Distributions of Forecast Variables in a Convective-scale 1000-member Ensemble

Matjaž Puh

Master Thesis
at the Faculty of Physics
of the Ludwig Maximilian University of Munich

Munich, September 9^{th} 2020

Verteilung von Vorhersagegrößen in einem 1000-member Ensemble auf der konvektiven Skala

Matjaž Puh

Masterarbeit
an der Fakultät für Physik
der Ludwig–Maximilians–Universität München

München, den 9. September 2020

First supervisor *Erstbetreuer*: Prof. Dr. George Craig
Second supervisor *Zweitbetreuer*: Dr. Christian Keil
Submission date *Datum der Abgabe*: 9.9.2020

Contents

Abstract

Ensemble prediction systems are a powerful tool to investigate forecast uncertainty in numerical weather prediction. However, computational resources limit the number of members in operational ensembles, which is not sufficient to accurately represent probability distributions of forecast variables. This thesis investigates these distributions using a large convective-scale, 1000-member ensemble. The inspected dataset consists of 14-hour forecasts of the SCALE-RM model with a resolution of 3 km over Germany, during a convectively active summertime period in 2016.

Distributions of forecast variables are categorized in 3 groups, based on their shape and properties: normally distributed, highly skewed and mixture distributions. It is shown that temperature and wind at 500 hPa and pressure at sea level are normally distributed variables, precipitation and reflectivity have highly skewed distributions and moisture-related variables, like specific or relative humidity, are mixtures.

In addition, spatial scale dependency and time evolution of probability distributions is studied. It is found that increasing the sampling region width more than 200 km leads to mixing of independent distributions and can cause non-Gaussianity and multi-modality. A conceptual model of time evolution of distributions is presented and the emerging results for lead times up to 14 hours are compared with it. Mixture variables appear to follow this model fairly accurately, although the last stage is not reached due to the short lead time. It is also shown that at least 320 to 600 ensemble members are needed for the distributions' properties to start converging towards the values of the full, 1000-member sample. Lastly, it is found that precipitation and vertical velocity extreme events depend more on ensemble size than other variables.

Chapter 1

Introduction

Increasing computational power has allowed for higher resolution numerical weather prediction (NWP) models to run operationally, simulating convective-scale phenomena. This brought new challenges in NWP: given the faster error growth at smaller scales, a probabilistic approach is required (Lorenz, 1969; Hohenegger et al., 2006; Clark et al., 2010). Ensemble prediction systems (i.e. *ensembles*) are of great help in representing this uncertainty, indicating a range of possible outcomes from small perturbations in initial and boundary conditions, as well as model physics.

Another challenge in higher resolution NWP is the accuracy of initial conditions at smaller scales (Stensrud et al., 2009; Sun et al., 2014). Current operationally used Data Assimilation (DA) systems were implemented under the assumption that error probability density functions (PDFs) are Gaussian and do not perform well when this assumption is not met (Evensen and van Leeuwen, 2000). This happens more frequently at smaller scales because of non-linear processes (Zhang, 2005).

Large ensembles are required to accurately investigate the shape of PDFs to improve both the understanding of forecast uncertainty and data assimilation. A small ensemble is likely to fail capturing non-Gaussian effects, such as multi-modality and outliers (Bannister et al., 2017). This study provides some new insights into this field using a unique dataset with a convective-scale 1000-member ensemble.

Chapter 1 introduces ensemble prediction in NWP and then presents the research questions and the outline of this thesis.

1.1 Ensemble prediction

Before the 1990s, NWP was prevalently treated as a well-defined deterministic problem, although several studies demonstrated the chaoticness of the atmospheric system (e.g. Lorenz, 1969). The intensity of instabilities and the error growth rate depend on the scale of motions, since spacial and temporal scales of features are connected: smaller-scale cumulus clouds develop as perturbations to the prevailing environment in the order of tens of minutes, while large-scale extratropical cyclones evolve over several days. For smaller systems, this results in much faster nonlinear perturbation growth. Faster perturbation growth corresponds with faster error growth and therefore a faster loss in predictability for smaller-scale features (Toth and Buizza, 2019). Lorenz (1969) investigated the "predictability of a flow which possesses many scales of motions" using a simple mathematical model and found that "each scale of motion possesses an intrinsic finite range of predictability". He concludes that "cumulus-scale" motions can be predicted up to about one hour, synoptic-scale motions a few days and the largest, planetary scales a few weeks in advance. This scale dependence of predictability shows up also in contemporary predictability studies (Boer, 2003).
Given the increased resolution of operational NWP models in the last decades, convective-scale features are now represented and simulated operationally. Since the predictability of these features is in the range of hours, a probabilistic approach is more suitable than a single deterministic forecast. Ensemble prediction is of significant help in representing this uncertainty.

The ensemble approach consists in generating a set of perturbed forecasts, designed to simulate the effect of possible uncertainties in the control (unperturbed) forecast. The resulting set of perturbed forecasts can then be used to estimate the uncertainty of the forecast, the most likely set of values of a certain variable, etc. Buizza and Leutbecher (2015) also defined the predictability horizon as the time when the PDF became indistinguishable from the climatological PDF.
In the last decades, different ensemble generation techniques have been tested. The first attempts in the 1980s, with random perturbations to initial conditions or lagged ensembles (forecasts initialized recently at different times) did not deliver the desired results: in the latter, short-term forecast quality was deteriorated by older forecasts, while the former did not provide sufficient forecast diversity (spread). More promising techniques have been developed in the 1990s, that can simulate initial condition and model-related uncertainties. The European Centre for Medium-Range Weather Forecasts (ECMWF) implemented the Singular Vectors (SVs) technique, which are the perturbations with the fastest growth over a finite time interval. In 2008, Ensembles of Data Assimilations were incorporated, i.e. perturbations from multiple DA cycles, which together with SVs remain today an essential component of the ECMWF ensemble (Toth and Buizza, 2019).

Ensembles provide not only the most likely scenario, but also the associated uncertainty and are therefore very useful in operational weather forecasting, as well as in predictability studies. Ensembles of short-range forecasts or DA cycles are also used in ensemble DA (such as in the Ensemble Kalman Filter) and to estimate the uncertainty of the analysis. In the context of higher resolution NWP and the simulation of smaller-scale features, ensemble prediction helps us to address some of the connected issues. Firstly, it applies the required probabilistic approach to NWP, giving better estimates of the forecast uncertainty. Secondly, it allows us to represent PDFs of forecast variables, which are essential to understand the evolution of forecast uncertainty and to develop data assimilation algorithms that provide accurate initial conditions.

Figure 1.1 shows how ensemble prediction systems can be used to investigate forecast uncertainty and the PDFs of forecast variables. Compared to the deterministic, control forecast, the ensemble represents several possible scenarios for the evolution of the state of the atmosphere, indicating which are potentially more likely if members are "close together". In the example in Figure 1.1, the control forecast deviates significantly from reality at the end, while half of the ensemble members are close to it. The resulting PDF has three peaks, the largest just around the real value of temperature.

However, to accurately represent PDFs and better resolve their details, a large number of ensemble members is required . Because of computational power limitations, current operational ensembles in NWP centers have sizes only up to 50 members for global models and even fewer for limited area models (LAM), which is not enough to accurately represent the nonlinear evolution of PDFs (Leutbecher, 2019). Furthermore, the problem of undersampling in ensemble prediction has been studied more for large-scale models, while few studies focused on LAM. Using an intermediate atmospheric general circulation model, Kondo and Miyoshi (2019) performed experiments with up to 10 240 ensemble members, suggesting that approximately 1000 ensemble members were necessary to represent the detailed structures of PDFs, such as multi-modality and tails. Focusing on the convective scale, Harnisch and Keil (2015) found that an increase of ensemble size from 20 to 40 members led to a more accurate analysis and more accurate 3-hour-forecast. Hagelin et al. (2017) showed a large improvement in forecast skill for precipitation with a doubling of ensemble size from 12 to 24 using the Met Office convective-scale ensemble (MOGREPS-UK). Raynaud and Bouttier (2017) compared the benefits of increasing ensemble size from 12 to 34 members, to the benefits of increasing horizontal resolution from 2.5 to 1.3 km. They showed that the increase in ensemble size is more beneficial than a resolution increase for lead times greater than about 12 h, when there is a larger uncertainty to be sampled, which agrees with Legrand et al. (2016), who found that non-Gaussianity increases with forecast lead time. Their and previous studies using big ensembles at both global scale and regional scale agree that the largest non-Gaussianity arises from highly non-linear processes in deep convective clouds (Miyoshi et al., 2014; Jacques and Zawadzki, 2015). Kawabata and Ueno (2020) examined the shape of forecast distributions with a 1000 member convective-scale ensemble using a particle filter approach. They found that non-Gaussianity can develop in

much less than an hour in deep moist convection, originating in the region of the convective updraft.
The question of how to increase ensemble size without increasing computing resources through alternative cheap approaches like neighbourhood techniques and lagged ensembles has also been studied (e.g. Ben Bouallègue et al., 2013; DelSole et al., 2017). Porson et al. (2020) recently showed that the time-lagging objectively improves the forecast at all lead times, with larger improvements in the first few hours. Increasing the size of a time-lagged ensemble through lagging over additional cycles leads to small but significant improvements, larger in most cases than those that can be obtained through neighbourhood processing.

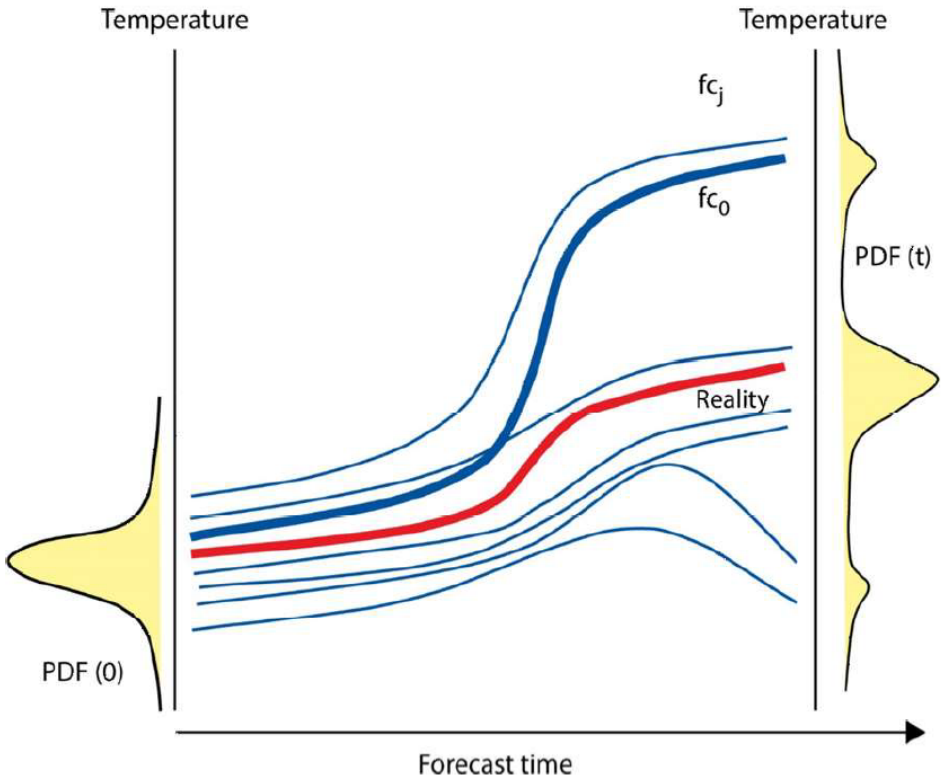


Figure 1.1: Schematic representation of the time evolution of an ensemble, showing the initial PDF of temperature and one at a later time. The single ensemble members (thin blue lines, fc_j) are compared with the control forecast (thick blue line, fc_0) and reality (thick red line). Figure 13 of Toth and Buizza (2019).

1.2 Research questions and outline

In this study, a convective-scale 1000-member ensemble is used to investigate forecast uncertainty in high resolution ensemble prediction in a convective environment. The main scientific questions of this study are the following:

1. Is it possible to classify PDFs of forecast variables in a limited number of categories based on their shape?

2. How do PDFs of forecast variables depend on the spatial scale they represent and how do they evolve with lead time?

3. How does the size of the ensemble influence PDFs of forecast variables and how large must the ensemble be to accurately represent the forecast distribution?

Preliminary results suggest 3 stages of PDF time evolution. Firstly, the error grows spatially and temporally from the initial Gaussian distribution coming from DA algorithms and the distribution broadens as the position and intensity of convective cells diverge. Secondly, when the initial convective cells decay and are replaced by new ones, the distribution develops non-Gaussian features such as heavy tails representing rare extreme events or multimodal structures associated with long-lived organized convective systems. Lastly, the distribution converges to the one determined by the synoptic environment and all predictability associated with the initial convective state is lost. This suggested conceptual model of PDF evolution, shown in Figure 1.2, is compared with new results. Differences between the time evolution of distributions in a stronger and a weaker synoptically-forced case are briefly discussed as well.
This thesis is structured as follows: Chapter 2 introduces the used dataset and the synoptic situation during the experimental period. Then the employed methods are described, followed by the presentation of the results and the discussion in chapter 3. Conclusions are proposed in chapter 4.

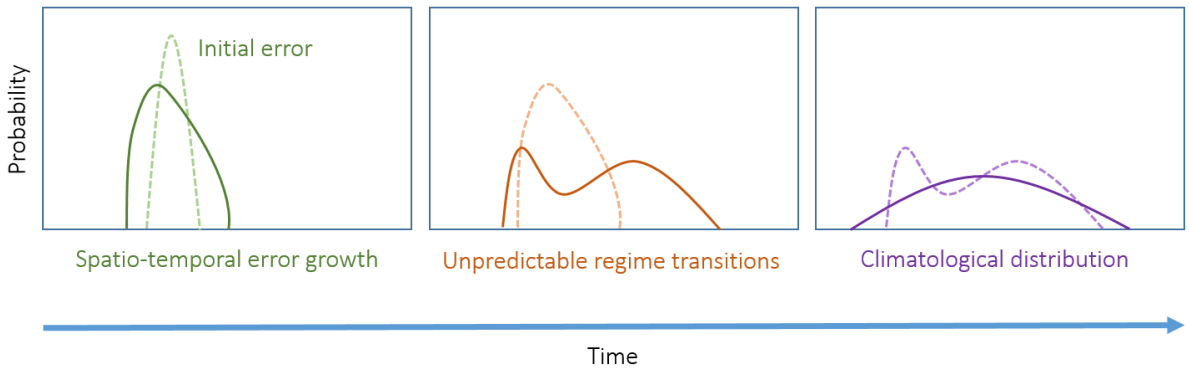


Figure 1.2: Conceptual model of the time evolution of PDFs of forecast variables.

Chapter 2

Data and methods

2.1 The SCALE-RM ensemble

The data used in this study was produced in the experiment performed by Necker et al. (2020a) with the SCALE-RM model, which consists of ten ensemble forecasts during summer 2016 over Germany. The simulation was computed on the K computer in collaboration with the RIKEN Institute for computational science in Japan.
The offline nesting approach was used by downscaling the initial conditions from a larger, 15-km grid-spaced domain to a smaller, 3-km domain (shown in figure 2.1) every 12 hours. The larger domain was used for the 15-km ensemble data assimilation 3h-cycles with the Global Ensemble Forecast System (GEFS) of the National Center for Environmental Prediction (NCEP) boundary conditions. The initial conditions were taken from a previous 15-km, 1000-member DA experiment over the same domain, spun up for a week. The boundary conditions for the 3-km domain were provided hourly by a parallel 15-km simulation on the outer domain, initialized every 12 hours. The boundary conditions for 15-km forecast were prepared every 6 hours, using a combination of randomly generated perturbations and the 20-member GEFS analyses. Please refer to Necker et al. (2020a) for more details about the ensemble setup.
The simulation used the Scalable Computing for Advanced Library and Environment Localized Ensemble Transform Kalman Filter (SCALE-LETKF) DA system (Lien et al., 2017). The SCALE-LETKF system combined the open source Scalable Computing for Advanced Library and Environment Regional Model (SCALE-RM: version 5.1.2; see Nishizawa et al., 2015; Sato et al., 2015; Nishizawa and Kitamura, 2018) and a Localized Ensemble Transform Kalman Filter (LETKF: Hunt et al., 2007). The LETKF assimilated conventional observations using a 3-hr assimilation window on the 15-km grid. All experiments used the Tomita (2008) single-moment bulk microphysics scheme, the Mellor–Yamada–Nakanishi–Niino 2.5 closure boundary-layer scheme (Nakanishi and Niino, 2004), the Model Simulation Radiation Transfer code for the representation of radiative fluxes (Sekiguchi and Nakajima, 2008), and the Beljaars-type surface model (Beljaars and Holtslag, 1991) for the computation of soil variables and surface fluxes.

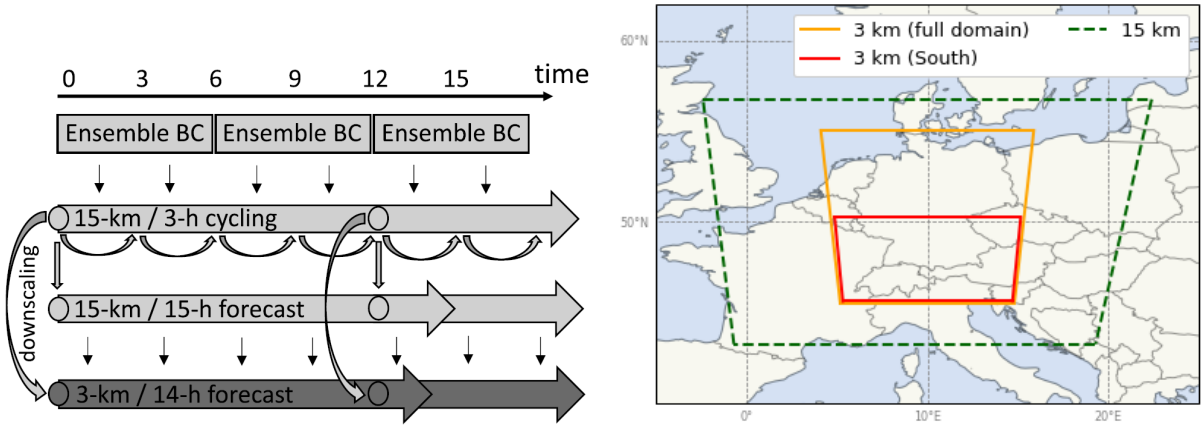
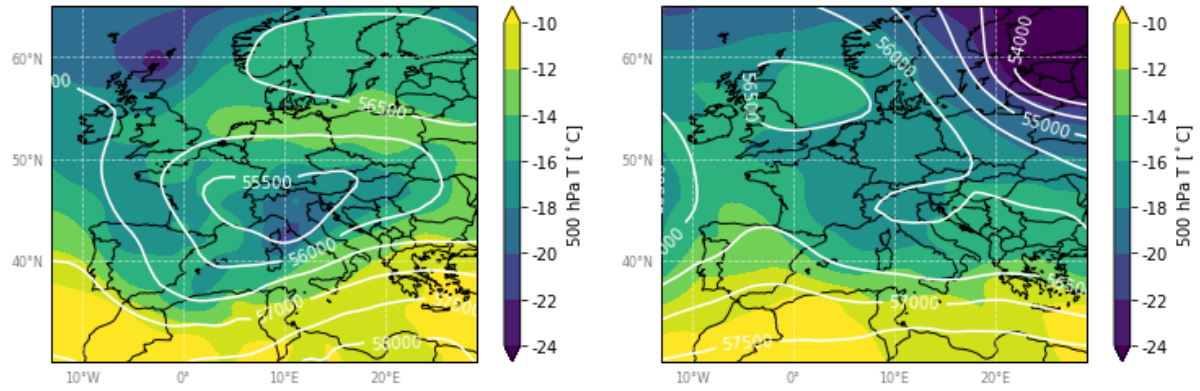


Figure 2.1: Left: flow chart of the simulation setup (from Necker et al. (2020a), Figure 1). Right: 15-km grid-spaced model domain used for data assimilation cycling and boundary conditions (dashed green), the full domain (solid yellow) and the Southern subdomain (solid red) of the 3-km forecast.

The forecasts used in this study are on the 3 km grid and span from 29 May, 12 UTC to 6 June, 02 UTC. A forecast was initialized every 12 hours (00 or 12 UTC) up to a lead time of 14h. The domain measures 350x250 grid-points with 30 vertical levels.

2.2 Synoptic situation

The analyzed period (from 29 May to 5 June 2016) was characterized by a fairly stationary weather pattern over central Europe with an upper-level trough over western Europe, accompanied by a shallow surface low in the first part of the simulation period, followed by a weaker pressure gradient synoptic pattern (figure 2.2). The mid-level winds backed from warm and moist southerly to easterly in the first period and to north-easterly in the second period.
Throughout the simulation, the environment was highly unstable with surface heating and convective instability triggering convective activity, which was particularly intense in the first few days and caused high impact weather in parts of Germany, also caused by low mid-level wind speeds, which kept the convective cells' motion slow.
Especially later in the simulation period, convection followed a clear diurnal cycle, with a peak later in the afternoon.
Because of the intense, extensive and recurring convection, this period has been the focus of many studies (e.g. Piper et al., 2016; Keil et al., 2019; Bachmann et al., 2020; Necker et al., 2020b; Scheck et al., 2020) and offers a good environment to investigate the peculiarities of PDFs of forecast variables in convective conditions.



(a) 31 May 2016, 18 UTC

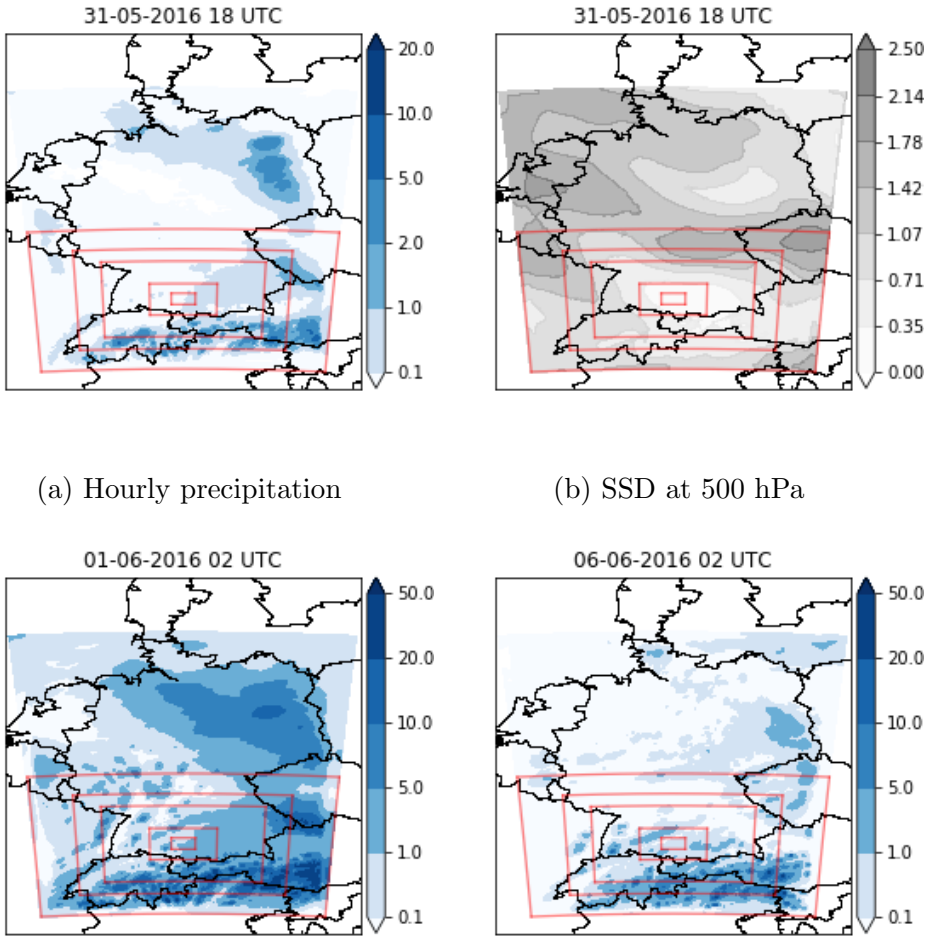
(b) 5 June 2016, 12 UTC

Figure 2.2: ECMWF ERA-Interim reanalysis of geopotential height (dam, white contours) and temperature at 500 hPa (°C, shaded) for 31 May 2016, 18 UTC (a) and 5 June 2016, 12 UTC (b).

2.3 Methods

2.3.1 Data selection and distribution classification

Probability distributions of data are easily explored by looking at histograms. In a chosen model domain, it aggregates values at all gridpoints inside the domain, for all the ensemble members. All 3-dimensional variables were only looked at on a constant pressure level. If the domain size is $n \times m$ gridpoints and n_{ens} is the number of ensemble members, $n \times m \times n_{ens}$ is the number of values included in the histogram.
The model domain was systematically divided into several subdomains to cover a range of environments as complete as possible. Given the geographical differences between the southern and northern part of Germany, namely the influence of the sea in the north, the southern half was chosen for investigation, as shown in picture 2.1. Convection was also generally more intense and widespread over and around the Alps in the south.
The shape of the subdomain was chosen to cover half of the model domain at full width. This ratio was used also for smaller subdomains, all centered above the same gridpoint (see figure 2.3). The smallest examined subdomain was 5 gridpoints wide and 7 gridpoints long (15×21km).
Histograms were plotted to analyse the large amount of distributions created. If reasonable, a Gaussian or other standard PDF was additionally plotted to compare it to the histogram. Variables were then categorized into groups with similar characteristics, based on their shape and this comparison. It has to be emphasized that the choice of variables was limited to temperature, horizontal and vertical wind velocities (u, v, w), specific humidity, relative humidity, hydrometeor mixing ratio, radar reflectivity, surface level pressure, surface precipitation rate, surface net longwave radiation flux, surface net shortwave radiation flux, surface downward longwave radiation flux, surface downward shortwave ra-



(a) Hourly precipitation

(b) SSD at 500 hPa

(c) 14-hour precipitation since 31 May 2016, 12 UTC

(d) 14-hour precipitation since 5 June 2016, 12 UTC

Figure 2.3: SCALE-RM ensemble mean hourly precipitation (a), specific saturation deficit at 500 hPa (b) and 14-hour accumulated precipitation (c, d) for indicated times. The red rectangles are the subdomains used in the analysis of spacial scale dependency (see section 2.3.2).

diation flux, TOA net longwave radiation flux and shortwave radiation flux. The selected variables for investigation were temperature, horizontal and vertical wind velocity, reflectivity, specific humidity, relative humidity, specific saturation deficit, surface precipitation rate and pressure at mean sea level. All variables were extracted at 500 hPa apart from the last two.

Specific saturation deficit (SSD) was defined as the difference between the saturation water vapor mixing ratio at the gridpoint temperature and the actual mixing ratio:

$$q_{def} = q_{sat} - q = q\left(f^{-1} - 1\right), \tag{2.1}$$

where f is the relative humidity. Results for 31 May 2016 are presented, because most of the specific features of distributions were visible. However, two other days were analysed, 29 May and 5 June 2016, the former exhibiting more intense and organised convective activity over the domain and the latter much weaker large-scale forcing and more scattered convection, mostly over orography.

2.3.2 Scale and time dependency

Six spatial scales were determined: 3, 45, 120, 285, 375 and 525 km of subdomain width. 3 km corresponds to one gridpoint, while 525 km covers the full southern domain. Distributions of variables were examined for these six subdomains, centered over the middle point of the largest subdomain (see figure 2.3) for one fixed point in time (31 May 2016, 18 UTC, +6h lead time). Histograms were produced in two ways: by aggregating data, as explained in the previous section, and by spatially averaging data over the examined domain, to investigate two different spacial scale dependencies.
Time dependency was studied by fixing the domain size (45 km wide) and plotting PDFs for lead times from the first timestep (a few seconds) to 14 hours. Histograms were then compared to the conceptual model of distribution evolution in figure 1.2.

2.3.3 Sampling error and extreme values

To quantify sampling error, sub-samples of size 20, 40, 80, 160, 320, 600 and 800 were selected, as in Necker et al. (2020a). The sub-samples were constructed considering the boundary conditions production with the 20 GEFS members, to avoid clustering. The Kolmogorov-Smirnov two-sample test was performed for each of the sub-samples against the full, 1000-member ensemble, which is considered to be a good estimate of the truth. The two-sample Kolmogorov-Smirnov test is a nonparametric hypothesis test that evaluates whether two samples come from the same distribution (Bonamente, 2017). The test statistic is defined as:

$$D_{mn} = \max_x \left|F_m(x) - G_n(x)\right| \tag{2.2}$$

where $F_m(x)$ is the sample cumulative distribution of a set of m observations, and $G_n(x)$ that of another independent set of n observations. The statistic D_{mn} therefore measures the maximum deviation between the two cumulative distributions. Its purpose in this analysis was to compare the statistic for sub-samples of different sizes, so the resulting values were not compared with critical values.
Additionally, the sample mean and variance were computed for all sub-samples. These values were then subtracted from the values for the full, 1000-member sample and divided

by these same values, to obtain a relative deviation of these distribution moments from the moments of the 1000-member distributions, i.e.

$$R_{m,1000} = \frac{M_m - M_{1000}}{M_{1000}} = \frac{\Delta M}{M_{1000}} \tag{2.3}$$

where M_m is the moment of the distribution obtained from the sub-sample of size m and M_{1000} the moment of the distribution obtained from the 1000-member sample.
Subsamples were chosen for three different subdomain sizes: 45, 120 and 375 km of width. For each subdomain size, several samples were chosen, to cover the largest as much as possible, without too many overlaps. This way, 71 samples for the 45 km, 20 samples for the 120 km and 1 sample for the 375 km wide subdomain were obtained. Sample means were calculated for all sampling error statistics.
The same procedure was adopted to evaluate the probability of extreme values depending on ensemble size. The 95^{th} percentile of distributions was calculated for this purpose and compared to the value for 1000 members according to equation 2.3.

Chapter 3

Results and discussion

3.1 Distributions classification

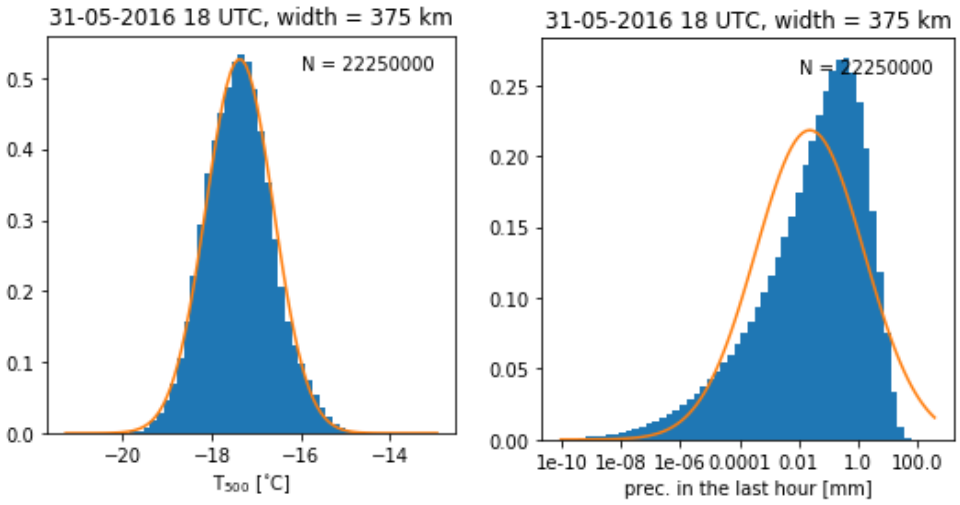
After a subjective graphical analysis, distributions are categorized into three groups: normally distributed, mixtures and highly skewed distributions. Table 3.1 shows how the examined variables are assigned to these groups. In this section, examples for each group are described and a physical interpretation for this classification is presented.

Normally distributed variables, such as temperature (see figure 3.1a), all wind components at 500 hPa and pressure at mean sea level, always stay close to a Gaussian distribution, independently of the aggregation region size and the forecast lead time.

The highly skewed category comprises distributions that are not Gaussian because of their large skewness. Distributions of precipitation (see figure 3.1b) and reflectivity, which are naturally connected through hydrometeor content, belong to this class. Moreover, precipitation distributions are close to log-normal because of underlying physics. Precipitation

Group	**Variable**
Normally distributed	Temperature
	Horizontal wind velocity
	Vertical wind velocity
	Pressure at mean sea level
Highly skewed	Precipitation
	Reflectivity
Mixtures	Specific humidity
	Specific saturation deficit
	Relative humidity

Table 3.1: Classification of variables in groups. All 3-dimensional variables are extracted at 500 hPa, apart from pressure, which is at mean sea level.

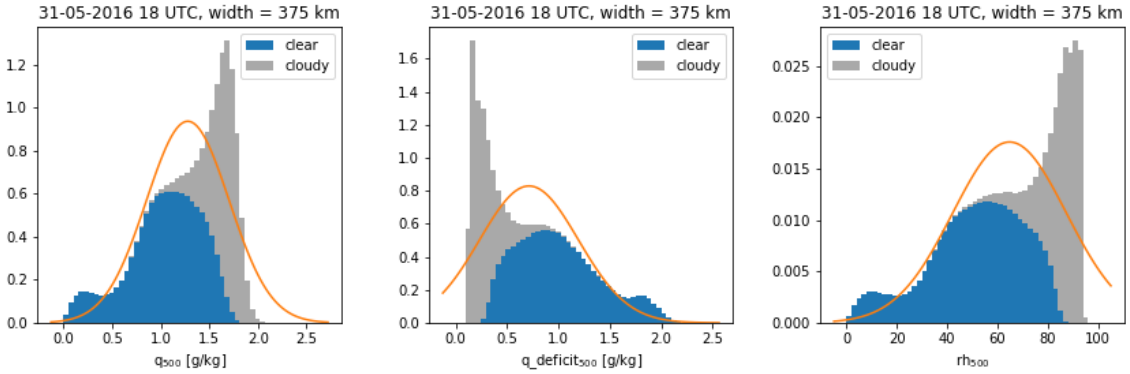


(a) (b)

Figure 3.1: Histograms of temperature [°C] at 500 hPa (a) and precipitation [mm] in the last hour (b) on 31 May at 18 UTC in a 375 km wide sub-domain. N is the used number of gridpoints. A Gaussian with the same mean and standard deviation (a) and a log-normal PDF (b) are shown for comparison.

development in clouds is a multiplicative process, which becomes additive on a log scale. This yields a log-normal distribution, following the law of proportionate effect (Kedem and Chiu, 1987). However, precipitation distributions show a significant number of values below 0.01 mm per hour. These values are not meaningful and are usually referred to as "model drizzle".

The last group, mixture distributions, appears to deviate significantly from the Gaussian assumption, especially in convective environments, when nonlinear processes are important. In this study, all moisture-related variables belong to this category. The origin of their distribution shape is the natural bimodality given by cloudy and clear gridpoints, as seen in figure 3.2. Cloudy gridpoints are moister, although they do not necessarily appear as saturated, which makes their distribution strongly skewed. The distribution for cloud-free gridpoints is considerably more Gaussian.
Moisture transport converges into convective clouds and diverges in clear regions, which together with subsiding dry air in clear sky areas between convective clouds enforces this bimodality. Moreover, specific humidity and specific saturation deficit are always positive, while relative humidity is also bounded by 100%, preventing distributions from having long tails.



(a) (b) (c)

Figure 3.2: Histograms of (a) specific humidity [g/kg], (b) specific saturation deficit [g/kg] and (c) relative humidity [%] at 500 hPa in the southern region with a width of 375 km on 31 May, 18 UTC. A Gaussian with the same mean and standard deviation is shown for comparison. The colors indicate contributions from cloudy and clear gridpoints at 500 hPa, respectively. The criterion to distinguish cloudy gridpoints is a simulated radar reflectivity higher than -19 dBZ at 500 hPa.

3.2 Scale and time dependency

Scale dependency is studied by analyzing distributions of variables for six concentric subdomains on 31 May 2016 at 18 UTC, as explained in section 2.3.2. Histograms are produced by aggregating data or by spatially averaging data over the examined domain.

3.2.1 Aggregation

A larger aggregation region (45 or 120 km instead of 3 km) reduces the noise, especially for precipitation, due to independent realizations of the same distribution (figure 3.3). However, if the aggregation region increases over around 200 km of width, different distributions start contributing to the histogram (e.g. cloudy and clear regions), so the SSD distribution develops a second peak. On the other hand, the temperature distribution remains Gaussian, but broadens.

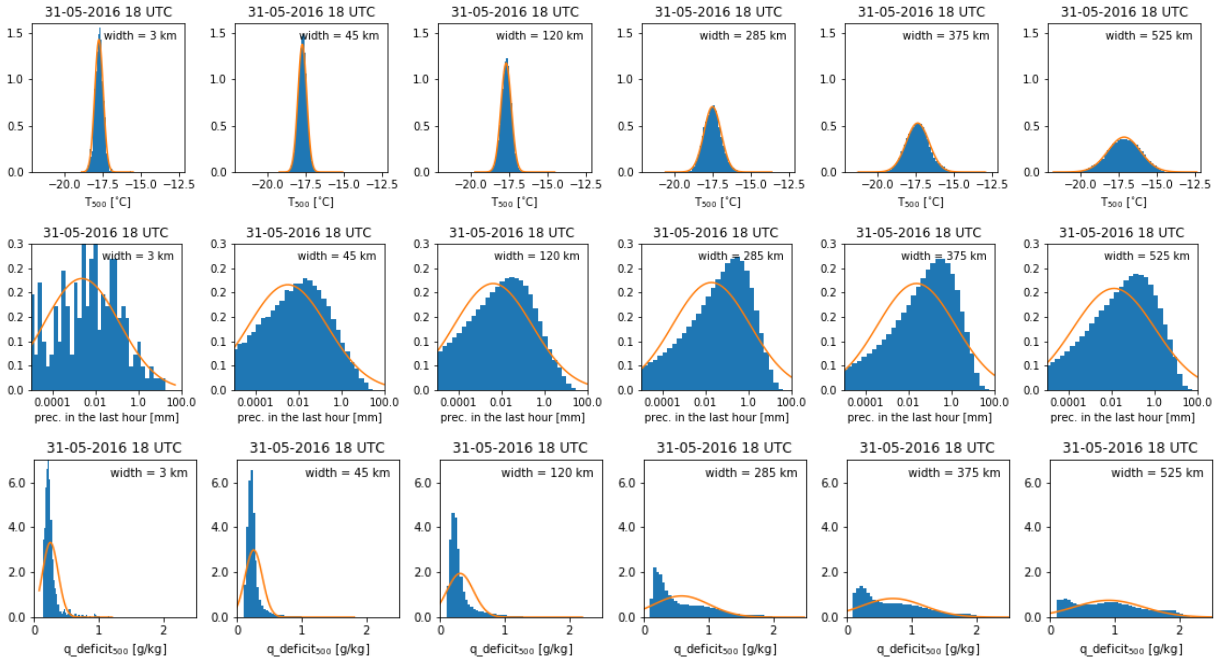


Figure 3.3: Histograms of temperature at 500 hPa (first row), precipitation (second row) and specific saturation deficit at 500 hPa (third row) on 31 May 2016 at 18 UTC in southern Germany. The histograms are computed by aggregating data. Sub-domain width increases from left to right (3, 45, 120, 285, 375 and 525 km). A Gaussian or log-normal distribution (for precipitation) with the same mean and standard deviation is shown for comparison.

3.2.2 Averaging

When averaging over small areas (up to around 200 km), distributions remain similar to the distribution for a single (central) gridpoint, because the realizations are not independent due to correlations in space. On the contrary to the aggregation approach, averaged distributions converge to a Gaussian over regions of more than around 200 km of width (figure 3.4). Sampled distributions are then independent, given the decrease in spacial correlations on longer distances, resulting in a Gaussian distribution according to the Central Limit Theorem, which says that the sum of a large number of independent random variables is approximately distributed as a Gaussian (Bonamente, 2017).
If the distribution has long tails (e.g. precipitation) it does not converge to a Gaussian but rather to a highly skewed distribution, like the log-normal distribution.

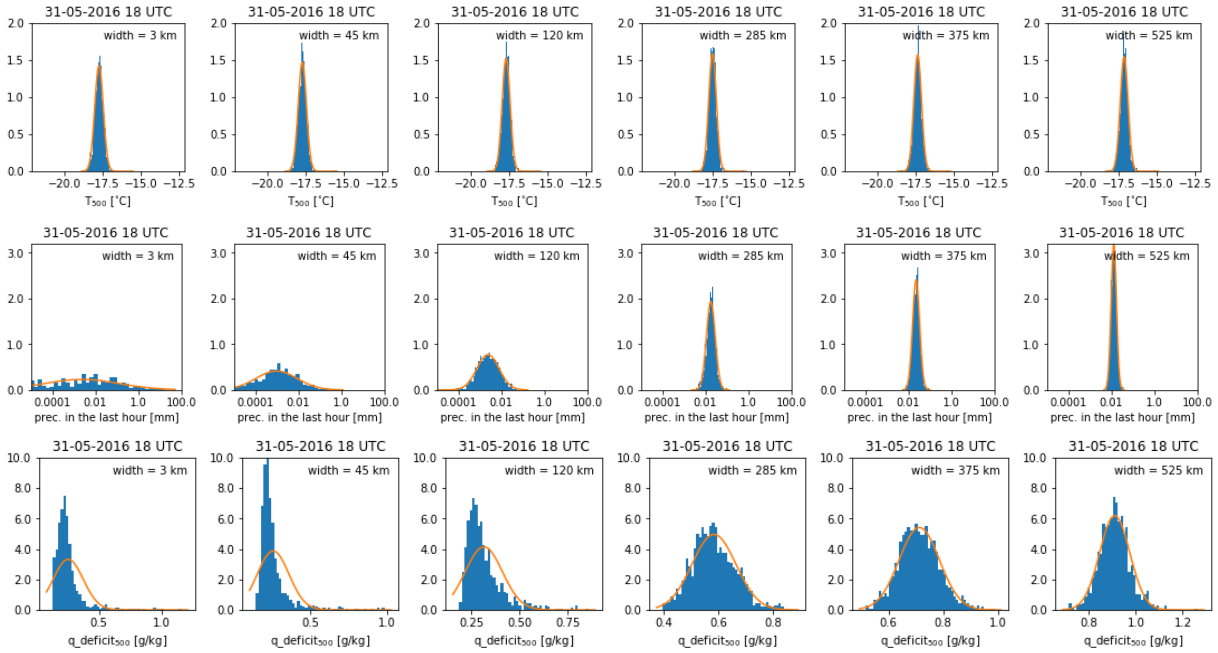
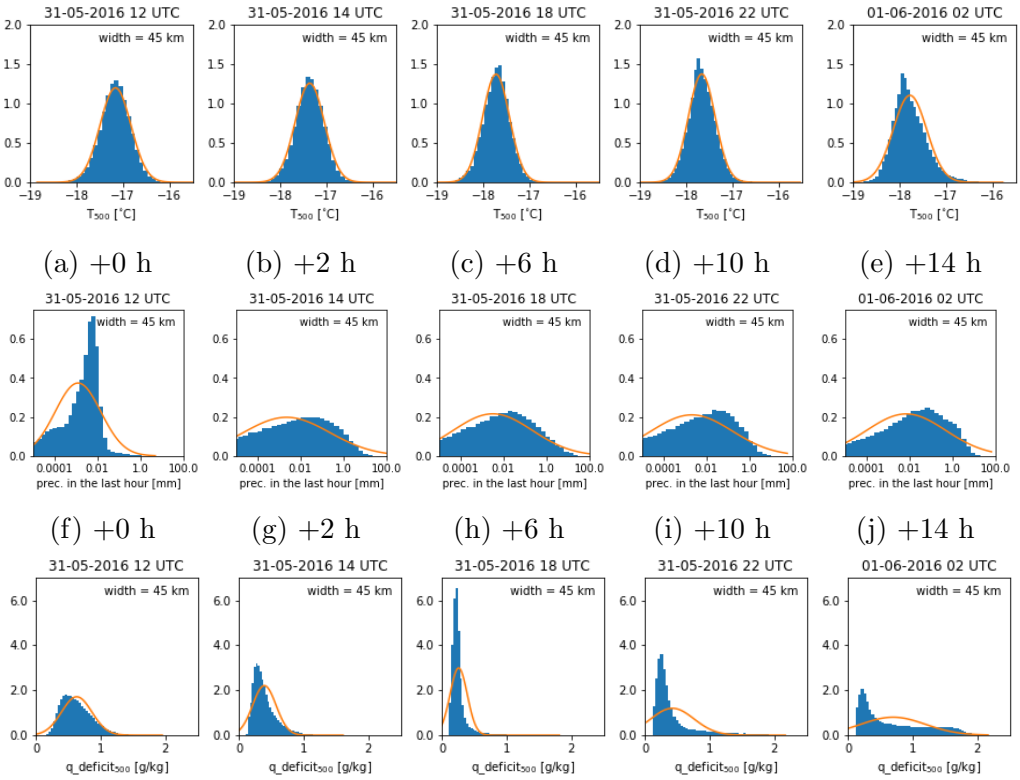


Figure 3.4: Histograms of temperature at 500 hPa (first row), precipitation (second row) and specific saturation deficit at 500 hPa (third row) on 31 May 2016 at 18 UTC in southern Germany. The histograms are computed by averaging over the sub-domain. Subdomain width increases from left to right (3, 45, 120, 285, 375 and 525 km). A Gaussian or log-normal distribution (for precipitation) with the same mean and standard deviation is shown for comparison.

3.2.3 Time evolution

The time evolution of distributions is compared with the conceptual model presented in the introduction (figure 1.2). Distributions are inspected on a 45km-wide subdomain. Figure 3.5 shows that SSD has the most variable distribution, which follows quite well the conceptual model. It starts close to Gaussian because of Data Assimilation, then it gets thinner and positively skewed and towards the end of the forecast period a second peak appears, when the distribution also gets broader. However, the last stage of the conceptual model is very likely not reached with a lead time of 14h. The other inspected variables' distributions do not experience such a time evolution. Precipitation starts forming already at the first time step, but the spin-up time amounts to more than 2 hours, which explains the initial different shape of the distribution.



(a) +0 h (b) +2 h (c) +6 h (d) +10 h (e) +14 h

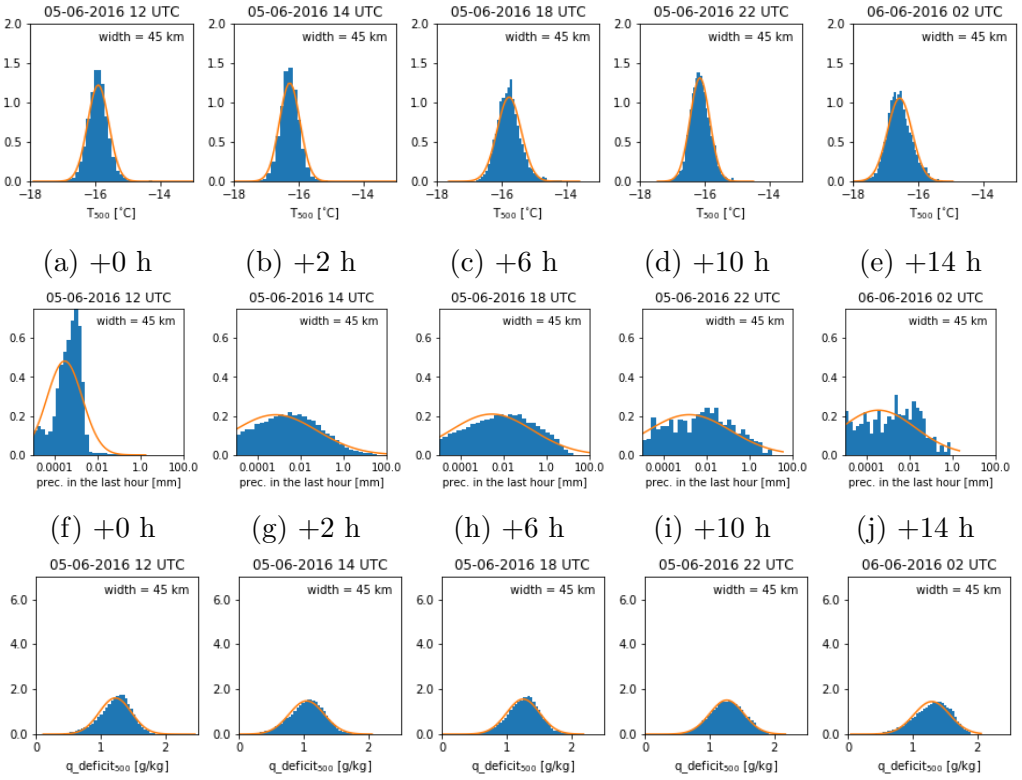
(f) +0 h (g) +2 h (h) +6 h (i) +10 h (j) +14 h

(k) +0 h (l) +2 h (m) +6 h (n) +10 h (o) +14 h

Figure 3.5: Histograms of temperature (top row, a-e), precipitation (middle row, f-j) and specific saturation deficit (bottom row, k-o) on 31 May 2016 in southern Germany, for a 45 km wide sub-domain. Forecast lead time increases from left to right from the first time step (12 UTC) to 14 hours (02 UTC, 1 June). A Gaussian or log-normal distribution (for precipitation) with the same mean and standard deviation is shown for comparison.

Some signals of the diurnal cycle might also be visible, but the studied forecast period covers only half of the day and starts at 12 UTC. The temperature distribution shape does

not change a lot with lead time, apart from getting slightly skewed and shifted to lower values, likely due to advection of colder air.
To briefly illustrate case dependency, results for 5 June 2016 are here presented as well. While 31 May is a synoptically forced convection case, with widespread large amounts of accumulated precipitation throughout the 14-hour forecast period, as shown in the ensemble mean accumulated precipitation map in figure 2.2, on 5 June, larger values only appear over orography (especially the Alps), where convection is triggered more successfully. Therefore, SSD distribution shape changes much less drastically both in time and space in this case, with more isolated convection. In fact, the PDF stays close to a Gaussian throughout the forecast period shown in figure 3.6 due to the prevalence of cloud-free gridpoints in the analyzed subdomain.
The precipitation PDF becomes noisier at later lead times due to a cessation of convective activity in the evening in many ensemble members.



(a) +0 h (b) +2 h (c) +6 h (d) +10 h (e) +14 h

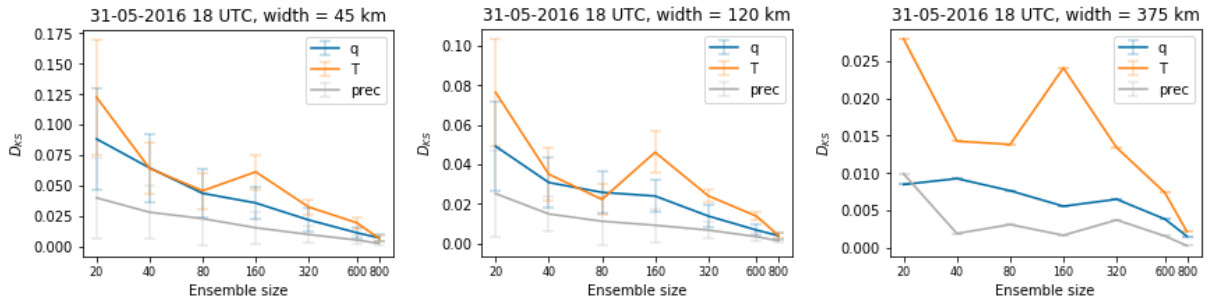
(f) +0 h (g) +2 h (h) +6 h (i) +10 h (j) +14 h

(k) +0 h (l) +2 h (m) +6 h (n) +10 h (o) +14 h

Figure 3.6: Histograms of temperature (top row, a-e), precipitation (middle row, f-j) and specific saturation deficit (bottom row, k-o) on 5 June 2016 in southern Germany, for a 45 km wide sub-domain. Forecast lead time increases from left to right from the first time step (12 UTC) to 14 hours (02 UTC, 6 June). A Gaussian or log-normal distribution (for precipitation) with the same mean and standard deviation is shown for comparison.

3.3 Sampling error and probability of rare events

An important question in ensemble prediction is how many members are needed for a realistic representation of forecast uncertainty. To address this question, samples of 20, 40, 80, 160, 320, 600 and 800 ensemble members are compared to the full, 1000-member sample (the truth). Precipitation seems to be the least dependent on ensemble size, according to the results of the Kolmogorov-Smirnov test of similarity between distributions (figure 3.7). The standard deviation within testing samples is indicated with errorbars (see section 2.3.3). For the values of all the considered variables' distributions to start converging towards the values for the truth, results suggest that 320 to 600 members are needed.
The dependency of sample mean and variance on ensemble size is investigated by comparing it for each sub-sample of smaller size to the moments of the 1000-member sample (figure 3.8). The temperature sample mean shows virtually no dependence on ensemble size, while the sample means of specific humidity and precipitation are more variable. The former is always lower for smaller samples, while the latter is always higher for the small and medium sub-regions of aggregation. The sample variance is more variable for specific humidity and temperature, without following a significant pattern with sample size.
The effect of ensemble size on extreme values is analyzed by comparing the 95^{th} percentile of distributions (figure 3.9). Precipitation and vertical velocity are the most variable with ensemble size and are correlated, since extremely strong updrafts are usually connected with strong convective cells that produce heavy precipitation. The 95^{th} percentile of specific humidity and temperature change less than 1% with smaller ensemble size, compared to the truth. It can be therefore concluded that the sample size mostly affects precipitation extreme events connected with strong convective storms, which was expected due to their intermittent nature and high spatio-temporal variability.



(a) (b) (c)

Figure 3.7: Statistic of the Kolmogorov-Smirnov two-sample test, where each sub-sampled ensemble of size 20, 40, 80, 160, 320, 600 and 800 is compared with the full 1000-member ensemble for specific humidity (q), temperature (T) and precipitation (prec.). Results are presented for three sub-domain widths: 45 km (a), 120 km (b) and 375 km (c), on 31 May 2016 at 18 UTC. The errorbars show the standard deviation of the tested samples of size 72 (a), 20 (b) and 1 (c).

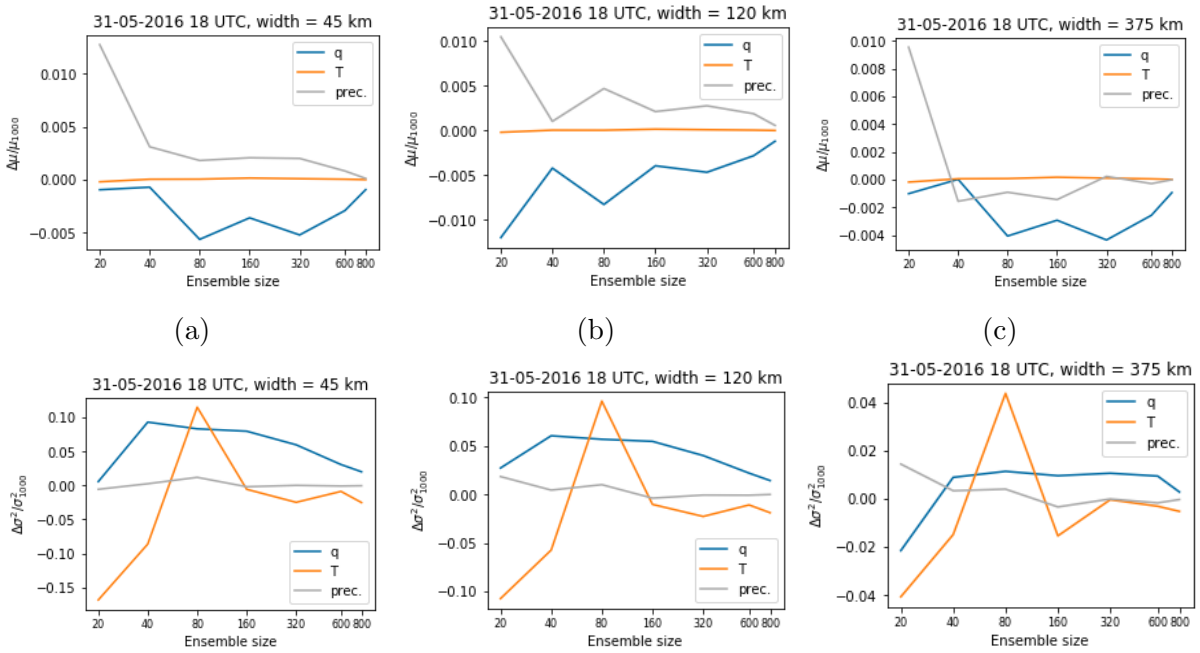


Figure 3.8: Relative deviation of the mean (first row) and of the variance (second row) from the moments of the 1000 member ensemble, for sub-sampled ensembles of size 20, 40, 80, 160, 320, 600 and 800, for specific humidity (q) and temperature (T). For precipitation (prec.), relative deviation from the moments of its logarithm are shown. Results are presented for three sub-domain widths: 45 km (a, d), 120 km (b, e) and 375 km (c, f), on 31 May 2016 at 18 UTC.

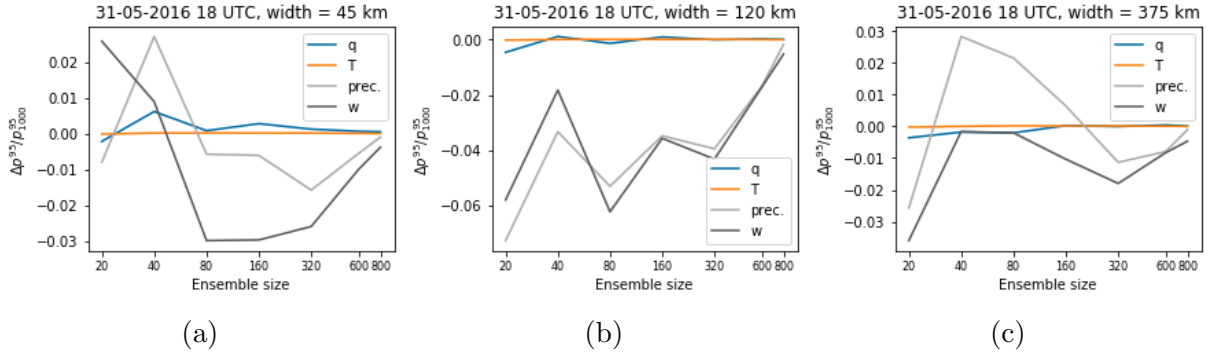


Figure 3.9: Relative deviation of the 95^{th} percentile from the value for the 1000-member ensemble, for sub-sampled ensembles of size 20, 40, 80, 160, 320, 600 and 800, for specific humidity (q), temperature (T), precipitation (prec.) and vertical velocity (w). Results are presented for three sub-domain widths: 45 km (a), 120 km (b) and 375 km (c), on 31 May 2016 at 18 UTC.

Chapter 4

Summary and conclusions

The steady increase in computational resources allows for convection-permitting simulations to be run operationally using high-resolution NWP models. Nevertheless, there are several challenges to overcome in smaller-scale weather prediction, from error growth to the accuracy of initial conditions. Ensemble prediction systems are very useful to investigate these uncertainties; however, due to limited computational resources, the number of members in operational ensembles in NWP is still not sufficient to accurately represent the nonlinear evolution of the PDFs of forecast variables.

A convective-scale 1000-member ensemble is used to study PDFs of forecast variables, in particular their shape, their spatial dependency and time evolution. The analyzed period between May and June 2016 was convectively very active in central Europe, offering a unique environment to study the effects of convection on PDFs of forecast variables. Previous studies (e.g. Miyoshi et al., 2014; Jacques and Zawadzki, 2015; Kawabata and Ueno, 2020) have shown that highly non-linear processes in convective clouds lead to stronger non-Gaussianity of PDFs, which is confirmed in this study.
Histograms of forecast variables are graphically inspected to categorize distributions in groups based on their shape and properties, which addresses the first of the posed scientific questions. Distributions are categorized in three groups: normally distributed, highly skewed and mixtures distributions. Temperature and wind at 500 hPa as well as pressure at sea level are normally distributed variables, since they are always close to a Gaussian. Precipitation and reflectivity, similar under certain conditions, are highly skewed variables, the former is mostly close to a log-normal distribution. Mixture distributions deviate the most from a Gaussian. Variables, whose distributions are mixtures, are all moisture-related in this study, mainly because of the cloudy/clear bimodality, but also because of physical constraints such as non-negativity of specific and relative humidity. These findings are in agreement with those of Legrand et al. (2016), who showed that specific humidity is the least Gaussian among the studied variables.
Scale dependency is investigated by aggregating and by averaging over subdomains of different width, from 3 up to 525 kilometers. A larger aggregation region reduces noise, which could be used to artificially increase ensemble size as a neighbourhood method. However, if

the region is too large (above around 200 km), the distribution becomes a mixture. When spatially averaged, distributions start converging to a Gaussian at about 200 km, the same scale as for the aggregation, according to the Central Limit Theorem of independent variables. In the case of precipitation the distributions converge to a skewed distribution close to a log-normal.
The time evolution of distributions is compared to the conceptual model presented in the introduction. Mixture variables follow the conceptual model in the first two stages, spatio-temporal error growth and unpredictable regime transitions, but the limited forecast lead time (14 hours) does not allow for the last stage to be reached. Other types of distributions are much less variable.
An estimate of sampling error is made with the use of sub-samples of the ensemble with a smaller ensemble size. Results suggest that at least 320 to 600 members are needed for the distributions to be robustly similar to the one given by the full ensemble. This is more than in Necker et al. (2020a), who showed that using 200 members instead of 40 substantially improves estimated covariances for convective-scale DA. In addition, the shape of the precipitation distribution seems to be the least dependent on ensemble size.
Finally, the dependence of the probability of rare, extreme events on ensemble size is investigated by computing the 95^{th} percentile of distributions. Precipitation and vertical velocity extremes are the most dependent on ensemble size, without a significant pattern, apart from the the two variables being correlated.

The dataset used in this work has a few limitations; data is available only hourly and on certain pressure levels, as well as only for some variables. Future research in this field should also consider larger ensembles to study sampling error more accurately. The vertical structure of PDFs of 3-dimensional forecast variables should be investigated, since all the analyses in this study are done only for a single pressure level. Distributions in the boundary layer are particularly interesting and complex. A longer forecast lead time has to be considered to study the evolution of forecast uncertainty and compare it to the later stages in the conceptual model of PDF evolution presented here.
A detailed understanding of the processes leading to error growth in weather forecasts should improve current prediction systems, so that they would represent the future state of the atmosphere and its uncertainty more accurately. A precise characterization of the properties of the distribution of uncertainty throughout the forecast period is a necessary contribution to this goal, and will ultimately provide a basis for evaluating other contributions to the uncertainty that are not represented in the current simulations.

Bibliography

Bachmann, K., Keil, C., Craig, G. C., Weissmann, M., and Welzbacher, C. A., Predictability of deep convection in idealized and operational forecasts: Effects of radar data assimilation, orography, and synoptic weather regime, *Monthly Weather Review*, 148 (1):63–81, 2020. doi: 10.1175/MWR-D-19-0045.1. URL https://doi.org/10.1175/MWR-D-19-0045.1.

Bannister, R. N., Migliorini, S., Rudd, A. C., and Baker, L. H., Methods of investigating forecast error sensitivity to ensemble size in a limited-area convection-permitting ensemble, *Geoscientific Model Development Discussions*, 2017:1–38, 2017. doi: 10.5194/gmd-2017-260. URL https://www.geosci-model-dev-discuss.net/gmd-2017-260/.

Beljaars, A. C. M. and Holtslag, A. A. M., Flux parameterization over land surfaces for atmospheric models, *Journal of Applied Meteorology*, 30(3):327–341, 1991. ISSN 0894-8763. doi: 10.1175/1520-0450(1991)030⟨0327:FPOLSF⟩2.0.CO;2. URL https://doi.org/10.1175/1520-0450(1991)030<0327:FPOLSF>2.0.CO;2.

Ben Bouallègue, Z., Theis, S. E., and Gebhardt, C., Enhancing cosmo-de ensemble forecasts by inexpensive techniques, *Meteorologische Zeitschrift*, 22(1):49–59, 2013. doi: 10.1127/0941-2948/2013/0374. URL http://dx.doi.org/10.1127/0941-2948/2013/0374.

Boer, G., Predictability as a function of scale, *Atmosphere-Ocean*, 41(3):203–215, 2003. doi: 10.3137/ao.410302. URL https://doi.org/10.3137/ao.410302.

Bonamente, M., *Statistics and Analysis of Scientific Data*, Springer New York, 2017. doi: 10.1007/978-1-4939-6572-4. URL https://doi.org/10.1007%2F978-1-4939-6572-4.

Buizza, R. and Leutbecher, M., The forecast skill horizon, *Quarterly Journal of the Royal Meteorological Society*, 141(693):3366–3382, 2015. doi: 10.1002/qj.2619. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2619.

Clark, A. J., Gallus Jr, W. A., Xue, M., and F., K., Growth of spread in convection-allowing and convection-parameterizing ensembles, *Weather and Forecasting*, 25:594–612, 2010. doi: 10.1175/2009WAF2222318.1. URL https://journals.ametsoc.org/doi/full/10.1175/2009WAF2222318.1.

DelSole, T., Trenary, L., and Tippett, M. K., The weighted-average lagged ensemble, *Journal of Advances in Modeling Earth Systems*, 9(7):2739–2752, 2017. doi: 10.1002/2017MS001128. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017MS001128.

Evensen, G. and van Leeuwen, P. J., An ensemble kalman smoother for nonlinear dynamics, *Monthly Weather Review*, 128(6):1852–1867, 2000. doi: 10.1175/1520-0493(2000)128⟨1852:AEKSFN⟩2.0.CO;2. URL https://doi.org/10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2.

Hagelin, S., Son, J., Swinbank, R., McCabe, A., Roberts, N., and Tennant, W., The met office convective-scale ensemble, mogreps-uk, *Quarterly Journal of the Royal Meteorological Society*, 143(708):2846–2861, 2017. doi: 10.1002/qj.3135. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3135.

Harnisch, F. and Keil, C., Initial Conditions for Convective-Scale Ensemble Forecasting Provided by Ensemble Data Assimilation, *Monthly Weather Review*, 143(5):1583–1600, 2015. ISSN 0027-0644. doi: 10.1175/MWR-D-14-00209.1. URL https://doi.org/10.1175/MWR-D-14-00209.1.

Hohenegger, C., Lüthi, D., and Schär, C., Predictability mysteries in cloud-resolving models, *Monthly Weather Reviews*, 134:2095–2107, 2006. doi: 10.1175/MWR3176.1. URL https://journals.ametsoc.org/doi/full/10.1175/MWR3176.1.

Hunt, B. R., Kostelich, E. J., and Szunyogh, I., Efficient data assimilation for spatiotemporal chaos: A local ensemble transform kalman filter, *Physica D: Nonlinear Phenomena*, 230(1):112–126, 2007. ISSN 0167-2789. URL http://www.sciencedirect.com/science/article/pii/S0167278906004647.

Jacques, D. and Zawadzki, I., The impacts of representing the correlation of errors in radar data assimilation. part ii: Model output as background estimates, *Monthly Weather Review*, 143(7):2637–2656, 2015. doi: 10.1175/MWR-D-14-00243.1. URL https://doi.org/10.1175/MWR-D-14-00243.1.

Kawabata, T. and Ueno, G., Non-gaussian probability densities of convection initiation and development investigated using a particle filter with a storm-scale numerical weather prediction model, *Monthly Weather Review*, 148(1):3–20, 2020. doi: 10.1175/MWR-D-18-0367.1. URL https://journals.ametsoc.org/doi/abs/10.1175/MWR-D-18-0367.1?mobileUi=0.

Kedem, B. and Chiu, L., On the lognormality of rain rate, *Proceedings of the National Academy of Sciences of the United States of America*, 84:901–5, 1987. doi: 10.1073/pnas.84.4.901.

Keil, C., Baur, F., Bachmann, K., Rasp, S., Schneider, L., and Barthlott, C., Relative contribution of soil moisture, boundary-layer and microphysical perturbations

on convective predictability in different weather regimes, *Quarterly Journal of the Royal Meteorological Society*, 145(724):3102–3115, 2019. doi: 10.1002/qj.3607. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3607.

Kondo, K. and Miyoshi, T., Non-gaussian statistics in global atmospheric dynamics: a study with a 10 240-member ensemble kalman filter using an intermediate atmospheric general circulation model, *Nonlinear Processes in Geophysics*, 26(3):211–225, 2019. doi: 10.5194/npg-26-211-2019. URL https://www.nonlin-processes-geophys.net/26/211/2019/.

Legrand, R., Michel, Y., and Montmerle, T., Diagnosing non-gaussianity of forecast and analysis errors in a convective-scale model, *Nonlinear Processes in Geophysics*, 23(1):1–12, 2016. doi: 10.5194/npg-23-1-2016. URL https://npg.copernicus.org/articles/23/1/2016/.

Leutbecher, M., Ensemble size: How suboptimal is less than infinity?, *Quarterly Journal of the Royal Meteorological Society*, 145(S1):107–128, 2019. doi: 10.1002/qj.3387. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3387.

Lien, G.-Y., Miyoshi, T., Nishizawa, S., Yoshida, R., Yashiro, H., Adachi, S. A., Yamaura, T., and Tomita, H., The near-real-time scale-letkf system: A case of the september 2015 kanto-tohoku heavy rainfall, *SOLA*, 13:1–6, 2017. doi: 10.2151/sola.2017-001.

Lorenz, E. N., The predictability of a flow which possesses many scales of motion, *Tellus*, 21:289–307, 1969. doi: 10.1111/j.2153-3490.1969.tb00444.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2153-3490.1969.tb00444.x.

Miyoshi, T., Kondo, K., and Imamura, T., The 10,240-member ensemble kalman filtering with an intermediate agcm, *Geophysical Research Letters*, 41(14):5264–5271, 2014. doi: 10.1002/2014GL060863. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014GL060863.

Nakanishi, M. and Niino, H., An improved mellor–yamada level-3 model with condensation physics: Its design and verification, *Boundary-Layer Meteorology*, 112(1):1–31, 2004. ISSN 1573-1472. doi: 10.1023/B:BOUN.0000020164.04146.98. URL https://doi.org/10.1023/B:BOUN.0000020164.04146.98.

Necker, T., Geiss, S., Weissmann, M., Ruiz, J., Miyoshi, T., and Lien, G.-Y., A convective-scale 1,000-member ensemble simulation and potential applications, *Quarterly Journal of the Royal Meteorological Society*, 146(728):1423–1442, 2020a. doi: 10.1002/qj.3744. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3744.

Necker, T., Weissmann, M., Ruckstuhl, Y., Anderson, J., and Miyoshi, T., Sampling Error Correction Evaluated Using a Convective-Scale 1000-Member Ensemble, *Monthly Weather Review*, 148(3):1229–1249, 2020b. ISSN 0027-0644. doi: 10.1175/MWR-D-19-0154.1. URL https://doi.org/10.1175/MWR-D-19-0154.1.

Nishizawa, S. and Kitamura, Y., A surface flux scheme based on the monin-obukhov similarity for finite volume models, *Journal of Advances in Modeling Earth Systems*, 10(12):3159–3175, 2018. doi: 10.1029/2018MS001534. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001534.

Nishizawa, S., Yashiro, H., Sato, Y., Miyamoto, Y., and Tomita, H., Influence of grid aspect ratio on planetary boundary layer turbulence in large-eddy simulations, *Geoscientific Model Development*, 8(10):3393–3419, 2015. doi: 10.5194/gmd-8-3393-2015. URL https://gmd.copernicus.org/articles/8/3393/2015/.

Piper, D., Kunz, M., Ehmele, F., Mohr, S., Mühr, B., Kron, A., and Daniell, J., Exceptional sequence of severe thunderstorms and related flash floods in may and june 2016 in germany – part 1: Meteorological background, *Natural Hazards and Earth System Sciences*, 16(12):2835–2850, 2016. doi: 10.5194/nhess-16-2835-2016. URL https://www.nat-hazards-earth-syst-sci.net/16/2835/2016/.

Porson, A. N., Carr, J. M., Hagelin, S., Darvell, R., North, R., Walters, D., Mylne, K. R., Mittermaier, M. P., Willington, S., and Macpherson, B., Recent upgrades to the met office convective-scale ensemble: An hourly time-lagged 5-day ensemble, *Quarterly Journal of the Royal Meteorological Society*, n/a(n/a), 2020. doi: 10.1002/qj.3844. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3844.

Raynaud, L. and Bouttier, F., The impact of horizontal resolution and ensemble size for convective-scale probabilistic forecasts, *Quarterly Journal of the Royal Meteorological Society*, 143(709):3037–3047, 2017. doi: 10.1002/qj.3159. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3159.

Sato, Y., Nishizawa, S., Yashiro, H., Miyamoto, Y., Kajikawa, Y., and Tomita, H., Impacts of cloud microphysics on trade wind cumulus: which cloud microphysics processes contribute to the diversity in a large eddy simulation?, *Progress in Earth and Planetary Science*, 2(1):23, 2015. ISSN 2197-4284. doi: 10.1186/s40645-015-0053-6. URL https://doi.org/10.1186/s40645-015-0053-6.

Scheck, L., Weissmann, M., and Bach, L., Assimilating visible satellite images for convective-scale numerical weather prediction: A case-study, *Quarterly Journal of the Royal Meteorological Society*, n/a(n/a), 2020. doi: 10.1002/qj.3840. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3840.

Sekiguchi, M. and Nakajima, T., A k-distribution-based radiation code and its computational optimization for an atmospheric general circulation model, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 109(17):2779–2793, 2008. ISSN 0022-4073. URL http://www.sciencedirect.com/science/article/pii/S0022407308001635.

Stensrud, D. J., Xue, M., Wicker, L. J., Kelleher, K. E., Foster, M. P., Schaefer, J. T., Schneider, R. S., Benjamin, S. G., Weygandt, S. S., Ferree, J. T., and Tuell,

J. P., Convective-scale warn-on-forecast system, *Bulletin of the American Meteorological Society*, 90(10):1487–1500, 2009. doi: 10.1175/2009BAMS2795.1. URL https://doi.org/10.1175/2009BAMS2795.1.

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

Tomita, H., New microphysical schemes with five and six categories by diagnostic generation of cloud ice, *Journal of the Meteorological Society of Japan. Ser. II*, 86A:121–142, 2008. doi: 10.2151/jmsj.86A.121.

Toth, Z. and Buizza, R. Chapter 2 - weather forecasting: What sets the forecast skill horizon? In Robertson, A. W. and Vitart, F., editors, *Sub-Seasonal to Seasonal Prediction*, pages 17 – 45. Elsevier, 2019. ISBN 978-0-12-811714-9. doi: https://doi.org/10.1016/B978-0-12-811714-9.00002-4. URL http://www.sciencedirect.com/science/article/pii/B9780128117149000024.

Zhang, F., Dynamics and structure of mesoscale error covariance of a winter cyclone estimated through short-range ensemble forecasts, *Monthly Weather Review*, 133(10):2876–2893, 2005. doi: 10.1175/MWR3009.1. URL https://doi.org/10.1175/MWR3009.1.

Acknowledgments

I would first like to thank my supervisors George Craig and Christian Keil for introducing me to the exciting field of forecast uncertainty and for providing guidance and feedback throughout this project.
Further, I am very grateful to Tobias Necker for his help with the data and all the valuable comments and suggestions. I would also like to thank Kirsten Tempest, Lucie Chabert and Jonas Späth for the stimulating discussions and their support.
Finally, a sincere *grazie/hvala* to my parents, who supported me throughout my studies and were there both in bad and good moments, even though not in person. This thesis is (in part) the result of your unconditional support.

The 1000-member ensemble simulation used computational resources of the K computer in Kobe, Japan provided by the RIKEN Center for Computational Science through the HPCI System Research project (Project ID:ra000015,ra001011).

Declaration *Erklärung*

I hereby declare that this thesis is my own work, and that I have not used any sources and aids other than those stated in the thesis.

Hiermit erkläre ich, die vorliegende Arbeit selbständig verfasst zu haben und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel benutzt zu haben.

München, 9. September 2020

Matjaž Puh