universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Objective Blending of Ensemble Information for Probabilistic Precipitation Nowcasts

verfasst von | submitted by

Simon Anton Köhldorfer BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 614
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Meteorologie
Betreut von | Supervisor:	Univ.-Prof. Mag. Dr. Martin Weißmann
Mitbetreut von | Co-Supervisor:	Dr. Tobias Necker M.Sc.

Abstract

Numerical weather prediction models have limited skill in the first few forecast hours. Therefore simple nowcasting methods, like the linear extrapolation of current radar information, are often used to predict the very near future more precisely. However, these extrapolation methods tend to rapidly decrease in skill after a few hours. Many efforts have been undertaken to combine recent observations or extrapolations of the latest states with numerical weather prediction (NWP) model computations. In order to achieve such a blending of current information and NWP forecasts, a data assimilation method is used in this thesis. The Local Ensemble Transform Kalman Filter (LETKF) is exploited to re-weight pre-existing ensemble forecasts according to recent observations, thus creating an adapted forecast, or rather, a nowcast. This LETKF-based nowcasting method is performed on the probability of hourly precipitation rates exceeding a threshold. The use of probabilities rather than precipitation intensities avoids unphysical jumps in the nowcasts.

For a period of six days in July 2020, this LETKF-based nowcasting is performed on GeoSphere Austria's 17-member C-LAEF ensemble. INCA analyses, which consist of a combination of recent radar observations and surface precipitation measurements, are used as observations. The nowcasts are, on average, able to outperform the underlying ensemble forecast for lead times of one to three hours, depending on the setting. For lead times of four hours and longer, however, the method in the presented setting can not outperform the ensemble forecast on average. Propagation of precipitation patterns outside of the localisation scale is found to be the main reason for decreasing nowcasting performance with increasing lead time.

The method is tested for three different precipitation rates and performs better for lower intensities. Sensitivity studies are carried out to assess the influence of both the spatial resolution of the probability field and observation-error covariances. While the exact setting of observation errors does not affect the outcome significantly, results are better for coarser spatial resolutions.

Kurzfassung

Wettervorhersagemodelle schneiden für die ersten Stunden des Vorhersagezeitraums oft nicht besonders gut ab. Deswegen werden simple Nowcasting-Methoden, wie die lineare Extrapolation von aktuellen Radarbildern zur Niederschlagsvorhersage, oft verwendet, um die nächsten Stunden besser vorhersagen zu können. Der Nachteil von solchen Kürzestfristvorhersagen ist wiederum, dass deren Vorhersaqualität innerhalb weniger Stunden drastisch sinkt. Deswegen werden immer wieder Bemühungen unternommen, aktuelle Beobachtungen oder die Extrapolation vom letztbekannten Zustand, mit den Berechnungen von numerischen Wettervorhersagemodellen zu vereinen.
Um solch eine Kombination von aktuellen Informationen mit numerischen Vorhersagen zu bewerkstelligen, wird in der zugrundeliegenden Arbeit eine Datenassimiliationsmethode verwendet. Der Local Ensemble Transform Kalman Filter (LETKF) wird benützt, um bereits existierende Ensemble-Vorhersagen anhand aktuellster Beobachtungen neu zu gewichten und dadurch eine adaptierte Vorhersage, oder besser gesagt einen Nowcast, zu berechnen. Die LETKF-basierte Nowcasting-Methode wird auf die Wahrscheinlichkeit für stündliche Niederschlagsraten größer gleich zu wählender Schwellwerte angewandt. Die Benützung von Wahrscheinlichkeiten statt Niederschlagsintensitäten vermeidet unphysikalische Sprünge in den Nowcasts.
Die Nowcasting-Methode wird über einen Zeitraum von sechs Tagen im Juli 2020 für das 17 Member umfassende C-LAEF Ensemble von GeoSphere Austria durchgeführt. Als Beobachtungen werden INCA Analysen herangezogen, welche eine Kombination aus Radarbeobachtungen und Niederschlagsmessungen am Boden sind. Je nach Wahl der Parameter kann die Nowcasting-Methode für ein- bis dreistündige Vorhersagen im Durchschnitt besser abschneiden als das zugrundeliegende Ensemble. Für Vorhersagen ab vier Stunden in die Zukunft bietet die Methode jedoch im Durchschnitt keine Verbesserung gegenüber der Referenz mehr, da sich die Niederschlagsfelder in diesem Zeitraum aus dem Lokalisierungsradius herausbewegen.
Die Methode wird für drei verschiedene Niederschlagsintensitäten durchgeführt und schneidet bei geringeren Intensitäten besser ab. Sensitivitätsstudien werden durchgeführt um den Einfluss der räumlichen Auflösung der Wahrscheinlichkeiten sowie verschiedener Beobachtungsfehlerkovarianzmatrizen festzustellen. Während die Einstellung der Fehlerkovarianzen keinen signifikanten Einfluss auf die Vorhersagegüte hat, werden mit gröber aufgelösten Wahrscheinlichkeitsfeldern bessere Resultate erzielt.

Contents

Abstract i

Kurzfassung iii

1. Introduction and research questions 1

2. Theoretical background 3
2.1. Data assimilation . . . 3
2.2. Ensemble forecasts . . . 4
2.3. The LETKF . . . 5
2.3.1. Step-by-step computation of the LETKF . . . 5
2.3.2. Local analyses and transformation . . . 7
2.3.3. The LETKF as a way around the main drawback of ensemble forecasts 8
2.4. Nowcasting methods . . . 8
2.5. State of the art in blending nowcasts and forecasts . . . 10

3. Datasets 13
3.1. Ensemble data . . . 13
3.2. Observation data . . . 14

4. Nowcasting Model 19
4.1. Creating upscaled probabilities from forecasted precipitation rates . . . 20
4.2. Comparing forecasts and observations . . . 21
4.3. Specific setup and tuning choices of the nowcasting model . . . 23
4.3.1. Resolution of upscaled probabilities . . . 23
4.3.2. Threshold of precipitation rates . . . 23
4.3.3. Localisation scale in the LETKF computation . . . 23
4.3.4. Assigning observation errors . . . 24
4.3.5. Symmetric square root . . . 24
4.3.6. Negative probabilities . . . 25
4.3.7. Using the most recent forecast datasets . . . 25
4.4. Verification method . . . 25

5. Results 27
5.1. Case studies . . . 27
5.1.1. 2h nowcast for the early afternoon on July 21st . . . 27
5.1.2. Second example: best 1h nowcast for strong precipitation . . . 28
5.1.3. Third example: worst 3h nowcast . . . 30

5.2. Statistical evaluation for all nowcasts with default setting 31
5.2.1. Correlation of spatial extent and nowcast skill 33
5.2.2. Dependence of RMSE on correct negatives 33
5.3. Sensitivity studies 35
5.3.1. Tuning the error-covariance matrix 35
5.3.2. Increased neighbourhood size 36

6. Conclusions **39**

7. Acknowledgements **41**

Bibliography **43**

List of Figures **47**

A. Appendix **49**

1. Introduction and research questions

NWP models provide predictions of the three-dimensional state of the atmosphere on a discrete grid and have improved greatly over the last few decades (see, for example [Magnusson and Källén, 2013] for the skill-evolution of the ECMWF model). Some situations, however, are still difficult to predict precisely. Extreme precipitation events prove to be particularly difficult for NWP models because their small-scale nature sometimes reaches the limit of what can be resolved on the grid. Furthermore, convection usually plays a role and poses a challenge due to its stochastic nature.
Therefore, even in modern weather forecasting, forecasters still often rely on nowcasting methods. This means observing the current and recent state of the atmosphere through measurements and extrapolating it into the future with more or less sophisticated methods. For short lead times this proves useful. The time-range where nowcasts tend to outperform NWP forecasts is often referred to as the nowcasting range and is said to be up to six hours.
NWP models do not start their computations from recent observations but incorporate observations continuously through a cycle of updating the model state and re-initializing new computations. The crucial process of accounting for observations in an optimal way is called data assimilation.
Ensemble products have gained more and more popularity among forecasters in the last few decades. Their major advantage is that they allow the forecaster to estimate the uncertainty of a forecast. Therefore they are a powerful tool for probabilistic forecasts. There are several different data assimilation methods to assimilate observations into an ensemble of forecasts efficiently while keeping desired properties like the spread of the ensemble members around the ensemble mean.
The main drawback of numerical weather prediction forecasts and even more so of ensemble forecasts is their long computation time. Their computations usually take a few hours (according to [Wastl et al., 2021] 4 hours for the C-LAEF ensemble). This means that no observations in that period can possibly be accounted for. Therefore there have been widespread efforts to join the advantages of nowcasts (using recent observations) and traditional NWP forecasts (full physical spatio-temporal evolution of the state of the atmosphere) into a blended forecast. However, blending two different forecasts into one may lead to unphysical fields. Another problem is that the blended forecast may end up as a too smooth field. The aim to avoid these problems while achieving a blended forecast is often referred to as „seamless blending“ or „seamless nowcasting“.
This thesis presents a novel approach that tries to achieve just that: combining an ensemble of forecasts with recent observations by exploiting an ensemble data assimilation method in a computationally cheap and therefore fast way. The underlying approach uses the Local Ensemble Transform Kalman Filter (LETKF) to re-weight the members of a

forecast ensemble for the probability of precipitation. This approach could be described as offline data assimilation, as the update step is performed but no new NWP forecasts are initialised. The hope is to bring the advantages of nowcasting into the ensemble forecast and thus improve it. In order to understand if and in what situations the nowcasting method is useful, the following questions are going to be answered:

1. Can an LETKF-based nowcast model for the probability of precipitation outperform the underlying ensemble forecast?
2. In which cases and for what lead times does the method perform well / not so well?
3. Does the rain rate or spatial extent of rain affect nowcasting performance?

The following Section 2 contains an oversight of the theoretical background of some core properties of data assimilation, ensemble forecasts, blending methods, and nowcasts. Furthermore, the LETKF and its features are discussed as it is the base for the underlying nowcasting method. Section 3 provides an overview of the datasets that were used for this thesis. In Section 4, the exact formulation and setting of a novel objective nowcasting model is explained and the verification method for both nowcast and ensemble forecast is discussed. Sections 5 and 6 contain the results of the nowcasting model and conclusions that can be drawn from them.

2. Theoretical background

Weather forecast models generally contain an estimate for the physical state of the atmosphere that is transferred through time by numerically solving physical equations such as the 3-dimensional motion equations, the conservation equations of mass and humidity, and the ideal gas equation directly. Solving all these equations requires a large amount of computation time, not only because millions of operations are involved, but also because some of them require iterative calculations to approximate their solution. In order to keep the computation time reasonable, all the equations are solved on a discrete 3-dimensional grid that is drawn through the atmosphere. The distance between those grid points (the model resolution) is usually roughly $10km$ for global weather models and for some local weather models (limited area models) often of the order of little more than $1km$. Some processes can not be represented on this scale because they happen on a smaller scale. The scale of cloud condensation processes, for example, is of the order of μm. Furthermore, some dynamic processes, like turbulence, can not be computed explicitly on the scale of km. Therefore it is necessary to introduce parametrisations to approximate the net effects of such physical processes. Parametrisations also serve to introduce empiric relationships for processes that are not precisely understood yet.
Parametrisations as approximations for real processes are one of the reasons why forecasts deviate from the later true state of the atmosphere. Other major error sources are uncertainties in the initial conditions of the state of the system atmosphere, the boundary conditions in Limited Area Models, as well as the chaotic nature of atmospheric processes. In order to minimise these errors, weather forecast models are continuously compared to and updated with recent observations, i. e. measurements that represent the true state of the atmosphere. The process of this update from a model estimate with observations is called data assimilation and it is key to the method in the underlying thesis.

2.1. Data assimilation

Data assimilation is the process of assimilating observations into an operational weather-forecast model. This is done in the so-called data assimilation cycle: Previous forecasts serve as an a-priori background estimate which is updated with observations. From the resulting analysis the next forecast is initialised and used as the next background estimate during the next data assimilation step. In operational weather models there are a few cycles (i. e. the model is re-run a few times) per day.
Rather than assuming observations as an undisputable truth, an uncertainty (error) is assigned to them. The observations and the background (with its own uncertainty) are

combined in an optimal way, that aims to minimize the error of the analysis. Mathematically, this coincides with minimizing a cost function.
One widely used approach in data assimilation is the so-called Kalman Filter ([Kalman, 1960]) which assumes Gaussian distributed variables and aims to minimize a 2-dimensional cost function, essentially by weighting forecast and observation with their inverse errors. Many adaptations of the Kalman Filter have been widely used in data assimilation, for example the non-linear Kalman filter ([Julier and Uhlmann, 1997]), the extended Kalman filter ([Jazwinski, 1970]) or the Ensemble Kalman filter ([Evensen, 1994]).

2.2. Ensemble forecasts

In operational weather forecasting products have emerged that make it vastly easier for forecasters to estimate the uncertainty of a forecast: ensemble forecasts. Ensemble forecasting means running a set of multiple forecasts, rather than only a single deterministic forecast. Different forecast members vary from each other through either different initial conditions, or small differences in the model formulations, i.e. in the computations of the physical development or through different parametrisations. For Limited Area Models (LAM) that are nested in a model with a larger domain, different boundary conditions can also be used for different members of the ensemble. The different members are within the boundaries of uncertainty, therefore they are similarly reasonable and correct forecasts. This also addresses the uncertainty that arises in weather forecasts due to the chaotic nature of the atmosphere. Ensembles are limited in size because of their large computational cost. Therefore it is difficult to portray all sources of uncertainty in an ensemble and there are too few members compared to the degrees of freedom in the system.
The spread of forecast members around the ensemble mean is a good indicator of the certainty of the forecast: if the forecasts of different members lead to vastly different outcomes (big spread), then the forecasted physical development can be considered uncertain. If different members compute similar forecasts and if therefore the spread of different forecasts is small, the forecast ensemble holds a high certainty. Generally, one can obtain a probability of a meteorological event by dividing the number of ensemble members that predict said event by the total number of ensemble members. For continuous variables, this is in practice done by applying a threshold for a forecast variable (e.g. temperature $T \geq 30°C$, wind gusts $v \geq 20\frac{m}{s}$, or precipitation rate $p \geq 10mm/h$. Since ensemble forecasts too have to be updated with observations frequently, data assimilation is naturally also performed on ensemble forecasts. Most of the data assimilation methods for ensemble forecasts are rooted in the Ensemble Kalman Filter [Evensen, 1994]. The core of this thesis is based on one particular Ensemble Kalman Filter that was proposed by [Hunt et al., 2007]: The Local Ensemble Transform Kalman Filter.

2.3. The LETKF

The Local Ensemble Transform Kalman Filter was designed as a data assimilation method for ensembles that takes little computation time compared to other ensemble data assimilation methods while performing similarly well.

2.3.1. Step-by-step computation of the LETKF

Along with their introduction of the LETKF, [Hunt et al., 2007] described a way to compute the LETKF as efficiently as possible. This method was used in the underlying thesis to minimize computation time.
Let p be a forecast variable of interest. Every member of a forecast ensemble serves as an independent a-priori background estimate. Forecasted values for p for the n closest grid points surrounding a grid point of interest are stacked in the background state vector of the dimensions $n \times 1$:

$$x_i{}^b = \begin{pmatrix} p^{f,1} \\ p^{f,2} \\ p^{f,3} \\ . \\ . \\ p^{f,n} \end{pmatrix} \tag{2.1}$$

Where the subscript i denotes the number of the ensemble member, the first superscript f stands for forecast and the second superscript $1, ..., n$ denotes the identity of the different grid points. Element-wise averaging of the state vectors from all ensemble members returns the background mean vector

$$\overline{x^b} = \frac{x_1{}^b + x_2{}^b + ... + x_k{}^b}{k} \tag{2.2}$$

where k is the total number of ensemble members. Analogically, the observations for a corresponding number of m grid points are stacked to form the observation vector:

$$y^o = \begin{pmatrix} p^{o,1} \\ p^{o,2} \\ p^{o,3} \\ . \\ . \\ p^{o,m} \end{pmatrix} \tag{2.3}$$

Usually the observation grid and the model grid are different in number ($n \neq m$) and exact location of the grid points. In order to correctly compare and update the background states with the observations, the model equivalent of the observations is obtained by applying the observation operator H, which transforms the background state estimates

from model space to observation space:

$$y_i^b = H(x_i^b) \tag{2.4}$$

The y_i^b are the model equivalents of the observations, different for every member of the ensemble. If the observation grid is identical to the model grid, then H becomes the identity matrix I, and Equation (2.4) simplifies to:

$$y_i^b = I(x_i^b) \longrightarrow y_i^b = x_i^b \tag{2.5}$$

In this case, it follows that $\overline{y^b} = \overline{x^b}$, otherwise $\overline{y^b}$ is computed from the y_i^b as in Equation 2.2. In the following steps of the computation following [Hunt et al., 2007] the background perturbation matrix X^b is required. Its entries are the deviations of the different background state vectors from the background ensemble mean, $x_i^b - \overline{x^b}$. For the explicit example of a variable p, X^b is written as:

$$X^b = \begin{pmatrix} {p_1}^{f,1} - \overline{p^{f,1}} & {p_2}^{f,1} - \overline{p^{f,1}} & .... & {p_k}^{f,1} - \overline{p^{f,1}} \\ {p_1}^{f,2} - \overline{p^{f,2}} & p_2^{f,2} - \overline{p^{f,2}} & .... & {p_k}^{f,2} - \overline{p^{f,2}} \\ .... & .... & .... & .... \\ . & & & \\ . & & & \\ {p_1}^{f,n} - \overline{p^{f,n}} & p_2^{f,n} - \overline{p^{f,n}} & .... & p_k^{f,n} - \overline{p^{f,n}} \end{pmatrix} \tag{2.6}$$

Here, again the second superscript denotes the n different points in model space that are part of the state vectors and the subscript denotes the identity of the ensemble member. Analogically to X^b, a perturbation matrix for the state vectors in model space, Y^b is formed from the deviations of $y_i^b - \overline{y^b}$. In the special case in which the observation grid and the model grid are identical, then Y^b and X^b are also identical. As the next step, which paves the way to the computation of the analysis error covariance matrix $\tilde{P^a}$, [Hunt et al., 2007] suggested to compute a matrix

$$C = {Y^b}^T R^{-1} \tag{2.7}$$

Where R is the observation error covariance matrix with size $n \times n$. The entries on the diagonal of R represent estimates of the errors of the observations. Off-diagonal elements represent the correlations of different observation errors. The superscript T denotes the transpose of Y^b and the superscript -1 denotes the inverse of R. The analysis error covariance matrix, is then calculated via:

$$\tilde{P^a} = \left[I * (k-1) + CY^b \right]^{-1} \tag{2.8}$$

Where I is again the identity matrix, here of the size $k \times k$. Next, the matrix W^a is computed via

$$W^a = \left[(k-1)\tilde{P^a} \right]^{\frac{1}{2}} \tag{2.9}$$

where the power of $\frac{1}{2}$ denotes the symmetric square root of the analysis error covariance matrix $\tilde{P}_a$. In the following step, the weight vector for the perturbation matrix, $\overline{w^a}$ is computed:

$$\overline{w^a} = \tilde{P^a} C(y^o - \overline{y^b}) \tag{2.10}$$

Adding this vector $\overline{w^a}$ to every single column of W^a turns the columns of W^a into the analysis vectors w^{a^i}. Lastly, the analysis-ensemble members are computed via:

$$x^{a,i} = X^b \cdot w^{a^i} + \overline{x^b} \tag{2.11}$$

Thus from an ensemble of k different background state vectors, an updated ensemble of k different analysis state vectors is computed. The ensemble mean $\overline{x^a}$ can be computed as an arithmetic mean of the analysis ensemble members $x^{a,i}$ as in Equation 2.2. However, if desired, rather than computing Equation 2.11 and averaging the ensemble members, the ensemble mean can (according to [Schraff et al., 2016], for example) be computed directly via the equation:

$$\overline{x^a} = X^b \cdot \overline{w^a} + \overline{x^b} \tag{2.12}$$

2.3.2. Local analyses and transformation

Two of the key features of the LETKF are the local analyses and the transformation into the ensemble space. Both serve to reduce computation time and make the LETKF unique among other Ensemble Kalman Filters. Local analyses means that the computations for the analyses are carried out at each point in the domain independently, using the background states for the n nearest grid points along with the grid point of interest, and the corresponding m local observations. After computing the analyses locally for all grid points, a global analysis is obtained by simply putting the local analyses together. [Hunt et al., 2007] stated that the number of observations that are taken into account in each local analysis should normally be similar or a bit larger than the number of ensemble members. Using not too many observations is what reduces the dimensions of all the vectors and matrices in the Equations 2.1 to 2.10 and this strongly reduces computation time. Another advantage is that spurious correlations are avoided. Spurious correlations are correlations over large distances that are the result of oversampling in the data assimilation step, rather than real physical correlations. On the other hand, if there are indeed real large-distance correlations, they can not be accounted for by the LETKF algorithm because of the local analyses. [Hunt et al., 2007] argued that there should be a large overlap of the set of observations that is used for neighbouring analyses. Thus the weight vectors $\overline{w^a}$ do not vary too strongly for neighbouring grid points and unphysical jumps in the global analysis are avoided.
Another important feature of the LETKF is the transformation that is performed from model space to a lower-dimension ensemble subspace. The transformation is performed in Equation 2.7 by using the increments $y_i^b - \overline{y^b}$ in the matrix Y^b. The analysis error covariance matrix P^a and the weight vector $\overline{w^a}$ are only computed in this subspace that is spanned by the ensemble perturbations. This also reduces the computational resources

needed. In Equation 2.11 (alternatively Equation 2.12) the transformation back into model space is performed. There the weight vector is imposed on the perturbation from the ensemble members around their mean. In other words, the LETKF-analysis is a linear combination of the original ensemble members.
The analysis error covariance matrix $\tilde{P}^a$ as computed in Equation is estimated from the ensemble itself (along with the observation error covariance matrix), as in all Ensemble Kalman Filters (in literature $\tilde{P}^a$ is often denoted as B). This is a big advantage compared to 3D-Var and 4D-Var data assimilation methods, where only a climatological estimate of B is used.

2.3.3. The LETKF as a way around the main drawback of ensemble forecasts

Timeliness is one of the crucial limitations of ensemble forecasts. They are computationally very expensive. Since ensembles consist of multiple deterministic NWP model runs, they also take multiple times longer to compute than a single deterministic forecast. Therefore, ensemble forecasts are typically available only a few hours after initiation. On afternoons of convective activity, the development of a few hours can, however, influence the next few hours greatly through processes like vertical transport of moisture and temperature, outflows of thunderstorm cells leading to a convergence in the ground wind field and thus triggering other convective cells, etc. Such new developments can not be accounted for in an ensemble forecast, as they only happen after the computation of the NWP model was started. The convective developments of a few hours can already be significantly different in forecasts and observations due to the stochastic nature of convection. Also, model spin-up often reduces the skill of NWP forecasts in the first timesteps.
Since data assimilation provides a feasible way to adjust forecast information with recent observations, and since the LETKF provides a fast way to compute such an update step for an ensemble of forecasts, the idea arose to exploit the features of the LETKF in order to overcome the challenges that convective activity poses for ensemble forecasts. In this thesis, the LETKF is used to update ensemble forecasts with recent observations. These LETKF „analyses “ for future time steps can be considered as nowcasts.

2.4. Nowcasting methods

Nowcasting methods are inherently different from NWP models. They mostly ignore the physical processes behind the development of weather regimes and instead try to predict the short-term future by extrapolating from the 'now'-state of the atmosphere. Nowcasting is mostly performed for precipitation, storms, or lightning [Pierce et al., 2012]. The current state can, for example, be obtained through direct measurements. Those can either be ground-based observations, for example of precipitation rates or storm gusts, or remote sensing data (for example detection of lightning strokes). Another possibility is the use of indirect measurements, for example wind speed derived from Doppler-Radar measurements, or precipitation rates derived from radar reflectivities. The latter uses

the so-called $Z - R$ relationship between radar reflectivity and rainfall rate and is the foundation of the most widely used methods for precipitation nowcasting.
The simplest approaches to radar-based precipitation nowcasting use linear extrapolation. There are a few different ways to perform linear extrapolation, as there are different approaches to derive the current motion from radar patterns. One approach is to obtain the motion of the whole field of radar patterns, which can be done in different ways. Often used systems for this purpose are for example COTREC ([Li et al., 1995]) or SWIRLS ([Li et al., 2000]). Another way to perform linear extrapolation is cell tracking: convective cells can be defined as independent objects, and the center of each object can be identified in consecutive radar images and be individually extrapolated. This was implemented, for example, in the TITAN ([Dixon and Wiener, 1993]) and SCIT ([Johnson et al., 1998]) algorithm. All of these methods try to advect observed radar reflectivities, but can neither account for errors that result from the $Z - R$-relationship, measurement errors in the radar reflectivity (for example radar beam blockage), nor for the growth and decay of convective systems. Therefore the skill of linear extrapolation is limited in lead time. [Germann and Zawadzki, 2004] found an extrapolation-based nowcast with the Local Eulerian method to be useful for lead times up to 2.3 hours, and Local Langragian nowcasts remained useful for a maximum of 4.2 hours. Usefulness was defined as forecasts that scored above 0.5 in the Conditional square root of the Ranked probability score (CSRR). A more practical approach is to define usefulness as the time range where the extrapolation nowcasts perform better than an available NWP forecast. Since local weather models improve steadily, the usefulness of linear extrapolation nowcasts is further reduced compared to them. However they remain relevant in the nowcasting range.
There are other precipitation nowcasting methods that try to account for the growth and decay of precipitation patterns. One example is the Analogues method ([Panziera et al., 2011]) where past developments of cells are categorised, linked to predictors, and then currently observed precipitation patterns are compared with the predictors, assigned to the respective category and the typical development of the respective category is assumed for the current cell. In the study by [Panziera et al., 2011] nowcasts from different initialisation times were used to form a nowcasting ensemble, which is a way to account for the uncertainty of nowcasts. Besides ensembles, there are other nowcasting methods that delve into probabilistic nowcasting, like neighbourhood methods or methods that introduce random numbers in order to create stochastic nowcasts.
An entirely different nowcasting method is NWP-based nowcasting. Recent information is used to re-run an NWP model frequently. The drawback is that this still requires the computation time of an NWP forecast. Therefore the main benefit of nowcasting methods is not fully exploited. There are efforts to further increase the intervals of re-initialisation through Rapid Update Cycling (RUC). In Austria the AROME-RUC is under development to produce new forecasts every hour. It assimilates ground-based observations, satellite information, radar information and more. Forecasts are available within one hour. Next to timeliness, another advantage is that more observations (e.g radar observations) can be accounted for in a RUC, and therefore they perform well on short time scales. A drawback is that in each cycle the analysis starts from the background

NWP forecast and not from the previous RUC-forecast.
In recent years machine learning methods like Convolutional Neural Networks and Vision Transformer Networks have been proposed for nowcasting (and weather forecasting in general). Some applications have been judged to increase skill significantly while reducing computation time. A detailed discussion of Machine Learning methods in weather forecasting is outside the scope of this thesis. Noteworthy research has been conducted (among many other contributions) for example by [Ko et al., 2022], [Leinonen et al., 2022], [Ayzel et al., 2020], [Lee et al., 2021], [Agrawal et al., 2019],[Ravuri et al., 2021], [Piran et al., 2024] and [Pulkkinnen et al., 2019]. An overview of Deep Learning methods in weather forecasting can be found in [Gao et al., 2021].

2.5. State of the art in blending nowcasts and forecasts

Since nowcasting methods perform best for short lead times, but tend to be less skillful than NWP forecasts after a few hours, there have been many efforts to combine NWP forecasts and nowcasts into one blended forecast. The goal of most researchers in this field is to obtain a forecast that performs at least as well as the better of the two forecasts at all lead times. Simple averaging of two independent fields leads to physically unrealistic fields. The effort to avoid unphysical states and get a smooth transition from the nowcasting range to longer lead times is referred to as „seamless nowcasting “. Efforts in seamless nowcasting have for example been undertaken by [Kober et al., 2011]. The nowcasting model that was used is the probabilistic module of the RAD-tram precipitation nowcast, which uses optimum-flow for field-based advection extrapolation and a neighbourhood method to generate probabilities. This probabilistic nowcast was combined with probabilistic precipitation forecasts derived from the COSMO-DE Ensemble Prediction System (EPS). Different methods for obtaining the probabilities from the EPS were used and compared. Probabilities from each method were calibrated via the reliability diagram statistics method before use. The lead time dependent skill of the nowcast was assessed through a verification method. A weighting function (with respect to the lead time) was designed empirically for the nowcast to give it a weight proportional to its skill, decreasing with increasing lead time. The NWP forecasts were given a weight so that the weight of the nowcast and the forecast would sum to unity. The blended forecast was computed as a linear combination of both forecast components with their respective weight at every lead time. This blended forecast was found to perform better than both respective components at all lead times and for all different probabilistic NWP forecasts.
This method's success was confirmed in a follow-up study for a longer period ([Kober et al., 2013]). The method's dependency on different weather regimes was evaluated. The dataset was classified into three different weather regimes and the weighting function as well as the NWP model calibration were done separately for all regimes and found to be significantly different. It was concluded that regime-dependent calibration improved the blended nowcast, but that after regime-dependent calibration a regime-dependent weighting did not make a significant difference for forecast quality compared to the general weighting function. This method of skill-dependent weights for RAD-Tram was

also used by [Scheufele, 2010] who blended it with a time-lagged ensemble that consisted of consecutive deterministic COSMO-DE runs. The weighting functions were different, but the result that the blended forecast outperformed both components was the same.
[Poletti et al., 2019] combined information from the phase-stochastic (PhaST) nowcasting ensemble and the deterministic NWP model MOLOCH to use it as input for a hydrological model. This hydrological model with different precipitation forecasts was assessed for three different extreme flooding events. The authors used (1) the nowcasting ensemble alone, (2) the nowcasting ensemble with the trend for total rainfall volume of the NWP imposed on it, and (3) a linear blending of (2) and the NWP model. The linear blending was performed with four different weighting functions, all of which gave a weight of 100% to the nowcast in the first 2 hours, then rapidly decreasing the weight with increasing lead time. The blended nowcasts were found to mostly perform at least as well as the nowcasting ensemble, with a few exceptions around the exact time when the components were blended.
[Nerini et al., 2019] also used ensemble information to adapt and improve an extrapolation nowcast method. Nowcast and forecast were blended with an Ensemble Kalman filter approach in a reversed manner: The observation-based nowcasts with stochastic realisations were used as background and precipitation intensities from an ensemble of NWP forecasts were used as pseudo-observations in a data assimilation step. The results varied for different weather regimes but were overall judged to be successful because they reproduced the skill of the nowcasting method alone for short lead times and converged to the NWP's skill for longer lead times.
All these methods try to achieve seamless nowcasting by blending nowcasts with NWP forecasts (or forecast ensembles). These blending methods proved to be overall successful, however, they do not provide an entirely objective way of combining the different forecasts. [Kober et al., 2011] and [Scheufele, 2010] assessed the skill of the nowcast model in order to develop an empirical function tailored to the behavior of the lead time dependent skill that was observed. The issue of obtaining unphysical fields after blending was mitigated by blending probabilities instead of precipitation intensities. In order to replicate their method effectively for other NWP models and nowcasting methods, a lot of verification is necessary to find the appropriate weighting functions. If the weighting function is less involved, like in the work of [Poletti et al., 2019], then it is not clear which lead time decreasing weight to use for the nowcast, at least in their work.
The underlying thesis proposes a very different approach for seamless nowcasting and a truly objective way of blending. Instead of trying to unite a nowcast and a forecast, observations are incorporated into high-resolution ensemble forecasts in an „offline-data-assimilation step“, which means that the new analysis according to the LETKF is not used to re-run the NWP models, but only for the adapted forecast. Similar work has been done by [Raynaud et al., 2014], who used the Bayes-theorem for non-linear filtering to compute adapted weights for a time-lagged ensemble. The weights were based on recent observations of precipitation probability. However, their re-weighting was found to not gain significant skill compared to an equal-weight combination of all members of the time-lagged ensemble. The underlying thesis gives the idea of offline-data assimilation for

nowcasting another try with the LETKF as assimilation method.

3. Datasets

Both the ensemble forecast data and the observation data were thankfully provided by courtesy of Austria's national weather service GeoSphere Austria, formerly known as „Zentralanstalt für Meteorologie und Geodynamik “ (ZAMG).

3.1. Ensemble data

The ensemble forecast that was used for this thesis is the so-called C-LAEF ensemble. It is nested in ECMWF's IFS model and consists of one main run plus 16 perturbed members, thus forming an ensemble of 17 members. The ensemble addresses uncertainties in the following ways ([Wastl et al., 2021]):

- Uncertainty in the lateral boundary conditions is introduced by using a different member of the IFS ensemble as boundary condition for each C-LAEF member.
- Errors in the initial conditions are represented by using a different perturbation for observations for each ensemble member and running separate 3D-Var analyses.
- Uncertainty in the model physics is represented by a hybrid approach that combines a parameter perturbation scheme and a tendency perturbation approach ([Wastl et al., 2019]).

Datasets for total precipitation of all ensemble members were provided in hourly resolution ranging from July 20th 2020, 0 UTC until July 26th 2020, 12 UTC. Ensemble runs are (for operational purposes) initiated every 12 hours, at 0 and 12 UTC. C-LAEF data was provided on a postprocessing grid with a grid-point distance of $\Delta lon = 0.028$ ° in the x-direction and $\Delta lat = 0.018$ ° in the y-direction. This roughly corresponds with a resolution of $2km$ in both the longitudinal and meridional direction. The domain extends from 5.498 °E to 22.102 °E and from 42.981 °N to 51.819 °N.
Along with the C-LAEF ensemble the deterministic AROME-AUT forecast model was also provided on the same grid for the same time period in hourly resolution. AROME-AUT is initiated every three hours. Because the C-LAEF ensemble is strongly based on the AROME-model, the control-run of C-LAEF was replaced with AROME-AUT in this thesis. Thus the total ensemble that was used consists of the 16 perturbed C-LAEF members and the deterministic AROME-AUT run.

3.2. Observation data

As observations, GeoSphere Austria's INCA analysis ([Haiden et al., 2011]) was used. The INCA analysis is based on a combination of radar measurements by the civil aviation administration („Austrocontrol “) and ground-based rain-gauge precipitation measurements of GeoSphere's semi-automatic „TAWES “(see [Kann and Haiden, 2011]) weather stations. As of 2024, the TAWES-observation network consists of roughly 280 stations across Austria.
The INCA analyses for total precipitation were provided from July 20th 2020, 0 UTC until July 25th 2020, 23:45 UTC in 15-minute resolution. The last full hourly precipitation rates can be obtained for July 25th, 23 UTC. So in total, there are 143 sets of hourly precipitation rates in both forecast and observation.

Coordinate transformation of INCA grid-points

The INCA dataset was provided on a horizontal grid with a resolution of exactly $1km$ on a projection on the Bessel-1841-ellipsoid. As the points in x- and y-direction were given in meters east and north of certain reference points, with values from $[\tilde{x}] = [20000, 21000, ..., 720000]$ and $[\tilde{y}] = [220000, 221000, ..., 720000]$ they needed to be transformed to their respective values of longitude and latitude. This section follows a guide from a website by the New Zealand government on how to transform from the Lambert conical projection with the Bessel-ellipsoid to a regular longitude-latitude grid: https://www.linz.govt.nz/guidance/geodetic-system/understanding-coordinate-conversions/projection-conversions/lambert-conformal-conic-transformations (last accessed on August 23rd, 2024).
The following table gives an oversight of the constants needed and their values, many of which were provided by GeoSphere, along with the datasets.

Variable	Name of variable	Value
a	semi-major axis of the reference ellipsoid	6377397.155
F	Ellipsoidal flattening	1/299.152815
Φ_1	Latitude of first standard parallel	46°N
Φ_2	Latitude of second standard parallel	49°N
Φ_0	Origin Latitude	47.5°N
λ_0	Origin Longitude	13.3333°E
N_0	False northing	400000
E_0	False easting	400000

Table 3.1.: Parameters needed for the transformation from a Lambert conical grid to a regular lon-lat-grid.

From these basic constants, a few more have to be computed. Some are denoted differently than in the above-mentioned guide to avoid confusion with other variables within this

thesis.

$$e = \sqrt{2F - F^2} \tag{3.1}$$

Solving the function

$$M(\Phi) = \frac{cos(\Phi)}{\sqrt{1 - e^2 sin^2(\Phi)}} \tag{3.2}$$

for the value $\Phi = \Phi_1$ yields the constant M_1, and solving for $\Phi = \Phi_2$ yields M_2. Next, the function

$$u(\Phi) = \frac{tan(\frac{\pi}{4} - \frac{\Phi}{2})}{(\frac{1-esin(\Phi)}{1+esin(\Phi)})^{\frac{e}{2}}} \tag{3.3}$$

has to be solved for $\Phi = \Phi_0$,$\Phi = \Phi_1$ and $\Phi = \Phi_2$ in order to obtain the constants u_0, u_1 and u_2. Furthermore, these are needed:

$$z = \frac{ln(M_1) - ln(M_2)}{ln(u_1) - ln(u_2)} \tag{3.4}$$

$$F = \frac{M_1}{z(u_1)^z} \tag{3.5}$$

$$d_0 = aFu_0^z \tag{3.6}$$

Now for every pair of $\tilde{x}$ and $\tilde{y}$ the deviations from the reference points are taken:

$$N' = \tilde{y} - N_0 \tag{3.7}$$

$$E' = \tilde{x} - E_0 \tag{3.8}$$

and then used to compute

$$d' = +\sqrt{E'^2 + (d_0 - N')^2} \quad if \quad z > 0 \tag{3.9}$$

or

$$d' = -\sqrt{E'^2 + (d_0 - N')^2} \quad if \quad z < 0 \tag{3.10}$$

Then for the final computation of the latitude, two more variables are needed:

$$u' = (\frac{d'}{aF})^{1/z} \tag{3.11}$$

$$\gamma' = arctan(\frac{E'}{d_0 - N'}) \tag{3.12}$$

Now the longitude of every grid point can be computed via

$$\lambda = \lambda_0 + \frac{\gamma'}{z} \tag{3.13}$$

The latitude has to be computed iteratively. As a start, a first approximation is calculated

via:

$$\Phi = \frac{\pi}{2} - 2arctan(u') \tag{3.14}$$

This expression is plugged into the equation

$$\Phi = \frac{\pi}{2} - 2arctan\left[u'\left(\frac{1 - esin(\Phi)}{1 + esin(\Phi)}\right)^{\frac{e}{2}}\right] \tag{3.15}$$

Iteratively, the result on the left-hand side of Equation 3.15 is to be put in as Φ on the right-hand side, until the results do not change anymore. For this thesis, ten iterations of Equation 3.15 were carried out.
On a regular lon-lat grid, the INCA grid is not aligned rectangular. According to the above transformation throughout Equations 3.1 to 3.15 the southernmost and northernmost points are located at 45.7727 °N and 49.4789 °N respectively. In the longitudinal direction, the most western grid point is located at 8.0982 °E and the most eastern point at 17.7437 °E.

Nowcasting domain

The nowcasting domain is chosen to be rectangular on a regular lon-lat grid like the C-LAEF domain and to be fully inside the INCA domain. This ensures that full sets of observations are available at every analysis point. The nowcasting domain is further cut off on the outsides so the closest n grid points around the analysis points on the C-LAEF grid still have full sets of corresponding observations. Thus the special case of the edge of the domain (observations not available in all directions) does not have to be dealt with.

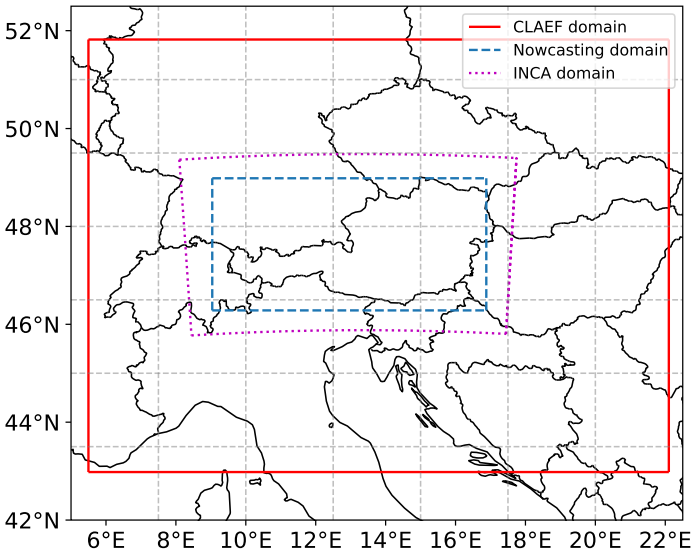


Figure 3.1.: Outlines of the C-LAEF domain that was provided, the INCA domain according to the transformation, and the nowcasting domain that was used in this thesis.

The resulting nowcasting domain spans from 8.774 °E to 17.146 °E and from 46.133 °N to 49.155 °N. It still covers the vast majority of Austria, except the very easternmost part (east of Austria's biggest lake, Neusiedlersee). Also within the domain are northeast Italy (mainly the province South Tyrol), Liechtenstein, parts of Switzerland (the northeast region as well as Graubunden), and in Germany a large part of Bavaria as well as a small part of Baden-Württemberg.

4. Nowcasting Model

Based on the LETKF a nowcasting model is derived in order to generate nowcasts for the variable p, which is the probability of the hourly precipitation rate exceeding a threshold. The essence of the idea for the nowcasting method is to introduce a timeshift in Equation 2.12. Recent INCA analyses are used as observations and the weights are calculated for the C-LAEF members through the Equations 2.1 to 2.10. Then the weight vector is applied to the background perturbation matrix X^b of future timesteps. Thus an adjusted forecast, or rather, a nowcast is computed.
So for a timestep $t - \Delta t$, the $\overline{w_a}$ are calculated at every grid point and applied on the ensemble perturbations of a future timestep t. Mathematically this is formulated as:

$$NOWC_t = X_t^b \cdot \overline{w_{t-\Delta t}^a} + \overline{x_t^b} \tag{4.1}$$

The minimal possible time shift Δt is one hour, as that is one step in the temporal resolution of the precipitation fields. The nowcasts are computed for time shifts of up to six hours. But since time shifts of $\Delta t > 4h$ are found to not produce useful nowcasts (see next section), only nowcasts of up to four hours are evaluated in this thesis.

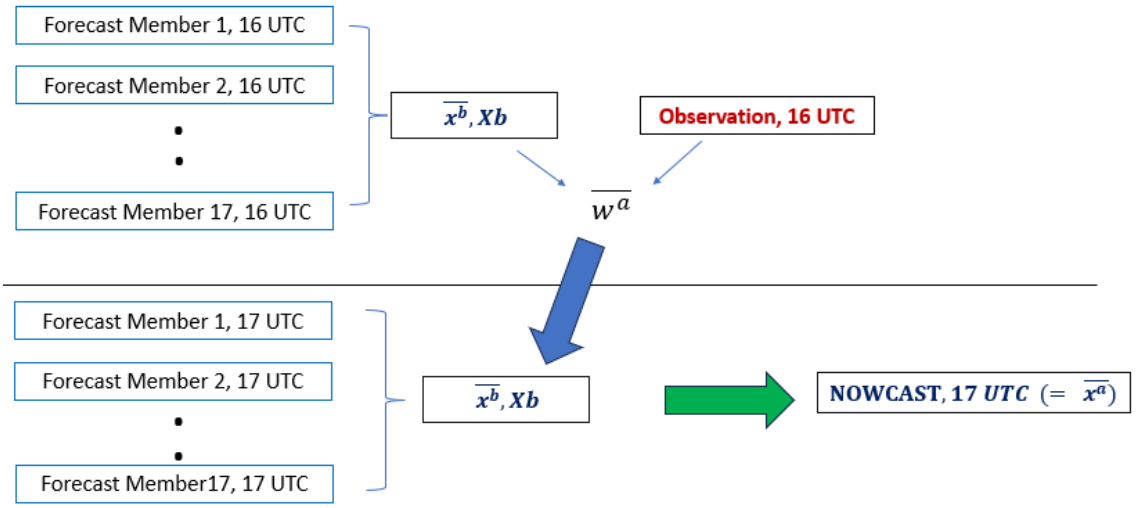


Figure 4.1.: This flowchart illustrates Equation 4.1. In this example, the ensemble forecast for 16 UTC is updated shortly after precipitation observations from 16 UTC in order to forecast precipitation probabilities for 17 UTC

The rest of this section contains information about the specific setup of the nowcasting

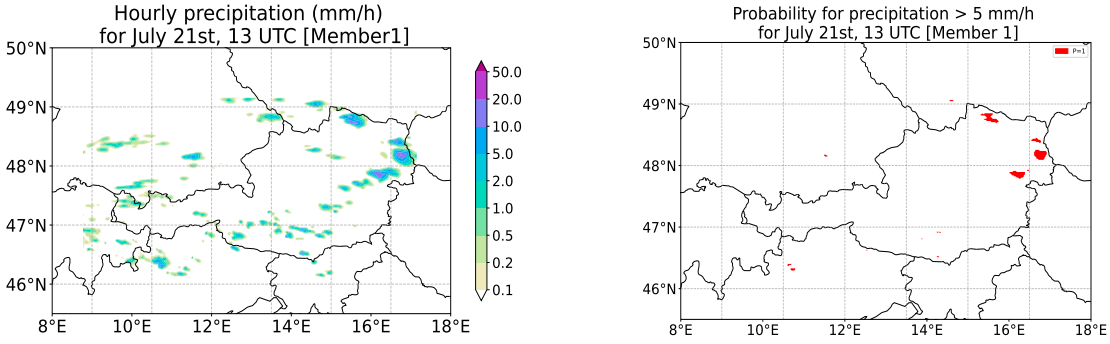
model.

4.1. Creating upscaled probabilities from forecasted precipitation rates

For this thesis, the probability of precipitation is chosen as variable, rather than precipitation intensities directly. This is done to get a variable that better fits the Gaussian assumption, which is one of the assumptions in the LETKF (or any Kalman Filter). Probabilities are formed from hourly precipitation intensities using an upscaling approach. First, the precipitation rates are converted into binary probabilities, using the condition:

$$\begin{aligned} \hat{p} &= 1 \quad if \quad r \geq \tau \\ \hat{p} &= 0 \quad if \quad r < \tau \end{aligned} \tag{4.2}$$

Where τ is a threshold for the hourly precipitation rate. With the above condition, a field of probabilities of 0 and 1 is created from a field of hourly precipitation rates.



(a) Example of a forecast for hourly precipitation

(b) Binary probabilities for a precipitation rate greater than $\tau_1 = 5mm/h$

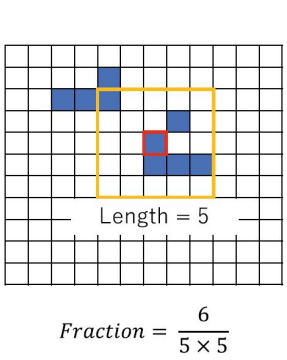
Figure 4.2.: This example visualizes the application of the condition in Equation 4.2 on one of the precipitation forecasts in order to create binary probabilities.

Secondly, a neighbourhood size is chosen for upscaling the probabilities. In order to obtain an upscaled probability field from the binary probabilities in Figure 4.2 b), a simple arithmetic mean is computed for a square of $\eta \times \eta$ grid points:

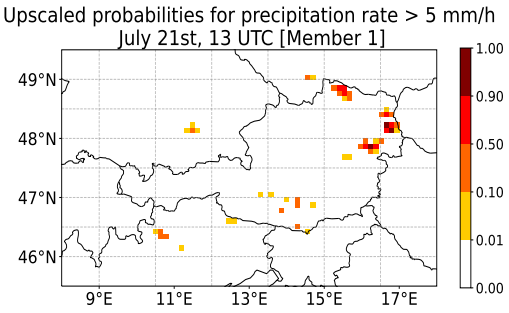
$$p = \frac{\sum_{K=1}^{\eta^2}(\hat{p}_K)}{\eta^2} \tag{4.3}$$

An advantage of using upscaled probabilities, rather than the grid point binary probabilities or probabilities computed from all ensemble members, is that the double penalty problem

is reduced. The double penalty problem is the issue, that if a precipitation pattern is forecasted and observed, but slightly dislocated, it will, by traditional scores or contingency tables, be penalized twice (it was forecasted where it was not observed and it was observed where it was not forecasted). Thus, a forecast with a close miss will perform worse than a forecast that does not predict any precipitation at all. Using the upscaling avoids double penalties on a small scale. The higher the value for η, the more the double penalty problem is reduced. On the other hand, the resulting probabilities are smoothed over larger areas, resulting in probability fields with a lower resolution and less sharp probabilities, compared to a smaller value of η.



(a) Illustration of upscaling: The blue boxes represent binary probabilities of 1 and the white boxes probabilities of 0. Averaging yields a probability of $\frac{6}{25} = 0.24$.Figure from [Park, 2023]



(b) Upscaled probabilities with $\eta = 5$

Figure 4.3.: Using the upscaling method of (a) on the field from Figure 4.2 b) yields the field of upscaled probabilities in (b)

Figures 4.2 and 4.3 together visualize the upscaling. The upscaling is carried out so that all the original C-LAEF grid points within the nowcasting domain are used and each grid point is used in exactly one $\eta \times \eta$ grid box.

4.2. Comparing forecasts and observations

As explained in Section 3.1 and Section 3.2, the INCA analyses that serve as observations are provided on a higher-resolution grid than the forecasts. In order to compare the observations to the forecasts, the observations are also converted into a probability field.

There are a few different possibilities as to how to deal with the different model resolutions at this point, including:

1. Averaging the INCA precipitation field onto the C-LAEF grid before computing the upscaled probabilities

2. Computing binary probabilities on the INCA grid according to Equation 4.2, binning the observations into the upscaled grid and computing the probability for every point on the upscaled grid according to Equation 4.3.

Option 1) would presumably reduce the bias that is caused by the different resolutions of the forecast and the analysis. The high-resolution INCA field is less smoothed than the C-LAEF field and therefore will have higher spikes of precipitation rates which can affect the probabilities. Option 2) remains more oriented toward the observations as it does not mathematically alter precipitation rates, and is, therefore, closer to the truth experienced on the ground. Since one of the intentions behind the method is to be impact-orientated, option 2) is chosen. It also appears to be mathematically cleaner, since essentially the probabilities are produced in the same way as the probabilities in the forecast.

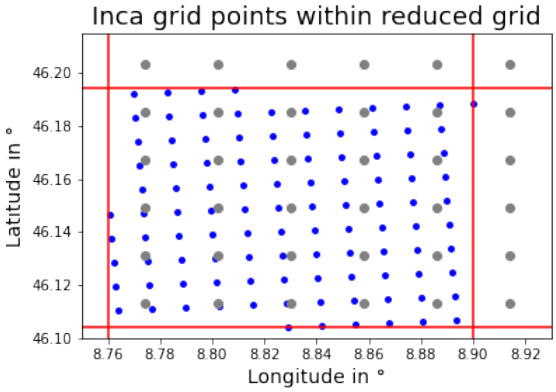


Figure 4.4.: The grey points represent grid points of the original C-LAEF grid. The red rectangle represents the grid box of the upscaled grid, spanning 5 by 5 grid points. The blue points show the INCA grid points within this particular upscaled grid box.

Figure 4.4 illustrates the upscaling process: Same as the values of p for a number of $\eta \times \eta$ C-LAEF grid points are averaged to obtain upscaled probabilities, the INCA grid-points within this area are also averaged in the same way. The number of INCA grid-points within upscaled grid boxes varies because of the slanted shape of the INCA domain compared to the rectangular longitude-latitude grid of the C-LAEF domain. Also, because of meridian convergence, the northern part of the regular lon-lat grid covers a shorter distance from west to east in km while the INCA grid always covers $700km$ from west to east (therefore it appears wider in the north than in the south on the lon-lat grid

in Figure 3.1). Therefore in the more southern latitudes of the domain, there are more observations within the boxes of the upscaled grid.

4.3. Specific setup and tuning choices of the nowcasting model

Several tuning choices have to be made. This section lists all of them and which setting is chosen, as well as the reasons for that.

4.3.1. Resolution of upscaled probabilities

The first choice to be made is the upscaling size, i.e. the side length η of the rectangle of observations that are used in Equation 4.3 in order to turn the precipitation field into a probability field. Here, empirically the value of $\eta = 5$ is chosen. This means that all the following computations are undertaken on a reduced grid that is five times coarser in each direction than the grid that was originally provided. In each box of the upscaled grid there are exactly 25 grid points of the provided C-LAEF grid, and roughly 105 points of INCA observations. The resulting grid is of the grid size $\Delta lon = 0.14$ °E in the x-direction and $\Delta lat = 0.09$ °N in the y-direction. The distance in real space is roughly $10km$. This grid size is of the order of the effective model resolution of C-LAEF (see [Bierdel et al., 2012] for more about the effective model resolution, there for the COSMO-DE model). A high-resolution field of probabilities is desirable for users because the probability of rainfall exceeding a threshold can then be forecasted very locally. With a small upscaling size η the probability forecast is more specific to every location and gives a sharper probability field.
In order to test the sensitivity of the nowcasting skill on the upscaling size, a sensitivity study is undertaken, where all nowcasts are also computed with a coarser resolution probability field of $\eta = 10$.

4.3.2. Threshold of precipitation rates

The threshold τ in Equation 4.2 can be chosen according to the user's desire. For the underlying thesis, the method is assessed for three different thresholds: $\tau_1 \geq 5mm/h$, which is referred to as strong precipitation, moderate precipitation with a threshold of $\tau_2 \geq 1mm/h$, and $\tau_3 \geq 0.1mm/h$, representing weak precipitation (or any precipitation).

4.3.3. Localisation scale in the LETKF computation

Another choice to be made is the number of observations (on the upscaled grid) to be considered in each local analysis when following the LETKF routine through Equations 2.1 to 2.10. As discussed in Section 2.3, the number of observations that is assimilated is usually similar to (or slightly higher than) than the number of ensemble members k. Since the ensemble used in this thesis consists of 17 members, it is decided to use not only neighbouring grid boxes but also the neighbours of the surrounding grid boxes, so

altogether a number of $m = n = 25$ observations in each local analysis. The choice of assimilating the rectangle of the closest 5×5 observations means that for neighbouring LETKF-computations many of the observations (20 out of 25) are shared. This meets the recommendation ([Hunt et al., 2007]) of a large overlap for a physically well-balanced field when putting the local analyses together in the end.
Because of the choice to compute the INCA probabilities on the same (upscaled) grid as the C-LAEF probabilities, the observation operator H becomes the Identity matrix (of dimensions 25×25), and the identity from Equation 2.5 follows. The state vectors y_i^b become n-dimensional vectors. Y^b becomes a $n \times k$ - sized matrix, which specifically means that its dimensions are 25×17. R becomes a 25×25-sized matrix.

4.3.4. Assigning observation errors

Another component that has to be tuned empirically is the observation error covariance matrix R. Since the upscaled probabilities are computed without any overlap, the observations and thus the observation errors are assumed to be uncorrelated. So off-diagonal elements in R are set to 0 and R becomes a diagonal matrix. R being diagonal reduces the computational cost, as the inversion in Equation 2.7 is computationally much cheaper for diagonal matrices than for non-diagonal ones.
The values of the diagonal elements in R are chosen to be larger for entries corresponding to observations further away from the center grid point of the local analysis ([Hunt et al., 2007] suggested either this or identical entries). The observation error at the center point is assumed to be $0.1 \cdot 1$, the neighbouring observations are assigned an error of $0.1 \cdot 2$, the furthest observations are assumed to have an error of $0.1 \cdot 4$.
To find out whether the choice of the values in R influences the performance of the nowcasts, the computations are also carried out with identical values in R for all observations in a sensitivity study. The results are presented in Section 5

4.3.5. Symmetric square root

The computationally most extensive part of the calculation is Equation 2.9 as it involves the symmetric square root of the non-diagonal matrix $\tilde{P}^a$ of the size 17×17. For this thesis, the symmetric square root is computed using the Python package *numpy* and its command 'root'. Generally, the symmetric square root of matrices can produce imaginary solutions. These imaginary solutions did occur in this computation and turned out to be a problem for the computation program, as they would have progressed in the computation and ultimately would have led to an imaginary part in the output probabilities. Therefore, the imaginary part of the symmetric square root is dropped. Only the real part of W_a is considered by using the command 'numpy.real()'

4.3.6. Negative probabilities

The nowcasting model treats the probabilities like any other variable, it merely adapts the forecasts by increasing or decreasing the values according to the weight vector $\overline{w^a}$. Rarely the nowcasting model calculates probabilities slightly larger than 1, and sometimes slightly smaller than 0. These probabilities are set to 1 and 0 respectively after computation and before statistical evaluation of skill.

4.3.7. Using the most recent forecast datasets

As mentioned above, the C-LAEF ensemble is initiated every 12 hours, AROME-forecasts are, however, initiated every 3 hours, at 0 UTC, 3 UTC, etc.
The choice of ensemble runs is made so that the most recent run of every model is used, with one exception: The probabilities for the very first hour of the forecasts are not used in order to avoid possible error due to model spin-up. This means, that for example at 13 UTC, 16 C-LAEF ensemble members from 0 UTC are used together with the 9 UTC AROME-AUT run. At 14 UTC a nowcast is started with the 9 UTC AROME-AUT run and the 12 UTC C-LAEF runs.
The benchmark of the nowcasting method is the underlying ensemble itself. This means, that it is investigated whether the nowcasting model is able to perform better than the ensemble that was available at the time of the nowcast initialisation. Future runs are not considered and not used as a benchmark. So, for example, a $3h$ nowcast that is started at 23 UTC and valid for 2 UTC is not compared to the ensemble runs from 0 UTC. Its reference forecast is the forecast for 2 UTC from the C-LAEF runs from 12 UTC on the previous day, together with the AROME-AUT run from 18 UTC.

4.4. Verification method

There are many different verification measures to assess the quality of forecasts. In this thesis, for reasons of simplicity and easy interpretation, the root mean square error (RMSE) is chosen to assess the skill of both the background ensemble mean and the nowcast. The INCA analyses serve as truth. The RMSE is computed as follows:

$$\overline{RMSE} = \sqrt{\frac{\sum_{l=1}^{L}\sum_{j=1}^{J}(p_{l,j}^{o} - p_{l,j}^{f})^2}{L \cdot J}} \tag{4.4}$$

Where $p_{l,j}^{o}$ is the observed probability at the grid point in the l-th column of the j-th row of the upscaled grid of the nowcasting domain, and $p_{l,j}^{f}$ is the corresponding forecasted probability at the same grid point. To get one verification measure, the RMSE is computed over all $L \cdot J$ grid points of the nowcasting domain.
In order to statistically assess the skill of the nowcasting method, rather than just the

skill of single nowcasts, the average RMSE over the whole period is also computed:

$$< \overline{RMSE} >= \frac{\sum_{t'=1}^{T}(\overline{RMSE}_{t'})}{T} \tag{4.5}$$

Comparing the RMSE values and computing either the absolute RMSE reduction

$$\Delta < \overline{RMSE} >=< \overline{RMSE} >_{Ens} - < \overline{RMSE} >_{Nowc} \tag{4.6}$$

or the relative RMSE-reduction in percent

$$\Delta < \overline{RMSE} >_{Rel} (\%) = \frac{< \overline{RMSE} >_{Ens} - < \overline{RMSE} >_{Nowc}}{< \overline{RMSE} >_{Ens}} \tag{4.7}$$

gives the final measure that is used to assess the usefulness of the nowcasts for the whole period of six days.
As precipitation never occurs at the whole domain in the same hour, and in some cases only occurs at a very small fraction of the whole domain, the RMSE as computed above, often is dominated by correct negatives. This means that no probability of precipitation is forecasted, and none is observed. Especially when choosing a high threshold like τ_1, this happens in the majority of cases. Therefore a separate evaluation is undertaken that discards the cases where both $P^o_{i,j} = 0$ and $P^f_{i,j}(Ens) = 0$. This results in higher RMSE-values for both the ensemble and the nowcast and changes the absolute (Equation 4.6) and relative (Equation 4.7)RMSE reduction.
Other verification scores like the Ranked Probability Skill Score, which is well suited for probability forecasts, or neighbourhood verification methods like the Fraction Skill Score ([Roberts and Lean, 2008]) also come to mind and might give an even better picture of the skill of the forecast. However the RMSE is deemed good enough to assess the general quality of the forecasts and is used because it is the most well-known and widespread measure of error quantity. The double penalty problem is already reduced because of the probability upscaling. And mathematically, the RMSE from Equation 4.4 equals the square root of the Brier Score ([Brier, 1950]) with the only difference that in the Brier Score the observations are usually used on grid point level with binary probabilities of either 1 (the event did occur) of 0 (the event did not occur) instead of neighbourhood probabilities. Furthermore in the Fraction Skill score the verification is undertaken by averaging precipitation in grid boxes ([Roberts and Lean, 2008]), just like the probabilities are formed in this thesis (as explained in Section 4.1). So computing the RMSE for the field of upscaled probabilities is similar to two different well-established verification methods.

5. Results

This section presents and evaluates the results. A few examples of nowcasts with the default setting are shown in order to understand where and how much the nowcast model deviates from the underlying ensemble and to get a better understanding of how the nowcasting model works in practice. After that, a full statistic evaluation for all nowcasts with standard setting is carried out by evaluating the RMSE statistics from Equations 4.4, 4.5, 4.6, and 4.7. In order to deepen the understanding of the nowcast model, the results of two sensitivity studies are presented: The results of full re-runs for the whole period and the thresholds τ_1, τ_2 and τ_3 but with different settings are presented. The first sensitivity study shows the influence of the setting of the observation error covariance matrix. The second sensitivity study shows the results for a coarser probability field with an upscaling of $\eta = 10$ instead of $\eta = 5$.

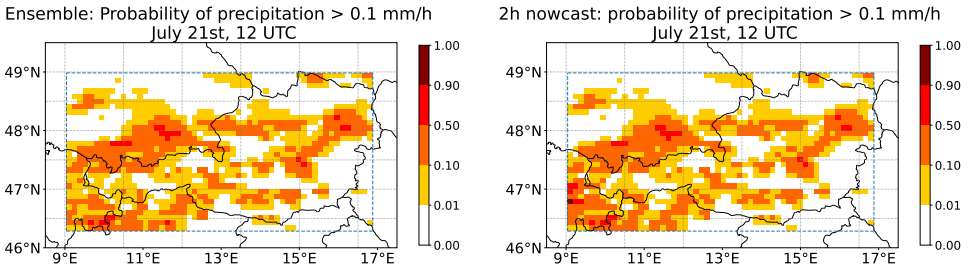
5.1. Case studies

The smaller the threshold τ is chosen, the more signals remain in the forecasts for the probability of precipitation exceeding that threshold. So first, a nowcast with a shift of $\Delta t = 2h$ is shown for a randomly chosen time in the period for a threshold of $\tau_3 = 0.1\frac{mm}{h}$.

5.1.1. 2h nowcast for the early afternoon on July 21st

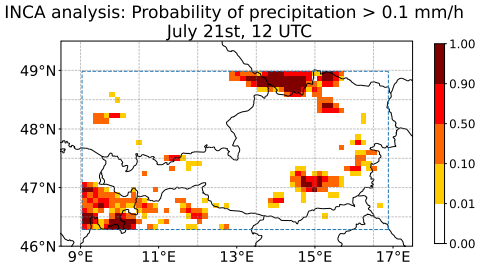
The nowcast that is valid for July 21st, 12 UTC (local time: 2 pm) and was initiated at 10 UTC is compared to the ensemble forecast. At first glance, the nowcast in Figure 5.1 (b) hardly deviates from the ensemble mean in (a). On second glance, however, probabilities are increased in the far west of the domain. Figure 5.2 shows the timestep two hours prior (10 UTC) that was used to calculate the weighting of the ensemble. Looking at this timestep of initialisation in Figure 5.2 it quickly becomes apparent, why not much was adapted in the future timestep. For almost the whole eastern half of the domain, the ensemble probability at initialisation time was 0. Therefore, every member was equally correct and is given the same weight by the LETKF algorithm. From Figure 5.1 and Figure 5.2 as well as from the derivation of the LETKF in Section 2.3 it can be concluded that the nowcast at any given point produces the exact same probability as the underlying ensemble, in the following cases:

1. If the local probability from the ensemble mean equals 0 at initialisation time. A mean probability of exactly 0 means that all members predicted 0 and are therefore equally correct. The nowcast will be an equal combination of the members and therefore identical to the ensemble mean.



(a) Background ensemble forecast

(b) Nowcast



(c) Observation

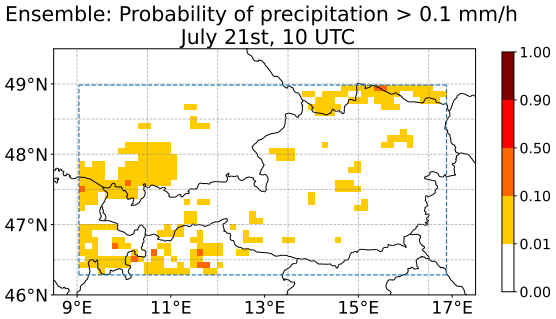
Figure 5.1.: Set of ensemble, nowcast, and observation (truth) for weak precipitation on the early afternoon of July 21st.

2. If the ensemble probability is 1, or if all members forecast any identical probability, for the same reason.

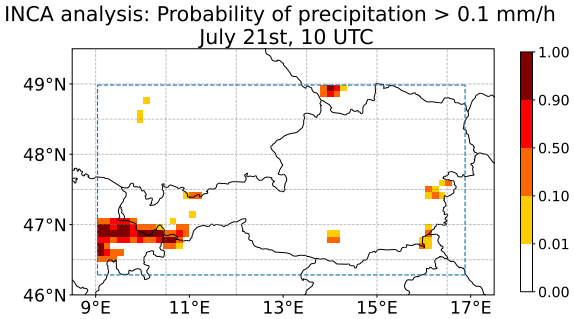
3. If the ensemble probability at the future timestep equals 0 (or 1 or any identical value for all members). That is because if all members hold a probability of 0 in a local area, any re-weighting of the different forecasts will again end up with a probability of 0.

5.1.2. Second example: best 1h nowcast for strong precipitation

Successful nowcasts of strong precipitation are particularly valuable due to the high impact of high precipitation intensities. This example features the 1h nowcast with the greatest absolute reduction in RMSE according to Equation 4.6. This most skillful nowcast is initiated on July 24th, 9 UTC, see Figure 5.3. The weights are applied to the background ensemble forecast for 10 UTC (in Figure 5.4 (a)). Computing the RMSE of the nowcast (the error of the field on the top right panel in Figure 5.4 compared to



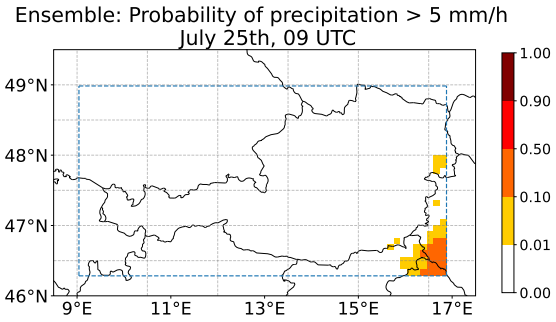
(a) Ensemble probabilities for the time of initialisation



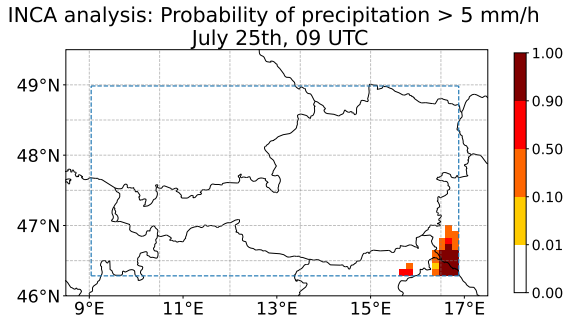
(b) And the corresponding observation

Figure 5.2.: Only a small part of the domain holds an overlap of forecasted and observed probability of precipitation

the observations at the bottom according to Equation 4.4) yields an RMSE of 0.0663. The RMSE for the ensemble for 10 UTC stands at 0.0834. Therefore the relative RMSE reduction as in Equation 4.7 is slightly above 20% for this particular nowcast. The time



(a) Background ensemble forecast

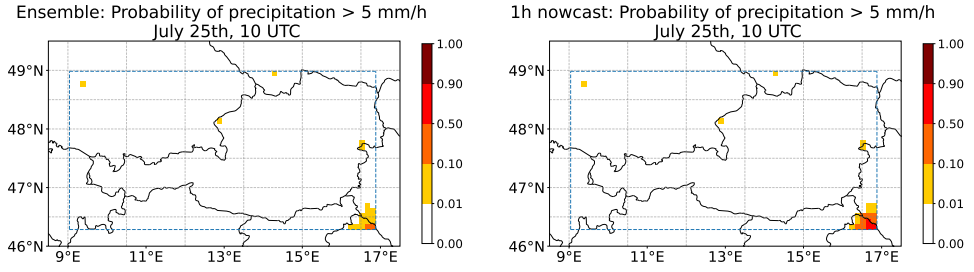


(b) Corresponding observation

Figure 5.3.: Initialisation of the most skillful nowcast

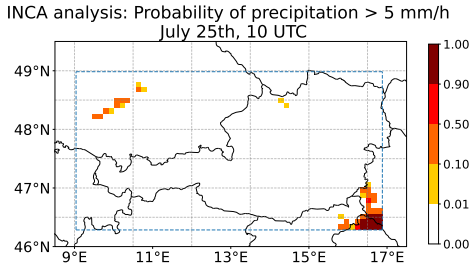
of initialisation shows why the nowcast is so successful. At 9 UTC (Figure 5.3) there is an overlap of forecasted and observed probability of precipitation. However, not all ensemble members but a fraction of them correctly predict precipitation in the southeast of Austria,

seeing as the probability of the ensemble mean is less than 50%. As a consequence, the LETKF algorithm gives more weight to those members that do predict a high probability of precipitation where precipitation was observed.



(a) C-LAEF ensemble forecast

(b) Nowcast



(c) Observation

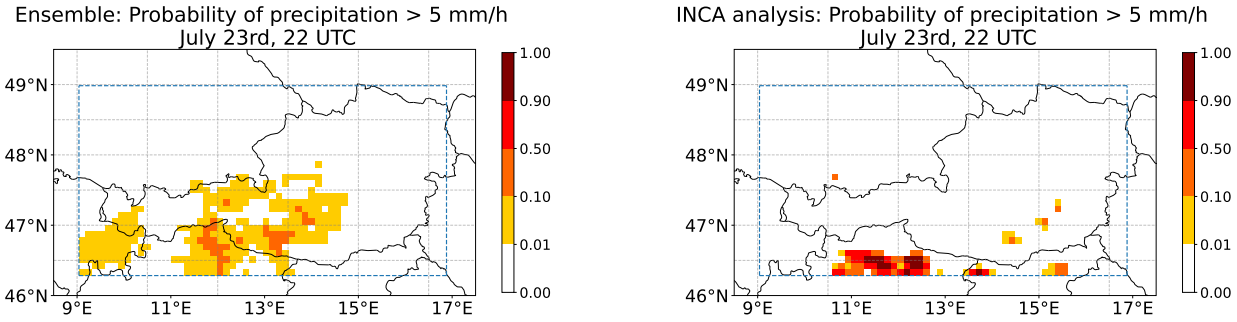
Figure 5.4.: The background forecast in (a) is adapted to the nowcast in (b). The bottom panel shows the corresponding observation.

The ensemble forecast for 10 UTC still holds a certain probability of precipitation in the same region where at the previous timestep in Figure 5.3 high probabilities were observed. Since the LETKF-based algorithm now increases the weight of the members that correctly predicted precipitation in the region and since the same members still predict a certain probability of precipitation, the probabilities in that area are increased by the nowcast. Since strong precipitation is still also observed in the same region, the nowcast is successful in increasing the skill.

5.1.3. Third example: worst 3h nowcast

In order to understand why nowcasts sometimes have less skill than the ensemble forecast, the next example shows the one particular 3h nowcast with the largest absolute increase in RMSE according to Equation 4.4. The nowcast is valid for July 24th, 1 UTC. The

weights were computed at 22 UTC on July 23rd. As Figure 5.5 shows, there already



(a) Background ensemble forecast

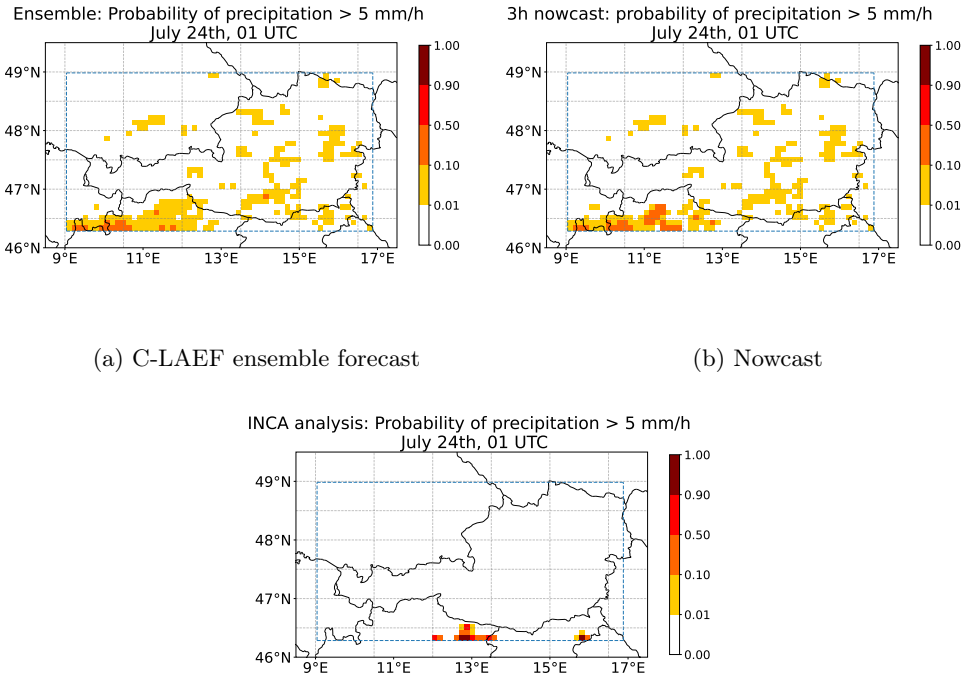
(b) Corresponding observation

Figure 5.5.: Initialisation of the least skillful 3h-nowcast

is a big discrepancy between forecast and observation. Areas with a probability of 0 in the background in 5.5 (b) will not be affected by the LETKF algorithm. There are, however, regions with an overlap of forecasted and observed probability, the largest one being in South Tyrol, south of Austria. This is where the algorithm will adapt the forecast if any members show a probability different from 0 in the future timestep. The nowcast increases the probabilities according to the assimilation of the observations into the background at initialisation time. Thus, probabilities are increased in South Tyrol / Trentino. However, as Figure 5.6 (c) shows, the observed probability, i. e. the precipitation field, barely overlaps with the observations from 22 UTC anymore. The precipitation pattern has moved further east, into the regions of Friuli-Venezia Giulia and Veneto. As the probabilities are increased in the wrong place (where precipitation was observed three hours before), the nowcast now holds a higher RMSE than the underlying ensemble forecast. This example shows that signal propagation of precipitation patterns is a major limitation. Precipitation patterns that cause the algorithm to adapt the probabilities in one place, after a few hours, often have propagated to another place outside of the localisation scale.
The RMSE for this worst 3h nowcast is 0.0723, the error of the background ensemble is 0.0609, which corresponds with a relative error increase of about 18.7%

5.2. Statistical evaluation for all nowcasts with default setting

Rather than the performance of single nowcasts, the average performance for all nowcasts (up to 143 for 1h nowcasts) is evaluated in this section for the three thresholds τ_1, τ_2 and



(a) C-LAEF ensemble forecast

(b) Nowcast

(c) Observation

Figure 5.6.: The background forecast in (a) is adapted to the nowcast in (b). The bottom panel shows the corresponding observation.

τ_3 and for nowcasts from one to four hours into the future. The point at $0h$ lead time in Figure 5.7 does not represent a nowcast, but the classic LETKF analysis, and is included here for comparison and in order to see how fast the skill reduces with the time shift. For strong precipitation rates $\geq 5mm/h$, the average RMSE for all probabilistic nowcasts with a lead time of one hour, as presented in Equation 4.7, was 2.99% smaller than the average RMSE of all the respective background ensemble forecasts. For moderate precipitation of $1mm/h$, the average RMSE was reduced by 10.06% in 1h nowcasts, and for the threshold τ_3 by 13.89%.
The skill drops drastically with increasing lead time, though. For all nowcasts with a shift of $2h$, the relative improvement compared to the underlying ensemble is 1.72% for weak precipitation (threshold τ_3), for the threshold τ_2 the nowcasts improved by 0.89% on average. For strong precipitation $\geq 5mm/h$ the relative error-reduction drops below zero at -0.62% which means that the error of the $2h$ nowcasts, on average, is higher than the error of the C-LAEF ensemble. For nowcasts three or four hours into the future, the nowcasting model produces, on average, slightly larger errors than the ensemble. In

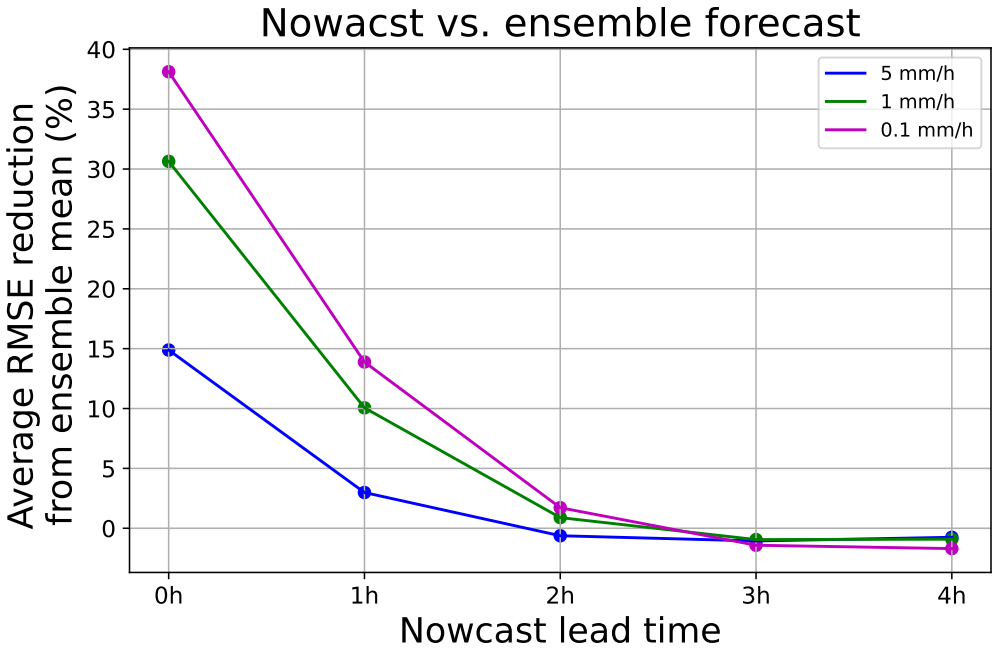


Figure 5.7.: RMSE reduction, averaged over all nowcasts, compared to the underlying ensemble forecasts.

other words it is - on average - less skillful than C-LAEF. While some nowcasts are still successful, the impact of nowcasts that increase the error outweighs them.

5.2.1. Correlation of spatial extent and nowcast skill

In order to learn more about the behaviour of the underlying nowcasting method than one overall RMSE-statistic, correlations between the observed average precipitation probability over the domain and the RMSE reduction are investigated. Here the absolute RMSE reduction as in Equation 4.6 was plotted on the y-axis. Figure 5.8 shows, that across all three precipitation intensities, the probability nowcasts perform better if a larger part of the whole domain features precipitation. This is an advantage, as nowcasts for precipitation are especially interesting if a lot of precipitation is expected. One reason why the performance is better for larger precipitation areas could be that large-scale precipitation patterns are more long-lived than small isolated convective cells.

5.2.2. Dependence of RMSE on correct negatives

Since the RMSE is strongly reduced by correct negatives (correctly predicted non-occurence of probability of precipitation) for both the ensemble and the nowcast, a different evaluation for the same nowcasts and ensemble forecasts was undertaken. Correct negatives were filtered out before computing the RMSE of the ensemble and the nowcast. Thus, RMSE

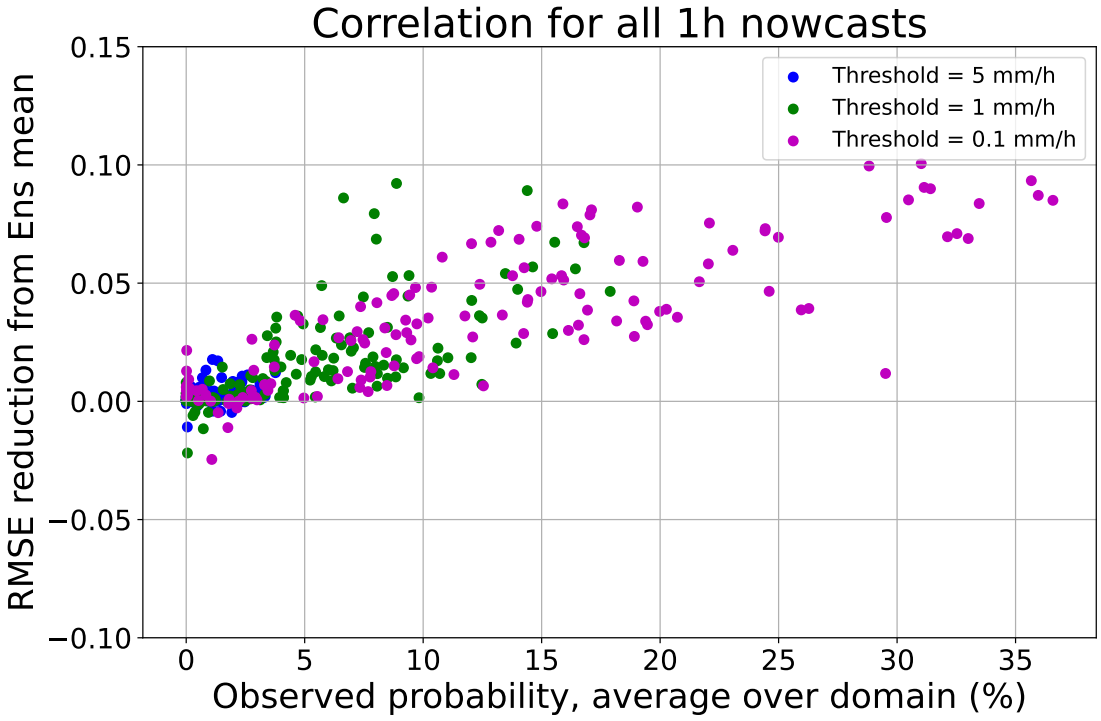


Figure 5.8.: Scatter plot for all 1h nowcasts: Absolute RMSE reduction vs. domain-mean of probability of precipitation for different thresholds. The domain-mean probability corresponds to the spatial extent of precipitation patterns.

values become higher for both the ensemble and the nowcast. The following table shows the results for all nowcasts with a threshold of $\tau_2 \geq 1mm/h$. The absolute error is reduced

	$\Delta t = 1h$	$\Delta t = 2h$	$\Delta t = 3h$	$\Delta t = 4h$
$RMSE_{C-LAEF}$	0.165	0.166	0.168	0.170
$RMSE_{Nowc}$	0.149	0.165	0.170	0.171
Relative reducion[%]	10.1	0.9	−0.9	−0.9
Absolute reduction	0.017	0.001	−0.002	−0.002
$\tilde{RMSE}_{C-LAEF}$	0.243	0.243	0.248	0.253
$\tilde{RMSE}_{Nowc}$	0.220	0.242	0.250	0.255
Relative reduction [%]	9.2	0.7	−1.0	−1.0
Absolute reduction	0.022	0.002	−0.003	−0.003

Table 5.1.: Average RMSE for all ensembles / nowcasts for time shifts of 1 to 4 hours. The $\tilde{RMSE}$ values in the bottom half are RMSE-statistics without correct negatives. Values are rounded to the third decimal place, relative reductions are rounded to the first decimal place.

more as the absolute values of both errors increase. The relative RMSE reduction in [%] however is not increased by filtering out the correct negatives. This stays the same for the other thresholds, the corresponding tables of the exact results for strong and weak precipitation can be found in the appendix.
In order to find out whether the correlation from Figure 5.8 was partly due to the correct negatives, the same correlation was plotted again for RMSE reduction without correct negatives. As Figure 5.9 shows, the correlation still prevails and therefore has nothing to

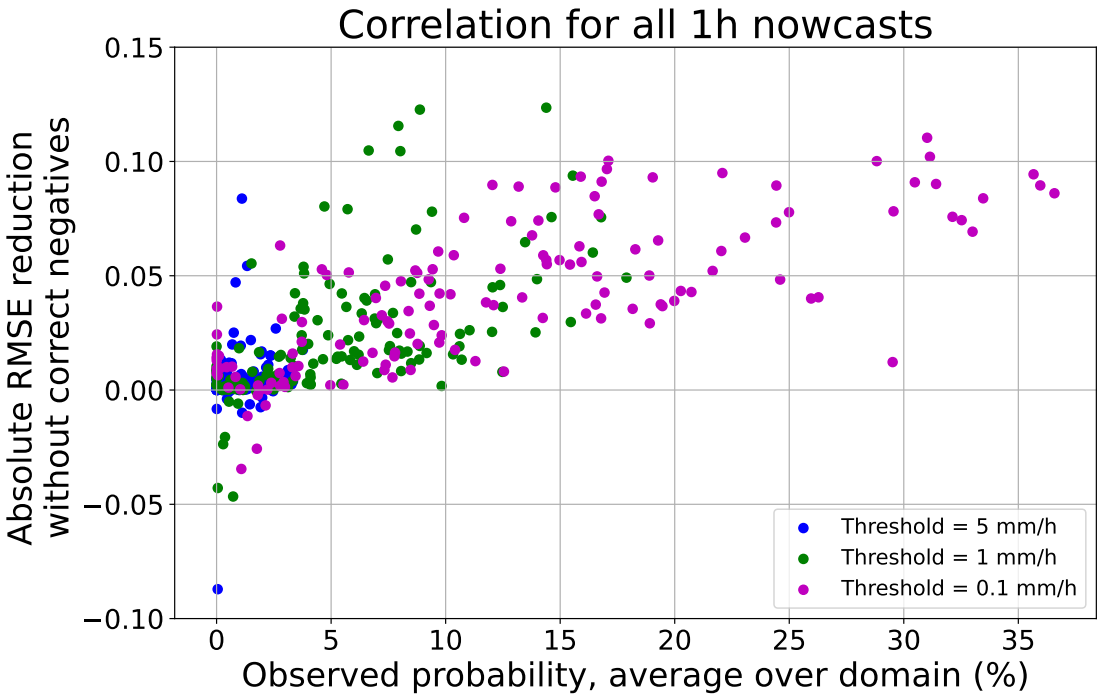


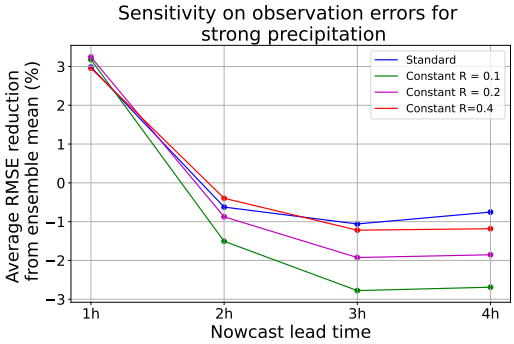
Figure 5.9.: Spatial extent of precipitation vs RMSE (without correct negatives)

do with correct negatives, but is due to the nowcasting model improving probabilities where they occurred.

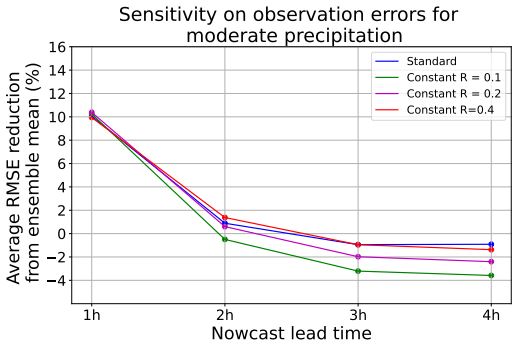
5.3. Sensitivity studies

5.3.1. Tuning the error-covariance matrix

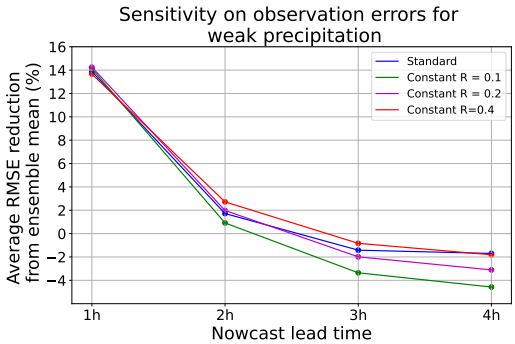
As explained in Section 4.3.4, the observation error was assumed to be larger for observations that are further away from the point of the analysis. However, as signal propagation plays a part in the performance of the nowcasts, it seems unjustified to give less credit to observations that are further away. Therefore, re-runs were computed with identical values for all the entries in the diagonal in R. Three experiments have been undertaken with the values $0.1, 0.2$, and 0.4 as assumed errors for all observations respectively. Covariances in R were still ignored and set to 0. The sensitivity study in



(a) Threshold $\tau_1 = 5mm/h$



(b) Threshold $\tau_2 = 1mm/h$



(c) Threshold $\tau_3 = 0.1mm/h$

Figure 5.10.: Skill evolution with lead time compared to the ensemble for different observation errors and all three thresholds

Figure 5.10 shows that the performance of the method is not too sensitive on the exact setting of the error-covariance-matrix. The general behaviour does not change. Among all three thresholds, the setting that performs slightly better than the others for $1h$ nowcasts is that with a constant observation error of 0.2, closely followed by the setting with the smallest observation error of 0.1. However, for larger time shifts of three or four hours, these two settings show the least skill. Assuming a big error of $R = 0.4$ reduces the skill a little for $1h$ nowcasts, compared to the other settings. For time shifts of $2h$, however, assuming larger errors for all the observations appears to pay off, as the setting with the largest assumed errors performs best. The standard setting shows the smallest decrease of skill with lead time, however it starts out with less skill than the two smaller constant errors for $1h$ nowcasts.

5.3.2. Increased neighbourhood size

As Section 5.1.3 shows, the propagation of precipitation patterns poses a major limitation to the skill of the nowcasting model. Therefore, a test with a lower-resolution probability field was undertaken. Instead of averaging 5×5 grid point in the upscaling that was

explained in Section 4.1, an alternative upscaled grid of 10×10 was used to form the probabilities before computing the nowcast. This means that four times as many points of the original C-LAEF grid are in the upscaled grid and that the horizontal distance of local observations is doubled in both the x-direction and the y-direction. This section presents the results for this 10×10 grid with standard settings. As Figure 5.11 shows,

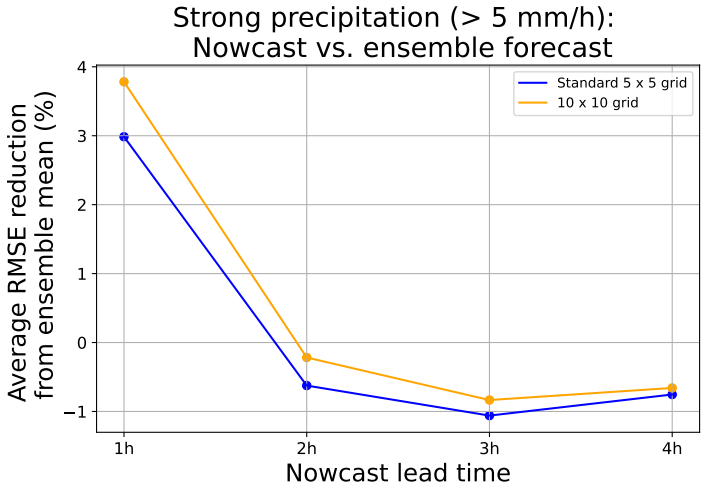
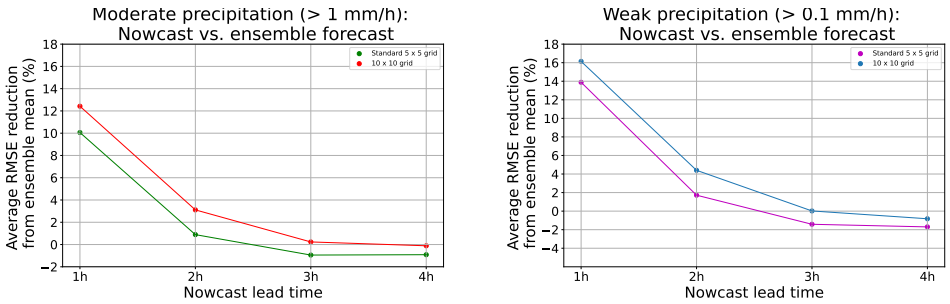


Figure 5.11.: Relative RMSE-reduction compared to the respective ensemble-mean on the standard grid, compared to the coarser 10×10 grid.

the nowcast model is now able to adjust the ensemble probabilities a little bit better. However, after 2 hours it is still less skillful than the ensemble. It is important to note here that the nowcasts are compared to different RMSE values, as the RMSE of the ensemble on the coarser grid is different than the RMSE on the standard grid. Typically for the coarser grid, RMSE values are lower for both the ensemble and the nowcast. For lower



(a) Relative RMSE-reduction for a threshold of $1 \frac{mm}{h}$

(b) ... and a threshold of $0.1 \frac{mm}{h}$

Figure 5.12.: Evaluation of equation 3.7 for all hours on thresholds τ_2 and τ_3,

thresholds of precipitation intensity, the nowcasting model also performs better on the

coarser grid, compared to the ensemble. Using this setting increases the usability of the nowcasting model to three hours, where it still on average holds a lower averaged RMSE than the background C-LAEF ensemble. This shows that increasing the neighbourhood size increases performance for larger time-shifts as precipitation patterns do not tend to move entirely out of the local vicinity as much. Also, the double penalty problem is reduced, but for both the ensemble and the nowcast.

6. Conclusions

The nowcasting model is a successful proof of concept for the idea of using the LETKF-based weighting to combine observations with an ensemble prediction system. With all the settings that are used in this thesis, the nowcasts are able to perform better than the C-LAEF ensemble for a lead time of $1h$. The nowcasting model performs better for lower intensities: For the thresholds τ_1 and τ_2, it is feasible to use for two or up to three hours with the presented settings, depending on the upscaling size.
One reason why the nowcast performs better for lower intensities can be seen in Figure 5.9. The larger the spatial extent of the precipitation patterns, the more corrections can be undertaken by the nowcasting model. This agrees with the work of [Kober et al., 2013] who found that their blended nowcast did not perform as well for small-scale precipitation as it did for larger precipitation patterns. Other studies also discuss that the scale of the precipitation pattern influences the predictability with nowcasting methods heavily, for example [Germann and Zawadzki, 2004]. Another reason why the nowcasting model does not perform as well for higher thresholds might be the typical duration of heavy precipitation in convective cells. While a single cell might produce precipitation for more than one hour, it is less likely to maintain heavy precipitation for over one hour, as after a while its vertical updraft is weakened by the downdraft of cold air that is dragged along with the falling precipitation. This reduced duration for larger intensities makes it more likely for the nowcasting model to wrongly increase the probabilities of precipitation for the following hours. It should be noted, however, that the results are less certain for higher precipitation rates because of the smaller sample size.
As the example of the least skillful 3h-nowcast has shown, signal propagation seems to be a major limitation for the usability of this nowcasting model. This is underlined by the fact that the nowcasting model was not skillful for lead times above three hours. A nowcasting method that on average performs worse than an ensemble forecast can be considered as not skillful (or not useful). Improvement can not be expected, even though it occurs in some cases. Therefore the results of the timeshifts $\Delta t = 5h$ and $\Delta t = 6h$ were not shown in this thesis.
One way to mitigate the effects of signal propagation would be to increase the radius of observations that are considered. Doing this by sampling more observations in the LETKF computation is unlikely to bring better results, as the LETKF is somewhat limited in the number of observations it can assimilate by the ensemble size. Therefore upscaling the probability field further proves promising and performs better with the underlying datasets in the sensitivity study. The results show that all statistics were more in favor of the nowcasting model compared to a higher-resolution grid. On the other hand, the resolution and sharpness of the probability field is reduced by a larger upscaling.
The problem of signal propagation could be avoided by trying to shift the weights not

only through time but also through space. One approach would be to shift the weights through space along with the mean wind during the time-shift. One way to obtain the mean flow could be the forecasted wind in the $700hPa$-layer for example. This was not implemented in the underlying thesis and gives room for further research. Different settings (for example correlated observation errors in R) might further increase the skill of the nowcasting model.
With INCA analyses being available in almost real-time, this nowcasting model is theoretically ready to go operational and produce skillful nowcasts for short lead times. What comes into play here is the computational efficiency of the LETKF. For this thesis, computations were undertaken on one of the servers of the University of Vienna. The computation time was typically around 23 seconds for one nowcast of the hourly precipitation probability for the whole nowcasting domain (i.e. for the whole of Austria).
The nowcasting model was in many cases not able to adjust the probabilities at all because all members computed a probability of 0. Using a larger ensemble like a multi-model ensemble or a time-lagged ensemble would be easy to implement and would more often result in probabilities different from 0. This would lead to more corrections by the nowcasting algorithm and might make the method overall more skillful. Even introducing a pseudo-member that holds a small probability everywhere in the domain might prove feasible as it might somewhat cover otherwise missed events. On the other hand, a larger ensemble again leads to more smoothed and less sharp probability fields.
Precipitation defies the Gaussian assumption more than other meteorological parameters (its distribution is usually modeled by Gamma functions, see [Wilks, 2011]). Precipitation was chosen for this thesis anyway because it is extremely desirable to have high-precision nowcasts for. The problem of the non-gaussianity was reduced by working with probabilities rather than intensities. The method is, however, in no way restricted to precipitation.
Another parameter that could be well suited for an LETKF-based nowcasting is wind gusts. For wind gusts out of convective cells, the same problem of signal propagation might occur. However, for longer-lasting storm events in the colder season due to large pressure differences on the synoptic scale, signal propagation should not pose such a problem. In the author's opinion, the LETKF-based nowcasting appears to be promising and could prove valuable in operational nowcasting if further research is done.

7. Acknowledgements

Massive thanks go out to my supervisors, Univ.-Prof. Mag. Dr. Martin Weißmann and Dr. M.Sc. Tobias Necker, who were infinitely patient with me when I took a long time-out from writing this thesis due to my employment as a weather forecaster,
and who were always very quick and helpful in answering my questions when I finally did get back to work on this. Also, thanks for motivating me to finish this thesis.
Thanks to GeoSphere Austria for providing the datasets that were used in this thesis. Regarding the provided information about how to process the data formats, special thanks go out to Mag. Alexander Kann and Maximilian Weissinger M.Sc.
Thanks to the University's Meteorology department for allowing me to use the servers to store and process the big amounts of data. Thanks in particular to Mag. PhD Michael Blaschek for helping me many times when I had problems connecting to the server.

Thanks to my family for more financial support than could be expected and for always believing in me. Thanks to everyone who asked repeatedly when this thesis will finally be finished. I hated the question but I also needed it.
And, last but not least, thanks to my partner of eight years, who gave me motivational speeches when I needed them most and who also never lost faith in my ability to finish this thesis while working.

Bibliography

[Agrawal et al., 2019] Agrawal, S., L., B., Bromberg, C., Burge, J., C., G., and Hickey, J. (2019). Machine learning for precipitation nowcasting from radar images. *arXiv preprint: arXiv:1912.12132 [cs.CV]*. DOI: 10.48550/arXiv.1912.12132.

[Ayzel et al., 2020] Ayzel, G., Scheffer, T., and Heistermann, M. (2020). RainNet v1.0: a convolutional neural network for radar-based precipitationnowcasting. *Geoscientific Model Development*, 13(6):2631–2644. DOI: 10.5194/gmd-13-2631-2020.

[Bierdel et al., 2012] Bierdel, L., Friederichs, P., and Bentzien, S. (2012). Spatial kinetic energy spectra in the convection-permitting limited-area NWP model COSMO-DE. *Meteorologische Zeitschrift*, 21(3):245–258. DOI: 10.1127/0941-2948/2012/0319.

[Brier, 1950] Brier, G. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1):1–3. DOI: 10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2.

[Dixon and Wiener, 1993] Dixon, M. and Wiener, G. (1993). TITAN: Thunderstorm Identification, Tracking, Analysis and Nowcasting - a radar-based methodology. *Journal of Atmospheric and Oceanic Technology*, 10(6):785–797. DOI: 10.1175/1520-0426(1993)010<0785:TTITAA>2.0.CO;2.

[Evensen, 1994] 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):10143–10162. DOI: 10.1029/94JC00572.

[Gao et al., 2021] Gao, Z., Shi, X., Wang, H., Yeung, D., Woo, W., and Wong, W. (2021). Deep learning and the weather forecasting problem: Precipitation nowcasting. In *Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences*. Wiley. DOI: 10.1002/9781119646181.

[Germann and Zawadzki, 2004] Germann, U. and Zawadzki, I. (2004). Scale dependence of the predictability of precipitation from continental radar images. part ii: Probability forecasts. *Journal of Applied Meteorology and Climatology*, 43(1):74–89. DOI: 10.1175/1520-0450(2004)043<0074:SDOTPO>2.0.CO;2.

[Haiden et al., 2011] Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., and Gruber, C. (2011). The Integrated Nowcasting through Comprehensive Analysis (INCA) system and its validation over the eastern alpine region. *Weather and forecasting*, 26(2):166–183. DOI: 10.1175/2010WAF2222451.1.

[Hunt et al., 2007] Hunt, B. R., Kostelich, E. J., and Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: a Local Ensemble Transform Kalman Filter. *Physica D (Nonlinear Phenomena)*, 230(1-2):112–126. DOI: 10.1016/j.physd.2006.11.008.

[Jazwinski, 1970] Jazwinski, A. H. (1970). Stochastic processes and filtering theory. *Academic Press*.

[Johnson et al., 1998] Johnson, J. T., MacKeen, P. L., Witt, A., De Wayne Mitchell, E., Stumpf, G. J., Eilts, M. D., and Thomas, K. W. (1998). The Storm Cell Identification and Tracking algorithm: An enhanced WSR-88D algorithm. *Weather and Forecasting*, 13(2):263–276. DOI: 10.1175/1520-0434(1998)013<0263:TSCIAT>2.0.CO;2.

[Julier and Uhlmann, 1997] Julier, S. J. and Uhlmann, J. K. (1997). New extension of the Kalman Filter to nonlinear systems. *Proc. SPIE 3068, Signal Processing, Sensor Fusion, and Target Recognition VI*. DOI: 10.1117/12.280797.

[Kalman, 1960] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45. DOI: 10.1115/1.3662552.

[Kann and Haiden, 2011] Kann, A. and Haiden, T. (2011). INCA – an operational nowcasting system for hydrology and other applications. *Berichte der Geologischen Bundesanstalt*, 88:10pp.

[Ko et al., 2022] Ko, J., Lee, K., Hwang, H., and Shi, K. (2022). Deep-learning-based precipitation nowcasting with ground weather station data and radar data. In *2022 IEEE International Conference on Data Mining Workshops (ICDMW)*. DOI:10.1109/ICDMW58026.2022.00138.

[Kober et al., 2013] Kober, K., Craig, G. C., and Keil, C. (2013). Aspects of short-term probabilistic blending in different weather regimes. *Quarterly Journal of the Royal Meteorological Society*, 140(681):1179–1188. DOI: 10.1002/qj.2220.

[Kober et al., 2011] Kober, K., Craig, G. C., Keil, C., and Dornbräck, A. (2011). Blending a probabilistic nowcasting method with a high-resolution numerical weather prediction ensemble for convective precipitation forecasts. *Quarterly Journal of the Royal Meteorological Society*, 138(664):755–768. DOI: 10.1002/qj.939.

[Lee et al., 2021] Lee, H., Kim, J., Kim, E. K., and Kim, S. (2021). A novel convective storm location prediction model based on machine learning methods. *Atmosphere*, 12(3). DOI: 10.3390/atmos12030343.

[Leinonen et al., 2022] Leinonen, J., Hamann, U., Germann, U., and Mecikalski, J. R. (2022). Nowcasting thunderstorm hazards using machine learning: the impact of data sources on performance. *Natural Hazards and Earth System Sciences*, 22(2):577–597. DOI: 10.5194/nhess-22-577-2022.

[Li et al., 1995] Li, L., Schmid, W., and Joss, J. (1995). Nowcasting of motion and growth of precipitation with radar over a complex orography. *Journal of Applied Meteorology and Climatology*, 34(6):1286–1300. DOI: 10.1175/1520-0450(1995)034<1286:NOMAGO>2.0.CO;2.

[Li et al., 2000] Li, P. W., Wong, W. K., Chan, K. Y., and Lai, E. S. T. (2000). SWIRLS - an evolving nowcasting system. Technical report, Hong Kong Observatory, Technical Note No. 100.

[Magnusson and Källén, 2013] Magnusson, L. and Källén, E. (2013). Factors influencing skill improvements in the ECMWF forecasting system. *Monthly Weather Review*, 141(9):3142–3153. DOI:10.1175/MWR-D-12-00318.1.

[Nerini et al., 2019] Nerini, D., Foresti, L., Leuenberger, D., Robert, S., and Germann, U. (2019). A reduced-space Ensemble Kalman Filter approach for flow-dependent integration of radar extrapolation nowcasts and NWP precipitation ensembles. *Monthly Weather Review*, 147(3):987–1006. DOI: 10.1175/MWR-D-18-0258.1.

[Panziera et al., 2011] Panziera, L., Germann, U., Gabella, M., and Mandapaka, P. V. (2011). NORA–nowcasting of orographic rainfall by means of analogues. *Quarterly Journal of the Royal Meteorological Society*, 137(661):2106–2123. DOI: 10.1002/qj.878.

[Park, 2023] Park, S. K. (2023). *Numerical Weather Predicion: East Asian Perspectives*. Springer Atmospheric Sciences, second edition.

[Pierce et al., 2012] Pierce, C., Seed, A., Ballard, S., Simonin, D., and Li, Z. (2012). Nowcasting. In *Doppler Radar Observations - Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications*. IntechOpen. DOI: 10.5772/2036.

[Piran et al., 2024] Piran, M. J., Wang, X., Kim, H. J., and Kwon, H. H. (2024). Precipitation nowcasting using transformer-based generative models and transfer learning for improved disaster preparedness. *International Journal of Applied Earth Observation and Geoinformation*, 132. DOI: 10.1016/j.jag.2024.103962.

[Poletti et al., 2019] Poletti, M. L., Silvestro, F., Davolio, S., Pignone, F., and Rebora, N. (2019). Using nowcasting technique and data assimilation in a meteorological model to improve very short range hydrological forecasts. *Hydrology and Earth System Sciences*, 23(9):3823–3841. DOI: 10.5194/hess-23-3823-2019.

[Pulkkinnen et al., 2019] Pulkkinnen, S., Nerini, D., Hortal, A. A. P., Velasco-Forero, C., Seed, A., Germann, U., and Foresti, L. (2019). Pysteps: an open-source python library for probabilisticprecipitation nowcasting (v1.0). *Geoscientific Model Development*, 12(10):4185–4219. DOI: 10.5194/gmd-12-4185-2019.

[Ravuri et al., 2021] Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., Prudden, R., Mandhane, A., Clark, A., Brock, A., Simonyan, K., Hadesll, R., Robinson, N. ., Clancy, E., Arribas,

A., and Mohames, S. (2021). Skilful precipitation nowcasting using deep generative models of radar. *Nature*, 597:672–677. DOI: 10.1038/s41586-021-03854-z.

[Raynaud et al., 2014] Raynaud, L., Pannekoucke, O., Arbogast, P., and Bouttier, F. (2014). Application of a Bayesian weighting for short-range lagged ensemble forecasting at the convective scale. *Quarterly Journal of the Royal Meteorological Society*, 141(687):459–468. DOI: 10.1002/qj.2366.

[Roberts and Lean, 2008] Roberts, N. M. and Lean, H. W. (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.

[Scheufele, 2010] Scheufele, K. (2010). *Probabilistic forecasting of precipitation by combining nowcasts and the time-lagged COSMO-DE ensemble*. PhD thesis, University of Munich.

[Schraff et al., 2016] Schraff, C., Reich, H., Rhodin, A., Schomburg, A., Stephan, K., Periáñez, A., and Potthast, R. (2016). Kilometre-Scale Ensemble Data Assimilation for the COSMO model (KENDA). *Quarterly journal of the Royal Meteorological Society*, 142(696):1453–1472. DOI: 10.1002/qj.2748.

[Wastl et al., 2021] Wastl, C., Wang, Y., Atencia, A., Weidle, F., Wittmann, C., Zingerle, C., and Keresturi, E. (2021). C-LAEF: Convection-permitting Limited-Area Ensemble Forecasting system. *Quarterly Journal of the Royal Meteorological Society*, 147(735):1431–1451. DOI:10.1002/qj.3986.

[Wastl et al., 2019] Wastl, C., Wang, Y., Atencia, A., and Wittmann, C. (2019). A hybrid stochastically perturbed parametrization scheme in a convection-permitting ensemble. *Monthly Weather Review*, 147(6):2217–2230. DOI: 10.1175/MWR-D-18-0415.1.

[Wilks, 2011] Wilks, D. S. (2011). *Statistical Methods in the Atmospheric Sciences*. Acedemic Press, third edition.

List of Figures

3.1. Outlines of the C-LAEF domain that was provided, the INCA domain according to the transformation, and the nowcasting domain that was used in this thesis. 17

4.1. This flowchart illustrates Equation 4.1. In this example, the ensemble forecast for 16 UTC is updated shortly after precipitation observations from 16 UTC in order to forecast precipitation probabilities for 17 UTC . 19
4.2. This example visualizes the application of the condition in Equation 4.2 on one of the precipitation forecasts in order to create binary probabilities. 20
4.3. Using the upscaling method of (a) on the field from Figure 4.2 b) yields the field of upscaled probabilities in (b) 21
4.4. The grey points represent grid points of the original C-LAEF grid. The red rectangle represents the grid box of the upscaled grid, spanning 5 by 5 grid points. The blue points show the INCA grid points within this particular upscaled grid box. 22

5.1. Set of ensemble, nowcast, and observation (truth) for weak precipitation on the early afternoon of July 21st. 28
5.2. Only a small part of the domain holds an overlap of forecasted and observed probability of precipitation 29
5.3. Initialisation of the most skillful nowcast 29
5.4. The background forecast in (a) is adapted to the nowcast in (b). The bottom panel shows the corresponding observation. 30
5.5. Initialisation of the least skillful 3h-nowcast 31
5.6. The background forecast in (a) is adapted to the nowcast in (b). The bottom panel shows the corresponding observation. 32
5.7. RMSE reduction, averaged over all nowcasts, compared to the underlying ensemble forecasts. 33
5.8. Scatter plot for all 1h nowcasts: Absolute RMSE reduction vs. domain-mean of probability of precipitation for different thresholds. The domain-mean probability corresponds to the spatial extent of precipitation patterns. 34
5.9. Spatial extent of precipitation vs RMSE (without correct negatives) . . . 35
5.10. Skill evolution with lead time compared to the ensemble for different observation errors and all three thresholds 36
5.11. Relative RMSE-reduction compared to the respective ensemble-mean on the standard grid, compared to the coarser 10×10 grid. 37

5.12. Evaluation of equation 3.7 for all hours on thresholds τ_2 and τ_3, ... 37

A. Appendix

	$\Delta t = 1h$	$\Delta t = 2h$	$\Delta t = 3h$	$\Delta t = 4h$
$RMSE_{C-LAEF}$	0.066	0.067	0.067	0.068
$RMSE_{Nowc}$	0.064	0.067	0.068	0.068
Relative reducion[%]	3.0	-0.6	-1.1	-0.8
Absolute reduction	0.002	0.000	-0.001	-0.001
$\tilde{RMSE}_{C-LAEF}$	0.152	0.153	0.156	0.158
$\tilde{RMSE}_{Nowc}$	0.148	0.154	0.157	0.159
Relative reduction [%]	2.7	-1.0	-1.0	-0.6
Absolute reduction	0.004	-0.001	-0.002	-0.001

Table A.1.: Average RMSE for all Ensembles / Nowcasts for time-shifts of 1 to 4 hours for a threshold of $\tau_1 \geq 5mm/h$. Top half: RMSE for all points, bottom half: RMSE without correct negatives

	$\Delta t = 1h$	$\Delta t = 2h$	$\Delta t = 3h$	$\Delta t = 4h$
$RMSE_{CLAEF}$	0.250	0.252	0.254	0.257
$RMSE_{Nowc}$	0.215	0.247	0.258	0.261
Relative reducion[%]	13.9	1.7	-1.4	-1.7
Absolute reduction	0.035	0.004	-0.004	-0.004
$\tilde{RMSE}_{CLAEF}$	0.308	0.309	0.313	0.317
$\tilde{RMSE}_{Nowc}$	0.267	0.304	0.317	0.323
Relative reduction [%]	13.2	1.6	-1.4	-1.7
Absolute reduction	0.041	0.005	-0.004	-0.005

Table A.2.: Average RMSE for all Ensembles / Nowcasts for time-shifts of 1 to 4 hours for a threshold of $\tau_3 \geq 0.1mm/h$. Top half: RMSE for all points, bottom half: RMSE without correct negatives