This three dimensional data set is usually collected during a time frame of approximately five minutes. In the following, the main functionality of a radar is described.

Two important parameters are the frequency and wavelength a radar operates. Wavelength and frequency are related via $f = c/\lambda$, where f is frequency in Hertz, c is the speed

Table 4.1.: Typical radar band designations depending on frequency and wavelength.

Band designation	Frequency	Wavelength
L	1-2 GHz	30-15 cm
S	2-4 GHz	15-8 cm
C	4-8 GHz	8-4 cm
X	8-12 GHz	4-2.5 cm
K_u	12-18 GHz	2.5-1.7 cm

of light[1], and λ is wavelength in m. The frequencies used in radar meteorology depend on the particular application; an overview over common frequencies and their corresponding band designation is shown in Table 4.1. The antenna is able to direct the signal into a specific direction by sending the signal towards a circular parabolic reflector, which reflects the signal out into the atmosphere. For weather radars, the diameter of the reflector can vary between 0.3 m and 9 m, depending on the transmitted frequency.

In the following, the radar measurement process is described in Sections 4.1 and 4.2, and the radar forward operator EMVORADO is presented in Section 4.3).

4.1. Beam propagation

If the propagation of a radar beam was perfectly straight, the radar beam's height above the Earth's surface would increase with increasing distance from the radar due to the curvature[2] of the Earth. The propagation of the radar beam is, however, usually not perfectly straight but is influenced by the composition of the atmosphere. The atmosphere's refractivity determines the speed of the electromagnetic radiation, which would be the speed of light in a vacuum, but is slightly slowed down in air or other materials. For the atmosphere, refractivity depends mainly on pressure, temperature and vapor pressure, which under normal atmospheric conditions differ stronger among vertical layers than in horizontal directions. Consequently, electromagnetic waves in different layers are traveling at different speeds. Since refractivity is larger close to the Earth's surface and decreases with height, the radar beam bends towards the surface. Thus, it is rather the gradient of refractivity than the absolute value of refractivity that influences beam propagation. According to Battan (1973), the curvature of the radar beam relative to the Earth's curved surface is given by

$$C = \frac{1}{R'} = \frac{1}{R} + \frac{\delta N \cdot 10^{-6}}{\delta H}, \tag{4.1}$$

where $R = 6374\,\text{km}$ is the Earth's radius, R' the radius of a circle corresponding to the curvature C, and $(\delta N \cdot 10^{-6})/\delta H$ is the gradient of refractivity with height. The assumption of a standard refraction allows to calculate the actual path a radar beam follows under the corresponding atmospheric conditions. A common choice is $\delta N/\delta H = -39\,\text{km}^{-1}$, yielding

$$C = \frac{1}{R'} = \frac{1}{6374\,\text{km}} - \frac{39 \cdot 10^{-6}}{\text{km}} = \frac{1.179 \cdot 10^{4}}{\text{km}}, \qquad \text{i.e.} \tag{4.2}$$

[1] $c = 299\,792\,458\,\text{m/s}$

[2] Curvature can be interpreted as the angular rate of change necessary to follow a curved path. The curvature C of a circle with radius R is thus the inverse of the radius $1/R$.

$$R' = \frac{1\,\mathrm{km}}{1.179 \cdot 10^4} \approx 8483\,\mathrm{km} \approx \frac{4}{3}R. \tag{4.3}$$

This means, that under normal atmospheric conditions, the radar beam propagates approximately along a circle with diameter of $4/3$ times the Earth's radius, as illustrated in Figure 4.1. Under these conditions, the height of the radar beam at elevation ϵ and range r_0 above the Earth's surface is given by (Blahak, 2004)

$$h(r_0, \epsilon) = \sqrt{(R')^2 + r_0^2 + 2 \cdot R' r_0 \sin(\epsilon)} - R'. \tag{4.4}$$

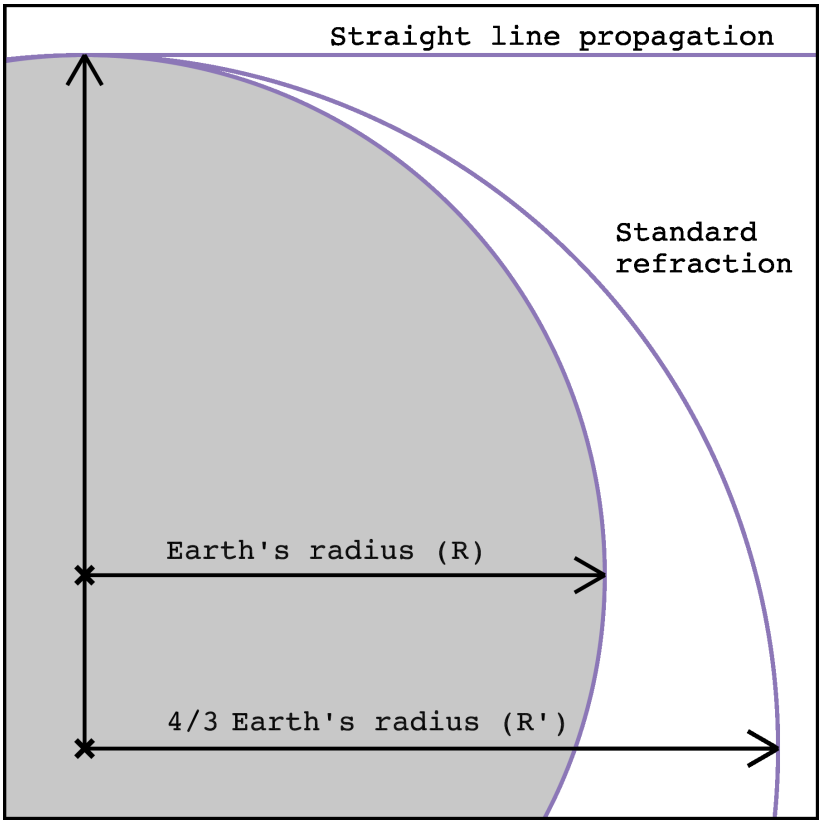


Figure 4.1.: Illustration of the $4/3$ Earth radius approximation showing the Earth (gray circle), its radius R, the straight-line propagation of the radar beam (under uniform atmospheric conditions), the propagation assuming standard refraction, and the corresponding effective radius $R' = 4/3 \cdot R$. The figure is adapted from Rinehart (2010).

4.2. Radar reflectivity

4.2.1. Backscattering cross section

Radars are not only able to detect and range precipitation, but are also capable of measuring the intensity of precipitation. For this purpose, the backscattering cross-sectional area σ

of a target is of interest. Geometrically, for a spherical point target, the cross-sectional area is given by

$$\sigma(D) = \pi(D/2)^2, \tag{4.5}$$

where D is the diameter of the drop. This equality, however, only holds in situations where the diameter of the drop is large compared to the wavelength of the radar, commonly in the order of $D/\lambda > 10$. If the diameter of the drop is small compared to the wavelength (approx. $D/\lambda < 0.1$), the so-called Rayleigh scattering theory applies:

$$\sigma(D) = \frac{\pi^5 |K|^2 D^6}{\lambda^4}, \tag{4.6}$$

i.e. in the Rayleigh region, the backscattering cross-sectional area is proportional to the sixth power of the drop diameter. $|K|^2$ is a parameter related to the complex refractive index of the hydrometeor. The Rayleigh scattering theory holds for many meteorological targets. Raindrops usually do not become larger than 6 mm without breaking up into smaller droplets, so for e.g. C-Band raindrops are within within the Rayleigh region ($D <$ 0.8 cm). However, hailstones can easily become larger than 8 mm, so there is still an important intermediate region where $0.1 \leq D/\lambda \leq 10$. This region is the so-called Mie-region, which has been named after Gustav Mie, who derived an analytical solution to calculate σ for spheres of all diameters (Mie, 1908).

When a radar sends its electromagnetic signal out into a precipitation system, there is usually not only a single point target but a large number of particles within the sample volume. The total backscattering cross-sectional area accumulates from the individual backscattering cross-sectional areas within a unit volume and is referred to as reflectivity η (Blahak, 2004; Zeng, 2013):

$$\eta = \sum_{i \in S} \sum_{j \in N_i} \sigma_{i,j}(D), \tag{4.7}$$

where S is the set of all hydrometeor types (rain drops, graupel and snow), and N_i the number of hydrometeors of type i within the volume. Equation (4.7) can also be expressed in continuous form via

$$\eta = \sum_{i \in S} \int_0^\infty \sigma_i(D) N_i(D) \mathrm{d}D, \tag{4.8}$$

where $N_i(D)$ is the particle size distribution of hydrometeor type i. Under Rayleigh condi-

tions, Equation (4.8) becomes

$$\eta = \sum_{i \in S} \frac{\pi^5 |K_i|^2}{\lambda^4} \int_0^\infty N_i(D) D^6 \mathrm{d}D. \tag{4.9}$$

Usually it is assumed that the scanned volume consists only of water, then the parameter related to the complex refractive index $|K|^2$ is replaced by that referring to water, $|K_w|^2 = 0.93$, yielding the so-called equivalent reflectivity

$$\eta_e = \frac{\pi^5 |K_w|^2}{\lambda^4} \int_0^\infty N_i(D) D^6 \mathrm{d}D. \tag{4.10}$$

Commonly, a different variable that does not depend on the wavelength is defined, called reflectivity factor

$$z = \int_0^\infty N_i(D) D^6 \mathrm{d}D. \tag{4.11}$$

In most applications, z is used since it has the advantage of being independent of the wavelength, and η and z are related through

$$z = \frac{\lambda^4}{\pi^5 |K|^2} \eta, \tag{4.12}$$

and inserting the equivalent reflectivity η_e defines the equivalent reflectivity factor z_e

$$z_e = \frac{\lambda^4}{\pi^5 |K_w|^2} \eta_e. \tag{4.13}$$

z_e has a large range of possible values, from 0 for no precipitation, 10 for light rain, 1000 to 10000 for moderate to strong rain, and 1000000 for severe rain and hail. Thus, the range of values covers several orders of magnitude. Therefore it has been found useful to compress z_e into logarithmic units via

$$Z_e = 10 \cdot \log_{10} \left(\frac{z_e}{mm^6/m^3} \right) \tag{4.14}$$

Z_e is the variable that is commonly used in meteorological applications, and it is often referred to as only "reflectivity".

4.2.2. Extinction

The power of electromagnetic radiation is reduced when traveling through any medium such as clouds or precipitation due to absorption and scattering of the signal in directions away from the radar. The amount of reduction depends on the density of the material and on the wavelength of the electromagnetic signal. Extinction is larger the smaller the wavelength: weather radars at S-Band (c.f. Table 4.1) suffer only little from extinction by precipitation, whereas extinction plays a significant role at X-Band. The extinction cross section σ_{ext} for a spherical target under Rayleigh conditions is given by (for a detailed derivation, see Blahak, 2004)

$$\sigma_{ext}(D) = \underbrace{\frac{\pi^2}{\lambda}\Im(-K)D^3}_{\text{absorption}} + \underbrace{\frac{2}{3}\frac{\pi^5}{\lambda^4}|K|^2 D^6}_{\text{scattering}}. \tag{4.15}$$

Generally, the extinction integrated along a path of the radar beam at range r is given by

$$\ell(r) = \exp \int_0^r \Lambda(r')\mathrm{d}r', \tag{4.16}$$

where $\Lambda(r')$ is called extinction coefficient and defined as

$$\Lambda(r') = \sum_{i \in S} \int_0^\infty \sigma_{ext,i}(D) N_i(D, r')\mathrm{d}D, \tag{4.17}$$

so the final form of the equivalent reflectivity factor z_e including extinction is given by

$$z_e = \frac{\lambda^4}{\pi^5 |K_w|^2} \frac{\eta_e}{\ell(r)^2}, \tag{4.18}$$

where squaring $\ell(r)$ accounts for two-way extinction.

4.3. Radar forward operator

As the philosophy of ensemble based Kalman filter methods is to compare model and observations in observation space, the observation forward operator H is a key feature in the assimilation process. In this study, the efficient modular volume radar operator (EMVORADO) (Blahak, 2008a; Blahak *et al.*, 2011; Zeng, 2013; Jerger, 2014; Zeng *et al.*, 2014), that is directly coupled to the COSMO-Model, is used. In the following, a brief overview is given.

4.3.1. Operator characteristics

EMVORADO works in two steps: first, radar reflectivity z_e and extinction coefficient Λ are computed for all hydrometeor types on the native model grid. Then, the grid point values are interpolated onto the radar coordinates depending on range r_0, azimuth α_0 and elevation angle ϵ_0. In the previous section, the relationship between backscattering objects and the measured reflectivity was explained. In order to simulate corresponding values based on numerical model fields, a formulation in radar coordinates depending on r_0, α_0 and ϵ_0 is needed. According to Blahak (2004); Zeng (2013), the pulse volume averaged radar reflectivity value at range r_0, azimuth α_0 and elevation angle ϵ_0 is given by

$$\langle z_e^{(R)} \rangle (r_0, \alpha_0, \epsilon_0) = \frac{\int_{r_0-c\tau/4}^{r_0+c\tau/4} \int_{\alpha_0-\pi}^{\alpha_0+\pi} \int_{\epsilon_0-\pi/2}^{\epsilon_0+\pi/2} \frac{z_e(r,\alpha,\epsilon)}{\ell_n^2(r,\alpha,\epsilon)} \frac{f_e^4(\alpha,\epsilon)}{r^2} \cos\epsilon \, \mathrm{d}\epsilon \, \mathrm{d}\alpha \, \mathrm{d}r}{\int_{r_0-c\tau/4}^{r_0+c\tau/4} \int_{\alpha_0-\pi}^{\alpha_0+\pi} \int_{\epsilon_0-\pi/2}^{\epsilon_0+\pi/2} \frac{f_e^4(\alpha,\epsilon)}{r^2} \cos\epsilon \, \mathrm{d}\epsilon \, \mathrm{d}\alpha \, \mathrm{d}r}, \tag{4.19}$$

where $z_e(r, \alpha, \epsilon)/\ell_n^2(r, \alpha, \epsilon)$ is the radar reflectivity accounting for two-way path integrated attenuation from the radar to location (r, α, ϵ). τ is the pulse duration, and f_e^4 is the approximate two-way effective beam weighting function of a scanning radar (Blahak, 2008b). Range weighting is approximated by a simple box-function. For the sake of efficiency, pulse volume averaging can be switched off, so the beam collapses to its central ray and the integrals in Equation (4.19) vanish:

$$\langle z_e^{(R)} \rangle (r_0, \alpha_0, \epsilon_0) = \frac{z_e(r, \alpha, \epsilon)}{\ell_n^2(r, \alpha, \epsilon)}. \tag{4.20}$$

The calculation of z_e is based on all hydrometeor types. For the one-moment bulk three-category ice scheme (Reinhardt and Seifert, 2006) used within COSMO, the modeled hydrometeor types are cloud drops q^c, rain drops q^r, cloud ice q^c, snow q^s and graupel q^g. Within the two-moment scheme (Seifert and Beheng, 2006), also hail is a prognostic model variable, which is however not used in this study. Radar reflectivity is simulated at the model grid points according to Blahak (2008b), where both Rayleigh and Mie scattering theory have been implemented. For Mie scattering, temperature-dependent refractive indices of the particles can be computed. Usually, Mie scattering would be computationally much more expensive than Rayleigh, but EMVORADO uses lookup-tables (Jerger, 2014) to make Mie scattering as efficient as Rayleigh. Rain drops are modeled as water spheres. Cloud ice and graupel are assumed to be one-layered spheres composed of a mixture of ice/air or ice/water/air in the case of partially melted particles. Snowflakes are modeled as two-layer

spheres with a denser core and a less denser shell. Melting of all ice categories is taken into account by a parameterized function depending on temperature and particle size. The drop size distributions needed for the calculation of z_e and Λ are chosen in a model-consistent way, i.e. for snow and graupel, an exponential drop size distribution is assumed, and for rain drops, a Gamma distribution is assumed (c.f. Section 2.3).

EMVORADO runs efficiently on parallel- and vector computers and its modular design allows to choose among different options useful for the particular application. In this study, Mie scattering theory has been used. Beam bending is approximated by the $4/3$ Earth radius approximation (c.f. Sec. 4.1), but more sophisticated options are possible, where beam bending is calculated according to the actual refractive index of the atmosphere determined during runtime (Zeng *et al.*, 2014).

4.3.2. Observation resolution

Due to the radar measurement strategy, data is irregularly distributed in space. For an azimuthal resolution of 1°, measurements from adjacent beams are approximately 0.17 km apart at a 10 km range, but diverge and at 180 km range, adjacent measurements are more than 3 km apart. So, the data density varies significantly with range and especially close to the radar site, the resolution is much finer than the model resolution. According to Liu and Rabier (2002), assimilating observations that are denser than the analysis grid is not necessarily advantageous and might even harm the analysis. The observations resolution can be reduced by different techniques, such as thinning or so-called superobbing. Thinning usually refers to the strategy to simply discard parts of the data, whereas superobbing combines information of several radar bins to obtain a so-called superobservation. EMVORADO implements both thinning and superobbing.

In the thinning strategy, only every ith radar bin in range from every jth azimuthal ray are kept, where the thinning parameters i and j can be set independently.

For superobbing, averaging is not based on areas defined by a fixed number of ranges and azimuths but is implemented in a quasi-cartesian way: A two-dimensional horizontal cartesian grid with a desired resolution L_x is defined. For each elevation separately, the radar bin centers are projected onto the cartesian plane. Then, for each cartesian point, the closest radar point is chosen. This point is the center of the superobservation, and it remains in the radar coordinate system and is not interpolated onto the cartesian grid. The cartesian grid only serves as auxiliary grid. Once the center point is chosen, radar bins surrounding the center bin are averaged, where the averaging area is a pie-wedge shaped area and its size is defined by a range and azimuth interval depending on the cartesian grid resolution. The number of radar bins averaged into one superobservation decreases with range. The width of the range interval is given by $L_x\sqrt{2}$, and the number of azimuths to

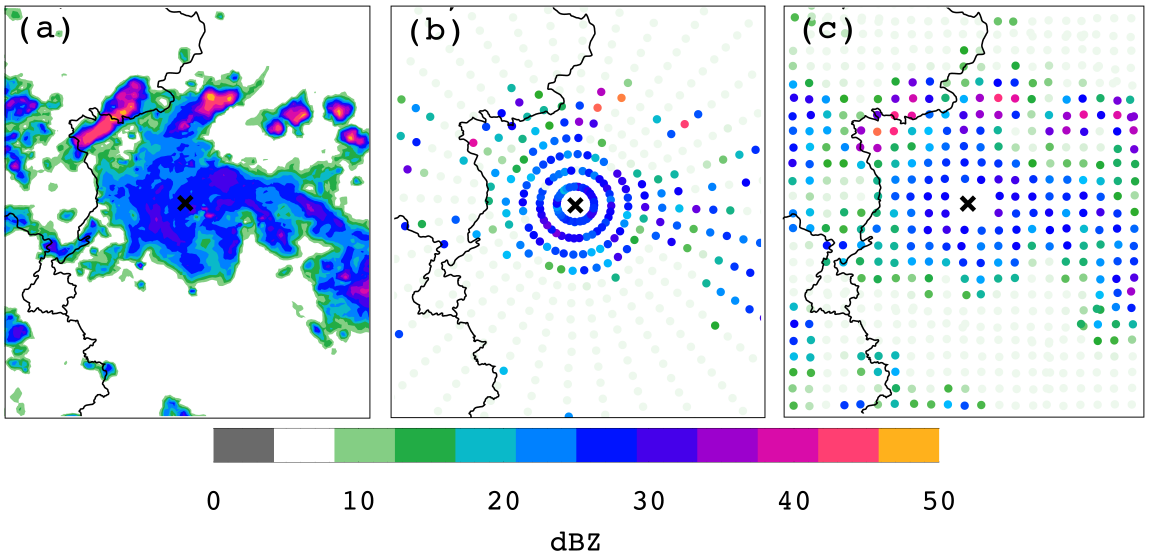


Figure 4.2.: Illustration of the thinning and superobbing procedure: a part of the COSMO-DE domain (approximately 220 km by 260 km) surrounding the radar site in Essen, Germany (black cross) is shown. A contour plot of the 0.5° PPI scan in its original resolution is shown in (a), the thinned observation (factor 10 in azimuthal and range direction) is shown in (b), and the superobbed observation with a resolution of 10 km is shown in (c). For superobbing, each averaging area (given in polar coordinates) approximates a quadratic area. This area is only well defined if it does not contain the radar site. This leads to small data gaps around the station.

be averaged at range r_0 is given by $2 \cdot \arctan\left(L_x\sqrt{2}/2 \cdot r_0^{-1}\right)$. Thus, for a cartesian grid with $L_x = 10$ km, close to the radar station, approximately 950 radar bins are averaged, whereas beyond a range of 100 km, less than 120 bins are averaged. In overlapping regions, observations from each radar site are so far treated as independent data.

In Figure 4.2, a measurement in its original resolution (Fig. 4.2 a) is compared with the thinned measurement (Fig. 4.2 b) with thinning factors 10 in both azimuth and range, and with the superobbed measurement (Fig. 4.2 c) with a cartesian grid resolution of 10 km. Thinning significantly reduces the amount of data. The observations are dense close to the radar site, but further away the thinned rays diverge and small cells are missed. By comparison, superobbing provides a homogeneous data distribution throughout the domain. Therefore, superobbing is used in this study. Since the analysis grid is coarsened by a factor of 3 (cf. Section 3.3), a superobbing resolution of 10 km is chosen for all experiments in order to ensure similar resolutions of model and observations.

4.3.3. Clear air information

The goal of radar data assimilation is not only to shift precipitation into the right location but it has been also found useful to suppress spurious precipitation by assimilating also signals in non-precipitating regions (Tong and Xue, 2005). This explicit treatment of no-

precipitation information is crucial: if only pairs where both the model and the observation show a valid value are handed to the LETKF, bins with clear air observed but precipitation forecasted are neglected and thus cannot be corrected in the assimilation step. Therefore, the value of these clear air bins is set to a value corresponding to no-precipitation.

Radar reflectivity values in logarithmic units usually range from -30 dBZ up to 60 dBZ. Since it is usually assumed that precipitation starts at values around 5-10 dBZ[3], a large range of values may originate from non-precipitating echoes. Therefore, deviations between no-precipitation observations and simulations might become very large without being linked to precipitation differences. In the assimilation step, these differences are translated into large, unphysical analysis increments. A common way to circumvent this issue is to threshold all Z_e values below a certain value. The absolute value chosen for this reference value is certainly a tunable parameter. Based upon preliminary experiments, a threshold of 5 dBZ has been chosen in this study; this threshold has previously been used in other studies, e.g. Aksoy *et al.* (2009). The threshold is applied before superobbing.

4.3.4. Observation errors

In this study, the observation error covariance matrix $\mathbf{R}$ (cf. Eqs. 3.4, or 3.10) is assumed to be diagonal, i.e. the observation errors of individual radar bins are assumed to be spatially uncorrelated. This assumption is a strong simplification, since individual radar bins are usually correlated, but due to the lack of appropriate knowledge, this assumption serves as a first approximation. The standard deviation of each observation bin is set to 10 dBZ. Although this value appears to be larger than what has been used in previous studies (e.g. 5 dBZ in Tong and Xue, 2005), this value has been found suited based on experimentation.

[3] A precipitation rate of 0.1 mm/h corresponds to approximately 10 dBZ according to the $Z - R$ relation by Marshall and Palmer (1948).

5. Experimental design and verification methods

In this section, the observation data sets and the study period are introduced (Sec. 5.1), the experimental design is defined (Sec. 5.2), and the verification scores are presented (Sec. 5.3).

5.1. Observation data sets and study period

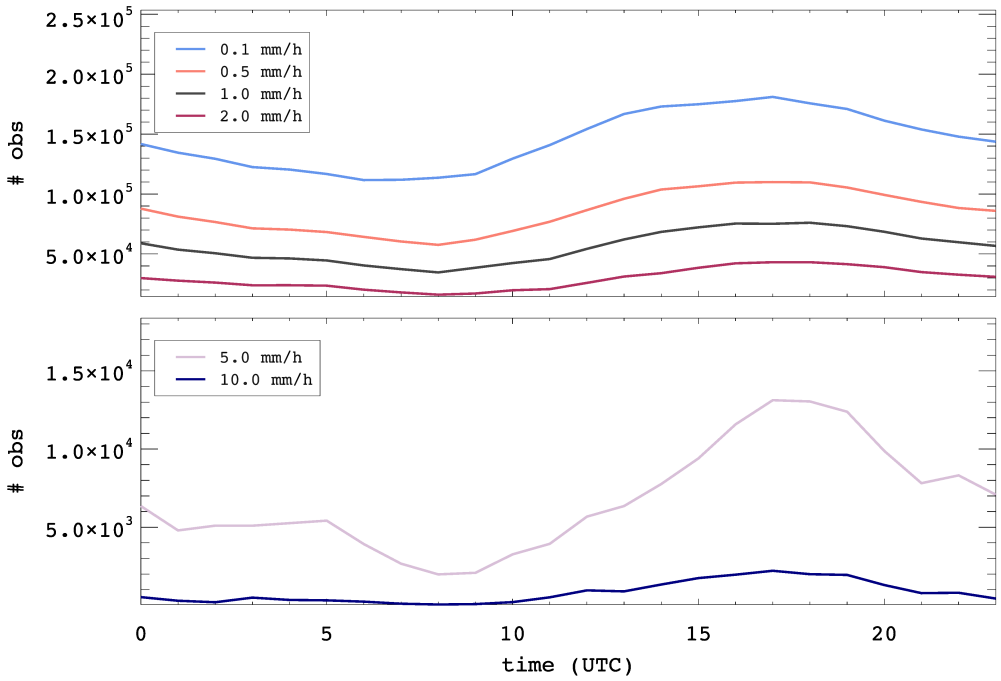


Figure 5.1.: Diurnal cycle of the number of observations of hourly radar derived precipitation rates exceeding different thresholds (0.1 mm/h, 0.5 mm/h, 1. mm/h, 2 mm/h, 5.0 mm/h, and 10.0 mm/h) observed during 22 May 2014 00 UTC and 29 May 2014 00 UTC.

The choice of the period is strongly limited by the availability of ICON ensemble boundary data. From the period with available boundary data (mid of May 2014 to mid of June 2014), the week 22 - 29 May 2014 covered the main precipitation events, and precipitation during this week had a clear diurnal cycle (Fig. 5.1 and 5.2). Further, the European

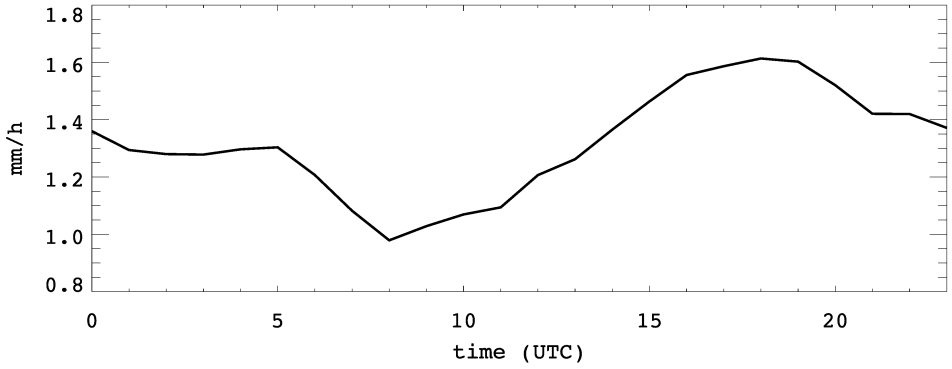


Figure 5.2.: Diurnal cycle of average radar derived precipitation intensity observed during 22 May 2014 00 UTC and 29 May 2014 00 UTC. The average is based on grid points exceeding 0.1 mm/h only.

Severe Weather Database (Dotzek *et al.*, 2009), a European database that collects severe thunderstorm reports, shows a large number of extreme hail and heavy rain reports during this week. Also severe lightnings caused damage and even tornados and funnel clouds were observed.

Conventional observations

Table 5.1.: Numbers of conventional observations (AMDAR, SYNOP, TEMP and wind profiler) assimilated in KENDA and number of radar bins after horizontal superobbing (c.f. Sec. 4.3.2). The numbers are average values based on the week 22-29 May 2014 and refer to the number of available observations within one hour. All numbers are approximate values. The number of available AMDAR and TEMP data varies throughout the day: due to night flying restrictions, there are less AMDAR measurements at night, and radiosondes are only launched 4 times a day. For each observation type, also the temporal resolution is given in column 3.

Observation type	Assimilated variables	Temporal resolution	Number of observations per hour
AMDAR	temperature	1 min	0-520
	horizontal wind		0-550
SYNOP	10 m horizontal wind	1 h	240
	surface pressure		600
TEMP	temperature	6 h	0-270
	horizontal wind		0-300
	rel. humidity		0-170
Wind profiler	horizontal wind	30 min	500
Radar	reflectivity	5 min	1.300.000

The set of conventional observations that is the basis for all data assimilation experiments in KENDA includes aircraft measurements (AMDAR) of temperature and horizontal wind, surface station measurements (SYNOP) of 10-meter horizontal wind and surface pressure,

radiosonde data (TEMP) of temperature, horizontal wind and humidity, and wind profiler measurements of horizontal wind. Table 5.1 summarizes the number of observations of each data type and variable.

DWD radar network

The radar network operated by DWD consists of 17 C-Band radar stations entirely covering Germany (see Fig. 5.3). The stations are operated in a volume mode with PPI scans at 10 elevation angles ranging from 0.5° to 25°. An entire volume scan is collected during a 5 minute time window. These exact measurement times are, however, not available, thus it is assumed that all observations are valid at discrete 5 minute intervals. In range, the measurements have a resolution of 1 km and in azimuthal direction, a resolution of 1°. Each station covers a radius of 180 km. In Table 5.1, the number of radar bins after horizontal superobbing (c.f. Sec. 4.3.2) is shown.

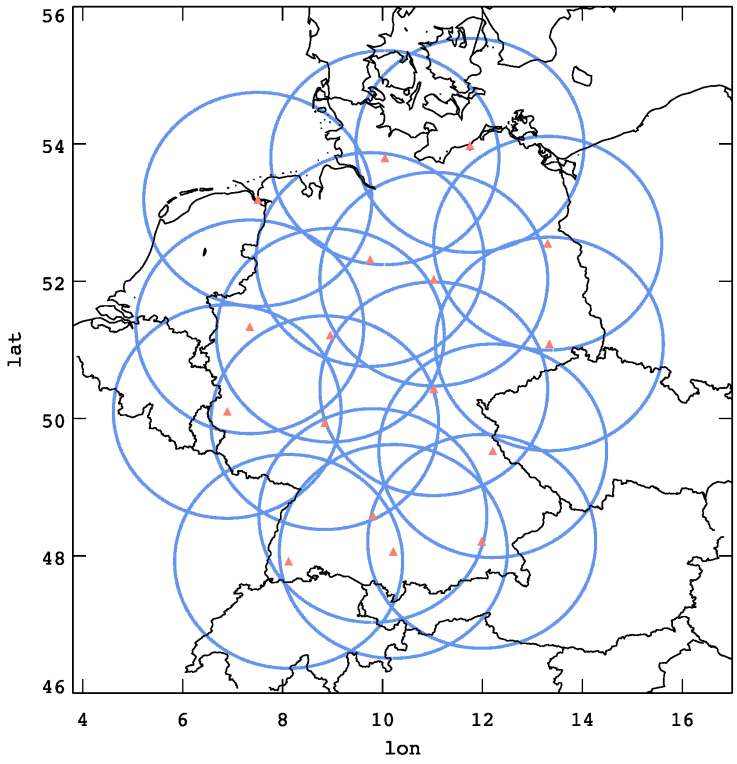


Figure 5.3.: The DWD radar network. For each of the 17 stations (shown by the red triangles), its range is shown by a circle with a radius of 180 km.

For data assimilation, quality control is an important aspect in order to avoid the assimilation of erroneous data. For radar data, this is particularly difficult, since there are numerous error sources in radar measurements, such as beam blockage, attenuation, ground clutter, anomalous propagation, and radar miscalibration (Villarini and Krajewski, 2010). For the

DWD radar network, a clutter filter is applied to the measurements at the radar site to remove ground clutter from the observations. This filter however only works successfully for static clutter and moving features such as wind turbines or insects and birds are not removed. The latter are of small intensity and are potentially removed by thresholding the data at 5 dBZ (c.f. Sec. 4.3.3). However, so far it cannot be assumed that all signals in the assimilated data are correct, which offers possibilities for further improvement.

5.2. Description of the experiments

Based on the data assimilation algorithms described in the preceding sections, different combinations of algorithms for assimilating conventional and precipitation observations are possible. The LETKF and LHN can be combined by assimilating conventional observations via KENDA and by using LHN to assimilate radar derived surface precipitation rates. LHN is applied to each ensemble member and the deterministic run individually. This setup is the current reference system at DWD and is scheduled for pre-operational testing. In a full LETKF system, all available observations are assimilated via KENDA, i.e. radar reflectivities are included via EMVORADO.

The combination of nudging and LHN is the current operational deterministic data assimilation system at DWD. The system assimilates conventional observations via nudging and radar derived surface precipitation rates via LHN. This setup is however not investigated in this study.

Experiment 1: Dependency of the update frequency

Radar data is available at a high temporal resolution. Thus, an obvious assumption is that frequent assimilation updates improve the analysis and subsequent forecasts. This hypothesis is examined in this first set of experiments. A one-day case study within the experimental period, 26 May 2014, with varying update frequency is performed and the quality of the analysis and forecast is studied. The experimental setup is as follows: The ensemble with 40 members is initialized at 00 UTC based on downscaled fields from the global model ICON. Prior to data assimilation, the ensemble is propagated for 12 hours to ensure that the ensemble develops a realistic approximation of the convective scale forecast error structures (following Zhang *et al.* 2009). In the assimilation cycle from 12 to 15 UTC, conventional and radar reflectivity observations are assimilated every 5, 15, 30 or 60 minutes with the full LETKF system. The setups are called CONV+RAD_5, CONV+RAD_15, CONV+RAD_30, and CONV+RAD_60. In each analysis step, all observations collected during the previous first guess window are assimilated by the 4D-LETKF. This means, for example, CONV+RAD_15 assimilates observations from three time steps: observations

Table 5.2.: Summary of the experimental setups used in experiments 1 and 2.

	Conventional (AMDAR, SYNOP, TEMP, Wind profiler)	Reflectivity	Rain rate (2D)	Update frequency	4D-weighting
CONV	✓			60	not applicable
CONV+RAD ($\hat{=}$ CONV+RAD_60)	✓	Entire volume		60	✓
CONV+RAD_30	✓	Entire volume		30	✓
CONV+RAD_15	✓	Entire volume		15	✓
CONV+RAD_5	✓	Entire volume		5	not applicable
CONV+RAD_1ELE	✓	Lowest elevation		60	✓
CONV+RAD_4ELE	✓	Every third elevation		60	✓
CONV+RAD_3DLETKF	✓	Entire volume		60	
CONV+LHN	✓		✓	60	not applicable

from the analysis time, and 5 and 10 minutes prior to the analysis. For verification, a reference simulation named CONV assimilates only conventional observations with hourly updates. The last analysis at 15 UTC provides the initial conditions for a 6 hour ensemble forecast and an additional deterministic forecast for the period 15 to 21 UTC.

Experiment 2: Evaluation over one week

This experiment is not restricted to a few test cases but assesses the impact of radar reflectivity assimilation on deterministic forecasts based on a simulation of seven days. The basic cycle of the experiments is as follows: The assimilation cycle is initialized on 21 May 2014 at 12 UTC by interpolating the global model fields provided by ICON to the COSMO-DE grid. With hourly update intervals, the model is continuously cycled until 29 May 2014, 00 UTC. Throughout this period, 29 deterministic forecasts with a lead time of 24 hours are initialized in 6 hourly intervals. The first forecast is initialized after a 12 hour spinup on 22 May, 00 UTC, the last forecast is initialized on 29 May, 00 UTC.

Firstly, the full LETKF system is compared with DWD's current reference system that uses the LETKF for conventional observations and LHN for radar derived precipitation rates, called CONV+LHN. As in experiment 1, a system CONV is set up that assimilates only conventional observations.

Secondly, this experiment assesses the impact of the amount of radar information assimilated. Intuitively, more observational data might improve analysis and forecast, however, the handling of large amounts of data requires many computational resources. Thus, for the full LETKF system, the value of the dense vertical information and the 4D-formulation of the LETKF (c.f. Sec. 3.1.4) is investigated in the following setups:

- CONV+RAD: As in CONV+RAD_60, entire radar reflectivity volumes are assimilated in addition to conventional observations. The trajectory of the entire first guess window is included.

- CONV+RAD_1ELE: Only the lowest PPI scan at 0.5° elevation angle is assimilated. Due to the propagation of the radar beam, there is still information on different vertical levels available in areas of station overlap (c.f. Section 4.1, Eq. 4.4). The trajectory of the entire first guess window is included.

- CONV+RAD_4ELE: Here, more and higher elevation scans are considered, namely every third PPI scan (elevations at 0.5°, 3.5°, 8.0°, and 25.0°). The trajectory of the entire first guess window is included.

- CONV+RAD_3DLETKF: Full radar reflectivity volumes are assimilated, but only measurements that are valid at the exact analysis time. Prior measurements during

the first guess window are not used.

A summary of the setups used in experiments 1 and 2 is given in Table 5.2.

5.3. Verification scores

The major goal of this study is to improve precipitation forecasts. For an objective verification, appropriate scores are needed. Traditional verification scores suffer, if applied to precipitation, from the so-called double penalty problem. In point-by-point comparisons, a small displacement of a precipitation event in the forecast is punished twice: first, for being in the wrong place, second, for missing the right place. In contrast, a forecast that does not predict the event at all is only punished once for not forecasting the event. The development of ensemble verification scores that do not suffer from the double penalty problem is however still a relatively new research area and thus, a score with this property is here only considered for deterministic forecasts. In the following, the scores used to verify the ensembles of experiment 1 and the deterministic forecasts of experiment 2 are described.

Ensemble based verification

In experiment 1, the ensemble forecast is verified using the Brier score (BS, Wilks, 2006) and the continuous ranked probability score (CRPS, Matheson and Winkler, 1976). The observational data used for verification is first the superobbed reflectivity data of the lowest PPI scan in dBZ and second, radar derived precipitation data in mm/h. For the first, in regions with station overlap, the data of the closest station are used.

The Brier score measures the accuracy of probabilistic forecasts in terms of the mean squared error between forecast and observation events. Events are defined by the exceedance of a threshold. Thus, observations and model output are transferred into binary fields: The observation event y_k at a particular grid point k is 1 if the event occurred or 0 otherwise. The probabilistic forecast of the event occurring is given by the fraction of the ensemble members exceeding the threshold, denoted $p(x_k)$. The Brier score is thus defined as

$$BS = \frac{1}{n}\sum_{k=1}^{n}(p(x_k) - y_k)^2, \tag{5.1}$$

where n is the number of model grid points.

When measuring the accuracy of forecasts, a common goal is to assess the skill of a forecast relative to a reference. Thus, forecast skill expresses the improvement obtained

by a particular forecast relative to the reference. Forecast skill in terms of a measure A is usually obtained via a skill score SS

$$SS(A) = 1 - \frac{A}{A_{ref}}. \tag{5.2}$$

The Brier skill score (BSS) is thus defined as

$$BSS = 1 - \frac{BSS}{BSS_{ref}}, \tag{5.3}$$

and in the context of experiment 1 calculated via

$$BSS = 1 - \frac{BS(\text{CONV+RAD\_5/15/30/60})}{BS(\text{CONV })} \tag{5.4}$$

with CONV taken as the reference. The BSS ranges from -1 to 1, where positive values indicate an improvement by the "new" system relative to the reference, and negative values indicate a degradation compared to the reference, and 0 refers to equal performance.

For continuous variables, the CRPS is used and defined as

$$CRPS(F, y) = \int_{-\infty}^{\infty} \left(F(x) - 1\{x \geq y\}\right)^2 \mathrm{d}x, \tag{5.5}$$

where F is the cumulative distribution function of the forecast, and the Heaviside function $1\{x \geq y\}$ is defined as

$$1\{x \geq y\} = \begin{cases} 1, & \text{if } x \geq y \\ 0, & \text{otherwise.} \end{cases} \tag{5.6}$$

The CRPS is commonly interpreted as the integral of the Brier score for all real-valued thresholds (Hersbach, 2000). According to Gneiting and Raftery (2007), an equivalent form of the CRPS is given by

$$CRPS(F, y) = \mathrm{E}_F|X - y| - \frac{1}{2}\mathrm{E}_F|X - X'|, \tag{5.7}$$

where X and X' are independent samples of a random variable with distribution F. This formulation allows the derivation of an explicit calculation of the CRPS in case of a discrete ensemble of size N representing the forecast distribution via (Grimit *et al.*, 2006)

$$CRPS = \frac{1}{N}\sum_{i=1}^{N}|x_i - y| - \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}|x_i - x_j|. \tag{5.8}$$

As for the BSS, the continuous ranked probability skill score (CRPSS) is obtained via

$$CRPSS = 1 - \frac{CRPS(\text{CONV+RAD\_5/15/30/60})}{CRPS(\text{CONV })}. \tag{5.9}$$

Deterministic verification

The deterministic forecast carried out within experiment 2 are verified with the fraction skill score (FSS, Roberts and Lean, 2008), the equitable threat score, and the frequency bias (ETS and FBI, Wilks, 2006).

The FSS is also a skill score of the form

$$FSS = 1 - \frac{MSE}{MSE_{ref}}. \tag{5.10}$$

where MSE denotes the mean squared error. Similarly to the Brier score (c.f. Eq. 5.1), the MSE is defined as

$$MSE = \frac{1}{n} \sum_{k=1}^{n} \left(p_s(x_k) - p_s(y_k)\right)^2, \tag{5.11}$$

where n is the number of model grid points and the probabilities $p_s(x_k)$ and $p_s(y_k)$ describe the probability of an event occurring based on the exceedance of a threshold for the model x_k and observations y_k. In contrast to the Brier score, where the forecast probability $p_s(x_k)$ is obtained from the ensemble, here the probability is obtained by averaging over a certain spatial scale or neighborhood s, which is done for the observation and the model. The spatial scale is usually defined by an area of adjacent model grid points. By using a scale s larger than grid point scale, the FSS relaxes the need of a direct match between model and observations and thus circumvents the double penalty problem. By varying s, the FSS provides information about how skill varies with spatial scale. The typical reference in the FSS is the worst possible MSE that can be obtained from the underlying model and observation fields and is calculated via

$$MSE_{ref} = \frac{1}{n} \sum_{k=1}^{n} p_s(x_k)^2 + \frac{1}{n} \sum_{k=1}^{n} p_s(y_k)^2, \tag{5.12}$$

which is motivated by neglecting all common non-zero grid points of model and observation (i.e. neglecting the term $2 \cdot p_s(x_k)p_s(y_k)$ that arises by expanding the right hand side of Eq. 5.11). The FSS ranges from 0 to 1, where the value 1 is assigned to a perfect forecast.

The ETS and FBI are verification measures for non-probabilistic forecasts and are based

Table 5.3.: 2×2 Contingency Table based on binary outcomes of observation and forecast with notations for hits (a), false alarms (b), misses (c) and correct negatives (d).

		Observation	
		Yes	No
Forecast	Yes	a	b
	No	c	d

on the contingency table (Table 5.3) that consists of frequencies of binary forecasts and observations, which is again defined by the exceedance of a threshold. The ETS relates the forecast *yes* outcomes (event simulated) to the observed *yes* outcomes (event observed) accounting for random hits:

$$ETS = \frac{a - a_{ref}}{a - a_{ref} + b + c}, \tag{5.13}$$

where

$$a_{ref} = \frac{(a+b)(a+c)}{a+b+c+d}. \tag{5.14}$$

The ETS ranges from $-1/3$ to 1 with perfect score equal to 1. The FBI is defined as the ratio of *yes* forecasts to *yes* observations:

$$FBI = \frac{a+b}{a+c} \tag{5.15}$$

which is 1 for unbiased forecasts. Values below 1 indicate underforecasting, values greater than 1 indicate overforecasting of precipitation events.

The prognostic model variables not directly linked to precipitation are compared against observations in order to ensure that the radar data assimilation does not have a negative impact on variables such as temperature, wind, or humidity. Therefore, the root mean squared error (RMSE) and bias against SYNOP station data are calculated according to

$$RMSE = \sqrt{\frac{1}{m}\sum_{k=1}^{m}(x_k - y_k)^2}, \tag{5.16}$$

and

$$bias = \frac{1}{m}\sum_{k=1}^{m}(x_k - y_k), \tag{5.17}$$

where m is the number of observations.

Bootstrapping

A common method to assess the uncertainty in forecast verification scores is the bootstrap method (Efron and Tibshirani, 1993). From the original data set, e.g. an ensemble or multiple deterministic forecasts, m samples are drawn with replacement, the so-called bootstrap samples. Each bootstrap sample is supposed to be of the same size as the original sample. The forecast verification score of interest (e.g. the Brier score) is then calculated for each of the bootstrap samples, which can then again be related to the bootstrap samples of a reference (for e.g. the BSS). Percentiles of these m realizations are used to estimate the uncertainty of the respective score. In this study, the bootstrap samples consist of 1000 realizations. In case of experiment 1, bootstrapping is applied to the ensemble, and in case of experiment 2, bootstrapping is applied to the set of multiple deterministic forecasts.

6. Results

6.1. Experiment 1: Dependency of the update frequency

The first experiment studies the influence of radar reflectivity data assimilation in addition to conventional data and assesses the impact of the data assimilation frequency on the quality of analysis and forecast. Figure 6.1 displays hourly accumulated precipitation from observations and from the model (deterministic run). The upper row shows CONV, the reference, and, exemplarily, CONV+RAD_60 at 15 UTC compared to the precipitation rate product derived from the radar measurement. Precipitation is accumulated between 14 and 15 UTC, i.e. the first row of Figure 6.1 corresponds to the model state immediately before the last assimilation step. By visual comparison, CONV+RAD_60 is clearly in better agreement with the observation (especially over the Netherlands or the southern Germany). The middle row of Figure 6.1 is valid after three hours forecast, i.e. at 18 UTC. Still, CONV+RAD_60 agrees better with the observation than CONV (e.g. over the western part of Germany). After six forecast hours (bottom row of Figure 6.1), it is difficult to judge simply by comparing visually whether one setup is superior. Both forecasts underestimate the intensity of the event.

Figure 6.2 shows the number of ensemble members exceeding 0.5 mm/h of accumulated precipitation for CONV+RAD_5/15/30/60 (Fig. 6.2 b, c, e, f) at the end of the assimilation window (15 UTC) and compares the four setups with the actual radar derived precipitation field (Fig. 6.2 d), and the observation transferred into a binary field (Fig. 6.2 a). For all update frequencies, the observations are very well reproduced. CONV+RAD_5/15/30 lead to very similar high numbers of ensemble members which reproduce precipitation rates above this particular threshold. Especially for CONV+RAD_5 and CONV+RAD_15, the red areas (i.e. areas where all 40 ensemble members exceed the threshold of 0.5 mm/h) coincide with the binary observation. The ensemble members differ only outside the radar observed area. In CONV+RAD_60, the ensemble exhibits more spread also within the radar observed area. Since the analysis increments are a linear combination of the ensemble perturbations, more spread is not necessarily a disadvantage. Quite the contrary, ensemble spread is required to successfully assimilate observations. However, the large red areas

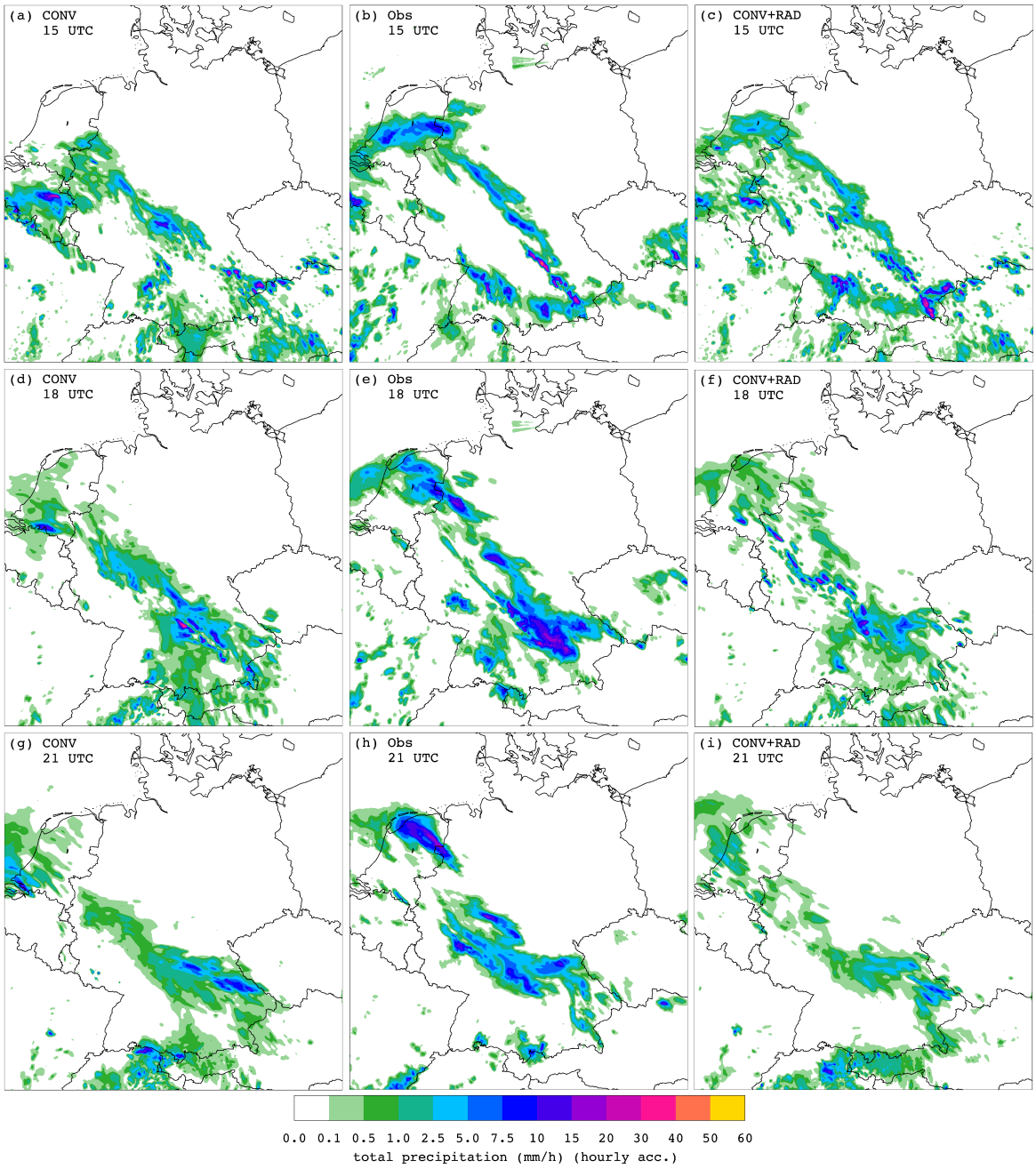


Figure 6.1.: Hourly precipitation accumulation for experiment 1 (26 May 2014). The radar-derived precipitation is shown in the middle column (b, e, h), the model-generated precipitation of setup CONV is shown in the left column (a, d, g), and the model-generated precipitation of setup CONV+RAD_60 is shown in the right column (c, f, i). Figures (a-c) are valid at 15 UTC, Figures (d-f) at 18 UTC and Figures (g-i) at 21 UTC.

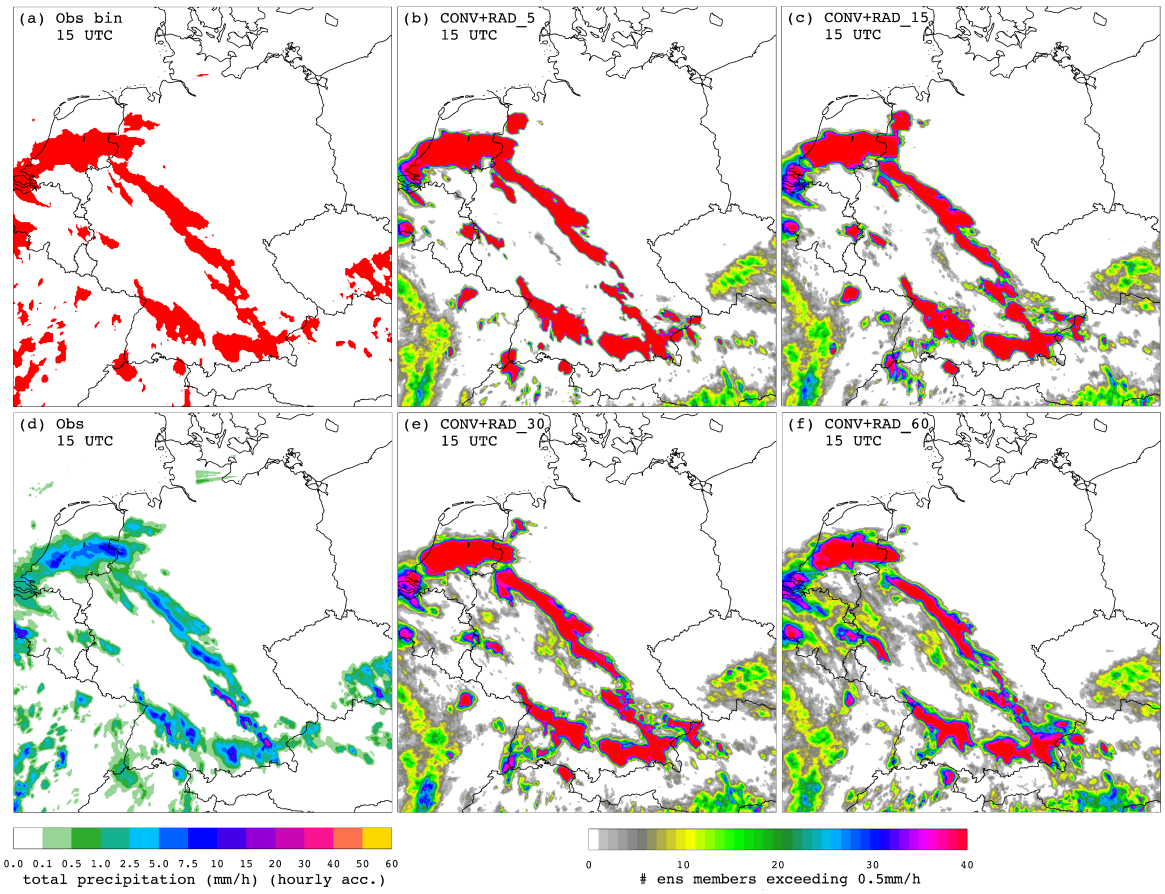


Figure 6.2.: First guess ensemble and radar derived precipitation rate (experiment 1, 26 May 2014) accumulated between 14 and 15 UTC. In (a), the observation is transformed to a binary field: grid points exceeding 0.5 mm/h are shown in red. For the ensemble, the number of ensemble members exceeding 0.5 mm/h is counted for all 4 update frequencies: (b) CONV+RAD_5, (c) CONV+RAD_15, (e) CONV+RAD_30, and (f) CONV+RAD_60. As a reference, the observed total precipitation is shown in (d).

indicate that the ensemble members of all four setups have similar precipitation system structures.

6.1.1. Quantitative verification

For objective comparison, the BSS and CRPSS are computed (c.f. Section 5.3). Due to technical constraints, precipitation is not contained in the analysis files, thus the model needs to be run for a short time interval (until the next observation is available; here for five minutes) to apply the forward operator and calculate the dBZ model equivalent. This 5-minute forecast is taken as a proxy for the analysis during the assimilation window. The BSS of CONV+RAD_60 with CONV as reference (solid red line in Fig. 6.3) strongly increases during the assimilation window showing a very beneficial influence of the radar reflectivity assimilation. This positive impact is visible when verifying against reflectivities (Fig. 6.3 a)

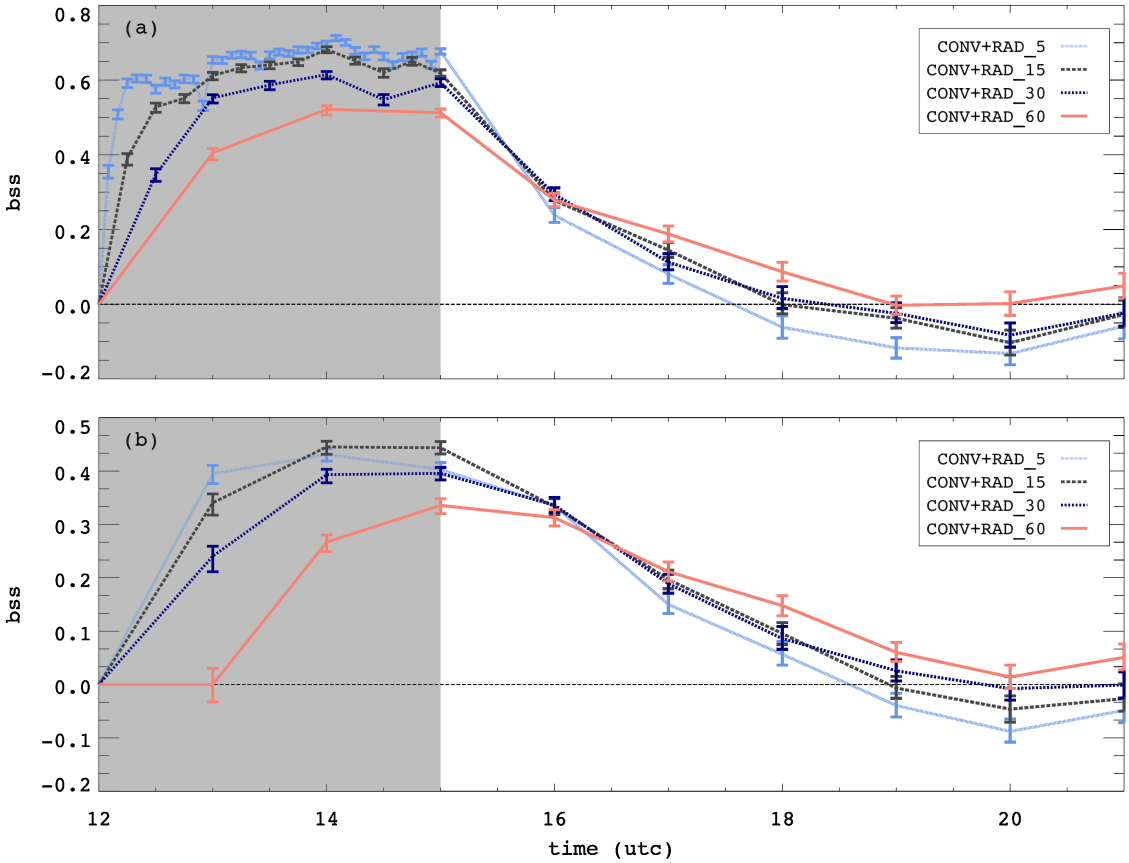


Figure 6.3.: The BSS for setups CONV+RAD_5 (light blue), CONV+RAD_15 (dark gray), CONV+RAD_30 (dark blue), and CONV+RAD_60 (red) with CONV as the reference. The BSS is plotted against time. The gray shaded area corresponds to the assimilation window. The error bars are obtained via bootstrapping (2.5% and 97.5% percentiles). In (a), the BSS is based on instantaneous reflectivity measurements with a threshold of 20 dBZ. During the assimilation window, the analysis is approximated by a 5 minute forecast. For graphical display, these approximations are shown at analysis times (i.e. every 5, 15, 30 or 60 minutes). In (b), the BSS is based on precipitation rates with a threshold of 0.5 mm/h.

and against radar derived precipitation rates (Fig. 6.3 b). For the more frequent updates, an even stronger increase of skill during the assimilation window is visible. In Figure 6.3 a, a clear hierarchy among the four setups is visible at the end of the assimilation window: the most frequently updated setup yields the best analysis, the least frequently updated setup the worst analysis. All four setups are able to improve the analyses compared to CONV. After the assimilation window, skill decreases for all four setups. After one forecast hour, the four systems are indistinguishable. For CONV+RAD_60, skill remains significantly positive for the first three forecast hours based on reflectivities. For accumulated precipitation, the results indicate even a positive impact throughout the entire forecast window. For the more frequently updated setups, a degradation of forecast quality compared to CONV is visible after 3 hours in Figure 6.3 a and after four forecast hours in Figure 6.3 b as shown

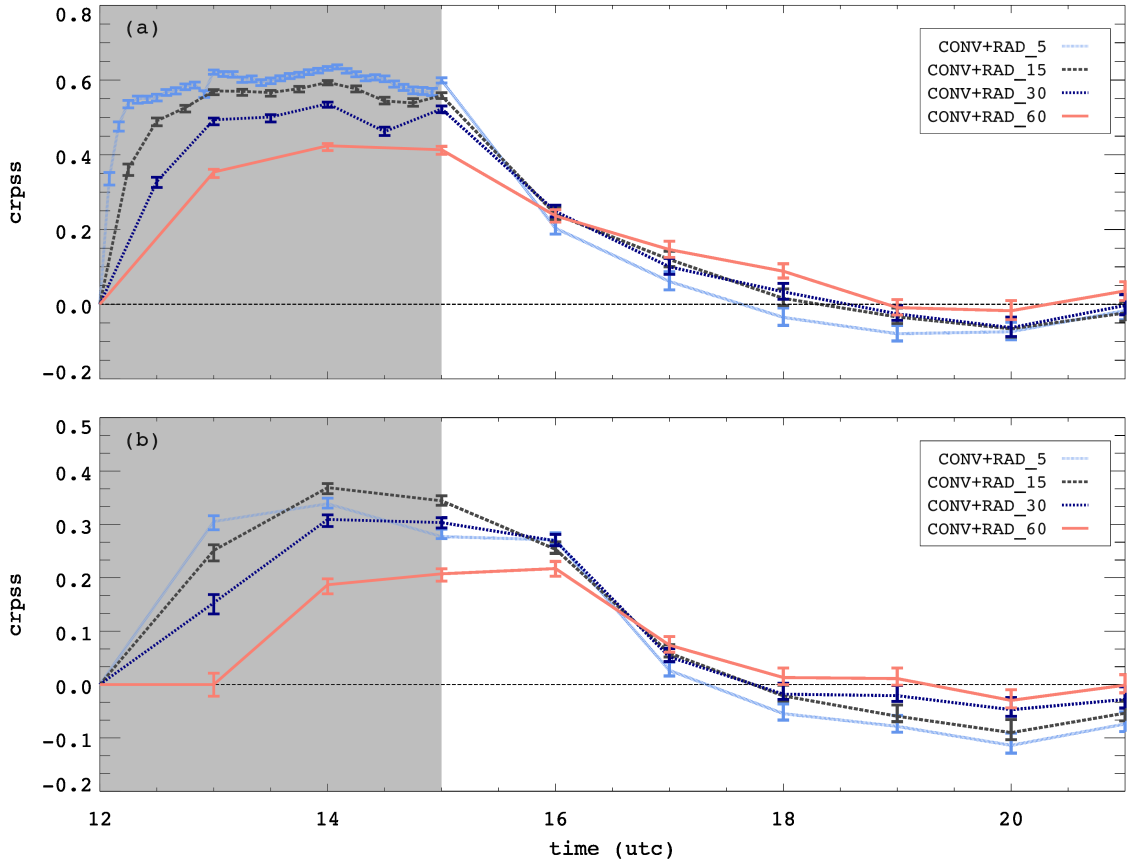


Figure 6.4.: The same as Figure 6.3, but for the CRPSS.

by the negative values of the BSS.

The CRPSS of the four CONV+RAD setups with CONV as reference (Fig. 6.4) evolves similarly to the BSS: during the assimilation window, a significant increase of skill is visible in terms of reflectivities (Fig. 6.4 a) and hourly precipitation rates (Fig. 6.4 b). More frequent updates lead to better analyses, but also suffer from the drawback of worse forecast quality. For CONV+RAD_60, a positive impact is visible for up to four hours (Fig. 6.4 a and b), whereas CONV+RAD_5 performs worse than CONV from forecast hour three.

6.1.2. Surface pressure tendencies

In data assimilation, analysis increments are added to the model state that may disrupt the dynamical balance of the model. Thus, numerical noise is introduced if structures in the analysis increments, especially small-scale structures, are incompatible with the dynamics of the model and if the model is not able to dissolve mass-momentum imbalances (Stauffer and Seaman, 1990). In the ensemble Kalman filter framework, gravity waves are a common problem (Mitchell *et al.*, 2002). A measure to quantify these imbalances is the first time

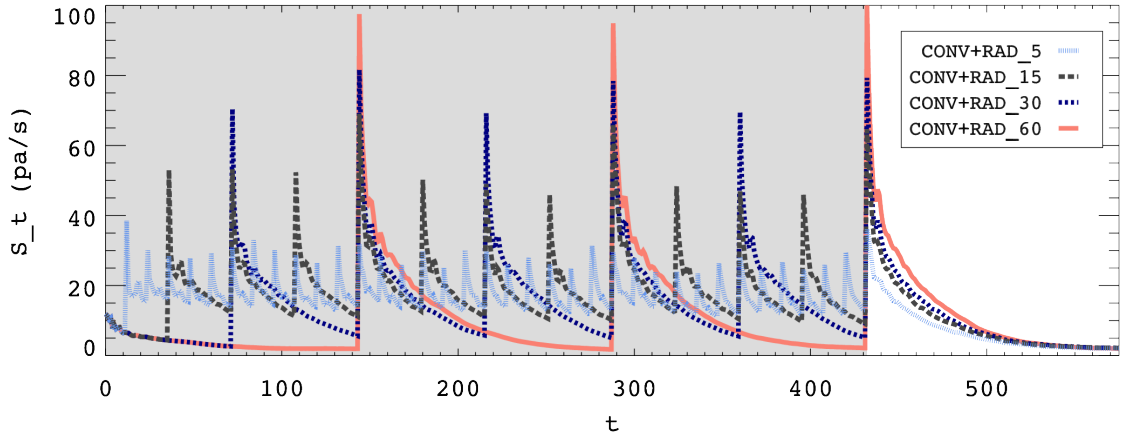


Figure 6.5.: Domain averaged surface pressure tendencies for each integration time step (1 time step $\hat{=}$ 25 s, i.e. 3 h $\hat{=}$ 432 time steps) during the assimilation window (shaded in gray) for CONV+RAD_5 (light blue), CONV+RAD_15 (dark grey), CONV+RAD_30 (dark blue), and CONV+RAD_60 (red).

derivative of surface pressure, averaged over the model domain (Stauffer and Seaman, 1990; Chen and Huang, 2006):

$$S_t = \frac{1}{n_1 \cdot n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left| \frac{\partial p_s}{\partial t} \right|_{ij}, \tag{6.1}$$

where p_s is the surface pressure and summation is done over the entire model domain. Figure 6.5 shows how S_t based on the deterministic runs of CONV+RAD_5/15/30/60 evolves during the assimilation window. When data is assimilated less frequently, the model is potentially further away from the observation and analysis increments are larger. Therefore, less frequent updates cause a more intense shock and produce higher noise than more frequent updates. Although the noise introduced by data assimilation is the strongest in CONV+RAD_60, only here the surface pressure tendency of the model state decays almost to zero before the next assimilation step, i.e. only for CONV+RAD_60 the model state is able to adjust to the changes introduced by data assimilation. Especially for CONV+RAD_5 and CONV+RAD_15, the noise does not peak as high as for CONV+RAD_60, but the model state is not able to recover completely from the analysis and remains at a particular noise level throughout the entire assimilation window. This persistent noise suggests that the analyses of the frequently updated setups are less physically consistent. Therefore, forecasts initialized from these potentially unbalanced model states are worse compared to the less frequently updated setups.

The data assimilation cycling framework BACY (cf. Section 3.3.3) runs COSMO-DE and KENDA via shell scripts, i.e. output files are written by each modules to pass the data.

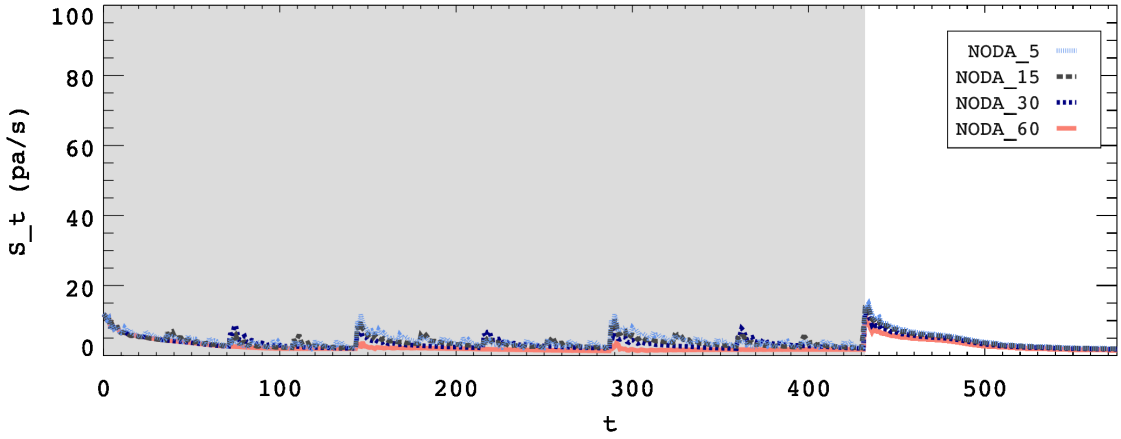


Figure 6.6.: As Figure 6.5, but for NODA_5 (light blue), NODA_15 (dark grey), NODA_30 (dark blue), and NODA_60 (red).

After each LETKF-step, COSMO is newly initialized based on a restricted set of model fields provided by the LETKF[1]. Therefore, noise arises by newly initializing COSMO. In order to quantify the noise caused by reinitialization, four runs are carried out without data assimilated but COSMO interrupted and newly initialized at frequencies corresponding to the assimilation frequencies (called NODA_5/15/30/60). The influence of reinitialization on the surface pressure tendency is approximately an order of magnitude smaller than what can be observed in the CONV+RAD setup (Fig. 6.6), thus the main contribution to the noise in Figure 6.5 can be attributed to the assimilation of observations.

6.2. Experiment 2: Evaluation over one week

6.2.1. Comparison against latent heat nudging

In experiment 2, 29 deterministic forecasts initialized throughout one week in May 2014 are evaluated. The FSS is calculated for a neighborhood region of 5 grid points in each direction, i.e. a total width of 11 grid points is used to measure the match between model and observations. The thresholds used are 0.5 mm/h and 2 mm/h. The higher the threshold, the rarer are the events, which impedes the verification of high thresholds for experiment 1. For this purpose, longer evaluation periods are needed to reduce uncertainty of the comparison metrics. However, even during the entire week in May, there are only few observations exceeding 5 mm/h or higher (see Fig. 5.1).

For 0.5 mm/h, a significant improvement in forecast skill is achieved by assimilating radar reflectivites (Fig. 6.7 a). During the first four forecast hours, the FSS of CONV+RAD is

[1] In contrast, restart setups offer the possibility to save information about the entire model state.

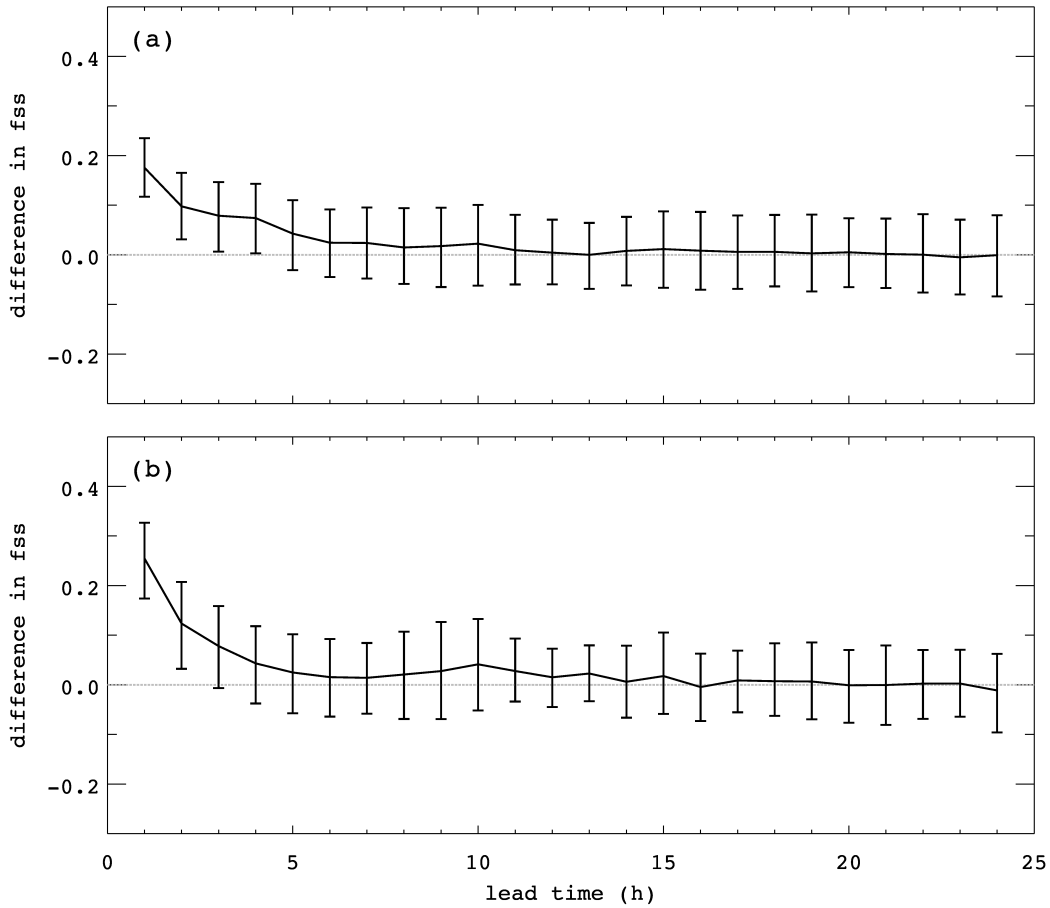


Figure 6.7.: FSS of 29 deterministic forecasts of CONV+RAD minus CONV against forecast lead time. The FSS is calculated for a neighborhood of 5 grid points (in each horizontal direction) and (a) a threshold of 0.5 mm/h, and (b) a threshold of 2 mm/h. The error bars are obtained via bootstrapping. 2.5% and 97.5% percentiles are the end parts of the bars.

significantly better than for CONV, as indicated by the positive values in Figure 6.7 a. A slight positive impact can be seen even for forecast hours six to ten but without statistical significance. For longer lead times, the effect is neutral, since the analysis field is progressively advected outside the domain and dynamics thus governed by the driving lateral boundary data. For the 2 mm/h threshold, the positive impact in the first forecast hour is even stronger compared to 0.5 mm/h. A significant improvement compared to CONV is achieved for the first three forecast hours. As for 0.5 mm/h, CONV+RAD does not become worse than CONV for longer forecast lead times.

The difference in FSS between CONV+RAD and CONV+LHN for 0.5 mm/h, (Fig. 6.8 a) documents a slight but yet significant advantage of CONV+RAD over CONV+LHN in the first forecast hour. Throughout the entire forecast lead time, CONV+RAD competes with CONV+LHN. For 2 mm/h, CONV+LHN seems slightly better than CONV+RAD during the first three hours, but this result is not significant. For forecast hours 4 to 24, both

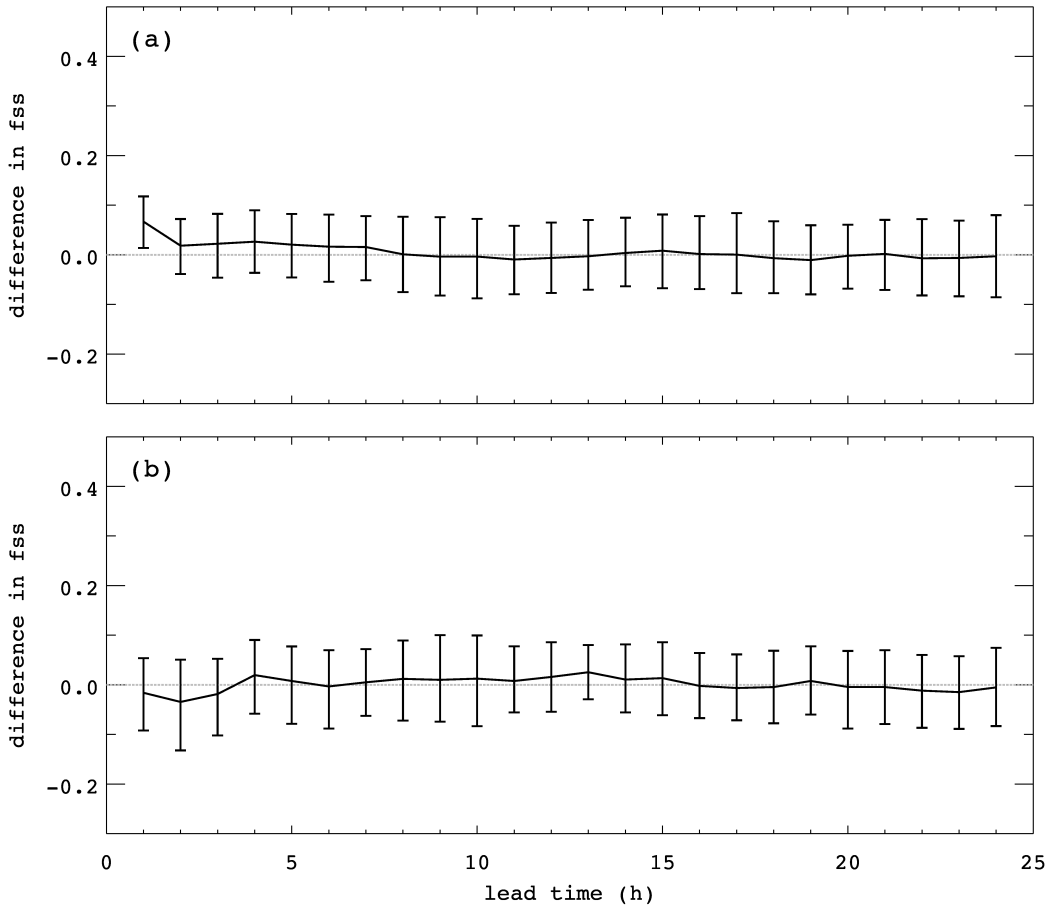


Figure 6.8.: As Figure 6.7, but for the difference CONV+RAD minus CONV+LHN.

setups perform mostly equally.

The equitable threat score (ETS, Fig. 6.9) decreases with forecast lead time for all three setups for 0.5 mm/h (Fig. 6.9 a) and for 2 mm/h (Fig. 6.9 b). Similarly to the FSS, a significant improvement over CONV is visible both for CONV+RAD and CONV+LHN during the first forecast hours, while the latter two only slightly deviate. For longer lead times, the ETS for all setups converges. The frequency bias (FBI, Fig. 6.10) shows a significant underestimation of events for 0.5 mm/h (Fig. 6.10 a) and for 2 mm/h (Fig. 6.10 b) throughout the entire forecast window, which is most pronounced during the first 10 forecasts hours, similarly for all three setups. For 2 mm/h, CONV+LHN performs slightly better during the first forecast hours.

In the following, it is examined if the systems are able to reproduce the diurnal cycle of precipitation. In Figure 6.11, the diurnal cycle of precipitation events exceeding the thresholds 0.5 mm/h (Fig. 6.11 a) and 2 mm/h (Fig. 6.11 b) during the assimilation cycle are shown. For 0.5 mm/h, all setups similarly overestimate the observed precipitation during night and do not properly capture the afternoon maximum. For 2 mm/h, CONV strongly underestimates the entire diurnal cycle, especially during afternoon. Dur-

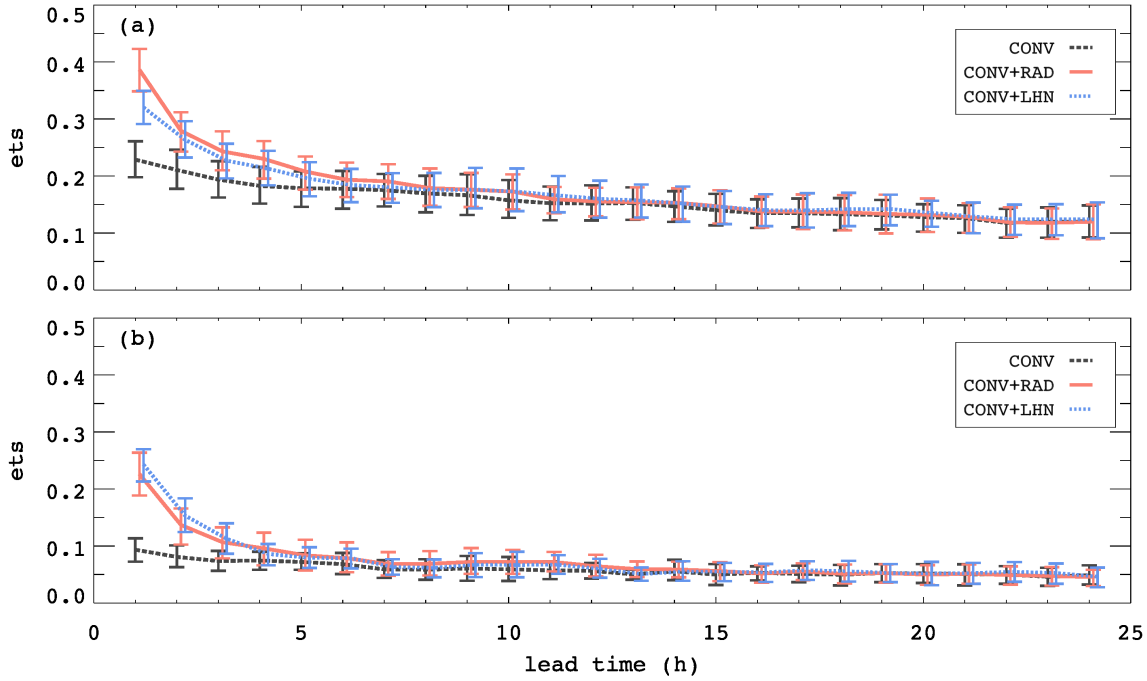


Figure 6.9.: ETS of CONV (gray, dashed), CONV+RAD (red, solid) and CONV+LHN (blue, dotted) for (a) 0.5 mm/h and (b) 2.0 mm/h

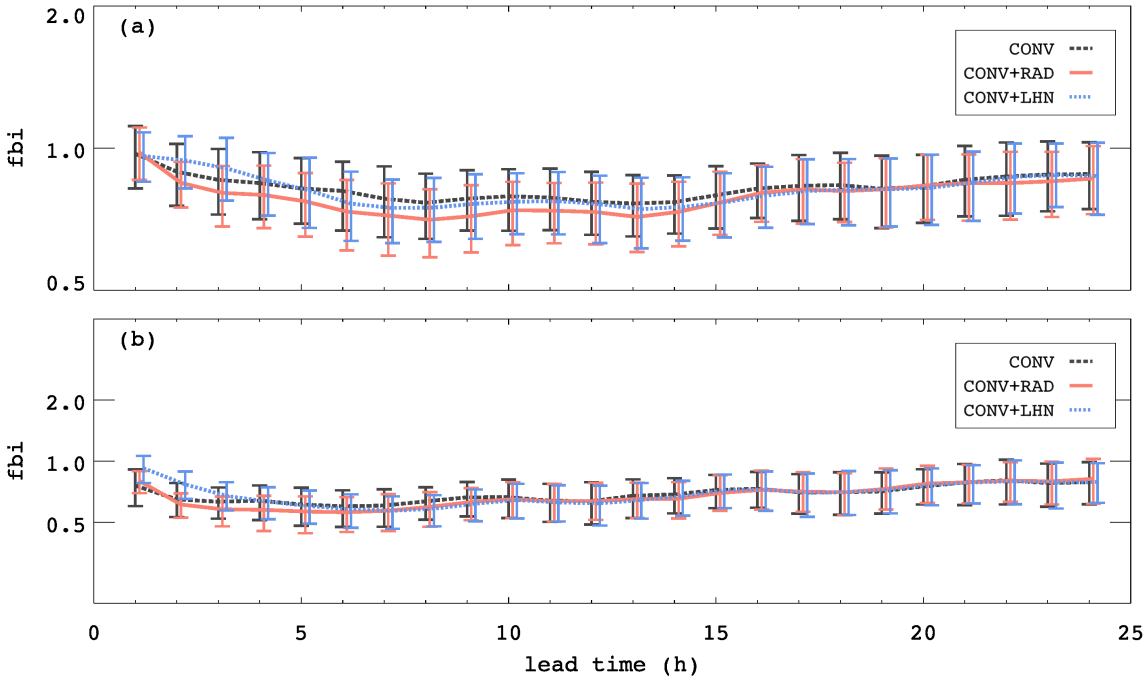


Figure 6.10.: FBIof CONV (gray, dashed), CONV+RAD (red, solid) and CONV+LHN (blue, dotted) for (a) 0.5 mm/h and (b) 2.0 mm/h

ing morning, CONV+RAD slightly underestimates the observation, whereas CONV+LHN overestimates the observations during night and morning. The afternoon maximum is best

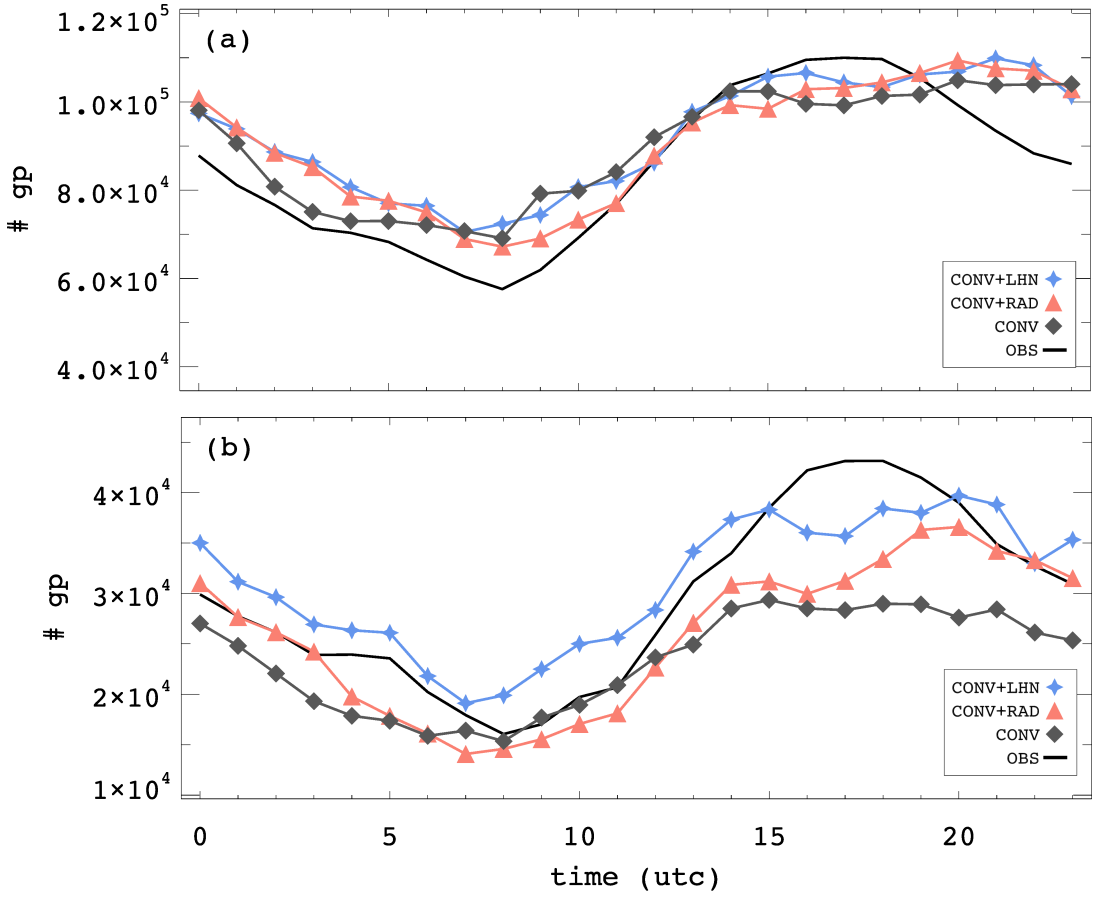


Figure 6.11.: Diurnal cycle of precipitation exceeding (a) 0.5 mm/h and (b) 2.0 mm/h for CONV+LHN (blue, stars), CONV+RAD (red, triangles), and CONV (gray, diamonds) during the assimilation window compared to the observation (black solid line).

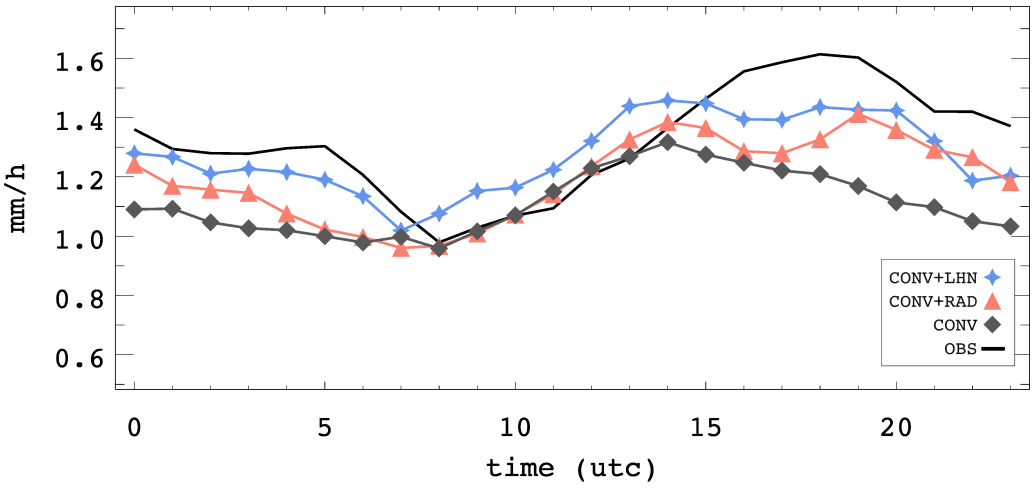


Figure 6.12.: Diurnal cycle of average precipitation intensities (only events above 0.1 mm/h) for CONV+LHN (blue, stars), CONV+RAD (red, triangles), and CONV (gray, diamonds) during the assimilation window compared to the observation (black solid line).

captured by CONV+LHN. CONV+RAD performs better than CONV, but shows a smaller

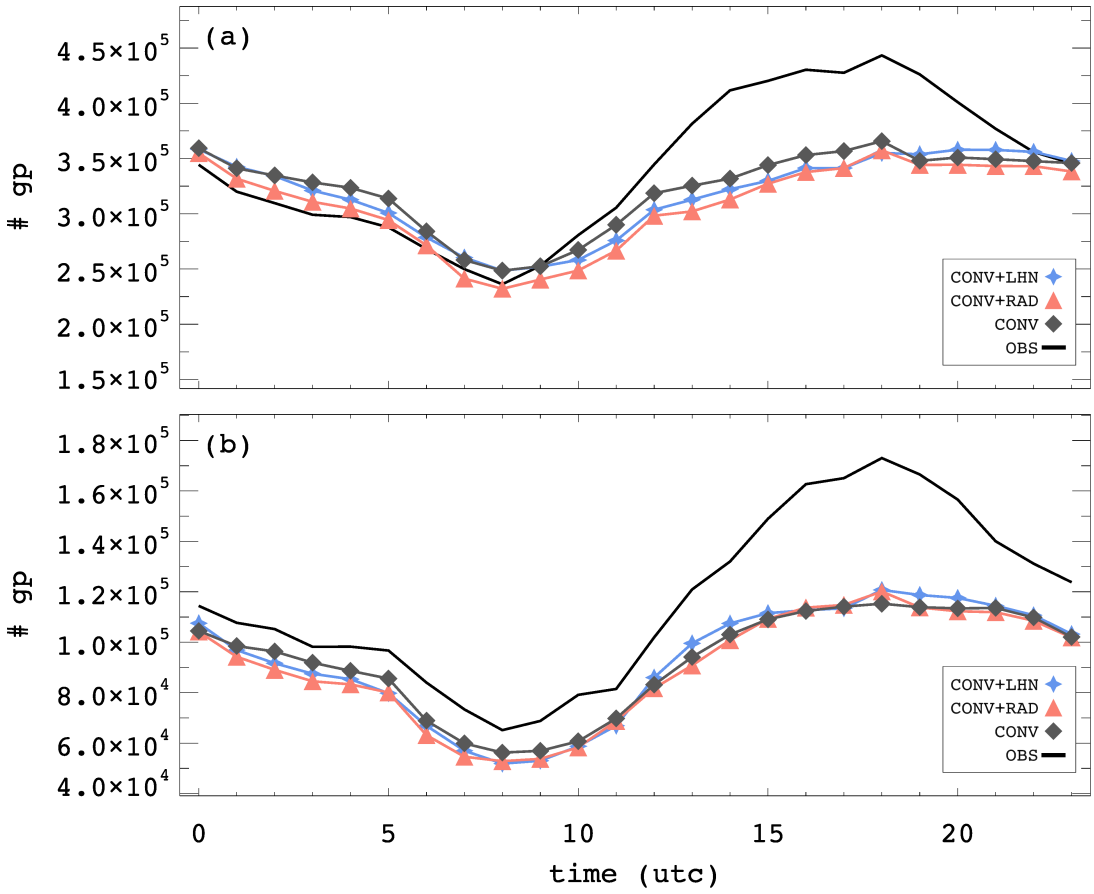


Figure 6.13.: Diurnal cycle of precipitation exceeding (a) 0.5 mm/h and (b) 2.0 mm/h based on forecasts initialized from CONV+LHN (blue, stars), CONV+RAD (red, triangles), and CONV (gray, diamonds) compared to the observation (black solid line).

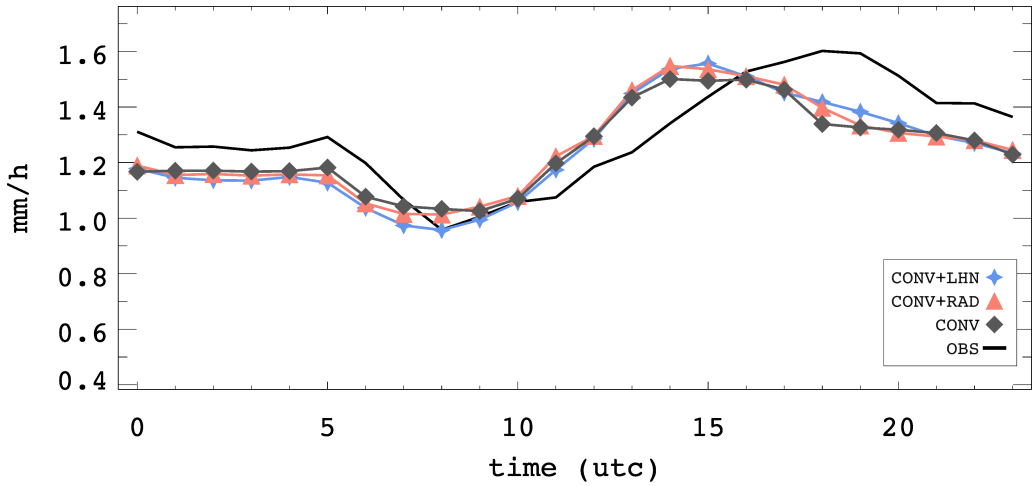


Figure 6.14.: Diurnal cycle of average precipitation intensities (only events above 0.1 mm/h) based on forecasts initialized from CONV+LHN (blue, stars), CONV+RAD (red, triangles), and CONV (gray, diamonds) compared to the observation (black solid line).

number of events during afternoon than CONV+LHN. Figure 6.12 shows the diurnal cycle of average precipitation intensities based on events above 0.1 mm/h. During noon,

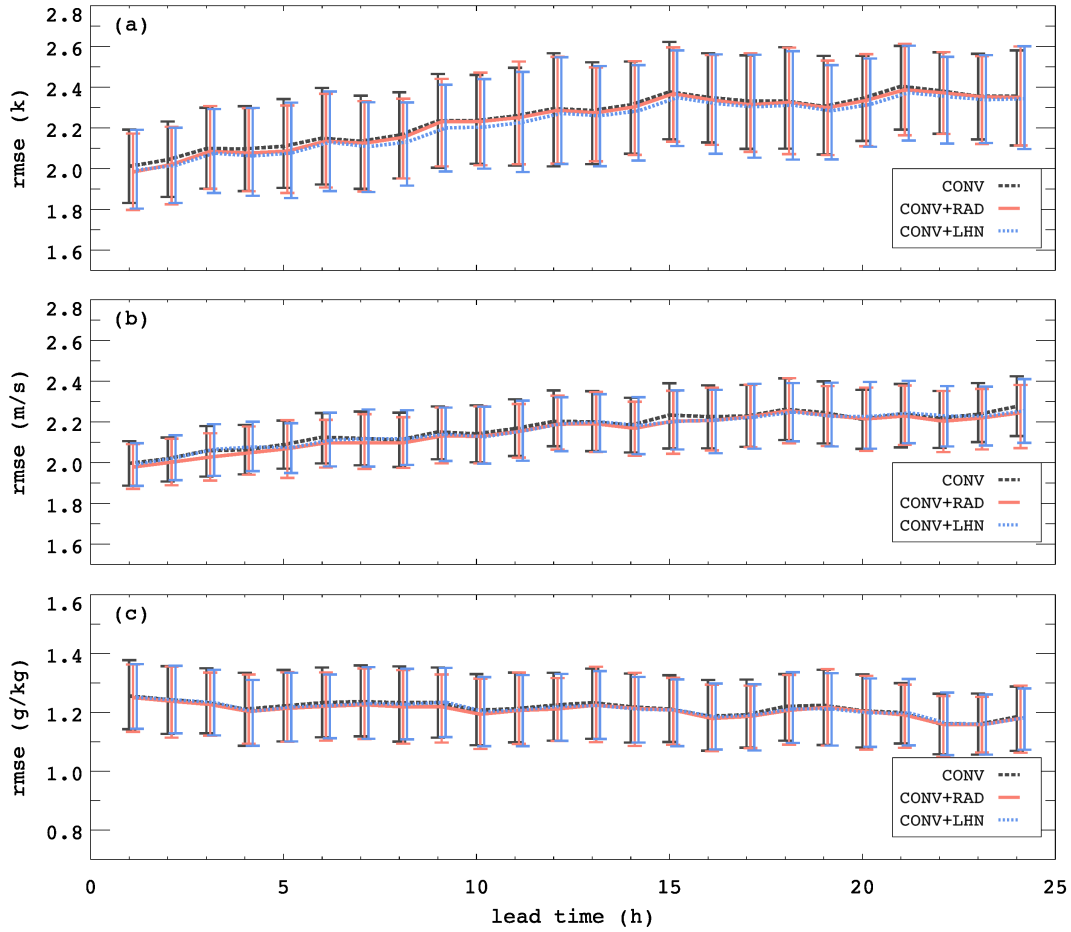


Figure 6.15.: RMSE of CONV (gray, dashed), CONV+RAD (red, solid) and CONV+LHN (blue, dotted) for (a) 2-meter temperature, (b) *u*-component of 10-meter wind, and (c) 2-meter specific humidity against approximately 840 SYNOP stations. The error bars are obtained via bootstrapping. 2.5% and 97.5% percentiles are the end parts of the bars.

CONV and CONV+RAD capture the diurnal cycle quite well, whereas CONV+LHN shows a slight overestimation. During the rest of the day, all three setups underestimate the mean observed intensities. This underestimation is most pronounced for CONV, and least pronounced for CONV+LHN. Results obtained during the assimilation cycle have however to be interpreted with caution: for CONV+LHN, observations are assimilated continuously during the COSMO forward integration, i.e. observations are assimilated during the accumulation of precipitation. Thus, the amount of precipitation accumulated is directly influenced by the observation. For CONV+RAD, observations are assimilated at hourly intervals, and precipitation is accumulated during a one hour free forecast. It thus can be expected that CONV+LHN performs better than CONV+RAD during the analysis cycle.

During free forecasts, all three setups show indeed similar numbers of events exceeding 0.5 mm/h and 2 mm/h (Fig. 6.13). For 0.5 mm/h (Fig. 6.13 a), there is only a slight overestimation of the number of events during night and morning, and the minimum of

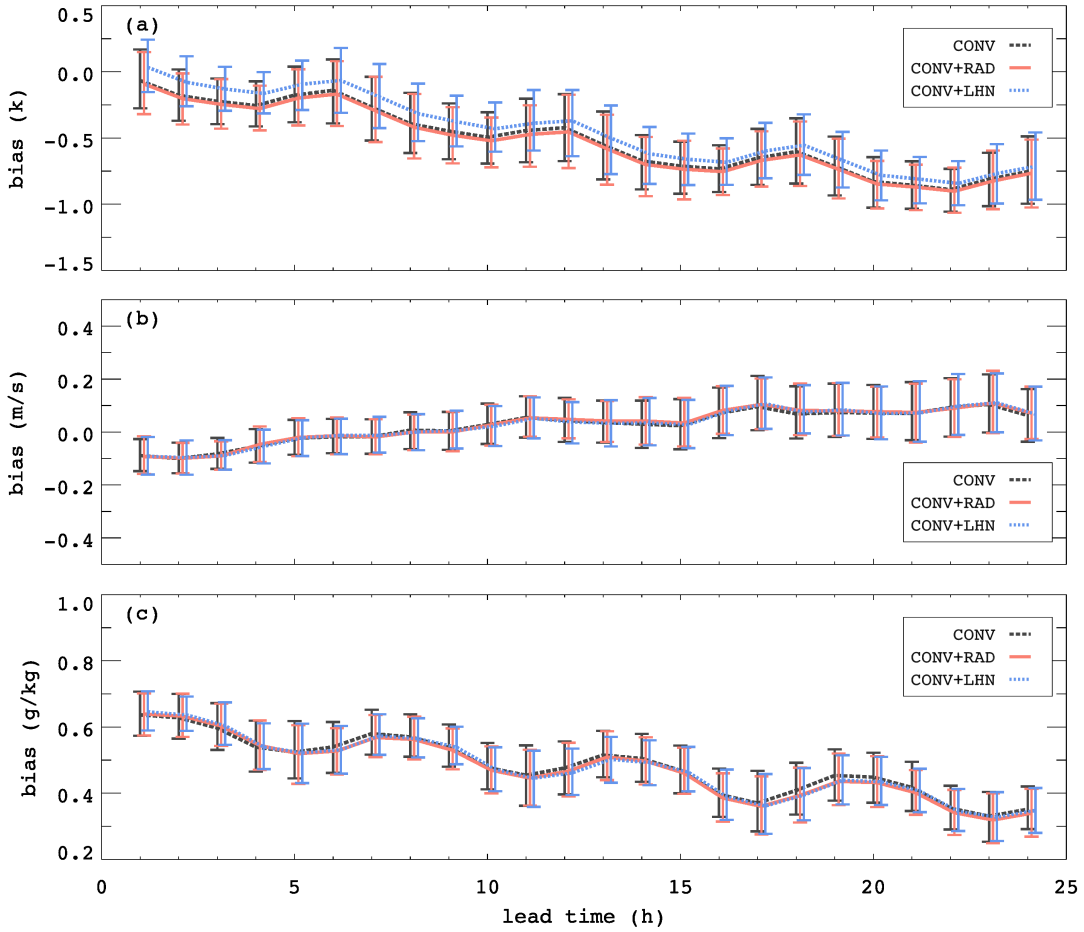
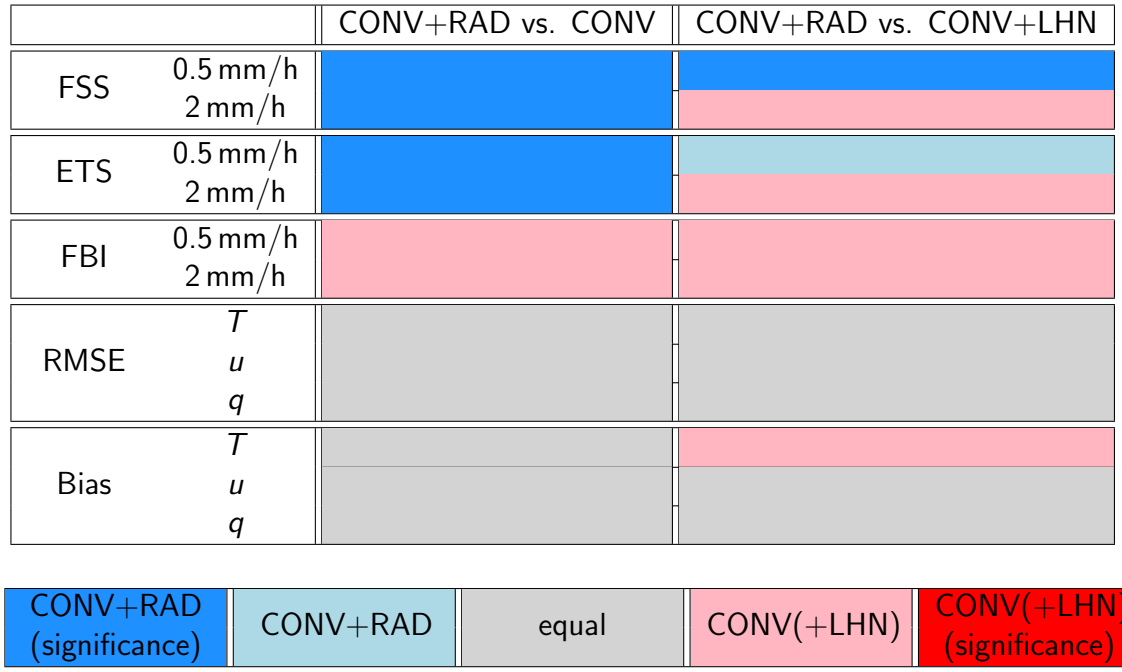


Figure 6.16.: Bias of CONV (gray, dashed), CONV+RAD (red, solid) and CONV+LHN (blue, dotted) for (a) 2-meter temperature, (b) u-component of 10-meter wind, and (c) 2-meter specific humidity against approximately 840 SYNOP stations. The error bars are obtained via bootstrapping. 2.5% and 97.5% percentiles are the end parts of the bars.

the diurnal cycle is captured well. The afternoon maximum is however underestimated by all three setups. Events exceeding 2 mm/h (Fig. 6.13 b) are underestimated throughout the entire day by all setups, while this underestimation is slight in the morning and more severe in the afternoon. For average precipitation intensities (Fig. 6.14), all three setups show a very similar behavior. During night and morning, the average intensity is slightly underestimated. The magnitude of the maximum of all three setups corresponds quite well to the observations but appears approximately four hours too early. Therefore, based on forecasts, CONV+LHN does not perform better than CONV+RAD.

Finally, the model variables not directly linked to precipitation are verified against SYNOP station measurements of 2-meter temperature T, u-component of 10-meter wind, and 2-meter specific humidity q by RMSE (Fig. 6.15) and bias (Fig. 6.16). The RMSE of temperature and wind increases with forecast lead time. Slight variations from hour to hour are visible for all three variables, which can be attributed to the diurnal cycle of precipitation.

Table 6.1.: Summary of the comparison of CONV+RAD against CONV and CONV+LHN (FSS, ETS, FBI, RMSE, and bias). The first column compares CONV+RAD with CONV, the second column compares CONV+RAD with CONV+LHN. Cells are colored dark blue if the performance of CONV+RAD is significantly better during a part of the forecast lead time and comparable during the remaining lead time. Similarly, cells are colored light blue if the results indicate an improvement by CONV+RAD during a part of the forecast time, which is, however, insignificant. Grey cells refer to comparable performance of both methods. Light red cells refer to a better performance of CONV or CONV+LHN during a part of the forecast lead time and comparable otherwise. For a significantly better performance of CONV or CONV+LHN, cells are colored dark red.



Since convective activity is larger during afternoon, errors are larger during noon, and averaging forecasts with different initialization times during day causes fluctuations in the overall RMSE. The differences in RMSE of all variables are small compared to the associated uncertainty and CONV+RAD does not harm forecast quality of 2-meter temperature, 10-meter horizontal wind, and 2-meter specific humidity. There is no bias for 2-meter temperature (Fig. 6.16 a) in the first forecast hour, but all three setups similarly develop a negative bias with forecast lead time. CONV and CONV+RAD perform very similar, whereas CONV+LHN is slightly but not significantly better. For the u-component of 10-meter wind (Fig. 6.16 b), all setups show a slight negative bias for shorter lead times, which evolves into a slight positive bias with forecast lead time. The curves for CONV, CONV+RAD and CONV+LHN are indistinguishable. For 2-meter specific humidity (Fig. 6.16 c), all setups show a positive bias in the beginning of the forecast time, which decreases with forecast lead time. All setups perform similarly. Especially for 2-meter temperature and 2-meter specific humidity, the diurnal cycle of the bias is visible as indicated by the 6

Table 6.2.: As Table 6.1, but for CONV+RAD against CONV+RAD_1ELE (first column), CONV+RAD_4ELE (second column), and CONV+RAD_3DLETKF (third column).

		CONV+RAD vs. CONV+RAD_1ELE	CONV+RAD vs. CONV+RAD_4ELE	CONV+RAD vs. CONV+RAD_3DLETKF
FSS	0.5 mm/h			
	2 mm/h			
ETS	0.5 mm/h			
	2 mm/h			
FBI	0.5 mm/h			
	2 mm/h			
RMSE	T			
	u			
	q			
Bias	T			
	u			
	q			

CONV+RAD (significance)	CONV+RAD	equal	CONV+RAD_1ELE/ 4ELE/3DLETKF	CONV+RAD_1ELE/4ELE/ 3DLETKF(significance)

hourly peaks in the curves.

To summarize, for many of the verification functions used, CONV+RAD is able to significantly improve forecast performance compared to CONV, and even competes successfully with CONV+LHN. In some instances, the results suggest a slightly better performance of CONV+LHN, though none of these are significant (see Table 6.1 for an overview).

6.2.2. Influence of the vertical and temporal resolution

In the following, CONV+RAD is compared against its vertical thinned variations CONV+RAD_1ELE, and CONV+RAD_4ELE, and its variation without 4-dimensional weighting (c.f. Sec. 3.1.4), CONV+RAD_3DLETKF. The results of this comparison are summarized in Table 6.2, and details are shown in Appendix B.

Based on FSS, ETS, FBI, RMSE, and bias, it is shown that reducing the vertical or temporal resolution of the volume radar data has no significant negative impact on forecast accuracy. The results indicate a slight improvement by CONV+RAD_3DLETKF in terms of FSS, ETS, and FBI. However, CONV+RAD_1ELE, CONV+RAD_4ELE, and CONV+RAD_3DLETKF increase the bias of specific humidity compared to CONV+RAD though further tests are necessary to corroborate statistical significancy.

6.3. Computational resources

For operational systems, the efficiency of the data assimilation and forecasting system is crucial. Since the BACY scripts are designed to let COSMO and the LETKF communicate via their respective output files (c.f. Section 3.3.3), input/output routines are necessary for each COSMO and LETKF run. In Table 6.3, the elapsed time[2] of each setup relative to CONV is shown. For the setups with varying update rate (CONV+RAD_5, CONV+RAD_15, CONV+RAD_30, CONV+RAD_60), runtime differs significantly. Only for the CONV+RAD experiments, the radar forward operator needs to be activated, which in CONV+RAD_60 increases the runtime of COSMO compared to CONV by a factor of 1.33. Although the forward operator was designed to run efficiently on supercomputers, it still slows down experiments if the operator needs to be active. The costs of the COSMO run increase with increasing update rate. For CONV+RAD_5, this increase in computational costs is most severe: CONV+RAD_5 needs six times as long as CONV to run COSMO for one hour. The costs of the LETKF are strongly increased in CONV+RAD_5/15/30/60 compared to CONV (ranging from factor 12.67 to 24) due to

[2]The simulations have been carried out at DWD's LINUX Cray system. COSMO uses six nodes with 20 cores each, and each core is assigned 1.5 GB memory. The LETKF uses 40 nodes, but only 5 processes are assigned to each node. This means that only every fourth core is used (160 cores in total) which increases memory of each core to 6 GB.

Table 6.3.: Elapsed runtime of COSMO-DE and LETKF runs relative to the runtime of CONV needed for 1 hour of cycled data assimilation. For CONV+RAD_60, this comprises a one hour COSMO forecast and one LETKF update. For CONV+RAD_30, it is two COSMO-DE integrations, each 30 minutes long, both followed by a LETKF step. Thus, the elapsed time is the sum over these two cycles. For CONV+RAD_15, the duration of 4 cycles, and for CONV+RAD_5 the duration of 12 cycles is accumulated. The number of assimilated radar observations within one hour is shown in column 5. For all CONV+RAD experiments, this number refers to radar reflectivity bins. For CONV+LHN, it refers to the radar derived precipitation rate on the COSMO-DE model grid (German radar sites).

	COSMO (relative to CONV)	LETKF (relative to CONV)	Total (relative to CONV)	Number of radar obs
CONV+RAD ($\hat{=}$ CONV+RAD_60)	1.33	12.67	3.22	1.343.148
CONV+RAD_30	1.87	13.33	3.78	1.343.148
CONV+RAD_15	2.4	14.67	4.44	1.343.148
CONV+RAD_5	6	24	9	1.343.148
CONV+RAD_1ELE	1.33	2.33	1.5	194.856
CONV+RAD_4ELE	1.33	5.33	2	469.512
CONV+RAD_3DLETKF	1	1.66	1.11	111.929
CONV+LHN	1	1	1	73.489

the large amount of radar observations assimilated. Again, CONV+RAD_5 is most expensive. In total, i.e. the costs of COSMO and LETKF together, CONV+RAD_60 needs 3.22 times as long as CONV, and CONV+RAD_5 even needs 9 times as long as CONV. Here, I/O consumes many resources and produces a large overhead.

The use of LHN does not increase the runtime of COSMO compared to CONV, and also the costs of the LETKF do not increase since the set of observations assimilated by the LETKF is the same in CONV and CONV+LHN. Thus, CONV+RAD(_60) runs 3.22 times slower than CONV+LHN. This is not surprising because CONV+RAD processes a huge number of observations compared to CONV+LHN. Decreasing the number of observations in CONV+RAD also decreases the elapsed time. Using only every third elevation (CONV+RAD_4ELE) or only the lowest elevation (CONV+RAD_1ELE), reduces the costs of LETKF, such that in total, CONV+RAD_4ELE needs only twice as long as CONV, CONV+RAD_1ELE even only 1.5 times as long as CONV. Based on the scores evaluated here, forecast quality did not suffer significantly. Excluding the intermediate observations during the first guess window (CONV+RAD_3DLETKF) decreases the costs of the COSMO and the LETKF step yielding a comparable runtime to CONV+LHN (1.11 times longer than CONV in total). Also for CONV+RAD_3DLETKF, there is no loss in forecast quality visible. Thus, vertical thinning or excluding 4D-weighting could be viable options for speeding up experiments.

7. Summary and Discussion

This study investigated the potential to improve short-term precipitation forecasts by assimilating radar reflectivity observations with an ensemble Kalman filter approach. For this purpose, volume measurements from the German radar network consisting of 17 C-band radar sites have been assimilated into the convection-permitting NWP model COSMO-DE with the ensemble Kalman filter framework provided by KENDA and the radar forward operator EMVORADO. As done by Snook *et al.* (2015), conventional observations (aircraft, radiosondes, synoptic stations and wind profiler observations) were assimilated in addition to radar reflectivities. Furthermore, radar reflectivities were not only assimilated in regions of precipitation but also in regions without precipitation in order to suppress spurious convection (Tong and Xue, 2005).

The first of two experiments addressed the impact of the data assimilation update frequency on the analysis quality and on the quality of a subsequent precipitation forecast. Generally, the assimilation of radar reflectivities in addition to conventional observations strongly improved the analysis quality compared to the assimilation of conventional data only.

Within this experiment, several update frequencies (5, 15, 30, and 60 minutes) have been applied. The model was dragged closer to the observation and the analysis quality was better with higher data assimilation update frequencies. However, the forecast quality for lead times beyond one hour suffered from higher update rates, as quantified by the Brier skill score and the continuous ranked probability skill score. In the choice the optimal update interval, two competing goals emerge, namely the balance of the model state and the closeness to the observation. Less frequently updated model states were physically more consistent as quantified by domain averaged surface pressure tendencies. The results of this study suggest that less accurate but more balanced model states eventually lead to better forecasts for lead times beyond one hour. In contrast, many previous radar data assimilation studies favored very frequent update rates (e.g. 2 minutes in Aksoy *et al.* 2009, or 5 minutes in Snook *et al.* 2015). Dong and Xue (2013) found update intervals of 10 and 30 minutes to yield better forecast quality than an update frequency of 60 minutes. The results of this study are, however, in line with experiments performed in an idealized setup of the COSMO-Model described in Lange and Craig (2014), where forecast error

growth was faster if the model was forced strongly onto the observation.

In the second experiment, observations were continuously assimilated over a period of seven days. Deterministic forecasts with a lead time of 24 hours were initialized every six hours during this period. The LETKF based assimilation setup proposed here included volume radar reflectivities and conventional data, and was compared against (1) the assimilation of conventional data only, and (2) the current reference system at DWD, which assimilates conventional observations with the LETKF and radar derived surface precipitation rates with latent heat nudging for each ensemble member.

Compared to the assimilation of conventional data only, the LETKF-based radar reflectivity assimilation significantly improved the quality of precipitation forecasts up to four hours in terms of the fraction skill score. Forecasts of longer lead times performed equally in both setups, which suggests that the lateral boundaries dominate, and that there is no longer term feedback to the synoptic scale from assimilating radar reflectivities. Compared to DWD's reference setup with latent heat nudging, the experiment showed that the LETKF-based radar reflectivity assimilation yields comparable results in the fraction skill score with a significant improvement for the first forecast hour for a 0.5 mm/h threshold, but a slight disadvantage for a 2 mm/h threshold, though not statistically significant.

The afternoon maximum of the diurnal cycle of precipitation was best reproduced by the latent heat nudging during the analysis cycle. This was expected because latent heat heat nudging was applied at every model time step during the model forward integration, whereas the LETKF was applied at prescribed time intervals larger than the model time step, in this case hourly. It was shown that the advantage of latent heat nudging during the analysis cycle did not result in an improved diurnal cycle for the free forecasts compared to the assimilation of radar reflectivities with the LETKF.

Comparing the three setups against measurements from synoptical stations showed that all setups perform similarly in terms of root mean squared error. This result suggests that the LETKF-based assimilation of radar reflectivities had neutral impact on the prognostic model variables such as temperature or wind, which is not clear per se since the LETKF relies on low-rank covariance approximations between the observed variables and the prognostic model variables.

When the vertical resolution of the volume radar data is varied, i.e. only PPI scans at particular elevations are used, no significant degradation in quality of precipitation forecasts were observed. The fraction skill score was similar for small and more intense precipitation events. Reducing the amount of data assimilated led to significant speed up of the experiments. Furthermore, the exclusion of the time trajectory information, i.e. the use of radar observations at the analysis times only instead of the 4D observation set, led as well to results comparable to those obtained with the full observational data set. Moreover, the

reproduction of the diurnal cycle was similar for all variations of the observation data set assimilated. Although intuitively larger amounts of observational information should lead to better analyses and forecasts, it seems reasonable that this thought cannot be applied here. The more observations are assimilated, the more difficult it is to match the model states with a strongly limited number of ensemble members.

The results of this study are extremely encouraging, and it is remarkable that even at this early development stage, the LETKF-based radar reflectivity data assimilation already challenges the well-tested and well-tuned latent heat nudging. Although basic Kalman filter assumptions are violated within the current setup, analyses and forecasts can significantly be improved by assimilating radar reflectivities with the LETKF. There is, however, room for improvement since many aspects still remained untouched. Crucial LETKF parameters such as the localization radius - especially with regard to the assimilation of two observation types that strongly differ in their spatial density - superobbing resolution, and the magnitude of the observation error have so far been chosen heuristically without extensive testing. Although the setup based on LHN was the most efficient, the LETKF setup has the advantage that radial velocities can easily be included in the assimilation process. Previous studies suggest that radial velocities provide valuable additional information: for example, Tong and Xue (2005) achieved best results when radar reflectivities and radial velocities were assimilated together. In KENDA, the assimilation of radial winds as prepared by Zeng (2013) still needs to be tested and combined with radar reflectivities. Further improvements could be achieved by applying a Gaussian transformation to the precipitation data as suggested by Lien *et al.* (2013), in order to better meet the requirements of the Kalman filter assumptions.

Bibliography

Aksoy A, Dowell DC, Snyder C. 2009. A Multicase Comparative Assessment of the Ensemble Kalman Filter for Assimilation of Radar Observations. Part I: Storm-Scale Analyses. *Monthly Weather Review* **137**(6): 1805–1824, doi:10.1175/2008MWR2691.1.

Aksoy A, Dowell DC, Snyder C. 2010. A Multicase Comparative Assessment of the Ensemble Kalman Filter for Assimilation of Radar Observations. Part II: Short-Range Ensemble Forecasts. *Monthly Weather Review* **138**(4): 1273–1292, doi:10.1175/2009MWR3086.1.

Anderson JL. 2001. An Ensemble Adjustment Kalman Filter for Data Assimilation. *Monthly Weather Review* **129**(12): 2884–2903, doi:10.1175/1520-0493(2001)129⟨2884:AEAKFF⟩2.0.CO;2.

Anderson JL, Anderson SL. 1999. A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts. *Monthly Weather Review* **127**(12): 2741–2758, doi:10.1175/1520-0493(1999)127⟨2741:AMCIOT⟩2.0.CO;2.

Anthes R. 1974. Data assimilation and initialization of hurricane prediction models. *Journal of Atmospheric Sciences* **31**: 702–719, doi:10.1175/1520-0469(1974)031⟨0702:DAAIOH⟩2.0.CO;2.

Baldauf M, Seifert A, Förstner J, Majewski D, Raschendorfer M, Reinhardt T. 2011. Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities. *Monthly Weather Review* **139**(12): 3887–3905, doi:10.1175/MWR-D-10-05013.1.

Battan LJ. 1973. *Radar Observations of the Atmosphere*. The University of Chicago Press, ISBN 0-226-03919-6.

Bick T, Simmer C, Wapler K, Hendricks Franssen HJ, Stephan K, Blahak U, Schraff C, Reich H, Zeng Y, Potthast R. 2016. Assimilation of 3D radar reflectivities with an ensemble Kalman filter on the convective scale. *Accepted for publication in Quarterly Journal of the Royal Meteorological Society* doi:10.1002/qj.2751.

Bishop CH, Etherton BJ, Majumdar SJ. 2001. Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Monthly Weather Review* **129**(3): 420–436, doi:10.1175/1520-0493(2001)129⟨0420:ASWTET⟩2.0.CO;2.

Blahak U. 2004. Analyse des Extinktionseffektes bei Niederschlagsmessungen mit einem C-Band Radar anhand von Simulation und Messung. Dissertation, University of Karlsruhe.

Blahak U. 2008a. An approximation to the effective beam weighting function for scanning meteorological radars with an axisymmetric antenna pattern. *Journal of Atmospheric and Oceanic Technology* **25**(7): 1182–1196, doi:10.1175/2007JTECHA1010.1.

Blahak U. 2008b. RADAR MIE LM and RADAR MIELIB — calculation of radar reflectivity from model output. Internal report, imk-tro, university/research center karlsruhe. available on request ulrich.blahak@dwd.de.

Blahak U, Zeng Y, Epperlein D. 2011. Radar forward operator for data assimilation and model verification for the COSMO-model. In: *35th Conference on Radar Meteorology*. pp. 26–30.

Bott A. 2012. *Synoptische Meteorologie*. Springer, ISBN 978-3-642-25121-4, doi:10.1007/978-3-642-25122-1.

Burgers G, van Leeuwen PJ, Evensen G. 1998. Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review* **126**(6): 1719–1724, doi:10.1175/1520-0493(1998)126⟨1719:ASITEK⟩2.0.CO;2.

Caya A, Sun J, Snyder C. 2005. A Comparison between the 4DVAR and the Ensemble Kalman Filter Techniques for Radar Data Assimilation. *Monthly Weather Review* **133**(11): 3081–3094, doi:10.1175/MWR3021.1.

Chang W, Chung KS, Fillion L, Baek SJ. 2014. Radar Data Assimilation in the Canadian High-Resolution Ensemble Kalman Filter System: Performance and Verification with Real Summer Cases. *Monthly Weather Review* **142**(6): 2118–2138, doi:10.1175/MWR-D-13-00291.1.

Chen M, Huang XY. 2006. Digital Filter Initialization for MM5. *Monthly Weather Review* **134**(4): 1222–1236, doi:10.1175/MWR3117.1.

Daley R. 1991. *Atmospheric Data Analysis*. Cambridge University Press.

Davies HC, Turner R. 1977. Updating prediction models by dynamical relaxation - an examination of the technique. *Quarterly Journal of the Royal Meteorological Society* **103**: 225–245, doi:10.1002/qj.49710343602.

Doms G, Baldauf M. 2015. A Description of the Nonhydrostatic Regional COSMO-Model, Part I: Dynamics and Numerics. *www.cosmo-model.org* .

Doms G, Förstner J, Heise E, Herzog HJ, Mironov D, Raschendorfer M, Reinhardt T, Ritter B, Schrodin R, Schulz JP, Vogel G. 2011. Consortium for Small-Scale Modelling A Description of the Nonhydrostatic Regional COSMO Model Part II : Physical Parameterization. *www.cosmo-model.org* .

Doms G, Schättler U. 2002. A description of the nonhydrostatic regional model LM, Part I: Dynamics and Numerics. *COSMO Newsletter* **2**: 225–235.

Dong J, Xue M. 2013. Assimilation of radial velocity and reflectivity data from coastal WSR-88D radars using an ensemble Kalman filter for the analysis and forecast of landfalling hurricane Ike (2008). *Quarterly Journal of the Royal Meteorological Society* **139**(671): 467–487, doi:10.1002/qj.1970.

Dong J, Xue M, Droegemeier K. 2011. The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteorology and Atmospheric Physics* doi:10.1007/s00703-011-0130-3.

Dotzek N, Groenemeijer P, Feuerstein B, Holzer AM. 2009. Overview of ESSL's severe convective storms research using the European Severe Weather Database ESWD. *Atmospheric Research* **93**(1-3): 575–586, doi:10.1016/j.atmosres.2008.10.020.

Dowell DC, Wicker LJ. 2009. Additive noise for storm-scale ensemble data assimilation. *Journal of Atmospheric and Oceanic Technology* **26**(5): 911–927, doi:10.1175/2008JTECHA1156.1.

Dowell DC, Wicker LJ, Snyder C. 2011. Ensemble Kalman Filter Assimilation of Radar Observations of the 8 May 2003 Oklahoma City Supercell: Influences of Reflectivity Observations on Storm-Scale Analyses. *Monthly Weather Review* **139**(1): 272–294, doi:10.1175/2010MWR3438.1.

Dowell DC, Zhang F, Wicker LJ, Snyder C, Crook NA. 2004. Wind and Temperature Retrievals in the 17 May 1981 Arcadia, Oklahoma, Supercell: Ensemble Kalman Filter Experiments. *Monthly Weather Review* **132**(8): 1982–2005, doi:10.1175/1520-0493(2004)132⟨1982:WATRIT⟩2.0.CO;2.

Duncan D, Horn S. 1972. Linear Dynamic Recursive Estimation from the Viewpoint of Regresion Analysis. *Journal of the American Statistical Association* **67**(340): 815–821.

Efron B, Tibshirani RJ. 1993. *An Introduction to the Bootstrap*. Chapman & Hall, ISBN 0-412-04231-2.

Ehrendorfer M, Tribbia JJ, Errico RM. 1996. Mesoscale predictability: An assessment through adjoint methods. In: *Seminar on Predictability*. pp. 157–183.

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research* **99**(C5): 10 143, doi:10.1029/94JC00572.

Field P, Hogan R, Brown P, Illingworth A, Choulartona T, Cotton R. 2005. Parametrization of ice-particle size distributions for mid-latitude stratiform cloud. *Quarterly Journal of the Royal Meteorological Society* **131**: 1997–2017, doi:10.1256/qj.04.134.

Freitag M, Potthast R. 2013. Synergy of inverse problems and data assimilation techniques. In: *Large Scale Inverse Problems: Computational Methods and Applications in the Earth Sciences*, July 2015, pp. 1–54, doi:10.1515/9783110282269.1.

Gao J, Xue M. 2008. An Efficient Dual-Resolution Approach for Ensemble Data Assimilation and Tests with Simulated Doppler Radar Data. *Monthly Weather Review* **136**(3): 945–963, doi:10.1175/2007MWR2120.1.

Gaspari G, Cohn SE. 1999. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society* **125**(April 1998): 723–757, doi:10.1002/qj.49712555417.

Gneiting T, Raftery AE. 2007. Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association* **102**(477): 359–378, doi:10.1198/016214506000001437.

Grasselt R, Schüttemeyer D, Warrach-Sagi K, Ament F, Simmer C. 2008. Validation of TERRA-ML with discharge measurements. *Meteorologische Zeitschrift* **17**(6): 763–773, doi:10.1127/0941-2948/2008/0334.

Grimit EP, Gneiting T, Berrocal VJ, Johnson Na. 2006. The continuous ranked probability score for circular variables and its application to mesoscale forecast ensemble verification. *Quarterly Journal of the Royal Meteorological Society* **132**(621C): 2925–2942, doi:10.1256/qj.05.235.

Hager WW. 1989. Updating the inverse of a matrix. *SIAM Review* **31**(2): 221–239, doi:10.1137/1031049.

Harnisch F, Keil C. 2015. Initial conditions for convective-scale ensemble forecasting provided by ensemble data assimilation. *Monthly Weather Review* **143**(5): 1583–1600, doi:10.1175/MWR-D-14-00209.1.

Hersbach H. 2000. Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. *Weather and Forecasting* **15**(5): 559–570, doi:10.1175/1520-0434(2000)015⟨0559:DOTCRP⟩2.0.CO;2.

Holton JR. 1979. *An Introduction to Dynamic Meteorology*. Academic Press.

Houtekamer P, Mitchell HL. 2005. Ensemble Kalman filtering. *Quarterly Journal of the Royal Meteorological Society* **131**(613): 3269–3289, doi:10.1256/qj.05.135.

Houtekamer PL, Mitchell HL. 1998. Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review* **126**(3): 796–811, doi:10.1175/1520-0493(1998)126⟨0796:DAUAEK⟩2.0.CO;2.

Houtekamer PL, Mitchell HL. 2001. A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review* **129**(1): 123–137, doi:10.1175/1520-0493(2001)129⟨0123:ASEKFF⟩2.0.CO;2.

Hunt B, Kostelich EJ, Szunyogh I. 2007. Efficient data assimilation for spatiotemporal chaos: a local transform Kalman filter. *Physica D* **230**(1–2): 112–126, doi:10.1016/j.physd.2006.11.008.

Janjić T, Cohn SE. 2006. Treatment of Observation Error due to Unresolved Scales in Atmospheric Data Assimilation. *Monthly Weather Review* **134**(10): 2900–2915, doi:10.1175/MWR3229.1.

Jerger D. 2014. Radar Forward Operator for Verification of Cloud Resolving Simulations within the COSMO-Model. Dissertation, IMK-TRO, Department of Physics, Karlsruhe Institute of Technology.

Jung Y, Xue M, Tong M. 2012. Ensemble Kalman Filter Analyses of the 29–30 May 2004 Oklahoma Tornadic Thunderstorm Using One- and Two-Moment Bulk Microphysics Schemes, with Verification against Polarimetric Radar Data. *Monthly Weather Review* **140**(5): 1457–1475, doi:10.1175/MWR-D-11-00032.1.

Kalman RE. 1960. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME-Journal of Basic Engineering* **82**(Series D): 35–45, doi:10.1115/1.3662552.

Kalnay E. 2010. Ensemble Kalman Filter: Current Status and Potential. In: *Data Assimilation - Making Sense of Observations*, Springer, pp. 69–92.

Lange H, Craig GC. 2014. The Impact of Data Assimilation Length Scales on Analysis and Prediction of Convective Storms. *Monthly Weather Review* **142**(10): 3781–3808, doi:10.1175/MWR-D-13-00304.1.

Lien GY, Kalnay E, Miyoshi T. 2013. Effective assimilation of global precipitation: Simulation experiments. *Tellus, Series A: Dynamic Meteorology and Oceanography* **65**(1): 1–16, doi:10.3402/tellusa.v65i0.19915.

Lin YL, Farley RD, Orville HD. 1983. Bulk Parameterization of the Snow Field in a Cloud Model. *Journal of Climate and Applied Meteorology* **22**(6): 1065–1092, doi:10.1175/1520-0450(1983)022⟨1065:BPOTSF⟩2.0.CO;2.

Liu ZQ, Rabier F. 2002. The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study. *Quarterly Journal of the Royal Meteorological Society* **128**: 1367–1386, doi:10.1256/003590002320373337.

Lorenc AC. 1986. Analysis methods for numerical weather prediction. doi:10.1256/smsqj.47413.

Lorenz EN. 1963. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences* **20**(2): 130–141, doi:10.1175/1520-0469(1963)020⟨0130:DNF⟩2.0.CO;2.

Lorenz EN. 1996. Predictability: a problem partly solved. *Predictability of Weather and Climate* : 40–58.

Marshall JS, Palmer WMK. 1948. The Distribution of Raindrops with Size. *Journal of Meteorology* **5**(4): 165–166, doi:10.1175/1520-0469(1948)005⟨0165:TDORWS⟩2.0.CO;2.

Matheson JE, Winkler RL. 1976. Scoring Rules for Continuous Probability Distributions. *Management Science* **22**(1): 1087–1096.

Mie G. 1908. Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen. *Annalen der Physik* **330**(3): 377–445, doi:10.1002/andp.19083300302.

Milan M, Venema V, Schüttemeyer D, Simmer C. 2008. Assimilation of radar and satellite data in mesoscale models: A physical initialization scheme. *Meteorologische Zeitschrift* **17**(6): 887–902, doi:10.1127/0941-2948/2008/0340.

Mitchell HL, Houtekamer PL, Pellerin G. 2002. Ensemble Size, Balance, and Model-Error Representation in an Ensemble Kalman Filter. *Monthly Weather Review* **130**(11): 2791–2808, doi:10.1175/1520-0493(2002)130⟨2791:ESBAME⟩2.0.CO;2.

Nakamura G, Potthast R. 2015. Programming of numerical algorithms and useful tools. In: *Inverse Modeling: An introduction to the theory and methods of inverse problems and data assimilation*, IOP Publishing, pp. 6–1—-6–18, doi:10.1088/978-0-7503-1218-9ch6.

Nichols NK. 2010. Mathematical Concepts of Data Assimilation. In: *Data Assimilation - Making Sense of Observations*, Springer, pp. 13–40.

Ott E, Hunt BR, Szunyogh I, Zimin AV, Kostelich EJ, Corazza M, Kalnay E, Patil DJ, Yorke JA. 2004. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus A* **56**: 415–428, doi:10.1111/j.1600-0870.2004.00076.x.

Perianez A, Reich H, Potthast R. 2014. Optimal Localization for Ensemble Kalman Filter Systems. *Journal of the Meteorological Society of Japan. Ser. II* **92**(6): 585–597, doi: 10.2151/jmsj.2014-605.

Raschendorfer M. 2001. The new turbulence parameterisation of LM. *COSMO Newsletter* **1**: 89–97.

Reinhardt T, Seifert A. 2006. A three-category ice scheme for LMK. *COSMO Newsletter* **6**: 115–120.

Rinehart RE. 2010. *Radar for Meteorologists*. Rinehart Publishing, 5th edn.

Ritter B, Geleyn JF. 1992. A Comprehensive Radiation Scheme for Numerical Weather Prediction Models with Potential Applications in Climate Simulations. *Monthly Weather Review* **120**(2): 303–325, doi:10.1175/1520-0493(1992)120⟨0303:ACRSFN⟩2.0.CO;2.

Roberts NM, Lean HW. 2008. Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Monthly Weather Review* **136**(1): 78–97, doi:10.1175/2007MWR2123.1.

Schättler U, Blahak U. 2015. A Description of the Nonhydrostatic Regional COSMO-Model. Part V: Preprocessing: Initial and Boundary Data for the COSMO-Model. *www.cosmo-model.org* .

Schiff SJ, Sauer T. 2008. Kalman filter control of a model of spatiotemporal cortical dynamics. *Journal of neural engineering* **5**(1): 1–8, doi:10.1186/1471-2202-9-S1-O1.

Schraff C, Hess R. 2012. A Description of the Nonhydrostatic Regional COSMO-Model Part III : Data Assimilation. *www.cosmo-model.org* : 93.

Schraff C, Reich H, Rhodin A, Schomburg A, Stephan K, Perianez A, Potthast R. 2016. Kilometre-Scale Ensemble Data Assimilation for the COSMO Model (KENDA). *Accepted*

for publication in Quarterly Journal of the Royal Meteorological Society doi:10.1002/qj.2748.

Schraff CH. 1997. Mesoscale data assimilation and prediction of low stratus in the Alpine region. *Meteorology and Atmospheric Physics* **64**(1-2): 21–50, doi:10.1007/BF01044128.

Schwitalla T, Wulfmeyer V. 2014. Radar data assimilation experiments using the IPM WRF Rapid Update Cycle. *Meteorologische Zeitschrift* **23**(1): 79–102, doi:10.1127/0941-2948/2014/0513.

Seifert A, Beheng KD. 2001. A double-moment parameterization for simulating autoconversion, accretion and selfcollection. *Atmospheric Research* **59-60**(June 2015): 265–281, doi:10.1016/S0169-8095(01)00126-0.

Seifert A, Beheng KD. 2006. A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description. *Meteorology and Atmospheric Physics* **92**(1-2): 45–66, doi:10.1007/s00703-005-0112-4.

Snook N, Xue M, Jung Y. 2011. Analysis of a Tornadic Mesoscale Convective Vortex Based on Ensemble Kalman Filter Assimilation of CASA X-Band and WSR-88D Radar Data. *Monthly Weather Review* **139**(11): 3446–3468, doi:10.1175/MWR-D-10-05053.1.

Snook N, Xue M, Jung Y. 2015. Multi-Scale EnKF Assimilation of Radar and Conventional Observations and Ensemble Forecasting for a Tornadic Mesoscale Convective System. *Monthly Weather Review* **143**(4): 1035–1057, doi:10.1175/MWR-D-13-00262.1.

Snyder C, Zhang F. 2003. Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Monthly Weather Review* **131**(8): 1663–1677, doi:10.1175//2555.1.

Sobash RA, Stensrud DJ. 2013. The impact of covariance localization for radar data on EnKF analyses of a developing MCS: Observing system simulation experiments. *Monthly Weather Review* **14**: 3691–3709, doi:10.1175/MWR-D-12-00203.1.

Stauffer DR, Seaman NL. 1990. Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model. Part I: Experiments with Synoptic-Scale Data. *Monthly Weather Review* **118**(6): 1250–1277, doi:10.1175/1520-0493(1990)118⟨1250:UOFDDA⟩2.0.CO;2.

Stephan K, Klink S, Schraff C. 2008. Assimilation of radar-derived rain rates into the convective-scale model COSMO-DE at DWD. *Quarterly Journal of the Royal Meteorological Society* **134**(May): 1315–1326, doi:10.1002/qj.

Sun J, Crook NA. 1997. 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, doi:10.1175/1520-0469(1997)054⟨1642:DAMRFD⟩2.0.CO;2.

Sun J, Crook NA. 1998. Dynamical and Microphysical Retrieval from Doppler Radar Observations Using a Cloud Model and Its Adjoint. Part II: Retrieval Experiments of an Observed Florida Convective Storm. *Journal of the Atmospheric Sciences* **55**(5): 835–852, doi:10.1175/1520-0469(1998)055⟨0835:DAMRFD⟩2.0.CO;2.

Sun J, Xue M, Wilson JW, Zawadzki I, Ballard SP, Onvlee-Hooimeyer J, Joe P, Barker DM, Li PW, Golding B, Xu M, Pinto J. 2014. Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bulletin of the American Meteorological Society* **95**(3): 409–426, doi:10.1175/BAMS-D-11-00263.1.

Talagrand O. 1997. Assimilation of observations, an introduction. *Journal of the Meteorological Society of Japan* **75**(1B): 191–209.

Tarantola A. 1987. *Inverse Problem Theory*. Elsevier.

Tiedtke M. 1989. A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models. *Monthly Weather Review* **117**(8): 1779–1800, doi:10.1175/1520-0493(1989)117⟨1779:ACMFSF⟩2.0.CO;2.

Tippett MK, Anderson JL, Bishop CH, Hamill TM, Whitaker JS. 2003. Ensemble Square Root Filters. *Monthly Weather Review* **131**(7): 1485–1490, doi:10.1175/1520-0493(2003)131⟨1485:ESRF⟩2.0.CO;2.

Tong M, Xue M. 2005. Ensemble Kalman Filter Assimilation of Doppler Radar Data with a Compressible Nonhydrostatic Model: OSS Experiments. *Monthly Weather Review* **133**(Evensen 2003): 1789–1807, doi:10.1175/MWR2898.1.

Villarini G, Krajewski WF. 2010. Review of the Different Sources of Uncertainty in Single Polarization Radar-Based Estimates of Rainfall. *Surveys in Geophysics* **31**(1): 107–129.

Whitaker JS, Hamill TM. 2002. Ensemble Data Assimilation without Perturbed Observations. *Monthly Weather Review* **130**(7): 1913–1924, doi:10.1175/1520-0493(2002)130⟨1913:EDAWPO⟩2.0.CO;2.

Whitaker JS, Hamill TM. 2012. Evaluating Methods to Account for System Errors in Ensemble Data Assimilation. *Monthly Weather Review* **140**(9): 3078–3089, doi:10.1175/MWR-D-11-00276.1.

Wicker LJ, Skamarock WC. 2002. Time-Splitting Methods for Elastic Models Using Forward Time Schemes. *Monthly Weather Review* **130**(8): 2088–2097, doi:10.1175/1520-0493(2002)130⟨2088:TSMFEM⟩2.0.CO;2.

Wilks DS. 2006. *Statistical methods in the atmospheric sciences*. Academic Press, second edn.

Xue M, Tong M, Droegemeier KK. 2006. An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting. *Journal of Atmospheric and Oceanic Technology* **23**(1): 46–66, doi:10.1175/JTECH1835.1.

Yang SC, Kalnay E, Hunt B, Bowler NE. 2009. Weight interpolation for efficient data assimilation with the Local Ensemble Transform Kalman Filter. *Quarterly Journal of the Royal Meteorological Society* **135**(January 2009): 251–262, doi:10.1002/qj.353.

Zängl G, Reinert D, Ripodas P, Baldauf M. 2015. The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society* **141**: 563–579, doi:10.1002/qj.2378.

Zeng Y. 2013. Efficient Radar Forward Operator for Operational Data Assimilation within the COSMO-model. Dissertation, IMK-TRO, Department of Physics, Karlsruhe Institute of Technology.

Zeng Y, Blahak U, Neuper M, Jerger D. 2014. Radar Beam Tracing Methods Based on Atmospheric Refractive Index. *Journal of Atmospheric and Oceanic Technology* **31**(12): 2650–2670, doi:10.1175/JTECH-D-13-00152.1.

Zhang F, Snyder C, Sun J. 2004. Impacts of Initial Estimate and Observation Availability on Convective-Scale Data Assimilation with an Ensemble Kalman Filter. *Monthly Weather Review* **132**(5): 1238–1253, doi:10.1175/1520-0493(2004)132⟨1238:IOIEAO⟩2.0.CO;2.

Zhang F, Weng Y, Sippel Ja, Meng Z, Bishop CH. 2009. Cloud-Resolving Hurricane Initialization and Prediction through Assimilation of Doppler Radar Observations with an Ensemble Kalman Filter. *Monthly Weather Review* **137**(7): 2105–2125, doi:10.1175/2009MWR2645.1.

Zhu Y, Toth Z, Wobus R, Richardson D, Mylne K. 2002. The economic value of ensemble-based weather forecasts. *Bulletin of the American Meteorological Society* (January): 73–83, doi:10.1175/1520-0477(2002)083⟨0073:TEVOEB⟩2.3.CO;2.

List of Figures

[1] According to Bick *et al.* (2016).

List of Tables

[2]According to Bick *et al.* (2016).

Acronyms

BACY basic cycling

BS Brier score

BSS Brier skill score

COSMO consortium for small scale modeling

CRPS continuous ranked probability score

CRPSS continuous ranked probability skill score

DWD Deutscher Wetterdienst

EAKF ensemble adjustment Kalman filter

EMVORADO efficient modular volume radar operator

EnKF ensemble Kalman filter

EnSRF ensemble square root filter

ETKF ensemble transform Kalman filter

ETS equitable threat score

FBI frequency bias

FSS fraction skill score

ICON icosahedral non-hydrostatic modeling framework

KENDA km-scale ensemble data assimilation

LETKF local ensemble transform Kalman filter

LHN latent heat nudging

MSE mean squared error

NWP numerical weather prediction

OSSE observing system simulation experiment

PPI plan position indicator

RMSE root mean squared error

RTPP relaxation to prior perturbations

RTPS relaxation to prior spread

A. Detailed derivation of the Kalman filter in the least-squares framework

In order to minimize the cost function $J(\mathbf{x})$ in Equation (3.4), its gradient is set to zero. The gradient (Eq. 3.6) can be derived by the following steps:

$$
\begin{aligned}
\nabla_{\mathbf{x}} J(\mathbf{x}) &= \nabla_{\mathbf{x}} \left\{ \frac{1}{2}(\mathbf{x}-\mathbf{x}^b)^T(\mathbf{P}^b)^{-1}(\mathbf{x}-\mathbf{x}^b) + \frac{1}{2}(\mathbf{H}\mathbf{x}-\mathbf{y})^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}-\mathbf{y}) \right\} \\
&= \nabla_{\mathbf{x}} \left\{ \frac{1}{2}\left[\mathbf{x}^T(\mathbf{P}^b)^{-1}\mathbf{x} - \mathbf{x}^T(\mathbf{P}^b)^{-1}\mathbf{x}^b - (\mathbf{x}^b)^T(\mathbf{P}^b)^{-1}\mathbf{x} + (\mathbf{x}^b)^T(\mathbf{P}^b)^{-1}\mathbf{x}^b\right] \right. \\
&\quad \left. + \frac{1}{2}\left[\mathbf{x}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{x} - \mathbf{x}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y} - \mathbf{y}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{x} + \mathbf{y}^T\mathbf{R}\mathbf{y}\right] \right\} \\
&= \frac{1}{2}\left[\nabla_{\mathbf{x}}(\mathbf{x}^T(\mathbf{P}^b)^{-1}\mathbf{x}) - \nabla_{\mathbf{x}}(\mathbf{x}^T(\mathbf{P}^b)^{-1}\mathbf{x}^b) - \nabla_{\mathbf{x}}((\mathbf{x}^b)^T(\mathbf{P}^b)^{-1}\mathbf{x}) \right. \\
&\quad \left. + \nabla_{\mathbf{x}}((\mathbf{x}^b)^T(\mathbf{P}^b)^{-1}\mathbf{x}^b)\right] + \frac{1}{2}\left[\nabla_{\mathbf{x}}(\mathbf{x}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{x}) - \nabla_{\mathbf{x}}(\mathbf{x}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}) \right. \\
&\quad \left. - \nabla_{\mathbf{x}}(\mathbf{y}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{x}) + \nabla_{\mathbf{x}}(\mathbf{y}^T\mathbf{R}\mathbf{y})\right] \\
&= \frac{1}{2}\left[2(\mathbf{P}^b)^{-1}\mathbf{x} - 2(\mathbf{P}^b)^{-1}\mathbf{x}^b\right] + \frac{1}{2}\left[2\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{x} - 2\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}\right] \\
&= (\mathbf{P}^b)^{-1}(\mathbf{x}-\mathbf{x}^b) + \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}-\mathbf{y}).
\end{aligned}
$$

In order to find an explicit solution of the least-squares problem, $(\mathbf{P}^b)^{-1}(\mathbf{x}-\mathbf{x}^b) + \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}-\mathbf{y}) = 0$ is solved for $\mathbf{x}$ (Equation 3.7):

$$
\begin{aligned}
& & 0 &= (\mathbf{P}^b)^{-1}(\mathbf{x}-\mathbf{x}^b) + \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}-\mathbf{y}) \\
\Leftrightarrow & & 0 &= (\mathbf{P}^b)^{-1}\mathbf{x} - (\mathbf{P}^b)^{-1}\mathbf{x}^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{x} - \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y} \\
\Leftrightarrow & (\mathbf{P}^b)^{-1}\mathbf{x} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{x} &= (\mathbf{P}^b)^{-1}\mathbf{x}^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y} \\
\Leftrightarrow & \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\mathbf{x} &= (\mathbf{P}^b)^{-1}\mathbf{x}^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y} \\
\Leftrightarrow & \mathbf{x} &= \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\ \right)^{-1}\left[(\mathbf{P}^b)^{-1}\mathbf{x}^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}\right] \\
\Leftrightarrow & \mathbf{x} &= \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}
\end{aligned}
$$

$$
\begin{aligned}
& \cdot \left[(\mathbf{P}^b)^{-1}\mathbf{x}^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{x}^b - \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{x}^b\right] \\
\Leftrightarrow \quad & \mathbf{x} = \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1} \\
& \cdot \left[\left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\mathbf{x}^b + \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^b)\right] \\
\Leftrightarrow \quad & \mathbf{x} = \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\mathbf{x}^b \\
& + \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^b) \\
\Leftrightarrow \quad & \mathbf{x} = \mathbf{x}^b + \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^b).
\end{aligned}
$$

Furthermore, the following equality holds:

$$
\begin{aligned}
\mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{R} + \mathbf{H}\mathbf{P}^b\mathbf{H}^T\right) &= \mathbf{H}^T\mathbf{R}^{-1}\mathbf{R} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{P}^b\mathbf{H}^T \\
&= \mathbf{H}^T + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{P}^b\mathbf{H}^T \\
&= (\mathbf{P}^b)^{-1}\mathbf{P}^b\mathbf{H}^T + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{P}^b\mathbf{H}^T \\
&= \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\mathbf{P}^b\mathbf{H}^T
\end{aligned}
$$

Therefore it can be concluded that

$$
\begin{aligned}
& \mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{R} + \mathbf{H}\mathbf{P}^b\mathbf{H}^T\right) = \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\mathbf{P}^b\mathbf{H}^T \\
\Leftrightarrow \quad & \left(\mathbf{R} + \mathbf{H}\mathbf{P}^b\mathbf{H}^T\right)\left(\mathbf{P}^b\mathbf{H}^T\right)^{-1} = \left(\mathbf{H}^T\mathbf{R}^{-1}\right)^{-1}\left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right) \\
\Leftrightarrow \quad & \left[\left(\mathbf{R} + \mathbf{H}\mathbf{P}^b\mathbf{H}^T\right)\left(\mathbf{P}^b\mathbf{H}^T\right)^{-1}\right]^{-1} = \left[\left(\mathbf{H}^T\mathbf{R}^{-1}\right)^{-1}\left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)\right]^{-1} \\
\Leftrightarrow \quad & \mathbf{P}^b\mathbf{H}^T\left(\mathbf{R} + \mathbf{H}\mathbf{P}^b\mathbf{H}^T\right)^{-1} = \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1},
\end{aligned}
$$

and the Kalman filter equation is updated as follows (Equation (3.8)):

$$
\begin{aligned}
\mathbf{x} &= \mathbf{x}^b + \left((\mathbf{P}^b)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^b) \\
&= \mathbf{x}^b + \mathbf{P}^b\mathbf{H}^T\left(\mathbf{R} + \mathbf{H}\mathbf{P}^b\mathbf{H}^T\right)^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^b).
\end{aligned}
$$

B. Detailed evaluation of the influence of the vertical and temporal resolution

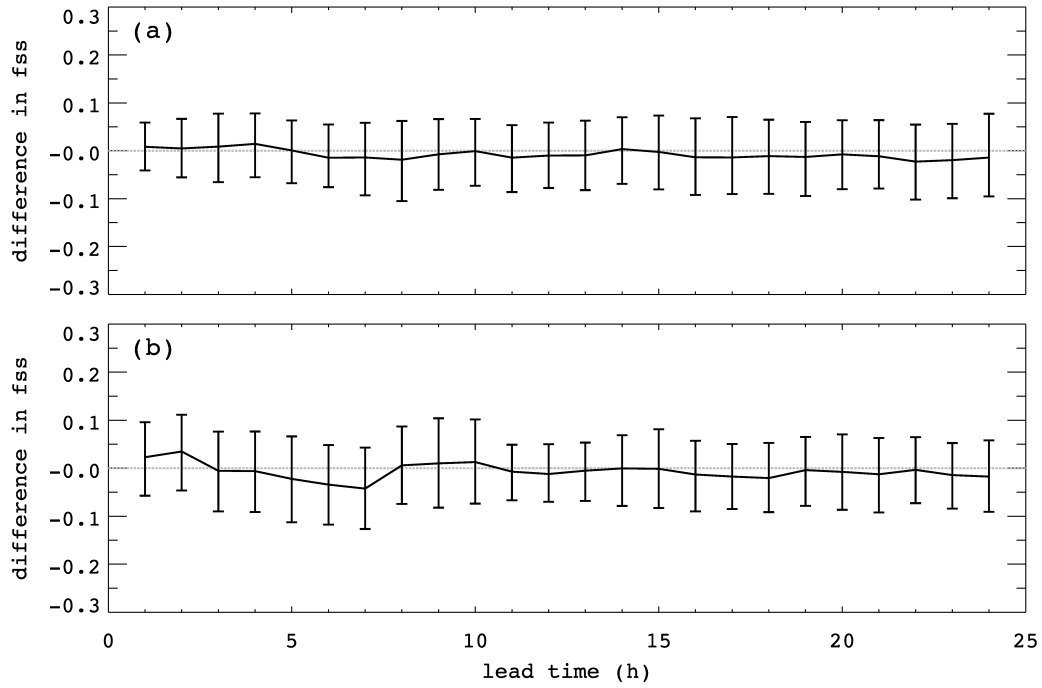


Figure B.1.: As Figure 6.7, but for the difference CONV+RAD minus CONV+RAD_1ELE.

For the summary of the results in Section 6.2.2, a more detailed insight into the results is given in this section. The performance of CONV+RAD_1ELE, CONV+RAD_4ELE, and CONV+RAD_3DLETKF is mostly equal to CONV+RAD in terms of FSS for 0.5 mm/h (Figs. B.1 a, B.2 a, and B.3 a). For the 2 mm/h threshold (Figs. B.1 b, B.2 b, and B.3 b), slight but insignificant fluctuations occur during the first forecast hours. None of the setups performs superior or inferior than the others.

The ETS for both thresholds (Fig. B.4) decreases with forecast time equally for all four setups. Also the FBI (Fig. B.5) behaves very similarly with an underestimation during the first forecast hours for both thresholds. CONV+RAD_3DLETKF appears slightly but

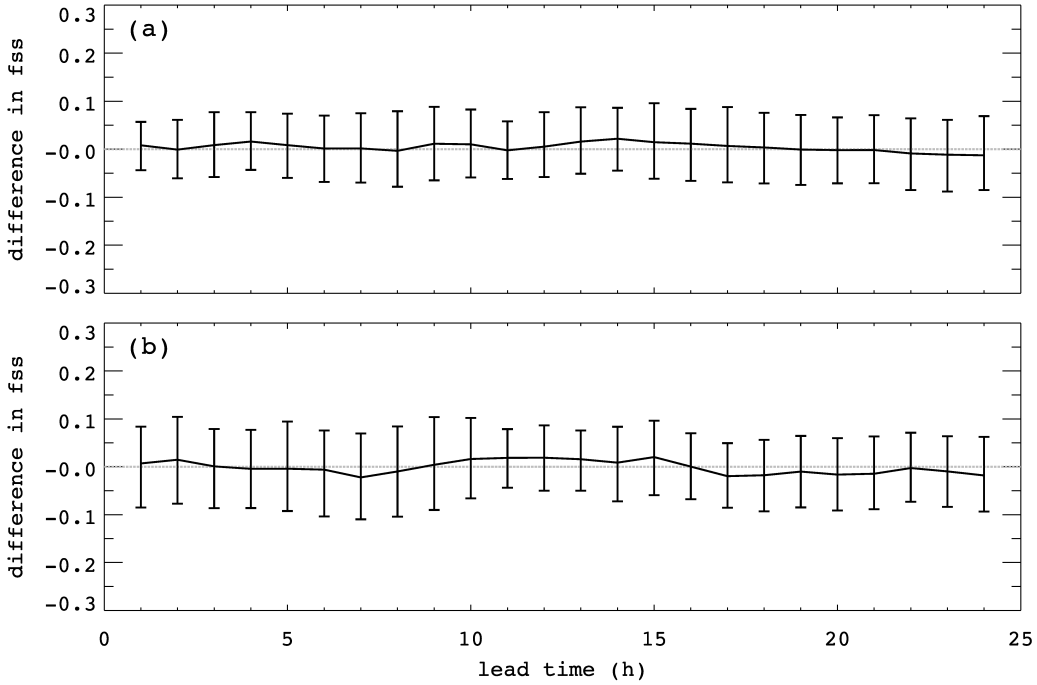


Figure B.2.: As Figure 6.7, but for the difference CONV+RAD minus CONV+RAD_4ELE.

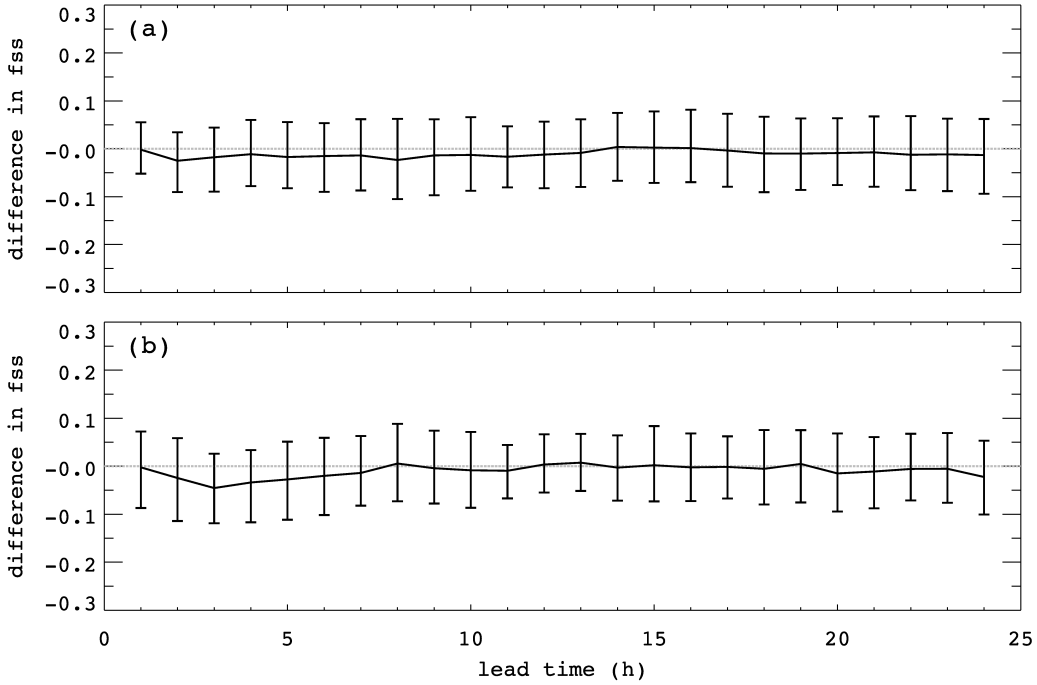


Figure B.3.: As Figure 6.7, but for the difference CONV+RAD minus CONV+RAD_3DLETKF.

insignificantly favorable.

The diurnal cycle of the number of precipitation events exceeding 0.5 mm/h and 2 mm/h (Fig. B.8) is similar for CONV+RAD, CONV+RAD_1ELE, CONV+RAD_4ELE, and CONV+RAD_3DLETKF. Reducing the vertical or temporal resolution does not significantly

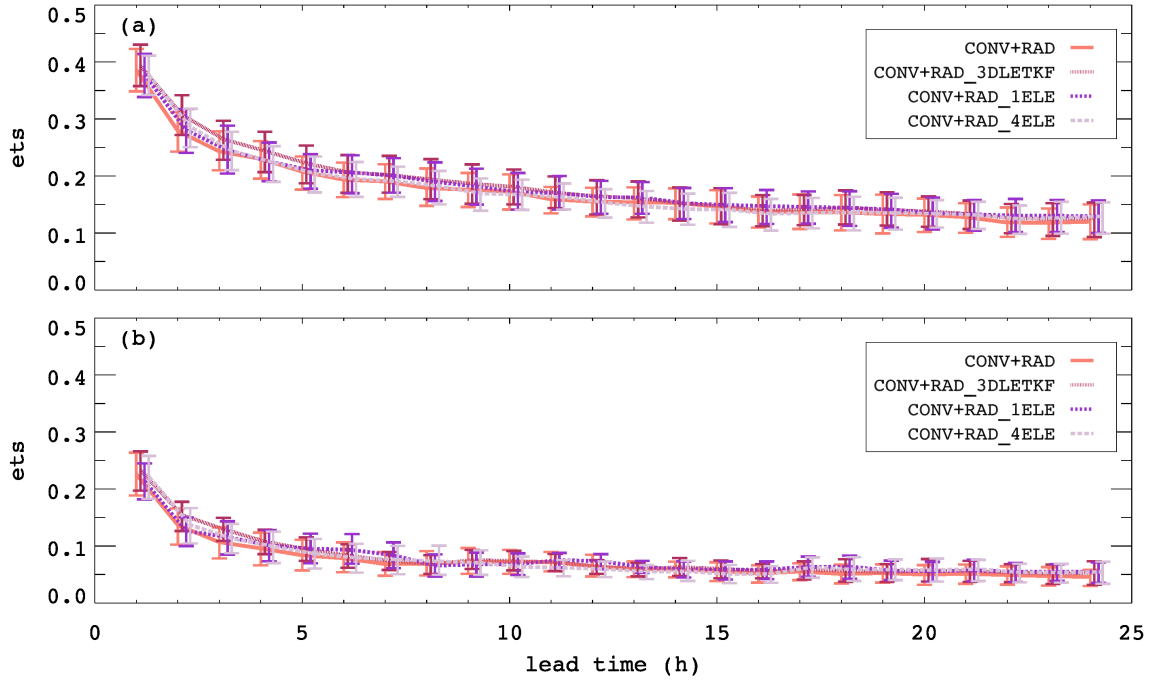


Figure B.4.: As Figure 6.9, but for CONV+RAD_1ELE (dark purple, dotted), CONV+RAD_4ELE (light purple, dashed), and CONV+RAD_3DLETKF (dark red, fine dashes).

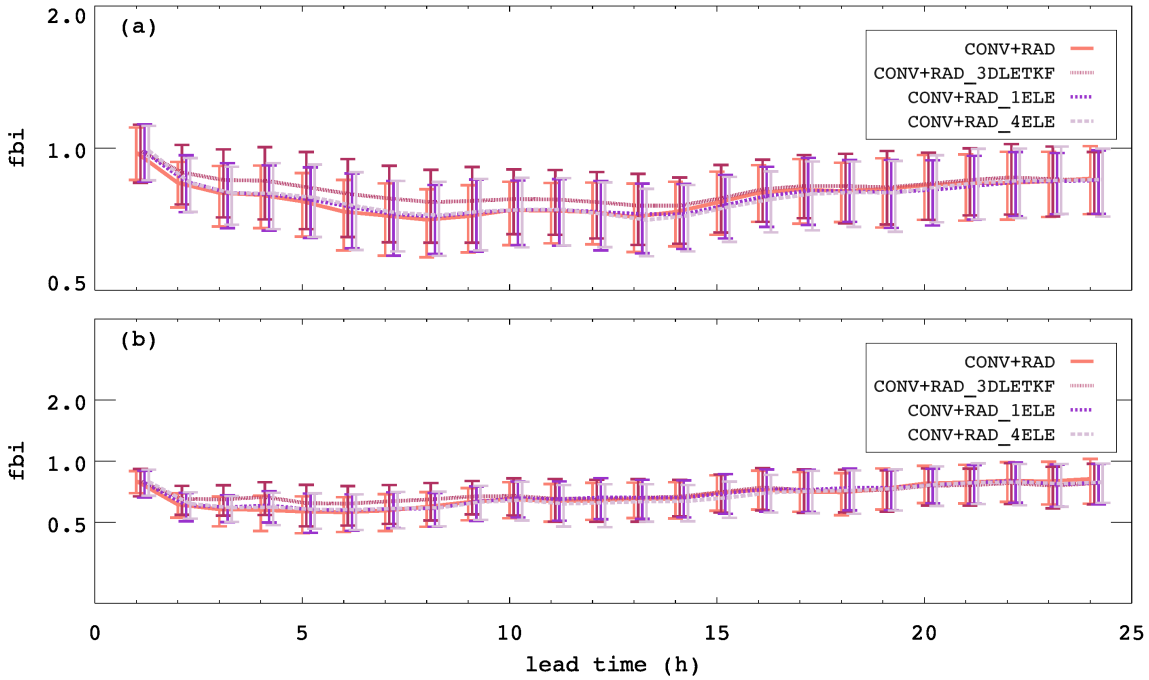


Figure B.5.: As Figure 6.10, but for CONV+RAD (light red, solid), CONV+RAD_1ELE (dark purple, dotted), CONV+RAD_4ELE (light purple, dashed), and CONV+RAD_3DLETKF (dark red, fine dashes).

correct the overestimation of events exceeding 0.5 mm/h (Fig. B.8 a) or underestimation of events exceeding 2 mm/h (Fig. B.8 b). CONV+RAD_4ELE seems to produce slightly less

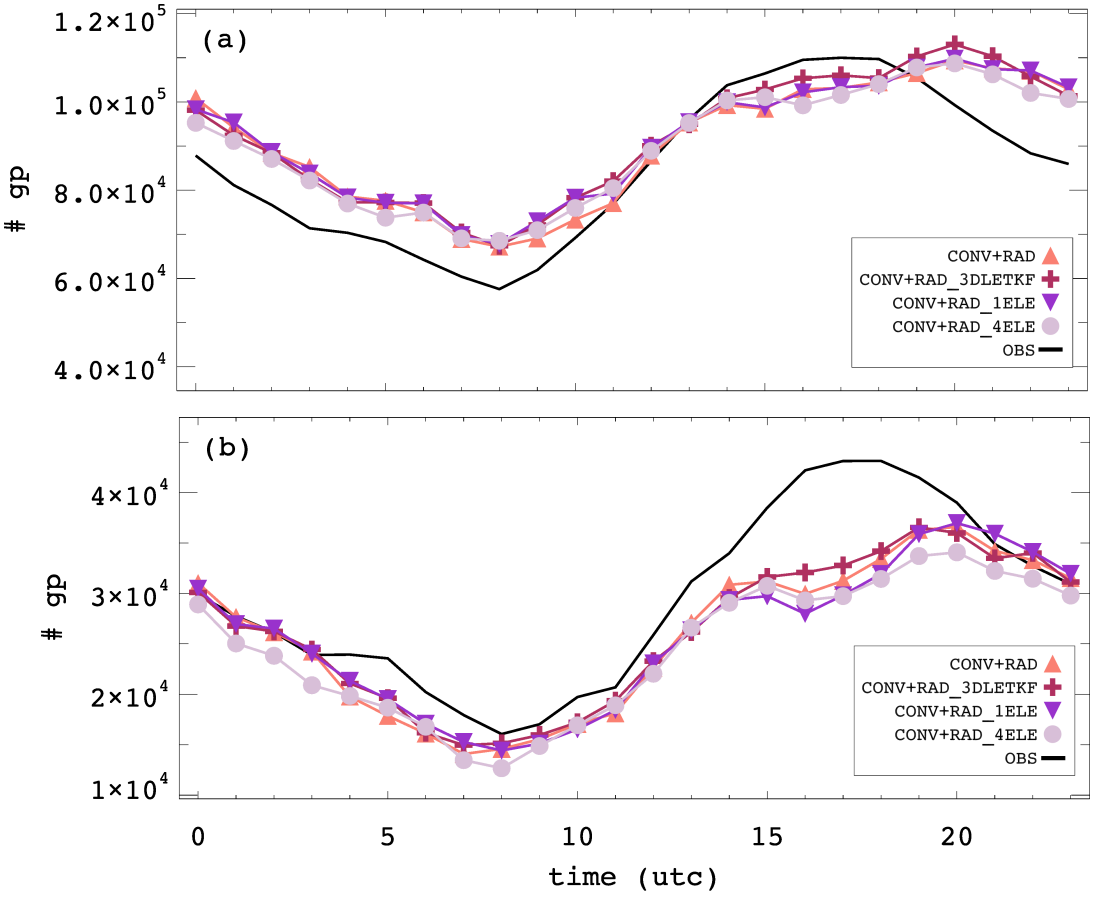


Figure B.6.: As Figure 6.11, but for CONV+RAD (red, upward triangles), CONV+RAD_3DLETKF (dark red, plus signs), and CONV+RAD_4ELE (purple, downward triangles).

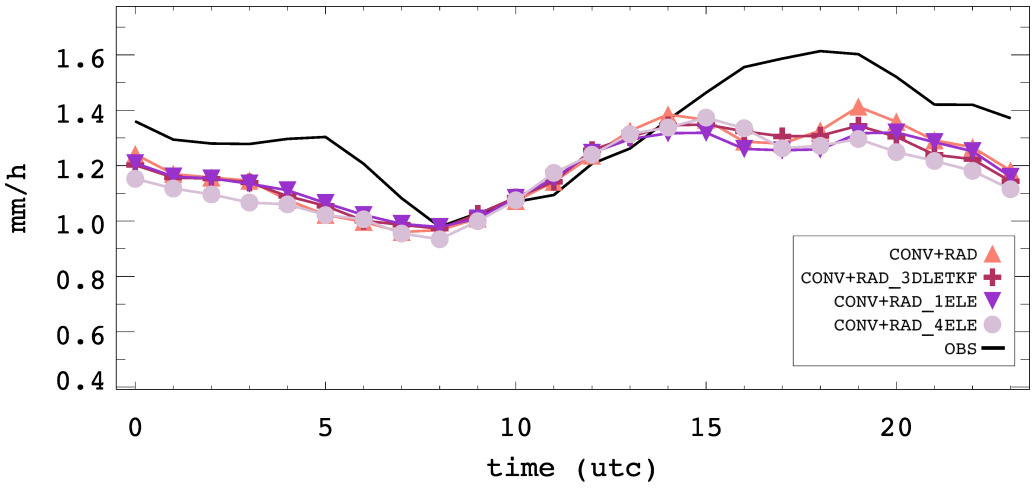


Figure B.7.: As Figure 6.12, but for CONV+RAD (red, upward triangles), CONV+RAD_3DLETKF (dark red, plus signs), and CONV+RAD_4ELE (purple, downward triangles).

events than CONV+RAD, and CONV+RAD_3DLETKF seems to produce slightly more

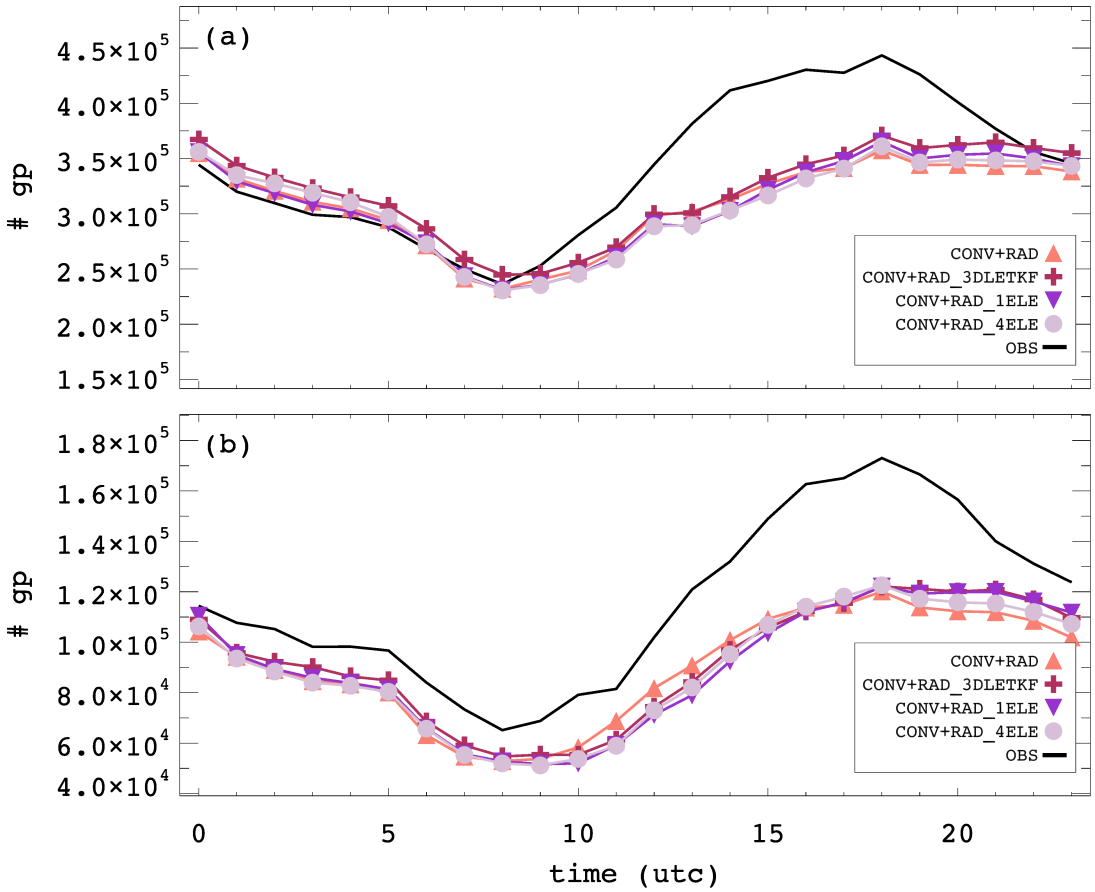


Figure B.8.: As Figure 6.13, but for CONV+RAD (red, upward triangles), CONV+RAD_3DLETKF (dark red, plus signs), and CONV+RAD_4ELE (purple, downward triangles).

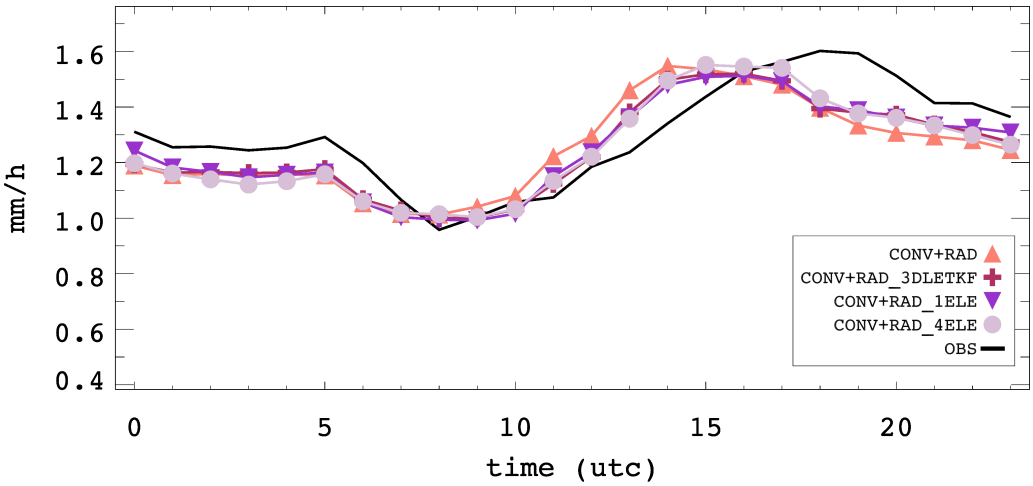


Figure B.9.: As Figure 6.14, but for CONV+RAD (red, upward triangles), CONV+RAD_3DLETKF (dark red, plus signs), and CONV+RAD_4ELE (purple, downward triangles).

events, although these differences are only very small. For the average precipitation intensities (Fig. B.9) the results are comparable: the differences between the four setups

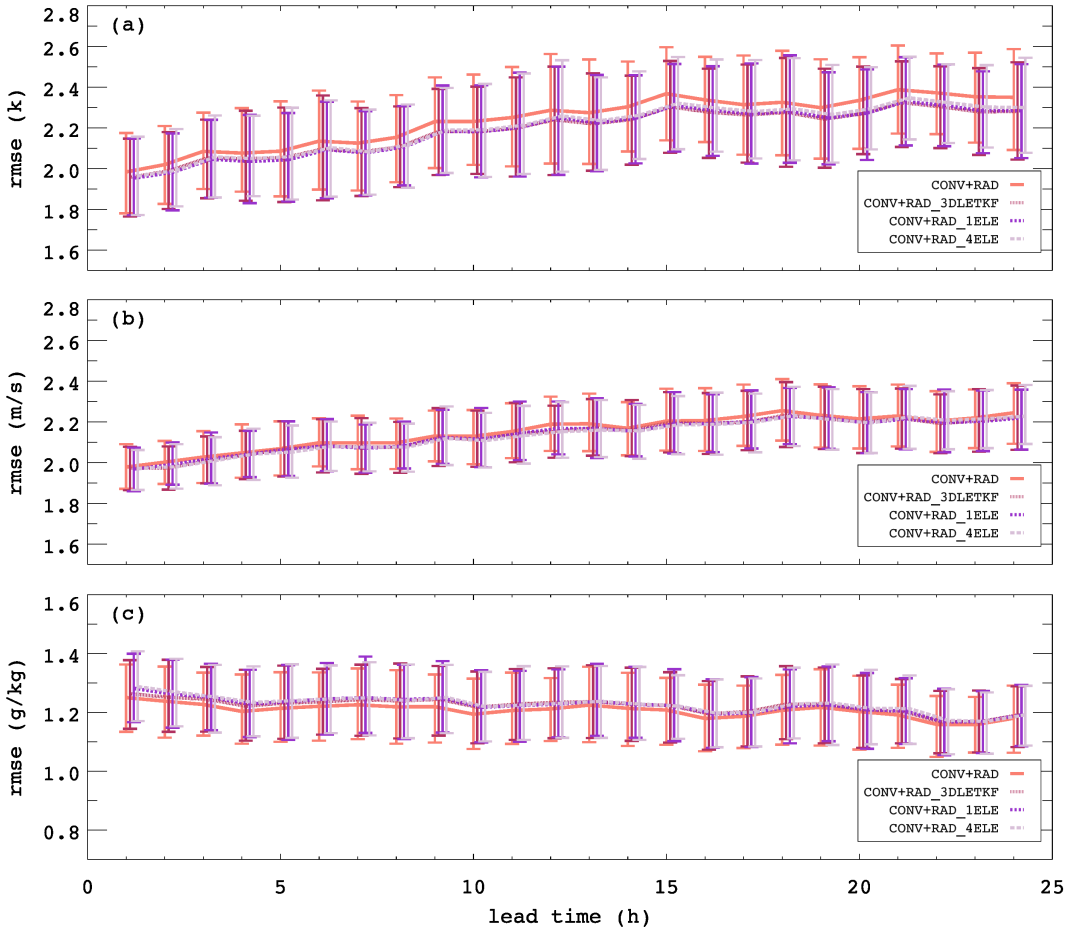


Figure B.10.: As Figure 6.15, but for CONV+RAD (light red, solid), CONV+RAD_1ELE (dark purple, dotted), CONV+RAD_4ELE (light purple, dashed), and CONV+RAD_3DLETKF (dark red, fine dashes).

are only small, where CONV+RAD_4ELE produces the least intense precipitation. All four setups strongly underestimate the afternoon maximum.

The RMSE (Fig. B.10) for all variables increases with forecast time and all for setups perform mostly equally. The differences among the four systems are small compared to the uncertainty. Also the bias (Fig. B.11) shows only small differences between the four systems.

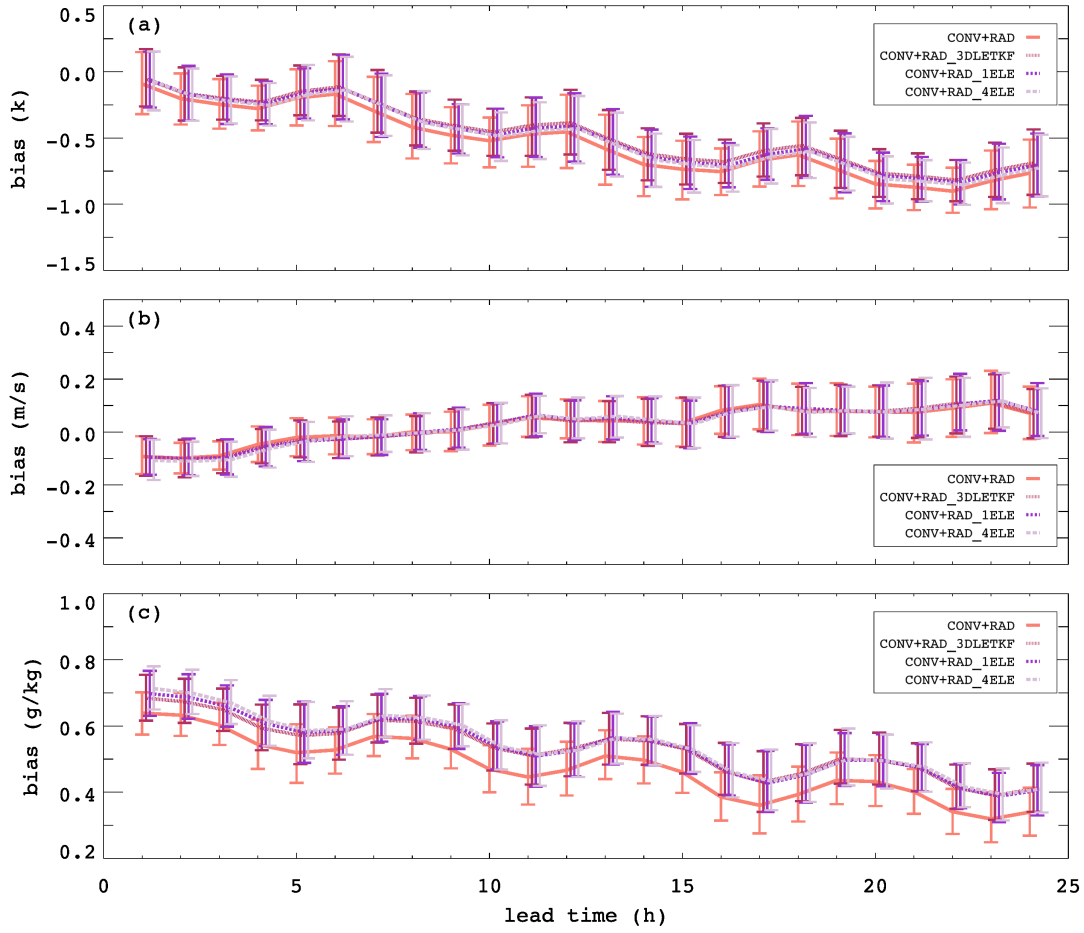


Figure B.11.: As Figure 6.16, but for CONV+RAD (light red, solid), CONV+RAD_1ELE (dark purple, dotted), CONV+RAD_4ELE (light purple, dashed), and CONV+RAD_3DLETKF (dark red, fine dashes).

BONNER METEOROLOGISCHE ABHANDLUNGEN

Herausgegeben vom Meteorologischen Institut der Universität Bonn durch Prof. Dr. H. FLOHN (Hefte 1-25), Prof. Dr. M. HANTEL (Hefte 26-35), Prof. Dr. H.-D. SCHILLING (Hefte 36-39), Prof. Dr. H. KRAUS (Hefte 40-49), ab Heft 50 durch Prof. Dr. A. HENSE.

Heft 1-49: siehe `http://www.meteo.uni-bonn.de/bibliothek/bma`

Heft 50: ***Petra Friederichs***: Interannuelle und dekadische Variabilität der atmosphärischen Zirkulation in gekoppelten und SST-getriebenen GCM-Experimenten. 2000, 133 S. + VIII. € 25

Heft 51: ***Heiko Paeth***: Anthropogene Klimaänderungen auf der Nordhemisphäre und die Rolle der Nordatlantik-Oszillation. 2000, 168 S.+ XVIII. € 28

Heft 52: ***Hildegard Steinhorst***: Statistisch-dynamische Verbundsanalyse von zeitlich und räumlich hoch aufgelösten Niederschlagsmustern: eine Untersuchung am Beispiel der Gebiete von Köln und Bonn. 2000, 146 S. + XIV. € 25

Heft 53: ***Thomas Klein***: Katabatic winds over Greenland and Antartica and their interaction with mesoscale and synoptic-scale weather systems: three-dimensional numerical models. 2000, 146 S. + XIV. € 25

Heft 54: ***Clemens Drüe***: Experimentelle Untersuchung arktischer Grenzschichtfronten an der Meereisgrenze in der Davis-Straße. 2001, 165 S. + VIII. € 28

Heft 55: ***Gisela Seuffert***: Two approaches to improve the simulation of near surface processes in numerical weather prediction models. 2001, 128 S. + VI. € 25

Heft 56: ***Jochen Stuck***: Die simulierte axiale atmosphärische Drehimpulsbilanz des ECHAM3-T21 GCM. 2002, 202 S. + VII. € 30

Heft 57: ***Günther Haase***: A physical initialization algorithm for non-hydrostatic weather prediction models using radar derived rain rates. 2002, 106S. + IV. € 25

Heft 58: ***Judith Berner***: Detection and Stochastic Modeling of Nonlinear Signatures in the Geopotential Height Field of an Atmospheric General Circulation Model. 2003, 157 S. + VIII. € 28

Heft 59: ***Bernd Maurer***: Messungen in der atmosphärischen Grenzschicht und Validation eines mesoskaligen Atmosphärenmodells über heterogenen Landoberflächen. 2003, 182 S. + IX. € 30

Heft 60: ***Christoph Gebhardt***: Variational reconstruction of Quaternary temperature fields using mixture models as botanical – climatological transfer functions. 2003, 204 S. + VIII. € 30

Heft 61: ***Heiko Paeth***: The climate of tropical and northern Africa – A statistical-dynamical analysis of the key factors in climate variability and the role of human activity in future climate change. 2005, 316 S. + XVI. € 15

Heft 62: ***Christian Schölzel***: Palaeoenvironmental transfer functions in a Bayesian framework with application to Holocene climate variability in the Near East. 2006, 104 S. + VI. € 15

Heft 63: ***Susanne Bachner***: Daily precipitation characteristics simulated by a regional climate model, including their sensitivity to model physics, 2008, 161 S. € 15

Heft 64: ***Michael Weniger***: Stochastic parameterization: a rigorous approach to stochastic three-dimensional primitive equations, 2014, 148 S. + XV. open access[1]

Heft 65: ***Andreas Röpnack***: Bayesian model verification: predictability of convective conditions based on EPS forecasts and observations, 2014, 152 S. + VI. open access[1]

Heft 66: ***Thorsten Simon***: Statistical and Dynamical Downscaling of Numerical Climate Simulations: Enhancement and Evaluation for East Asia, 2014, 48 S. + VII. + Anhänge open access[1]

Heft 67: ***Elham Rahmani***: The Effect of Climate Change on Wheat in Iran, 2014, [erschienen] 2015, 96 S. + XIII. open access[1]

Heft 68: ***Pablo A. Saavedra Garfias***: Retrieval of Cloud and Rainwater from Ground-Based Passive Microwave Observations with the Multi-frequency Dual-polarized Radiometer ADMIRARI, 2014, [erschienen] 2015, 168 S. + XIII. open access[1]

Heft 69: ***Christoph Bollmeyer***: A high-resolution regional reanalysis for Europe and Germany - Creation and Verification with a special focus on the moisture budget, 2015, 103 S. + IX. open access[1]

Heft 70: ***A S M Mostaquimur Rahman***: Influence of Subsurface Hydrodynamics on the Lower Atmosphere at the Catchment Scale, 2015, 98 S. + XVI. open access[1]

Heft 71: ***Sabrina Wahl***: Uncertainty in mesoscale numerical weather prediction: probabilistic forecasting of precipitation, 2015, 108 S. open access[1]

Heft 72: ***Markus Übel***: Simulation of mesoscale patterns and diurnal variations of atmospheric CO_2 mixing ratios with the model system TerrSysMP-CO_2, 2015, [erschienen] 2016, 158 S. + II open access[1]

Heft 73: ***Christian Bernardus Maria Weijenborg***: Characteristics of Potential Vorticity anomalies associated with mesoscale extremes in the extratropical troposphere, 2015, [erschienen] 2016, 151 S. + XI open access[1]

Heft 74: ***Muhammad Kaleem***: A sensitivity study of decadal climate prediction to aerosol variability using ECHAM6-HAM (GCM), 2016, 98 S. + XII open access[1]

Heft 75: ***Theresa Bick***: 3D Radar reflectivity assimilation with an ensemble Kalman filter on the convective scale, 2016, [erschienen] 2017, 96 S. + IX open access[1]

[1]Available at `http://hss.ulb.uni-bonn.de/fakultaet/math-nat/`

METEOROLOGISCHES INSTITUT
MATHEMATISCH NATURWISSENSCHAFTLICHE FAKULTÄT
UNIVERSITÄT BONN

Meteorologisches Institut
Universität Bonn