https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

Nonlinear Processes in Geophysics
Discussions
Open Access

Ensemble-based statistical interpolation with Gaussian anamorphosis for the spatial analysis of precipitation

Cristian Lussana[1], Thomas N. Nipen[1], Ivar A. Seierstad[1], and Christoffer A. Elo[1]

[1]Norwegian Meteorological Institute, Oslo, Norway

Correspondence: Cristian Lussana (cristianl@met.no)

Abstract. Hourly precipitation over a region is often simultaneously simulated by numerical models and observed by multiple data sources. An accurate precipitation representation based on all available information is a valuable result for numerous applications and a critical aspect of climate. Inverse problem theory offers an ideal framework for the combination of observations with a numerical model background. In particular, we have considered a modified ensemble optimal interpolation scheme, that takes into account deficiencies of the background. An additional source of uncertainty for the ensemble background has been included. A data transformation based on Gaussian anamorphosis has been used to optimally exploit the potential of the spatial analysis, given that precipitation is approximated with a gamma distribution and the spatial analysis requires normally distributed variables. For each point, the spatial analysis returns the shape and rate parameters of its gamma distribution. The Ensemble-based Statistical Interpolation scheme with Gaussian AnamorPhosis (EnSI-GAP) is implemented in a way that the covariance matrices are locally stationary and the background error covariance matrix undergoes a localization process. Concepts and methods that are usually found in data assimilation are here applied to spatial analysis, where they have been adapted in an original way to represent precipitation at finer spatial scales than those resolved by the background, at least where the observational network is dense enough. The EnSI-GAP setup requires the specification of a restricted number of parameters and specifically the explicit values of the error variances are not needed, since they are inferred from the available data. The examples of applications presented provide a better understanding of the characteristics of EnSI-GAP. The data sources considered are those typically used at national meteorological services, such as local area models, weather radars and in-situ observations. For this last data source, measurements from both traditional and opportunistic sensors have been considered.

1 Introduction

Precipitation amounts are measured or estimated simultaneously by multiple observing systems, such as networks of automated weather stations and remote sensing instruments. At the same time, sophisticated numerical models simulating the evolution of the atmospheric state provide a realistic precipitation representation over regular grids with spacing of a few kilometers. An unprecedented amount of rainfall data is nowadays available at very short sampling rates of one hour or less. Nevertheless, it is common experience within national meteorological services that the exact amount of precipitation, to some extent, eludes our knowledge. There may be numerous reasons for this uncertainty. For example, a thunderstorm triggering a landslide may have occurred in a region of complex topography where in-situ observations are available but not exactly on the landslide spot, weather radars cover the region in a patchy way because of obstacles blocking the beam, and numerical weather prediction forecasts are likely misplacing precipitation maxima. Another typical situation is when an intense and localized summer thunderstorm hits a city. In this case, several observation systems are measuring the event and more than one numerical model may provide precipitation totals. From this plurality of data, a detailed reconstruction of the event is possible, provided that the data agrees both in terms of the event intensity and on its spatial features. This is not always the case and sometimes meteorologists and hydrologists are left with a number of slightly different but plausible scenarios.

The objective of our study is the precipitation reconstruction through the combination of numerical model output with observations from multiple data sources. The aim is that the combined fields will provide a more skillful prediction than any of the original data sources. As remarked above, any improvement in the accuracy and precision of precipitation can be of great help for monitoring the weather, but not only that. Snow- and hydrological- modeling will benefit from improvements in the quality of precipitation, which is one of the atmospheric forcing variables (Magnusson et al., 2019; Huang et al., 2019). Climate applications that make use of reanalysis (e.g. Hersbach et al., 2020; Jermey and Renshaw, 2016) or observational gridded datasets (e.g. Lussana et al., 2018), as for instance the evaluation of regional climate model (Kotlarski et al., 2017) or the calculation of climate indices (Vicente-Serrano et al., 2015), may also benefit from datasets combining model output and observations, as shown by Fortin et al. (2018). Besides, the intensity-duration-frequency curve (IDF curve) derived from precipitation datasets are widely used in civil engineering for determining design values and the quality of the reconstruction of extremes has a strong influence on IDF curves (Dyrrdal et al., 2015).

The data source considered in our study are precipitation ensemble forecasts, observations from in-situ measurement stations and estimates derived from weather radars. Numerical model fields are available everywhere and the quality of their output is constantly increasing over the years. The weather-dependent uncertainty is often delivered in the form of an ensemble. At present for data-sparse regions, such as in the mountains, the quality of numerical model output in modeling rain and snow is comparable or even better than those of observational networks (Lundquist et al., 2019). However, because model fields represent areal averages, the characteristics of simulated precipitation depend significantly on the model resolution, as remarked for global and regional reanalyses over the Alps by Isotta et al. (2019). In particular, Jermey and Renshaw (2016) demonstrates that increasing resolution via downscaling improves precipitation representation, though they also point out that assimilating observations at high resolution in numerical models is important for reconstructing high-threshold/small-scale events. The sources of model errors and their treatments in data assimilation (DA) schemes have been studied extensively. For instance, in the introduction of the paper by Raanes et al. (2015), a list of model errors is reported together with several references to other studies addressing them. Regarding precipitation forecasts, model errors often encountered in applications are (Müller et al., 2017): systematic under- or overestimation of amounts; spatial errors in the placement of events; underestimation of uncertainty. With reference to spatial analysis, we consider observed precipitation data to be more accurate than model estimates. In fact,

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

model outputs are evaluated in their ability to reconstruct observed values. The most important disadvantage of observational networks is that often they do not cover the region under consideration, moreover observations may be irregularly distributed in space and present missing data over time. Each observational data source has its own characteristics that have been extensively studied in literature and that we will address here only superficially, since our objective is the combination of information. For example, rain gauges are possibly the most accurate precipitation measurement available at present (CIMO, 2014), apart from when the observations are affected by gross measurement errors, as they have been defined by Gandin (1988). There are multiple sources of uncertainty for gauge measurements (Zahumensky, 2004), such as catching and counting (Pollock et al., 2014). The undercatch of solid precipitation due to wind (Wolff et al., 2015) is a significant problem in cold climates. Radar-derived estimates are affected by several issues, such as blocking and non-uniform attenuation of the radar beam due to obstacles along the path, especially in complex terrain. A statement in the Introduction of the book by Germann and Joss (2004) is illuminating in this sense ”To put a weather radar in a mountainous region is like pitching a tent in a snowstorm: the practical use is obvious and large — but so are the problems”. In addition, weather radars do not directly measure precipitation, instead they measure reflectivity, which is then transformed into precipitation rate. The transformation itself contributes to increasing the uncertainty of the final estimates. Another important aspect of observational data that will be treated only marginally here is data quality control, in this work we will consider only quality controlled observations. To sum up, in-situ data are the more accurate observations of precipitation we will consider. Then, radar estimates, which are calibrated using gauges as references, are less accurate than in-situ data. They are spatially correlated with the actual precipitation and they are affected by less uncertainty than the simulations carried out by numerical models. Numerical model output is the basic information available everywhere and the one we consider more uncertain.

Inverse problem theory (Tarantola, 2005) provides the ideal framework for the combination of observations with a numerical model background. The marginal distribution of the precipitation analysis is assumed to be a gamma distribution and we aim at estimating its shape and rate parameters for each grid point. The statistical interpolation method presented is similar to the ensemble Optimal Interpolation (Evensen, 1994). The hourly precipitation over the grid is regarded as the realization of trans-gaussian random field (Frei and Isotta, 2019). The Gaussian anamorphosis (Bertino et al., 2003) transforms data such that precipitation better complies with the assumptions of normality that are required by the analysis procedure. The non-stationary covariance matrices are approximated with locally stationary matrices, as in the paper by Kuusela and Stein (2018). In addition, the background error covariance matrix includes a scale matrix that accounts for deficiencies in the background ensemble as in hybrid ensemble optimal interpolation (Carrassi et al., 2018). The term scale matrix has been used by Bocquet et al. (2015). The ensemble-based statistical interpolation with Gaussian anamorphosis for the spatial analysis of precipitation is referred to in the following with the acronym of EnSI-GAP.

The innovative part of the presented approach to statistical interpolation is in the application to spatial analysis of concepts that are usually encountered in DA. The formulation of the problem is adapted to our aim, which is improving precipitation representation instead of providing initial conditions for a physical model, as it is for DA. In the literature, there are a number of articles describing similar approaches applied to precipitation analysis, such as Mahfouf et al. (2007); Soci et al. (2016); Lespinas et al. (2015). However, our statistical interpolation is the first one, that we know of, where the background error

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

covariance matrix is derived from numerical model ensemble and where Gaussian anamorphosis is applied directly to precipitation data. An additional innovative part of the method is that EnSI-GAP does not require the explicit specification of error variances for the background or observations, as most of the other methods (Soci et al., 2016). In fact, those error variances are often difficult to estimate in a way that is general enough to cover a wide range of cases. Our approach is to specify the reliability of the background with respect to observations, in such a way that error variances can vary both in time and space. An additional innovative part of our research is that we consider opportunistic sensing networks of the type described by de Vos et al. (2020) within the examples of applications proposed. Thanks to those networks, for some regions we can rely on an extremely dense spatial distribution of in-situ observations.

The remaining of the paper is organized as follows. Sec. 2 describes the EnSI-GAP method in a general way, without references to specific data sources. Sec. 3 presents the results of EnSI-GAP applied to three different problems: an idealized situation, then two examples where the method is applied to real data, such as those mentioned above. The results are discussed in Sec. 3.3.

2 Methods: EnSI-GAP, Ensemble-based statistical interpolation with Gaussian anamorphosis for precipitation

We assume that the marginal probability density function (PDF) for the hourly precipitation at a point in time follows a gamma distribution (Wilks, 2019). This marginal PDF is characterized through the estimation of the gamma shape and rate for each point and hour.

Precipitation fields are regarded as realizations of locally-stationary, trans-Gaussian random fields, where each hour is considered independently from the others. Trans-Gaussian random fields are used for the production of precipitation observational gridded datasets by Frei and Isotta (2019). A random field is said to be stationary if the covariance between a pair of points depends only on how far apart they are located from each other. Precipitation totals are nonstationary random fields because the covariance between a pair of points in space depends not only on the distance between them but it varies also when considered in different directions. In our method, precipitation is locally modeled as a stationary random field. The covariance parameter estimation and spatial analysis are carried out in a moving-window fashion around each grid point. A similar approach is described by Kuusela and Stein (2018) and the elaboration over the grid can be carried out in parallel for several grid points simultaneously.

A particular implementation of EnSI-GAP is reported in Algorithm 1. The mathematical notation and the symbols used are described in two tables: Tab. 1 for global variables and Tab. 2 for local variables, which are those variables that vary from point to point. As in the paper by Sakov and Bertino (2011), upper accents have been used to denote local variables. If $\mathbf{X}$ is a matrix, $\mathbf{X}_i$ is its ith column (column vector) and $\mathbf{X}_{i,:}$ is its ith row (row vector). The Bayesian statistical method used in our spatial analysis is optimal for Gaussian random fields. Then, a data transformation is applied as a pre-processing step before the spatial analysis. The introduction of a data transformation compels us to inverse transform the predictions of the spatial analysis back into the original space of precipitation values. Algorithm 1 can be divided into three parts, that are described in the next sections: the data transformation in Sec. 2.1, the Bayesian spatial analysis in Sec. 2.2 and the inverse transformation in Sec. 2.3.

2.1 Data transformation via Gaussian anamorphosis

The data transformation chosen is a Gaussian anamorphosis (Bertino et al., 2003), that transforms a random variable following a gamma distribution into a standard Gaussian. This pre-processing strategy has been used in several studies in the past, e.g. Amezcua and Leeuwen (2014); Lien et al. (2013). A visual representation of the transformation process can be found in Fig.1 of the paper by Lien et al. (2013).

The hourly precipitation background and observations, $\tilde{\mathbf{X}}^{\mathrm{f}}$ and $\tilde{\mathbf{y}}^{\mathrm{o}}$ respectively, are transformed into those used in the spatial analysis by means of the Gaussian anamorphosis $g()$:

$$\mathbf{X}^{\mathrm{f}} = g(\tilde{\mathbf{X}}^{\mathrm{f}}) \quad (1)$$

$$\mathbf{y}^{\mathrm{o}} = g(\tilde{\mathbf{y}}^{\mathrm{o}}) \quad (2)$$

The gamma shape and rate are derived from the precipitation values by a fitting procedure based on maximum likelihood. The maximum likelihood estimators are calculated iteratively by means of a Newton-Raphson method as described by Wilks (2019), Sec. 4.6.2. In particular, the gamma distribution parameters are fitted to each ensemble member field of precipitation separately. Then, the averaged shape and rate are used in $g()$ for Eqs. (1)-(2).

In Gaussian anamorphosis, zero precipitation values must be treated as special cases, as explained by Lien et al. (2013). The solution we adopted is to add a very small amount to zero precipitation values, e.g. 0.0001 mm, then to apply the transformation $g()$ to all values. The same small amount is then subtracted after the inverse transformation. This is a simple but effective solution for spatial analysis, as shown in the example of Sec. 3.1.1. In principle, the statistical interpolation is sensitive to the small amount chosen, such that using 0.01 mm instead of 0.0001 mm will return slightly different analysis values in the transition between precipitation and no-precipitation. In practice, we have tested it and we found negligible differences when values smaller than e.g. 0.05 mm (half of the precision of a standard rain gauge measurement) have been used.

2.2 Spatial analysis

The spatial analysis inside Algorithm 1 has been divided into three parts. In Sec. 2.2.1, global variables have been defined. Then, as stated in the introduction of Sec. 2, the analysis procedure is performed on a gridpoint-by-gridpoint basis. In Sections 2.2.2-2.2.3, the procedure applied at the generic ith gridpoint is described. In Sec. 2.2.2, the specification of the local error covariance matrices is described. In Sec. 2.2.3, the standard analysis procedure is presented together with the treatment of a special case.

2.2.1 Definitions

In Bayesian statistics, according to Savage (1972), a state is ”a description of the world, which is the object which we are concerned, leaving no relevant aspect undescribed” and ”the true state is the state that does in fact obtain”. The object of our study is the hourly precipitation field $x()$, that is the hourly total precipitation amount over a continuous surface covering a spatial domain in terrain-following coordinates $\mathbf{r}$. Our state is the discretization over a regular grid of this continuous field. The true state (our "truth", $\mathbf{x}^{\mathrm{t}}$) at the ith grid point is the areal average:

$$\mathbf{x}_i^{\mathrm{t}} = \int\limits_{V_i} x(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{3}$$

where V_i is a region surrounding the ith grid point. The size of V_i determines the effective resolution of $\mathbf{x}^{\mathrm{t}}$ at the ith grid point. Our aim is to represent the truth with the smallest possible V_i. The effective resolution of the truth will inevitably vary across the domain. In observation-void regions, the effective resolution will be the same as that of the numerical model used as the background, then approximately $o(10-100\,km^2)$ for high-resolution local area models (Müller et al., 2017). In observation-dense regions, the effective resolution should be comparable to the average distance between observation locations, with the model resolution as the upper bound.

The analysis is the best estimate of the truth, in the sense that it is the linear, unbiased estimator with the minimum error variance. The analysis is defined as $\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{t}} + \boldsymbol{\eta}^{\mathrm{a}}$, where the column vector of the analysis error at grid points is a random variable following a multivariate normal distribution $\boldsymbol{\eta}^{\mathrm{a}} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}^{\mathrm{a}})$. The marginal distribution of the analysis at the ith grid point is a normal random variable and our statistical interpolation scheme returns its mean value $\mathbf{x}_i^{\mathrm{a}}$ and its standard deviation $\sigma_i^{\mathrm{a}} = \sqrt{\mathbf{P}_{ii}^{\mathrm{a}}}$.

As for linear filtering theory (Jazwinski, 2007), the analysis is obtained as a linear combination of the background (a priori information) and the observations. The background is written as $\mathbf{x}^{\mathrm{b}} = \mathbf{x}^{\mathrm{t}} + \boldsymbol{\eta}^{\mathrm{b}}$, where the background error is a random variable $\boldsymbol{\eta}^{\mathrm{b}} \sim N\left(\mathbf{0}, \mathbf{P}^{\mathrm{b}}\right)$. The background PDF is determined mostly, but not exclusively, by the forecast ensemble, as described in Sec. 2.2.1. The forecast ensemble mean is $\mathbf{x}^{\mathrm{f}} = k^{-1}\mathbf{X}^{\mathrm{f}}\mathbf{1}$, where $\mathbf{1}$ is the m-vector with all elements equal to 1. The background expected value is set to the forecast ensemble mean, $\mathbf{x}^{\mathrm{b}} = \mathbf{x}^{\mathrm{f}}$. The forecast perturbations are $\mathbf{A}^{\mathrm{f}}$, where the ith perturbation is $\mathbf{A}_i^{\mathrm{f}} = \mathbf{X}_i^{\mathrm{f}} - \mathbf{x}^{\mathrm{f}}$. The covariance matrix:

$$\mathbf{P}^{\mathrm{f}} = (k-1)^{-1}\mathbf{A}^{\mathrm{f}}\mathbf{A}^{\mathrm{f}\,\mathrm{T}} \tag{4}$$

plays a role in the determination of $\mathbf{P}^{\mathrm{b}}$, as defined in Sec. 2.2.2.

The p observations are written as $\mathbf{y}^{\mathrm{o}} = \mathbf{H}\mathbf{x}^{\mathrm{t}} + \varepsilon^{\mathrm{o}}$, where the observation error is $\varepsilon^{\mathrm{o}} \sim N(\mathbf{0}, \mathbf{R})$ and $\mathbf{H}$ is the observation operator, that we consider as a linear function mapping $\mathbb{R}^m$ onto $\mathbb{R}^p$.

2.2.2 Specification of the observation and background error covariance matrices

Our definitions of the error covariance matrices follow from a few working assumptions, WAn indicates the nth working assumption and the abbreviations will be used in the text. WA1, background and observation uncertainties are weather- and location- dependent. WA2, the background is more uncertain where either the forecast is more uncertain or observations and forecasts disagree the most. WA3, observations are a more accurate estimate of the true state than the background. We want to specify how much more we trust the observations than the background in a simple way, such as e.g. "we trust the observations twice as much as the background". WA4, the local observation density must be used optimally to ensure a higher effective resolution, as it has been defined in Sec. 2.2.1, where more observations are available. WA5, the spatial analysis at a particular hour does not require the explicit knowledge of observations and forecasts at any other hour. However, constants in the covariance matrices can be set depending on the history of deviations between observations and forecasts. WA5 makes the procedure more robust and easier to implement in real-time operational applications.

A distinctive feature of our spatial analysis method is that the background error covariance matrix $\overset{i}{\mathbf{P}}^{\mathrm{b}}$ is specified as the sum of two parts: a dynamical component and a static component. The dynamical part introduces nonstationarity, while the static part describes covariance stationary random variables. This choice follows from WA1 and it has been inspired by hybrid data assimilation methods (Carrassi et al., 2018). The dynamical component of the background error covariance matrix is obtained from the forecast ensemble. Because the ensemble has a limited size, and often the number of members is quite small (order of tenths of member), a straightforward calculation of the background covariance matrix will include spurious correlations between distant points. Localization is a technique applied in DA to fix this issue (Greybush et al., 2011). The static component has also been introduced to remedy the shortcomings of using numerical weather prediction as the background. There are deviations between observations and forecasts that cannot be explained by the forecast ensemble. A typical example is when all the ensemble members predict no precipitation but rainfall is observed. In those cases, we trust observations, as stated through WA3. Then, the static component adds noise to the model-derived background error, as in the paper by Raanes et al. (2015). In Bocquet et al. (2015), the static component is referred to as a scale matrix, since it is used to scale the noise component of the model error, and we adopt the same term here. We will also refer to this matrix, and its related quantities, with the letter *u* to emphasize that this component of the background error is ”unexplained” by the forecast.

$\overset{i}{\mathbf{P}}^{\mathrm{b}}$ is written as:

$$\overset{i}{\mathbf{P}}^{\mathrm{b}} = \overset{i}{\mathbf{\Gamma}} \circ \mathbf{P}^{\mathrm{f}} + \overset{i}{\sigma}{}_u^2 \overset{i}{\mathbf{\Gamma}}{}^{\mathrm{u}} \tag{5}$$

The first component on the right-hand side of Eq. (5) is the dynamical part. $\mathbf{P}^{\mathrm{f}}$ is the forecast uncertainty of Eq. (4), $\overset{i}{\mathbf{\Gamma}}$ is the localization matrix and $\circ$ is the Schur product symbol. The localization technique we apply is a combination of local analysis and covariance localization, as they have been defined by Sakov and Bertino (2011). In the local analysis, only the closest observations are used and we have implemented it by considering only observations within a predefined spatial window surrounding each grid point, up to a pre-set maximum number of p_{mx}. The covariance localization is implemented through the element-wise multiplication of $\mathbf{P}^{\mathrm{f}}$ by $\overset{i}{\mathbf{\Gamma}}$, which has the form of a correlation matrix that depends on distances and that is used to suppress long-range correlations. The second component on the right-hand side of Eq. (5) is the static part. The scale matrix is expressed through a constant variance $\overset{i}{\sigma}{}_u^2$, that modulates the noise, and the correlation matrix $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{u}}$ defining the spatial structure of that noise. In the examples of applications presented in Sec. 3, both $\overset{i}{\mathbf{\Gamma}}$ and $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{u}}$ are obtained as analytical functions of the spatial coordinates. In Algorithm 1, $\overset{i}{\mathbf{\Gamma}}$ and $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{u}}$ have been specified through Gaussian functions, other possibilities for correlation functions have been described for instance by Gaspari and Cohn (1999). We have chosen not to inflate or deflate $\mathbf{P}^{\mathrm{f}}$ directly and to modulate the amplitude of background covariances only through the terms of Eq. (5), this way we reduce the number of parameters that need to be specified. As a matter of fact, for the combination of observations and background in the analysis procedure, the m by m covariance matrices are never directly used. Instead, the matrices used are: the covariances between grid points and observation locations, $\overset{i}{\mathbf{G}}{}^{\mathrm{b}} = \overset{i}{\mathbf{P}}{}^{\mathrm{b}}\overset{i}{\mathbf{H}}{}^{\mathrm{T}}$, specifically only the ith row of this matrix is used; and the covariances between observation locations $\overset{i}{\mathbf{S}}{}^{\mathrm{b}} = \overset{i}{\mathbf{H}}\overset{i}{\mathbf{P}}{}^{\mathrm{b}}\overset{i}{\mathbf{H}}{}^{\mathrm{T}}$. $\overset{i}{\mathbf{H}}$ is the local observation operator, that is a linear function: $\mathbb{R}^{m} \rightarrow \mathbb{R}^{p_i}$.

The local observation error covariance matrix $\overset{i}{\mathbf{R}}$ is written as the constant observation error variance $\overset{i}{\sigma}{}^{2}_{o}$ multiplying the correlation matrix $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{o}}$:

$$\overset{i}{\mathbf{R}} = \overset{i}{\sigma}{}^{2}_{o}\,\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{o}} \tag{6}$$

$\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{o}}$ often is the identity but other choices are possible. For instance, if some observations are know to be more accurate than the average of the others, then the corresponding diagonal elements of $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{o}}$ can be set to values smaller than 1. The observation uncertainty can vary in time and space, accordingly to WA1, however its spatial structure is fixed and depends on the analytical function chosen for $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{o}}$. Note that the observation error is not determined by the instrumental error only but it includes the representativeness error (Lussana et al., 2010; Lorenc, 1986), which is often the largest component of the observation error. The representative error is a consequence of the mismatch between the spatial supports of the areal averages reconstructed by the background and the almost point-like observations.

The spatial structures of the error covariance matrices are determined through Eqs. (5)- (6). At this point, we need to scale the covariances to satisfy our WA2 and the first step is to define the background error variance in a way that is in line with WA3. We prefer to specify the ratio between variances instead of the variances themselves. Then, we introduce the global variable ε^2 as:

$$\varepsilon^2 = \overset{i}{\sigma}{}^{2}_{o} / \overset{i}{\sigma}{}^{2}_{b} \tag{7}$$

where $\overset{i}{\sigma}{}^{2}_{b}$ is the average background error variance in the surroundings of the ith grid point. ε^2 is used to express the level of confidence we have in the observations with respect to the background and it should be set to a value smaller than 1 (WA3). For example, $\varepsilon^2 = 0.1$ means that we believe the observations to be ten times more precise an estimate of the true value than the background.

The definition of $\overset{i}{\sigma}{}^{2}_{b}$ is an adaptation from Eq. (5):

$$\overset{i}{\sigma}{}^{2}_{b} = \overset{i}{\sigma}{}^{2}_{f} + \overset{i}{\sigma}{}^{2}_{u} \tag{8}$$

where the average forecast error variance at the ith grid point $\overset{i}{\sigma}{}^{2}_{f}$ is defined as:

$$\overset{i}{\sigma}{}^{2}_{f} = \alpha \; \langle \mathrm{diag}\left(\overset{i}{\mathbf{S}}{}^{\mathrm{f}}\right) \rangle = \alpha \; \frac{1}{p_i} \sum_{j=1}^{p_i} \overset{i}{\mathbf{S}}{}^{\mathrm{f}}{}_{jj} \tag{9}$$

The mean over an ensemble of similar realizations $\langle \ldots \rangle$ is interpreted as the mean over the diagonal elements of $\overset{i}{\mathbf{S}}{}^{\mathrm{f}}$. $\mathrm{diag}\,(\ldots)$ stands for the vector composed by the diagonal elements of the matrix in parentheses. α is a stabilization factor introduced to make the estimation of $\overset{i}{\sigma}{}^{2}_{f}$ more robust. A proper estimation of $\overset{i}{\sigma}{}^{2}_{f}$ would require more than just one case, the ideal situation would be to include in the average of Eq. (9) numerous situations characterized by similar weather conditions. For the reasons discussed in WA5, we prefer to introduce α, which can be optimized as described in Sec. 3.

Considering the definition of $\overset{i}{\sigma}{}^2_b$ given by Eq. (8), we can proceed with the implementation of WA2. The background values at observation locations are $\overset{i}{\mathbf{y}}{}^{\mathrm{b}} = \overset{i}{\mathbf{H}}\mathbf{x}^{\mathrm{b}}$. WA2 is realized by imposing that the sum of observation and background error variances in the surroundings of a point is proportional to the local mean squared innovation (i.e. observations minus background), that is:

$$\overset{i}{\sigma}{}^2_{ob} = \overset{i}{\sigma}{}^2_o + \overset{i}{\sigma}{}^2_b = \alpha \left\langle \left(\overset{i}{\mathbf{y}}{}^{\mathrm{o}} - \overset{i}{\mathbf{y}}{}^{\mathrm{b}}\right)^2 \right\rangle \tag{10}$$

a similar equation is used for diagnostics in DA (Desroziers et al., 2005). In Eq. (10), the same stabilization factor α introduced for Eq. (9) has been used, since the reasons that lead to the use of a stabilization factor are the same in both equations and we want to have as few parameters as possible to optimize. By substituting Eqs. (7)- (8)- (9) into Eq. (10), the value of $\overset{i}{\sigma}{}^2_u$ that satisfies Eq. (10) is obtained as:

$$\overset{i}{\sigma}{}^2_u = \frac{\overset{i}{\sigma}{}^2_{ob}}{1+\varepsilon^2} - \overset{i}{\sigma}{}^2_f, \quad \text{if} \quad \overset{i}{\sigma}{}^2_f < \overset{i}{\sigma}{}^2_{ob}/(1+\varepsilon^2) \tag{11}$$

$$\overset{i}{\sigma}{}^2_u = 0, \quad \text{if} \quad \overset{i}{\sigma}{}^2_f \geq \overset{i}{\sigma}{}^2_{ob}/(1+\varepsilon^2) \ \text{ and } \ \overset{i}{\sigma}{}^2_f > 0 \tag{12}$$

when the forecast ensemble spread is too small, then Eq. (11) ensures that Eq. (10) is valid by setting $\overset{i}{\sigma}{}^2_b$ to $\overset{i}{\sigma}{}^2_{ob}/(1+\varepsilon^2)$, as can be seen by substituting Eq. (11) into Eq. (8). On the other hand, when $\overset{i}{\sigma}{}^2_f$ is positive and large enough, then we have decided that the background uncertainty is defined only through the ensemble, thus Eq. (12) is valid and the scale matrix, that is the second component on the right-hand side of Eq. (5), is not needed anymore.

The special case of $\overset{i}{\sigma}{}^2_f = 0$ and $\overset{i}{\sigma}{}^2_{ob}/(1+\varepsilon^2) = 0$ (or $\left\langle \left(\overset{i}{\mathbf{y}}{}^{\mathrm{o}} - \overset{i}{\mathbf{y}}{}^{\mathrm{b}}\right)^2 \right\rangle = 0$) will be considered in Sec. 2.2.3.

2.2.3 Analysis procedure

The expressions for the analysis and its error variance are direct results of the linear filter theory (Jazwinski, 2007) and they are derived in several books based on different formulations (e.g., Tarantola (2005); Kalnay (2003); Carrassi et al. (2018)). The analysis at the ith grid point is equal to the background plus a weighted average of the $\overset{i}{p}$ innovations, while the analysis error variance is derived from the error covariance matrices:

$$\mathbf{x}^{\mathrm{a}}_i = \mathbf{x}^{\mathrm{b}}_i + \overset{i}{\mathbf{G}}{}^{\mathrm{b}}_{i,:} \left(\overset{i}{\mathbf{S}}{}^{\mathrm{b}} + \overset{i}{\mathbf{R}}\right)^{-1} \left(\overset{i}{\mathbf{y}}{}^{\mathrm{o}} - \overset{i}{\mathbf{y}}{}^{\mathrm{b}}\right) \tag{13}$$

$$(\boldsymbol{\sigma}^2)^{\mathrm{a}}_i = \overset{i}{\mathbf{P}}{}^{\mathrm{b}}_{ii} - \overset{i}{\mathbf{G}}{}^{\mathrm{b}}_{i,:} \left(\overset{i}{\mathbf{S}}{}^{\mathrm{b}} + \overset{i}{\mathbf{R}}\right)^{-1} \left(\overset{i}{\mathbf{G}}{}^{\mathrm{b}}_{i,:}\right)^{\mathrm{T}} \tag{14}$$

Eq. (13)- (14) are also typical of Optimal Interpolation and the formulation used is similar to the one adopted by Uboldi et al. (2008), which follows from Ide et al. (1997). Because an ensemble of fields is used to determine the background error, the method is similar to Ensemble OI (Evensen, 1994).

The special case of $\overset{i}{\sigma}{}^2_f = 0$ and $\overset{i}{\sigma}{}^2_{ob}/(1+\varepsilon^2) = 0$ still needs to be considered, since it does not belong to the cases of Eqs. (11)- (12). This is the case when both observations and forecast ensemble means at observation locations have exactly the same values and, in addition, all the ensemble members have the same values too. In practice, this might happen only when no precipitation is observed and forecasted at the same time. Since the innovation is a vector of zeros, from Eq. (13) it follows that

$\mathbf{x}_i^{\mathrm{a}} = \mathbf{x}_i^{\mathrm{b}}$. Because all the information available shows an exceptional level of agreement, we have chosen to set the analysis error variance to zero, such that for those points the analysis PDFs are Dirac's delta functions.

2.3 Data inverse transformation

The direct inverse transformation at the ith grid point is written as:

$$\tilde{\mathbf{x}}_i^{\mathrm{a}} = g^{-1}(\mathbf{x}_i^{\mathrm{a}}) \tag{15}$$

however, we need to back-transform a Gaussian PDF and not a scalar value. Eq. (15) returns an estimate of the median of the gamma distribution associated to the ith grid point. Our goal is to obtain the gamma shape and rate. To achieve that, the direct inverse transformation g^{-1} is applied to 400 quantiles of the univariate gaussian PDF defined by $\tilde{\mathbf{x}}_i^{\mathrm{a}}$ and $(\boldsymbol{\sigma}^2)_i^{\mathrm{a}}$, a similar approach is used by Erdin et al. (2012). Then, a least-mean-square optimization procedure is used to obtain the optimal shape and rate that better fits the back-transformed quantiles. Given the optimal shape and rate, it is possible to obtain the statistics that better represent the distribution for a specific application. In Sec. 3, the analysis value chosen is the mean as it is the value that minimizes the spread of the variance. However, other choices may be more convenient depending on the applications, as discussed by Fletcher and Zupanski (2006) where, for instance, the mode was chosen as the best estimate. In Sec. 3, we will also consider selected quantiles of the gamma distribution to represent analysis uncertainty.

3 Results

EnSI-GAP is applied to two case studies in Sec. 3.1, one on synthetic and one on real-world data. In addition, its performances are evaluated over several months in Sec. 3.2. Then, a discussion of the results is presented in Sec. 3.3.

In Sec. 3.1.1, EnSI-GAP is applied over a one-dimensional grid and in a "controlled environment", that is over synthetic data specifically generated for testing EnSI-GAP on precipitation. The spatial analysis is performed with and without data transformation to assess its effects. Furthermore, we have compared different ways of specifying the scale matrix and we have investigated the sensitivity of EnSI-GAP to variations of α and ε^2.

In Sec. 3.1.2, a second more realistic example of application for EnSI-GAP is reported, where the spatial analysis is performed for a case study of intense precipitation, which happened in July 2019 over South Norway. The data are those used in the operational daily routine at MET Norway. The forecasts are from the MetCoOp Ensemble Prediction System (MEPS, Frogner et al., 2019). MEPS has been running operationally four times a day (00 UTC, 06 UTC, 12 UTC, 18 UTC) since November 2016 and its ensemble consists of 10 members. The hourly precipitation fields are available over a regular grid of 2.5 km. The observational dataset of hourly precipitation is composed of two data sources: precipitation estimates derived from the composite of MET Norway's weather radar and meteorological weather stations equipped with ombrometers, such as rain gauges or other devices. The hourly precipitation in-situ observations have been retrieved from MET Norway's climate database frost.met.no (last time the website was checked is 2020-05-13). In addition to MET Norway's official weather stations, the database includes data collected by several Norwegian public institutions, such as for example: universities (e.g., the

Norwegian Institute of Bioeconomy Research - Nibio), the Norwegian Water Resources and Energy Directorate (NVE), the Norwegian Public Roads Administration (Statens vegvesen). As described in the recent paper by Nipen et al. (2020), MET Norway is successfully integrating amateur weather stations temperature data into its operational routine. In this study, hourly precipitation observations from the same network of opportunistic sensors are considered and used both in Sec. 3.1.2 and in Sec. 3.2. The majority of data measured by stations managed by citizens have been collected thanks to collaboration between MET Norway and Netatmo, a manufacturer of private weather stations. The observations used in Sec. 3.1.2 and in Sec. 3.2 have been quality controlled by MET Norway, therefore they are considered as correct data.

Section 3.2 describes the results of cross-validation experiments over South Norway. The results presented give an idea of the EnSI-GAP performances over a period of five months centered over summer 2019, that is from May 2019 to September 2019. The domain, data sources and grid settings of the cross-validation experiments are the same as for the case study of intense precipitation of Sec. 3.1.2. We are aiming at evaluating EnSI-GAP from both a deterministic and probabilistic viewpoint. The verification scores considered are commonly used in forecast verification and described in several books, such as for example Jolliffe and Stephenson (2012). A further useful reference for the scores is the website of the World Meteorological Organization https://www.wmo.int/pages/prog/arep/wwrp/new/jwgfvr.html (last time checked 2020-05-13).

3.1 Examples of application

3.1.1 One-dimensional simulations

EnSI-GAP is shown over a one-dimensional grid with 400 points and spacing of 1 spatial unit, or $1\,u$, such that the domain covers the region from 0.5 u to 400.5 u and the generic ith grid points is placed at the coordinate i u. A simulation begins with the creation of a true state, then observations and ensemble background are derived from it. The statistical interpolation scheme is applied with different configurations in order to illustrate the behaviour of the method.

The simulation presented here is shown in Fig. 1a. For each grid point, the true value (black line) is generated by a random extraction from the gamma distribution with shape and rate set to 0.2 and 0.1, respectively. To ensure spatial continuity of the truth, an anamorphism is used to link a 400-dimensional multivariate normal (MVN) vector with the gamma distribution. The samples from the MVN distribution, with a prescribed continuous spatial structure, are obtained as described by Wilks (2019), chapter 12.4. The MVN mean is a vector with 400 components all set to zero and the covariance matrix is determined using a Gaussian covariance function with 10 u as the reference length used for scaling distances. The effective resolution (Sec. 2.2.1) of the truth is then 10 u.

The ensemble background (gray lines in Fig. 1a) on the grid is obtained by perturbing the truth. The background values at observation locations are obtained using nearest neighbour interpolations applied to the ensemble members. The observation operator $\mathbf{H}$ (Sec. 2.2.1) is the nearest neighbour interpolation, that is a matrix with all zeros except for a single element on each line equal to 1 at the element corresponding to the closest grid point. The truth is perturbed considering four main error sources, which are typically found in precipitation fields simulated by numerical models. The first typical error is the misplacement of precipitation events, that is implemented here independently for each ensemble member by shifting the true values along the grid by a random number between -10 u and +10 u. Second, the effective resolutions of the background members are set to be coarser than the truth. For each member, the coarser resolution is obtained by multiplying the true values by a coefficient derived from a uniform distribution with values between 0.05 and 2 and a spatial structure function given by a MVN with Gaussian covariance function with a reference length greater than that of the truth, which is 10 u. The exact reference length scale varies between each member as it is extracted from a Gaussian distribution with mean of 50 u and a standard deviation of 5 u. Third, as previously stated in Sec. 2, the challenging special case of the background showing no-precipitation while the true state reports precipitation is considered. For this purpose, all the ensemble members for the grid points between 200 u and 300 u are set to 0 mm. The fourth typical error is a variation of the third one, the background at grid points between 50 u and 150 u follows an alternative truth but different from 0 mm. It is possible to recognize the four regions described above on the grid in Fig. 1 by means of the coordinate on the x-axis. Because we had to ensure continuity of the background, we have enforced smooth transitions between the regions mentioned above and their surroundings.

The number of observations (blue dots in Fig. 1*a*) is set to 40. The observation locations are randomly chosen, such that the observation density is variable across the grid. The observed value at a location is obtained as the true value of the nearest grid point, plus a random noise that is determined as a random number between -0.02 and 0.02 that multiplies the true value. The procedure is consistent with the fact that observation precipitation errors should follow a multiplicative model (Tian et al., 2013). The observation distribution is denser in the central part of the domain and sparser closer to the borders. The constraints on the number of observations are: 5 between 1 u and 100 u; 30 between 101 u and 300 u; 5 between the 301 u and 400 u. The panel *d* in Fig. 1 shows the Integral Data Influence (IDI Uboldi et al., 2008), that is a parameter that stay close to 1 for observation-dense regions, while it is exactly equal to 0 in observation-void regions. In practice, the IDI at the ith grid point is computed here as the analysis in Eq. (13) when all the observations are set to 1 and the background to 0, moreover only the static component $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{u}}$ of $\overset{i}{\mathbf{P}}{}^{\mathrm{b}}$ is considered in the IDI elaboration. Where the IDI is close to zero, the analysis is as good as the background. The simulation has been configured such that the domain is well covered by the observations. In this way, it can provide insights on the combination process itself. As it follows from Eq. (13) and Eq. (5), IDI depends on the error covariance matrices. We have used the equations of Algorithm 1, then the only additional parameter we have to set is D_i. In this example, D_i is estimated as the distance between the ith grid point and its closest 3rd observation location. This way, IDI stays close to 1, meaning that the observations do have an impact on the analysis, even for observation-sparse regions. D_i is shown in Fig. 1*c* and it has been constrained to vary between 5 u and 20 u. More details on how to set D_i are discussed in Sec. 3.3.

The effect of the Gaussian anamorphism is shown in Fig. 1*b*. The transformed precipitation varies within a smaller range than the original precipitation, thus effectively shortening the tail of the distribution, reducing its skewness and making it more similar to a Gaussian distribution.

The transformed precipitation analysis is shown in Fig. 2 for different configurations of the statistical interpolation scheme. The localization matrix $\overset{i}{\mathbf{\Gamma}}$, of Eq. (5), for all panels is specified using Gaussian functions, of the form of those used in Algorithm 1 for $\overset{i}{\mathbf{Z}}$ and $\overset{i}{\mathbf{V}}$, with L_i constant and set to 25 u for all the grid points. The value of ε^2 is set to 0.1, which means that we trust much more the observations than the background. The parameters that are allowed to vary are those determining the scale matrix. As stated above, D_i is determined adaptively at each grid point as shown in Fig. 1*c*. α is set to two different values. In

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

the left column (panels *a* and *c*) $\alpha = 0.1$, while in the right column (panels *b* and *d*) $\alpha = 1$. The background uncertainty is a trade off between the ensemble spread and the averaged innovation, with $\alpha = 1$ the weight of each component is determined by the simulation presented in Fig. 1, without considering that is just one possible realization of the ensemble that should be used to derive robust parameter estimates. In the case of $\alpha = 0.1$, the scale matrix is multiplied by a smaller value of $\overset{i}{\sigma}{}^2_u$, as can be seen from Eq. (11). The panels in the top row (panels *a* and *b*) are obtained using a Gaussian function for $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{u}}$, as in Algorithm 1, while the bottom row (panels *c* and *d*) shows the analyses when an exponential function is used. The exponential correlation function for the spatial analysis of precipitation is used for example by Mahfouf et al. (2007); Lespinas et al. (2015); Soci et al. (2016).

In Fig. 2, the analysis mean (or expected value, red line) fits the observations and the observed value is often within the analysis PDF envelope shown (pink region), despite the background deficiencies. The same comments hold true also when $\varepsilon^2 = 1$ (not shown here), with an increase of the analysis residuals (analysis minus observation) and a slightly larger spread of the analysis PDF. The analysis means, in all panels, follow the observed values closely and variations in the scale matrix seem to have limited effects on the analysis means, at least in regions with a dense observational network, such as between 100 u and 300 u. The analysis ensemble spread is more sensitive to variations in the scale matrix. In the case of $\alpha = 1$, the spread is larger than for $\alpha = 0.1$ because of the increased background uncertainty. In all panels, the analysis presents the higher uncertainties between 50 u and 100 u, where observations are sparse and the background follows an alternative truth. The analysis uncertainty is large also between 200 u and 300 u, where the background ensemble is set to no-precipitation while the rather dense observational network reports yes-precipitation. Interestingly enough, the region between 200 u and 300 u is also where the analyses derived from Gaussian and exponential functions differ the most. In this region, the ensemble-dependent component of $\overset{i}{\mathbf{P}}{}^{\mathrm{b}}$ in Eq. (5) underestimates the actual uncertainty, since all members are equals to zero while the observations are different from it, therefore the scale matrix determines almost completely the background error covariance. This can be also seen in Fig. 4*a*, $\overset{i}{\sigma}{}^2_u$ (dashed black line) reaches its maximum between 200 u and 300 u, furthermore $\overset{i}{\sigma}{}^2_b$ (thick black line) is equal to $\overset{i}{\sigma}{}^2_u$. With respect to Fig. 4*a*, it is also interesting to note that $\overset{i}{\sigma}{}^2_b$ reaches a second maximum between 50 u and 150 u, however there $\overset{i}{\sigma}{}^2_b$ is not determined solely by $\overset{i}{\sigma}{}^2_u$, since also $\overset{i}{\sigma}{}^2_f$ (difference between the thick and the dashed black lines) contributes significantly. The corresponding analysis ensemble spread, which is shown in Fig. 2*b*, is larger between 50 u and 150 u than in the region between 200 u and 300 u. This depends on the fact that $\overset{i}{\sigma}{}^2_f$ is just an averaged value summarizing the forecast uncertainty, while the complete characterization of the forecast uncertainty is given by the all elements of the covariance matrix $\overset{i}{\mathbf{P}}{}^{\mathrm{f}}$. As a result, situations like the large variability of the background between 50 u and 150 u tends to have a larger impact on the analysis error spread than situations like those occurring between 200 u and 300 u.

The precipitation analyses in the original space are shown in Fig. 3. Panels *a-d* are derived from the corresponding panels of Fig. 2, after the inverse transformation. Panels *e* and *f* show the analyses obtained without any data transformation. Note that in Fig. 4, the diagnostic shown in panel *a* corresponds to Fig. 3*b*, while panel *c* corresponds to Fig. 3*f*. The comparison of the spatial analysis with and without data transformation offers interesting ideas for discussion. The most important point is that data transformation helps in avoiding unrealistically large values of the analysis spread as it happens between 50 u and 100 u (note that the scales of the y-axes are different between panels *e-f* and panels *a-d*). In the case of the spatial analysis with the data transformation, the spread associated to the analysis PDF is proportional to the observed value, moreover the skewness of the PDF allows for values higher than the mean to happen with a higher probability than in the case of no data transformation. Those two characteristics are positive achievements of EnSI-GAP, since they are typical of a gamma PDF that is regarded as a good statistical model for precipitation uncertainties (Wilks, 2019). The impact of the data transformation seems to be less significant over the analysis means, which, given the EnSI-GAP settings used, are principally constrained to fit the observations. The panels *a-d* of Fig. 3 illustrate how the analysis uncertainty is modified when passing from the transformed to the original space. The regions with the higher uncertainties in Figs. 2- 3 coincide, except for the noticeable difference of the points between 250 u and 300 u. In Fig. 2, the analysis PDFs for all those grid points show Gaussian functions with some spread. However, the same points in Fig. 3 show a much narrower analysis PDF, often without any significant spread even though we are not in the special case of $\overset{i}{\sigma}_f^2 = 0$ and $\overset{i}{\sigma}_{ob}^2/(1+\varepsilon^2) = 0$ presented in Sec. 2.2.3. This is a direct effect of the inverse transformation, which usually groups together the analysis values with light or no rain more than those with heavy rain. In fact, in the same region, the analyses without data transformation show more spread, as can be seen in Fig. 3 *e* and, especially, *f*. From a qualitative point of view, in Fig. 3 the best results are obtained with $\alpha = 0.1$ with respect to $\alpha = 1$ because of the better fit of the analysis expected value with the truth, in addition the analysis spread with $\alpha = 0.1$ is smaller and the PDF envelope includes the truth. When comparing the effect of using a Gaussian or an exponential function in the specification of $\overset{i}{\mathbf{\Gamma}}^{\mathrm{u}}$, the results are similar and it is difficult to choose among them.

The observation and background error variances are shown in Fig. 4 for different configurations of the analysis procedure. For all panels, $\alpha = 1$, $L = 25\,u$ and Gaussian covariance functions are used for localization and the specification of the scale matrix. D_i is the same shown in Fig. 1*c*. The effect on the error variances of using a data transformation against not using it is investigated here. EnSI-GAP variances are shown in panels *a* and *b*, with $\varepsilon^2 = 0.1$ and $\varepsilon^2 = 1$ respectively. When $\varepsilon^2 = 1$, $\overset{i}{\sigma}_b^2$ is equal to $\overset{i}{\sigma}_o^2$ everywhere but their values vary along the one dimensional grid. The scheme without data transformation is shown in panels *c* and *d* for the same ε^2 values. As a consequence of the data transformation, the scales on the y-axis differ between panels *a-b* and *c-d*. For all panels, the analysis PDF spread is higher between 50 u and 100 u, as can be seen e.g. in Fig. 3*b* that correspond to Fig. 4*a*. When the data transformation is applied, $\overset{i}{\sigma}_b^2$ reaches its maximum between 200 u and 300 u, with a second maximum between 50 u and 105 u. In the case without data transformation, the two maxima of $\overset{i}{\sigma}_b^2$ are still in the same regions, though the principal one is between 50 u and 150 u. Besides, when using a data transformation the two maxima do have rather similar values, while without using data transformations the highest maximum can be four times the smallest one (Fig. 4*d*). In all panels, $\overset{i}{\sigma}_u^2$ is different from zero in the region between 180 u and 300 u. In addition, for $\varepsilon^2 = 0.1$, $\overset{i}{\sigma}_u^2$ is different from zero also in the region between 50 u and 150 u because the variation of $\varepsilon^2 = 0.1$ modifies the threshold value of $\overset{i}{\sigma}_{ob}^2/(1+\varepsilon^2)$ in Eqs. (11)- (12). Then, one interpretation is that between 200 u and 300 u, when the background fails to predict the occurrence of precipitation, all spatial analysis scheme recognize the importance of using the additional term modulated by $\overset{i}{\sigma}_u^2$ for the specification of the background error. On the other hand, for the type of background failure taking place between 50 u and 150 u, that is the background is following an alternative truth with respect to observations, the behaviour of the spatial analysis scheme is less predictable and it is more sensitive to the EnSI-GAP settings. Furthermore, this example indicates that a data transformation is useful to keep the variances at reasonable values.

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

3.1.2 Intense precipitation case over South Norway

A mass of moist air from the ocean moving towards the Norwegian mountains was at the origin of several intense showers over western Norway on the 30th of July 2019. South Norway is the domain under consideration and it is shown in Fig. 5, the domain measures 373 km in the meridional direction and 500 km in the zonal direction. The measurements from the weather stations managed by MET Norway show hourly precipitation values with more than $20\,\mathrm{mm}$, which is extremely intense given the climatology of the region. In addition, thousands of lightning strikes have been recorded (not shown here), thus confirming the convective nature of the precipitation. Intense events have been observed in the afternoon along the coast and over the nearby mountains, especially in Sogn og Fjordane. This region is shown as the black box in Fig. 5, it extends for 80 km in both meridional and zonal directions. Point A corresponds to the center of a grid box where a maximum of precipitation has been observed. Point B is the center of a grid box that is not covered by observations and where a maximum of precipitation has been reconstructed by the analysis. The distance between points A and B is 14 km. In Sogn og Fjordane damages have been reported, they were caused by the heavy rain that also triggered a series of landslides. One of them caused a fatality when a driver was caught in the debris flow.

On both domains, the focus is on the representation of hourly precipitation patterns at the mesoscale, as it has been defined by Thunis and Bornstein (1996); Stull (1988), though over different domains we will focus on different parts of the mesoscale. South Norway is used to show that the variability of the fields represented by the forecast ensemble members involves mostly the Meso-β part of the mesoscale (i.e. spatial scales from 20 km to 200 km). Sogn og Fjordane is a domain where high-resolution information is needed to support fine-scale analysis by e.g. civil protection authorities. In this case, we will study precipitation patterns at the Meso-γ scale (i.e. from 2 km to 20 km). The EnSI-GAP Algorithm 1 has been used over a grid with 2.5 km of spacing, which is the resolution of the MEPS grid (see Sec. 3). The parameters are $\varepsilon^2 = 0.1$, $\alpha = 0.1$, $p_{mx} = 200$, $L_i = 50\,\mathrm{km}$ constant, D_i is estimated adaptively on the grid as the distance between the grid point and the 10th closest observation location with upper and lower bounds of 3 km and 10 km, respectively. The EnSI-GAP settings are such that the analyses would stay much closer to the observations than to the forecasts, where observations are available. The analysis uncertainty will reflect locally both the forecast ensemble spread and the averaged innovation, but the weight of this last component is damped by α. The two parameters p_{mx} and D_i are used to limit the number of observations that can influence the analysis at a grid point. The localization parameter L_i is set to a rather large value, such that the dynamics of the forecasts ensemble are evident in the results. The observation error covariance matrix of Eq. (6) is defined with a diagonal $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{o}}$, that is a situation where radar-derived and in-situ observations are assumed to have the same precision, moreover we are ignoring the spatial correlation of radar-derived observation errors. An investigation of spatially correlated radar-derived observation errors is outside the scope of this study. Note that those settings are useful for the illustration of the method, while for operational applications other settings may be more appropriate, such as a smaller value of L_i or a more sophisticated characterization of the observation errors, for example.

Figure 6 shows the hourly precipitation data for 2019-07-30 15:00 UTC over South Norway. The observational data are shown in panel a. For each grid box, the average of radar-derived precipitation and in-situ measurements within that box is

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

shown. Note that the box-averaged observations are used only for illustration because the analysis is using each observation. Grid points that are not covered by observations are marked in gray. In panel *b*, the background ensemble mean derived from a 10-member ensemble forecast is shown, while six of the ten ensemble members are shown in Fig. 7. The 10-member ensemble shows realistic precipitation fields, moreover they are rather similar, at least in terms of weather situation at the Meso-β scale. Weather forecasters can be quite confident in stating that heavy precipitation is likely to occur over western and southern Norway, while is less likely over eastern Norway. The forecast uncertainty is large enough that it is difficult to predict exactly which subregion will be affected by the most intense showers. The observations confirm that showers occur along the coast of western Norway and that the most intense precipitation event is located in Sogn og Fjordane, that is the black box in Fig. 5. Note that approximately half of the box is not covered by observations. Panel *c* of Fig. 6 shows the analysis, specifically the analysis mean at each grid point. In this case, the spatial analysis acts almost as a ”gap filling” procedure to fill in empty spaces in between observations with the most likely precipitation values. As prescribed by our EnSI-GAP settings, the analyses over observation-dense regions are not that different from the observed values.

It is interesting to have a look at the results over Sogn og Fjordane and to focus on the Meso-γ scale. The evolution in time of the hourly precipitation fields is shown in Fig. 8 for observations, background and analysis at three different hours, as clearly indicated in the figure. The two crosses mark the locations of points A and B (see Fig. 5). As stated above, point A is in a densely observed part of the domain, while point B is almost in the middle of the observation-void region and the closest observations are located at a distance of approximately 10 km. As a consequence, in the EnSI-GAP algorithm, D_i at point A is closer to 3 km, while at point B is closer to 10 km. The observed fields show a large variability over short distances and the difference between two adjacent points can be as large as 30 mm/h. The background is smoother than the observed field and shows scattered showers for 14:00 UTC and 15:00 UTC, then a wider precipitation cell almost centered over point B is shown at 17:00 UTC. At 14:00 UTC, the observed value at point A (from radar-derived estimates) is over 30 mm/h and it is evident a sharp gradient from south-west to north-east. The gradient is so intense that the nearby points south-west of point A, only 3 km apart, shows almost no precipitation. The background indicates that a maximum of the field can occur between point A and B. The analysis matches the observations, though smoothing out their spatial variability, such that at point A the analysis value is less than 10 mm/h, that is representative of the areal average of the nearby points. A precipitation maximum of more than 30 mm/h has been reconstructed in the analysis between points A and B, that is consistent with the gradient in the observations and the pattern in the background. The situation at 15:00 UTC in the observations around point A is a bit different. The radar-estimated precipitation is again over 30 mm/h but there are several grid points in the surroundings of point A with similar values, such that the local gradient of the field is less steep and it shows a decrease of precipitation east of point A. The background also shows that it is more likely to find intense precipitation immediately to the west of point A than to the east. A second precipitation maximum is found in the background, north of point B. The analysis ignores this second precipitation maximum, since it is not supported by observations. The analysis around point A matches closely both the observed values and the gradient, such that the precipitation field in the observational-void area does not show significant local extremes. The precipitation at 17:00 UTC shows a different situation from the other two hours, because all the observations report values smaller than 20 mm/h and the analysis reconstruct a maximum of over 30 mm/h at point B. In this case, both the observations

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

and, more markedly, the background show a south-east to north-west gradient in the precipitation field and their corresponding precipitation yes/no patterns are not too different. Then, the analysis has a narrow band of precipitation around point B where values of more than 20 mm/h and up to 30 mm/h are extrapolated.

The time series of hourly precipitation at points A and B are shown in Fig. 9. At point A, the graphs show the time series of (aggregated) observation, background and analysis, together with the estimated uncertainties. Note that the observation is used in the analysis at point A. At point B, observations are not available. For the background, the percentiles are derived from the 10-member forecast ensemble through a linear interpolation of the empirical cumulative distribution function. For the analysis, the percentiles are derived from the estimated parameters of the gamma distribution representing the marginal probability density function of the analysis at the points. In general, EnSI-GAP forces the analysis to follow more closely the observations than the background and the analysis uncertainty is smaller than those of the background. As a consequence, the timing of the precipitation onset is also better represented in the analysis. At point A, the PDF of the precipitation analysis between 10:00 UTC and 13:00 UTC is a Dirac's delta function and it has the value of 0 mm/h. From 14:00 UTC onward, the analysis PDF is a gamma. During this period, on average the interquartile range (i.e., the difference between the 75th and the 25th percentiles) is 20% of the analysis value, the difference between the 90th and the 10th percentiles is 38% of the analysis and the difference between the 99th and the 1st percentiles is 70% of the analysis. From 14:00 UTC to 23:00 UTC, the observed values are within the analysis envelopes shown in Fig. 9 for 50% of the hours, that is a consistent improvement compared to the background. For the other 50% of the hours, the observed values lie outside the envelopes and 14:00 UTC and 19:00 UTC are the two hours when the deviations between observations and analyses are the most evident. For those two hours, the local variability of the precipitation field is extremely large, as shown in Fig. 8 for 14:00 UTC, and the observed values at point A are sort of outliers, if compared to their neighbours. The spatial analysis finds the best estimates of true values, which are areal averages, as discussed in Sec. 2.2.1 and defined in Eq. (3), with spatial supports determined by the EnSI-GAP settings. At 14:00 UTC and 19:00 UTC, the representativeness errors of the observations at point A are particularly large with respect to the spatial supports of the true values, such that the corresponding observations get ”filtered out” by the analysis and their values are unlikely to occur according to the analysis PDF. Note that by fine-tuning the EnSI-GAP settings, the analysis PDF can be modified such that the analysis spread would become larger, which in this case would correspond to a reduction of the spatial support for the true values, and the analysis envelope would be more likely to include the observations. The trade-off between e.g. accuracy and precision of the analysis at a point ultimately depends on the objective of an application. With respect to the precipitation yes/no distinction, from 14:00 UTC to 23:00 UTC, the analysis clearly shows that precipitation is occurring at the point, while the background is more uncertain. At point B, the analysis uncertainties between 10:00 UTC and 12:00 UTC is so small that the analysis PDF is a Dirac's delta function with 0 mm/h, despite there are no observations exactly located at that point. From 13:00 UTC onward, the analysis follows a gamma PDF and the spread is wider at point B that at point A. On average, the interquartile range is 82% of the analysis value, the difference between the 90th and the 10th percentiles and the 99th and the 1st percentiles are both higher than 100%. However, the values depend on the weather situation, for example at the precipitation peak of 32 mm/h at 17:00 UTC, the interquartile range is 35% of the analysis value and the difference between the 90th and the 10th percentiles is 68% of the analysis. The increase in the analysis spread at point B compared to point

A reflects the increase in the uncertainty in predicting the tails of the PDF where no observations are available. It is perhaps remarkable that even for observational dense regions, such as at point A, the analysis spread remains quite large. This is due to the large variability of the observations at the Micro-scale, which by definition includes spatial scales smaller than the Meso-γ. With respect to the EnSI-GAP theory, this large variability is interpreted as a large observation representativeness error. A very dense observational network, that is with observations that are closer than the effective resolution of the background, has two effects on EnSI-GAP: (i) it improves the accuracy of the analysis; (ii) it may increase the analysis uncertainty, such that the tails are more representative of the extremes that occur in the observational dense region, possibly on scales that are not properly resolved by the analysis grid.

One of the main innovation of EnSI-GAP compared to traditional spatial analysis methods (Hofstra et al., 2008) is the specification of anisotropic background error covariances between grid points through non-stationary covariance matrices. Two visual representations of the correlations associated with those covariances are shown in Fig 10 for points A and B. The correlations are shown instead of the covariances because we are interested in the shape of the covariance patterns and correlation is a quantity which is then more correct to compare between the two points. The domain considered is Sogn og Fjordane (see Fig. 5) and the covariances are computed over the same 2.5 km grid used in Fig. 8. Then, for visualization purposes, the correlations have been downscaled over a finer resolution grid to highlight asymmetries. The rather large localization length $L_i = 50\,\mathrm{km}$ allows for the dynamics of the forecast ensemble to be visible in the analysis. The corresponding observations, background and analysis are shown in the middle row of Fig. 8. With reference to $\overset{i}{\mathbf{P}}{}^{\mathrm{b}}$, the background error correlations between the generic ith grid point and the other grid points, evaluated at the ith grid point, are the ith row (or column) of the correlation matrix $\overset{i}{\boldsymbol{\Gamma}}{}^{\mathrm{b}}$, which is obtained as:

$$\overset{i}{\boldsymbol{\Gamma}}{}^{\mathrm{b}}{}_{i,:} = \frac{\overset{i}{\mathbf{P}}{}^{\mathrm{b}}{}_{i,:}}{\sqrt{\overset{i}{\mathbf{P}}{}^{\mathrm{b}}{}_{i,i}}\sqrt{\mathrm{diag}\left(\overset{i}{\mathbf{P}}{}^{\mathrm{b}}\right)}} \tag{16}$$

In Fig. 10, the background error correlations between the points A and B and their surroundings within Sogn og Fjordane are displayed. The closest two hundreds observations are shown with different symbols, depending on the rain occurrence. The two points are only 14 km apart, nonetheless the two maps in Fig. 10 are rather different. For point A, the correlation extends westward, while it decays faster moving eastward. The area where the correlation is higher than 0.6 is confined within approximately 5 km in any direction from point A. In the corresponding analysis in Fig. 8, the shape of the area with precipitation rate higher than 30 mm/h looks like the pattern of correlation higher than 0.6 of Fig. 10. At point B, the correlations tend to be more isotropic than for point A and it looks like the correlation extends more eastward than westward. The observations 20 km north-east of point B, that are reporting no precipitation, do have correlations with point B that are comparable to those of the observations at 10 km west of it. The analysis at point B, as shown in Fig. 8, takes into account those correlations and the predicted values in the region between points A and B decrease gradually when moving from the west to the east.

3.2 Validation over South Norway through cross-validation experiments

A series of cross-validation experiments have been performed over South Norway from the 1st of May to the 30th of September 2019. The South Norway domain shown in Fig. 5 and used in Sec. 3.1.2 has been used also in this section. The locations of the 57 weather stations directly managed by MET Norway are shown as red triangles. They are distributed all over the domain, the station network density is higher along the coast and sparser on the mountains, because of the inherent difficulties of operating weather stations there. The observations from MET Norway's stations have not been used in the spatial analysis. Instead, because of the expected better quality of those measurements, they have been reserved as independent observations for verification. This cross-validation strategy is widely used in atmospheric sciences (Wilks, 2019).

As for Sec. 3.1.2, the EnSI-GAP Algorithm 1 has been used. For this application, the spatial analysis predicts values at those station locations used for cross-validation. Some of the parameters are kept fixed, while others are allowed to vary. The fixed parameters are: $p_{mx} = 200$, $L_i = 50\,\mathrm{km}$. Then, as in Sec. 3.1.2, D_i is estimated adaptively at each location as the distance between that point and the 10th closest observation location with upper and lower bounds of 3 km and 10 km, respectively. The parameters that are allowed to vary and that are the objective of the sensitivity analysis that follows are: ε^2 and α. There is an important difference here, in this example the radar-derived estimates are assumed to be less precise than the in-situ observations but more precise then the background. The in-situ observations are assumed to be ten times more precise than the background, then ε^2 is set to 0.1 as in Sec. 3.1.2. However, the radar-derived observations are assumed to be only two times more precise than the background, or in other words they are five times less precise than the in-situ observations, and the elements of the diagonal matrix $\overset{i}{\mathbf{\Gamma}}{}^{o}$ corresponding to radar observations are set to 5, instead of 1 as for the in-situ observations. The background ensemble and analysis PDF values considered are those extracted at the locations of stations used for cross-validation.

Figure 11 shows the distribution of values for selected percentiles of the background ensemble and analysis PDF as a function of the independent observations. The distribution of the observed values has been divided into intervals, they are (units mm/h): 0-0.1, 0.1-0.5, 0.5-1, 1-2, 2-3, 3-5, 5-10, 10-35. The number of samples within each interval is shown in panel *a*, note the logarithmic scale on the y-axis. Most of the observations are smaller than 1 mm/h, nonetheless there are still more than one thousand values that are greater than 1 mm/h. Consider an arbitrary observation interval, for each probabilistic prediction in it, either background or analysis, we have computed the percentiles: 10th, 25th, 50th, 75th and 90th. The black line in Fig. 11 shows the average median within each interval, while the regions between the 90th and the 10th percentiles, and the 75th and the 25th are shown with gray shades. The diagonal (1:1) ideal line is shown as a dashed line. The background is shown in panel *a*. The background envelope deviates significantly from the diagonal, especially for values greater than 2-3 mm/h. The analysis PDFs are shown in the other panels for different EnSI-GAP configurations: panel *b* shows a configuration where the observation error variance is equal to the background error variance ($\varepsilon^2 = 1$) and the ensemble spread is not modified in the variance estimation procedure ($\alpha = 1$); panel *c* shows a configuration where the observations are assumed to be much accurate than the background ($\varepsilon^2 = 0.1$) and at the same time the background uncertainty is assumed to be on average smaller than those derived from the ensemble corresponding to the hour under consideration ($\alpha = 0.1$); panel *d* shows a configuration where the observations are more accurate than the background ($\varepsilon^2 = 0.1$) and $\alpha = 1$. In all cases, the medians are closer to the diagonal for the analyses than for the background. The analysis biases conditional to the observations are always smaller than that of the background. As expected, by giving more weight to the observations, with $\alpha = 1$ and $\varepsilon^2 = 0.1$, the analysis bias conditional to the observations decreases. If we compare different analysis configurations, the medians vary less than the other percentiles and this indicates that variations in the EnSI-GAP configuration impacts more on the spread of the analysis PDF (i.e. analysis uncertainty) than on its central moment. Panels *b* and *c* show the two extreme situations, while panel *d* displays an intermediate situation. The uncertainty is more sensitive to variations over α than over ε^2. In the case of $\alpha = 0.1$ and $\varepsilon^2 = 0.1$, the analysis spread is small and the independent observations have percentiles higher than the 90th percentile. For the two cases with $\alpha = 1$, the independent observations have percentiles around the 90th percentile.

Figure 12 shows the Equitable Threat Score (ETS) for the background and analysis means. Four different analysis configurations are shown. The independent observations are used to judge if an event does in fact obtain. The condition defining the ”yes” event for either observations or predictions is that the corresponding value must be higher than the precipitation threshold specified on the x-axis. For all predictions, it is more likely that a predicted ”yes” event corresponds to an observed ”yes” event for smaller thresholds than for the higher ones. The added value of the analysis over the background is evident for all configurations. The two configurations with $\alpha = 1$ present similar ETS curves, though the one with $\varepsilon^2 = 0.1$ performs better. The same holds true when $\alpha = 0.1$, though in this case the ETS is more sensitive to variations in ε^2 and the analysis performance decreases faster with the increase of ε^2.

The relative skill of the analysis probabilistic predictions over the background, used as reference, is shown by means of the Brier Skill Score (BSS) in Fig. 13. As for the ETS, also the BSS is shown as a function of the threshold used to define a ”yes” event. The BSS is higher for light precipitation, then it gradually decays with the increase of the precipitation rate. The analyses with $\alpha = 1$ show better performances over a wider range of amounts than those obtained with $\alpha = 0.1$.

The reliability diagrams plotting the observed frequency against the predicted background and analysis probabilities are shown in Fig. 14. The ”yes” events for independent observation and prediction are when the values are greater or equal to 1 mm/h. In addition to the thick line showing the reliability curve, each panel has also a box with the number of samples within the different probability classes (frequencies of the predictions). The three dashed lines are: the no-resolution horizontal line at the climatological frequency of occurrence for the event ”yes”; the diagonal line corresponding to the perfect reliability, deviations from the diagonal indicate conditional bias of the probabilistic prediction; points between the no-skill line and the diagonal contribute positively to the BSS. The background line shown in panel *a* lies below the diagonal, then the background is overpredicting the observed frequencies. The analyses tend to cross the diagonal, which means that they are overpredicting some probabilities and underpredicting other probabilities. When $\alpha = 1$, the analyses are underpredicting low probabilities and overpredicting high probabilities. The analysis configuration with $\alpha = 1$ and $\varepsilon^2 = 0.1$ stays rather close to the diagonal, overpredicting low probabilities and underpredicting high probabilities.

3.3 Discussion

For the presented implementation of EnSI-GAP, the Gaussian anamorphosis $g()$ of Eq. (16) is based on the same gamma distribution parameters for the whole domain. This assumption might be too restrictive for very large domains, such as for all Europe for instance. In this case, different solutions may be explored such as slowly varying the gamma parameters in space, or time, based on the climatology. The data transformation is important in preventing the analysis PDF spread to become too big, especially when the background and the observations, for the same hour, refer to different true states, as in the simulation of Sec. 3.1.1. The analysis expected values are less influenced by the data transformation than the analysis spread.

The combination of the ensemble-based error covariance matrix and the scale matrix provides a flexible and general method to characterize background uncertainty, even when and where the numerical weather forecasts are an inadequate background. The method is not particularly sensitive with respect to the function used to compute correlations in the scale matrix. In Sec. 3.1.1 experiments have been carried out with two functions: Gaussian and exponential. There are of course differences among them, but those are less significant than those obtained with variations in ε^2 or α. The inclusion of model ensembles in the background error covariance matrix allows the specification of a flow-dependent component of the error. The analysis is more likely to closely follow the patterns of precipitation structures simulated by the numerical model where the innovations are not too large.

The EnSI-GAP implementation in Algorithm 1 requires the specification of four parameters: D, the length scale of the scale matrix; L, the localization length scale, which is governing the covariance suppression rate with the distance; α, the stabilization factor; ε^2, used to determine the error variances. EnSI-GAP is designed such that the four parameters are location dependent. It is important to avoid abrupt variations in space of these parameters, otherwise the analysis field will show unrealistic patterns due to the sudden change in the statistical properties of the field. The type of studies presented in Secs. 3.1.2- 3.2 can provide a guideline for the adaptation of EnSI-GAP to specific applications. In fact, it is not possible to propose a set of parameters that fits all the possible applications. Ultimately, the ideal set of parameters depends on several domain- and application- specific requirements. For instance, different EnSI-GAP settings would be required to obtain the analysis PDFs when the aim is the most likely precipitation estimate compared to when trying to predict the possibility of extreme precipitation. The characteristics of the data sources used, such as the quality of the forecast ensemble and the density of the observational network, will condition the parameter values. In the examples of applications presented in Secs. 3.1.2- 3.2, we have followed a few rules. L depends on the characteristics of the numerical model used and it should reflect the size of typical precipitation events occurring in a region. D depends on the observation spatial distribution, because it is used to ensure that the scale matrix is based on the interaction of some observations and not only just one or two. If we assume that it is reasonable to use the observational network to refine the effective resolution of the background, then we can imagine that L should be larger than D. ε^2 varies between 0.1 and 1, for the reasons discussed in Sec. 2.2.2. In Sec. 3.2, we have also specified the observation errors as a function of the data source: radar data has a different quality than in-situ. It should also be possible to setup the analysis such that citizen observations have a different quality than observations measured by professional stations. We have assumed that the background ensemble is more likely to overestimate the spread than to underestimate it, for this reason we have assumed values of α between 0.1 and 1. Note that there is a further parameter that can be specified in EnSI-GAP, that is p_{mx}, which is used to keep the analysis procedure local. However, p_{mx} has not been considered among the four parameters optimized in Secs. 3.1.1- 3.2. Instead, it has been set to 200, that is a large number for our applications. If p_{mx} is too small, the analysis field will include discontinuities due to the sudden influence of observations that have been given significantly different weights in the analysis at a grid point with respect to the analyses at its neighbouring grid points. p_{mx} is important for operational application over computational systems with limited resources, because it can be regarded as a parameter that limits the use of computational resources within a predefined range.

4 Conclusions

The ensemble-based statistical interpolation with Gaussian anamorphosis (EnSI-GAP) applies inverse problem theory to the spatial analysis of hourly precipitation. Numerical model output provides the prior information, and specifically we have considered ensemble forecasts, that have been combined with radar-derived estimates and in-situ observations. EnSI-GAP has been applied on datasets that are typically available within national meteorological services. In addition, opportunistic sensing networks based on citizen observations have been considered. The precipitation representation is a synthesis of all the data available. Thanks to the diffusion of open data policies, the same datasets are also nowadays available in real-time to the general public. For instance, the Norwegian Meteorological Institute provides free access to the weather forecasts and the radar data used in this article via thredds.met.no, while in-situ observations, except the citizen observations, are available via frost.met.no.

EnSI-GAP assumes the precipitation fields to be locally stationary, trans-gaussian random fields. The marginal distribution of precipitation at a point is a gamma distribution. Gaussian anamorphosis is used to pre-process data in order to better comply with the requirements of linear filtering. The inverse transformation returns the gamma shape and rate. A special case is considered where uncertainties are so small that the returned analysis values have delta functions as their marginal distributions.

EnSI-GAP considers each hour independently and it requires the specification of four parameters that can vary across the domain. The implementation is designed to run in parallel on a grid point by grid point basis. Despite the small number of parameters to optimize, the spatial analysis scheme is flexible enough that it can be applied also when the background ensemble is not representing the truth satisfactorily. An important case is when, in a region, all the ensemble members show no precipitation, while the observations report precipitation. By adding a scale matrix to the flow-dependent background error covariance matrix, the analysis can predict precipitation even where the background is sure that it is not occurring.

The examples of applications presented allow for a better understanding of the characteristics of EnSI-GAP and they show how the statistical interpolation can be adapted to meet specific requirements. It can be used to fill in the gaps between observation-rich regions to obtain a continuous precipitation field. The analysis expected value is available everywhere, as it is the background, and in observation-dense regions it can be as accurate as the observations. Thanks to the data transformation, the spread of the analysis PDF is less likely to become unrealistically large because of either large model errors or large variability of observed small-scale precipitation. Within certain limits, determined by the spatial distribution of the observational network, the analysis envelope at a point can be tuned such that it is representative of the distribution of precipitation values determined by atmospheric processes occurring at smaller spatial scales than those resolved by the background. For instance, in an observation-void region, the EnSI-GAP analysis PDF at a point provides a better estimate than the background for the probability of precipitation exceeding a threshold by an observation hypothetically placed at that point. This is an important result, especially when high-impact weather is involved.

Data availability. Some of the datasets used in Secs. 3.1.2- 3.2 are freely available online. MET Norway provides free access to: the weather forecasts at https://thredds.met.no/thredds/catalog/meps25epsarchive/catalog.html; the hourly precipitation derived from the Norwegian composite of weather radars https://thredds.met.no/thredds/catalog/remotesensingradaraccr/catalog.html; the archive of Norwegian historical weather and climate in-situ observations frost.met.no. Due to distribution restrictions imposed by some of the providers, opportunistic sensing networks, such as citizen observations, are not freely available online.

Author contributions. CL developed EnSI-GAP, tested it on the case studies and prepared the manuscript with contributions from all co-authors. TN and IS configured EnSI-GAP to work with MET Norway's datasets, collected in-situ observations from opportunistic sensing networks and quality controlled them. CE prepared the radar data.

Competing interests. The authors declare that no competing interests are present.

Acknowledgements. This research was partially supported by: RADPRO (Radar for Improving Precipitation Forecast and Hydropower Energy Production), an innovative industry project funded by the Research Council of Norway (NFR) and partnering hydropower industries; the collaboration between MET Norway and "The Norwegian Water Resources and Energy Directorate" (NVE) within the national project "Felles aktiviteter NVE-MET 2019-2020 tilknyttet nasjonal flom- og skredvarslingstjeneste".

References

Amezcua, J. and Leeuwen, P. J. V.: Gaussian anamorphosis in the analysis step of the EnKF: a joint state-variable/observation approach, Tellus A: Dynamic Meteorology and Oceanography, 66, 23 493, https://doi.org/10.3402/tellusa.v66.23493, https://doi.org/10.3402/tellusa.v66.23493, 2014.

Bertino, L., Evensen, G., and Wackernagel, H.: Sequential Data Assimilation Techniques in Oceanography, International Statistical Review, 71, 223–241, https://doi.org/10.1111/j.1751-5823.2003.tb00194.x, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1751-5823.2003.tb00194.x, 2003.

Bocquet, M., Raanes, P. N., and Hannart, A.: Expanding the validity of the ensemble Kalman filter without the intrinsic need for inflation, Nonlinear Processes in Geophysics, 22, 645–662, https://doi.org/10.5194/npg-22-645-2015, https://www.nonlin-processes-geophys.net/22/645/2015/, 2015.

Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: an overview of methods, issues, and perspectives, Wiley Interdisciplinary Reviews: Climate Change, 9, e535, https://doi.org/10.1002/wcc.535, https://onlinelibrary.wiley.com/doi/abs/10.1002/wcc.535, 2018.

CIMO: WMO Guide to Meteorological Instruments and Methods of Observation, Tech. rep., World Meteorological Organization, 2014.

de Vos, L. W., Droste, A. M., Zander, M. J., Overeem, A., Leijnse, H., Heusinkveld, B. G., Steeneveld, G. J., and Uijlenhoet, R.: Hydrometeorological Monitoring Using Opportunistic Sensing Networks in the Amsterdam Metropolitan Area, Bulletin of the American Meteorological Society, 101, E167–E185, https://doi.org/10.1175/BAMS-D-19-0091.1, https://doi.org/10.1175/BAMS-D-19-0091.1, 2020.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Quarterly Journal of the Royal Meteorological Society, 131, 3385–3396, https://doi.org/10.1256/qj.05.108, http://dx.doi.org/10.1256/qj.05.108, 2005.

Dyrrdal, A. V., Lenkoski, A., Thorarinsdottir, T. L., and Stordal, F.: Bayesian hierarchical modeling of extreme hourly precipitation in Norway, Environmetrics, 26, 89–106, https://doi.org/10.1002/env.2301, https://onlinelibrary.wiley.com/doi/abs/10.1002/env.2301, 2015.

Erdin, R., Frei, C., and Künsch, H. R.: Data transformation and uncertainty in geostatistical combination of radar and rain gauges, Journal of Hydrometeorology, 13, 1332–1346, 2012.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research: Oceans, 99, 10 143–10 162, 1994.

Fletcher, S. J. and Zupanski, M.: A data assimilation method for log-normally distributed observational errors, Quarterly Journal of the Royal Meteorological Society, 132, 2505–2519, https://doi.org/10.1256/qj.05.222, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.05.222, 2006.

Fortin, V., Roy, G., Stadnyk, T., Koenig, K., Gasset, N., and Mahidjiba, A.: Ten Years of Science Based on the Canadian Precipitation Analysis: A CaPA System Overview and Literature Review, Atmosphere-Ocean, 56, 178–196, https://doi.org/10.1080/07055900.2018.1474728, https://doi.org/10.1080/07055900.2018.1474728, 2018.

Frei, C. and Isotta, F. A.: Ensemble Spatial Precipitation Analysis from Rain-Gauge Data – Methodology and Application in the European Alps, Journal of Geophysical Research: Atmospheres, 0, https://doi.org/10.1029/2018JD030004, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD030004, 2019.

Frogner, I.-L., Singleton, A. T., Køltzow, M. Ø., and Andrae, U.: Convection-permitting ensembles: challenges related to their design and use, Quarterly Journal of the Royal Meteorological Society, 0, https://doi.org/10.1002/qj.3525, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3525, 2019.

Gandin, L. S.: Complex Quality Control of Meteorological Observations, Monthly Weather Review, 116, 1137–1156, https://doi.org/10.1175/1520-0493(1988)116<1137:CQCOMO>2.0.CO;2, https://doi.org/10.1175/1520-0493(1988)116<1137:CQCOMO>2.0.CO;2, 1988.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Quarterly Journal of the Royal Meteorological Society, 125, 723–757, 1999.

Germann, U. and Joss, J.: Operational Measurement of Precipitation in Mountainous Terrain, pp. 52–77, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-662-05202-0_2, https://doi.org/10.1007/978-3-662-05202-0_2, 2004.

Greybush, S. J., Kalnay, E., Miyoshi, T., Ide, K., and Hunt, B. R.: Balance and ensemble Kalman filter localization techniques, Monthly Weather Review, 139, 511–522, 2011.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 Global Reanalysis, Quarterly Journal of the Royal Meteorological Society, n/a, https://doi.org/10.1002/qj.3803, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803, 2020.

Hofstra, N., Haylock, M., New, M., Jones, P., and Frei, C.: Comparison of six methods for the interpolation of daily, European climate data, Journal of Geophysical Research: Atmospheres, 113, 2008.

Huang, S., Eisner, S., Magnusson, J. O., Lussana, C., Yang, X., and Beldring, S.: Improvements of the spatially distributed hydrological modelling using the HBV model at 1km resolution for Norway, Journal of Hydrology, 577, 123 585, https://doi.org/https://doi.org/10.1016/j.jhydrol.2019.03.051, http://www.sciencedirect.com/science/article/pii/S0022169419302495, 2019.

Ide, K., Courtier, P., Ghil, M., and Lorenc, A.: Unified notation for data assimilation: operational, sequential and variational, Practice, 75, 181–189, 1997.

Isotta, F. A., Begert, M., and Frei, C.: Long-Term Consistent Monthly Temperature and Precipitation Grid Data Sets for Switzerland Over the Past 150 Years, Journal of Geophysical Research: Atmospheres, 124, 3783–3799, https://doi.org/10.1029/2018JD029910, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD029910, 2019.

Jazwinski, A. H.: Stochastic processes and filtering theory, Courier Dover Publications, 2007.

Jermey, P. and Renshaw, R.: Precipitation representation over a two-year period in regional reanalysis, Quarterly Journal of the Royal Meteorological Society, 2016.

Jolliffe, I. T. and Stephenson, D. B.: Forecast verification, Wiley Oxford, 2012.

Kalnay, E.: Atmospheric modeling, data assimilation and predictability, Cambridge university press, 2003.

Kotlarski, S., Szabó, P., Herrera, S., Räty, O., Keuler, K., Soares, P. M., Cardoso, R. M., Bosshard, T., Pagé, C., Boberg, F., Gutiérrez, J. M., Isotta, F. A., Jaczewski, A., Kreienkamp, F., Liniger, M. A., Lussana, C., and Pianko-Kluczyńska, K.: Observational uncertainty and regional climate model evaluation: A pan-European perspective, International Journal of Climatology, 0, https://doi.org/10.1002/joc.5249, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.5249, 2017.

Kuusela, M. and Stein, M. L.: Locally stationary spatio-temporal interpolation of Argo profiling float data, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 474, 20180 400, https://doi.org/10.1098/rspa.2018.0400, https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2018.0400, 2018.

Lespinas, F., Fortin, V., Roy, G., Rasmussen, P., and Stadnyk, T.: Performance Evaluation of the Canadian Precipitation Analysis (CaPA), Journal of Hydrometeorology, 16, 2045–2064, https://doi.org/10.1175/JHM-D-14-0191.1, https://doi.org/10.1175/JHM-D-14-0191.1, 2015.

Lien, G.-Y., Kalnay, E., and Miyoshi, T.: Effective assimilation of global precipitation: simulation experiments, Tellus A, 65, 2013.

Lorenc, A. C.: Analysis methods for numerical weather prediction, Quarterly Journal of the Royal Meteorological Society, 112, 1177–1194, https://doi.org/10.1002/qj.49711247414, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49711247414, 1986.

Lundquist, J., Hughes, M., Gutmann, E., and Kapnick, S.: Our skill in modeling mountain rain and snow is bypassing the skill of our observational networks, Bulletin of the American Meteorological Society, 100, 2473–2490, https://doi.org/10.1175/BAMS-D-19-0001.1, https://doi.org/10.1175/BAMS-D-19-0001.1, 2019.

Lussana, C., Uboldi, F., and Salvati, M. R.: A spatial consistency test for surface observations from mesoscale meteorological networks, Quarterly Journal of the Royal Meteorological Society, 136, 1075–1088, 2010.

Lussana, C., Saloranta, T., Skaugen, T., Magnusson, J., Tveito, O. E., and Andersen, J.: seNorge2 daily precipitation, an observational gridded dataset over Norway from 1957 to the present day, Earth System Science Data, 10, 235–249, https://doi.org/10.5194/essd-10-235-2018, https://www.earth-syst-sci-data.net/10/235/2018/, 2018.

Magnusson, J., Eisner, S., Huang, S., Lussana, C., Mazzotti, G., Essery, R., Saloranta, T., and Beldring, S.: Influence of Spatial Resolution on Snow Cover Dynamics for a Coastal and Mountainous Region at High Latitudes (Norway), Water Resources Research, 55, 5612–5630, https://doi.org/10.1029/2019WR024925, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019WR024925, 2019.

Mahfouf, J.-F., Brasnett, B., and Gagnon, S.: A Canadian precipitation analysis (CaPA) project: description and preliminary results, Atmosphere-Ocean, 45, 1–17, 2007.

Müller, M., Homleid, M., Ivarsson, K.-I., Køltzow, M. A., Lindskog, M., Midtbø, K. H., Andrae, U., Aspelien, T., Berggren, L., Bjørge, D., et al.: AROME-MetCoOp: a nordic convective-scale operational weather prediction model, Weather and Forecasting, 32, 609–627, 2017.

Nipen, T. N., Seierstad, I. A., Lussana, C., Kristiansen, J., and Hov, : Adopting Citizen Observations in Operational Weather Prediction, Bulletin of the American Meteorological Society, 101, E43–E57, https://doi.org/10.1175/BAMS-D-18-0237.1, https://doi.org/10.1175/BAMS-D-18-0237.1, 2020.

Pollock, M., Dutton, M., Quinn, P., O'connell, P., Wilkinson, M., and Colli, M.: Accurate rainfall measurement: The Neglected Achilles Heel of hydro-meteorology, in: WMO technical conference on meteorological and environmental instruments and methods of observation, pp. 7–9, 2014.

Raanes, P. N., Carrassi, A., and Bertino, L.: Extending the Square Root Method to Account for Additive Forecast Noise in Ensemble Methods, Monthly Weather Review, 143, 3857–3873, https://doi.org/10.1175/MWR-D-14-00375.1, https://doi.org/10.1175/MWR-D-14-00375.1, 2015.

Sakov, P. and Bertino, L.: Relation between two common localisation methods for the EnKF, Computational Geosciences, 15, 225–237, 2011.

Savage, L. J.: The foundations of statistics, Courier Corporation, 1972.

Soci, C., Bazile, E., Besson, F., and Landelius, T.: High-resolution precipitation re-analysis system for climatological purposes, Tellus A: Dynamic Meteorology and Oceanography, 68, 29 879, https://doi.org/10.3402/tellusa.v68.29879, https://doi.org/10.3402/tellusa.v68.29879, 2016.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, Springer, https://doi.org/https://doi.org/10.1007/978-94-009-3027-8, 1988.

Tarantola, A.: Inverse Problem Theory and methods for model parameter estimation, SIAM, Philadelphia, 2005.

Thunis, P. and Bornstein, R.: Hierarchy of mesoscale flow assumptions and equations, Journal of the atmospheric sciences, 53, 380–397, 1996.

Tian, Y., Huffman, G. J., Adler, R. F., Tang, L., Sapiano, M., Maggioni, V., and Wu, H.: Modeling errors in daily precipitation measurements: Additive or multiplicative?, Geophysical Research Letters, 40, 2060–2065, 2013.

Uboldi, F., Lussana, C., and Salvati, M.: Three-dimensional spatial interpolation of surface meteorological observations from high-resolution local networks, Meteorological Applications, 15, 331–345, 2008.

Vicente-Serrano, S. M., Van der Schrier, G., Begueria, S., Azorin-Molina, C., and Lopez-Moreno, J.-I.: Contribution of precipitation and reference evapotranspiration to drought indices under different climates, Journal of Hydrology, 526, 42 – 54, https://doi.org/https://doi.org/10.1016/j.jhydrol.2014.11.025, http://www.sciencedirect.com/science/article/pii/S0022169414009305, drought processes, modeling, and mitigation, 2015.

Wilks, D. S.: Statistical methods in the atmospheric sciences (Fourth Edition), Elsevier, fourth edition edn., https://doi.org/https://doi.org/10.1016/B978-0-12-815823-4.00009-2, http://www.sciencedirect.com/science/article/pii/B9780128158234000092, 2019.

Wolff, M. A., Isaksen, K., Petersen-Øverleir, A., Ødemark, K., Reitan, T., and Brækkan, R.: Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation: results of a Norwegian field study, Hydrology and Earth System Sciences, 19, 951–967, https://doi.org/10.5194/hess-19-951-2015, https://www.hydrol-earth-syst-sci.net/19/951/2015/, 2015.

Zahumensky, I.: World Guidelines on Quality Control Procedures for Data from Automatic Weather Stations, 2004.

Algorithm 1 EnSI-GAP: Ensemble-based Statistical Interpolation scheme with Gaussian AnamorPhosis. In this implementation, Gaussian functions have been used for correlation (and correlation-like) matrices. The symbols used are listed in Tables 1- 2.

```
Require: forecast ensemble $\tilde{\mathbf{X}}^{\mathrm{f}}$; observations $\tilde{\mathbf{y}}^{\mathrm{o}}$; parameters $\varepsilon^2$, $\alpha$, $\mathbf{L}$, $\mathbf{D}$
  Gaussian anamorphosis: $\mathbf{X}^{\mathrm{f}} = g(\tilde{\mathbf{X}}^{\mathrm{f}})$; $\mathbf{y}^{\mathrm{o}} = g(\tilde{\mathbf{y}}^{\mathrm{o}})$
  Define additional global variables: $\mathbf{x}^{\mathrm{b}} = \mathbf{x}^{\mathrm{f}}$; $\mathbf{A}^{\mathrm{f}}:\ \mathbf{A}^{\mathrm{f}}_i = \mathbf{X}^{\mathrm{f}}_i - \mathbf{x}^{\mathrm{f}}$
  for all grid points $\{i = 1,\dots,m\}$ do
    select the closest $p_i$ observations $\overset{i}{\mathbf{y}}^{\mathrm{o}}$ and obtain $\overset{i}{\mathbf{y}}^{\mathrm{b}} = \overset{i}{\mathbf{H}}\mathbf{x}^{\mathrm{b}}$
    {Dynamical background error covariance matrices, $\overset{i}{\mathbf{S}}^{\mathrm{f}} \approx (k-1)^{-1}(\overset{i}{\mathbf{H}}\overset{i}{\boldsymbol{\Gamma}}\overset{i}{\mathbf{H}}^{\mathrm{T}}) \circ [(\overset{i}{\mathbf{H}}\mathbf{A}^{\mathrm{f}})(\overset{i}{\mathbf{H}}\mathbf{A}^{\mathrm{f}})^{\mathrm{T}}]$}
    $\overset{i}{\mathbf{Z}}:\ \overset{i}{\mathbf{Z}}_{jl} = \exp\left\{-0.5\left[d(\mathbf{r}_j,\mathbf{r}_l)/L_i\right]^2\right\}, j = 1,\dots,p_i$ and $l = 1,\dots,p_i$; $d()$ horizontal distance
    $\overset{i}{\mathbf{S}}^{\mathrm{f}}_{jl} \approx (k-1)^{-1}\overset{i}{\mathbf{Z}}_{jl}[(\overset{i}{\mathbf{H}}\mathbf{A}^{\mathrm{f}})_{j,:}(\overset{i}{\mathbf{H}}\mathbf{A}^{\mathrm{f}})_{l,:}], j = 1,\dots,p_i$ and $l = 1,\dots,p_i$
    { $\overset{i}{\mathbf{G}}^{\mathrm{f}}_{i,:} \approx (k-1)^{-1}(\overset{i}{\boldsymbol{\Gamma}}\overset{i}{\mathbf{H}}^{\mathrm{T}})_{i,:} \circ [\mathbf{A}^{\mathrm{f}}_{i,:}(\overset{i}{\mathbf{H}}\mathbf{A}^{\mathrm{f}})^{\mathrm{T}}]$}
    $\overset{i}{\mathbf{V}}_{i,:}:\ \overset{i}{\mathbf{V}}_{il} = \exp\left\{-0.5\left[d(\mathbf{r}_i,\mathbf{r}_l)/L_i\right]^2\right\}, l = 1,\dots,p_i$
    $\overset{i}{\mathbf{G}}^{\mathrm{f}}_{i,l} \approx (k-1)^{-1}\overset{i}{\mathbf{V}}_{il}[\mathbf{A}^{\mathrm{f}}_{i,:}(\overset{i}{\mathbf{H}}\mathbf{A}^{\mathrm{f}})_{l,:}], l = 1,\dots,p_i$
    {Background error covariance matrices}
    definition of $\langle\dots\rangle$: $\langle\mathbf{c}\rangle = \sum_{l=1}^{p_i}(\overset{i}{\mathbf{V}}_{il}\mathbf{c}_l)/\sum_{l=1}^{p_i}(\overset{i}{\mathbf{V}}_{il})$, where $\mathbf{c}$ is a generic $p_i$ vector
    $\overset{i}{\sigma}{}^2_f = \alpha\langle\mathrm{diag}(\overset{i}{\mathbf{S}}^{\mathrm{f}})\rangle$; $\overset{i}{\sigma}{}^2_{ob} = \alpha\,\langle(\overset{i}{\mathbf{y}}^{\mathrm{o}} - \overset{i}{\mathbf{y}}^{\mathrm{b}})^2\rangle$
    if $(\overset{i}{\sigma}{}^2_{ob} = 0)$and$(\overset{i}{\sigma}{}^2_f = 0)$ then
      $\mathbf{x}^{\mathrm{a}}_i = \mathbf{x}^{\mathrm{b}}_i$, then apply the direct inverse data transformation $\tilde{\mathbf{x}}^{\mathrm{a}}_i = g^{-1}(\mathbf{x}^{\mathrm{a}}_i)$ and STOP
    else if $[\overset{i}{\sigma}{}^2_{ob}/\left(1+\varepsilon^2\right)] \leq \overset{i}{\sigma}{}^2_f$ then
      $\overset{i}{\sigma}{}^2_u = 0$; $\overset{i}{\mathbf{S}}^{\mathrm{b}} = \overset{i}{\mathbf{S}}^{\mathrm{f}}$; $\overset{i}{\mathbf{G}}^{\mathrm{b}}_{i,:} = \overset{i}{\mathbf{G}}^{\mathrm{f}}_{i,:}$
    else
      $\overset{i}{\sigma}{}^2_u = \overset{i}{\sigma}{}^2_{ob}/\left(1+\varepsilon^2\right) - \overset{i}{\sigma}{}^2_f$
      {add the scale matrix $\overset{i}{\sigma}{}^2_u\overset{i}{\boldsymbol{\Gamma}}^{\mathrm{u}}$ to the background error covariance matrices}
      $\overset{i}{\mathbf{S}}^{\mathrm{b}}:\ \overset{i}{\mathbf{S}}^{\mathrm{b}}_{jl} = \overset{i}{\mathbf{S}}^{\mathrm{f}}_{jl} + \overset{i}{\sigma}{}^2_u\exp\left\{-0.5\left[d(\mathbf{r}_j,\mathbf{r}_l)/D_i\right]^2\right\}, j = 1,\dots,p_i$ and $l = 1,\dots,p_i$
      $\overset{i}{\mathbf{G}}^{\mathrm{b}}_{i,:}:\ \overset{i}{\mathbf{G}}^{\mathrm{b}}_{il} = \overset{i}{\mathbf{G}}^{\mathrm{f}}_{il} + \overset{i}{\sigma}{}^2_u\exp\left\{-0.5\left[d(\mathbf{r}_i,\mathbf{r}_l)/D_i\right]^2\right\}, l = 1,\dots,p_i$
    end if
    Observation error covariance matrix: first $\overset{i}{\sigma}{}^2_b = \overset{i}{\sigma}{}^2_f + \overset{i}{\sigma}{}^2_u$, then $\mathrm{diag}(\overset{i}{\mathbf{R}}) = \varepsilon^2\,\overset{i}{\sigma}{}^2_b$
    {Analysis}
    $\mathbf{x}^{\mathrm{a}}_i = \mathbf{x}^{\mathrm{b}}_i + \overset{i}{\mathbf{G}}^{\mathrm{b}}_{i,:}(\overset{i}{\mathbf{S}}^{\mathrm{b}} + \overset{i}{\mathbf{R}})^{-1}\left(\overset{i}{\mathbf{y}}^{\mathrm{o}} - \overset{i}{\mathbf{y}}^{\mathrm{b}}\right)$
    $(\boldsymbol{\sigma}^2)^{\mathrm{a}}_i = \mathbf{P}^{\mathrm{f}}_{ii} + \overset{i}{\sigma}{}^2_u - \overset{i}{\mathbf{G}}^{\mathrm{b}}_{i,:}(\overset{i}{\mathbf{S}}^{\mathrm{b}} + \overset{i}{\mathbf{R}})^{-1}(\overset{i}{\mathbf{G}}^{\mathrm{b}}_{i,:})^{\mathrm{T}}$
    {Data back transformation}
    direct inverse transformation $g^{-1}$ of 400 quantiles of the distribution $N\left(\mathbf{x}^{\mathrm{a}}_i,(\boldsymbol{\sigma}^2)^{\mathrm{a}}_i\right)$, then the optimal gamma distribution parameters $\alpha$ and $\beta$, shape and rate respectively, are obtained through a least squares fitting method.
  end for
```

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

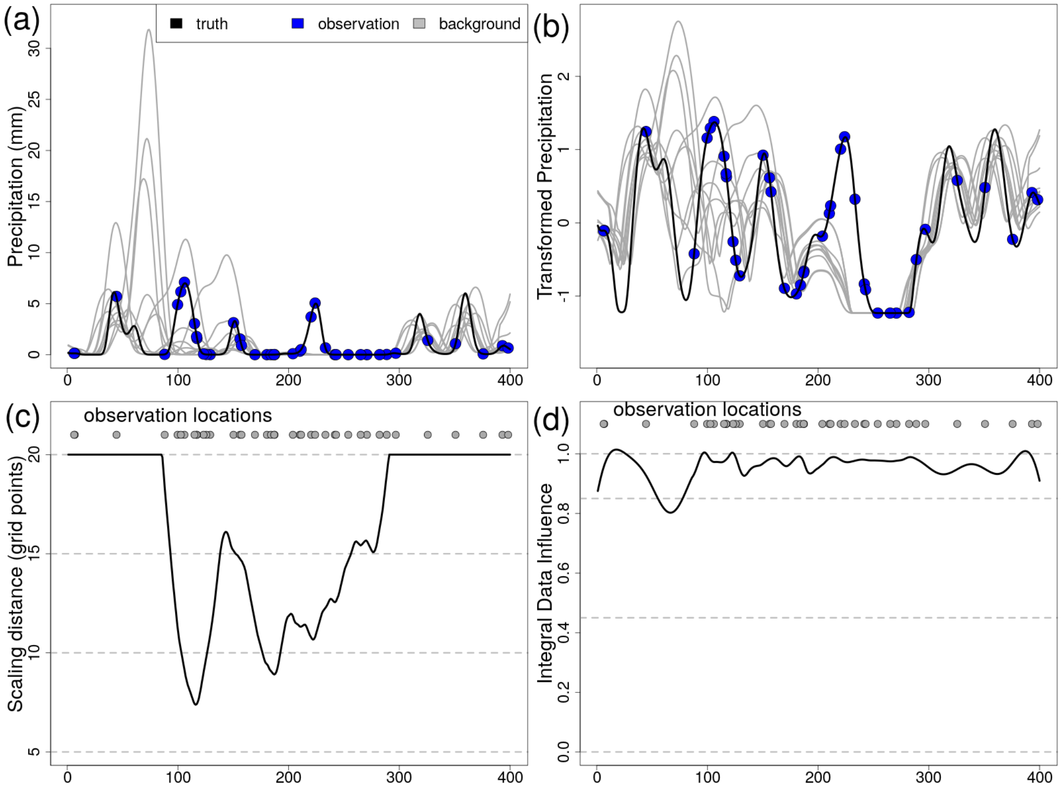

Figure 1. One-dimensional simulation. Panel *a*, precipitation (mm): truth (black line), observations (blue dots) and background (gray lines). Panel *b*, transformed values. Panel *c*, reference length scale for the scale matrix D_i (units u, as defined in Sec. 3.1.1), D_i is bounded within 3 u and 20 u. Panel *d*, Integral Data Influence (IDI) based on D_i from panel *c*.

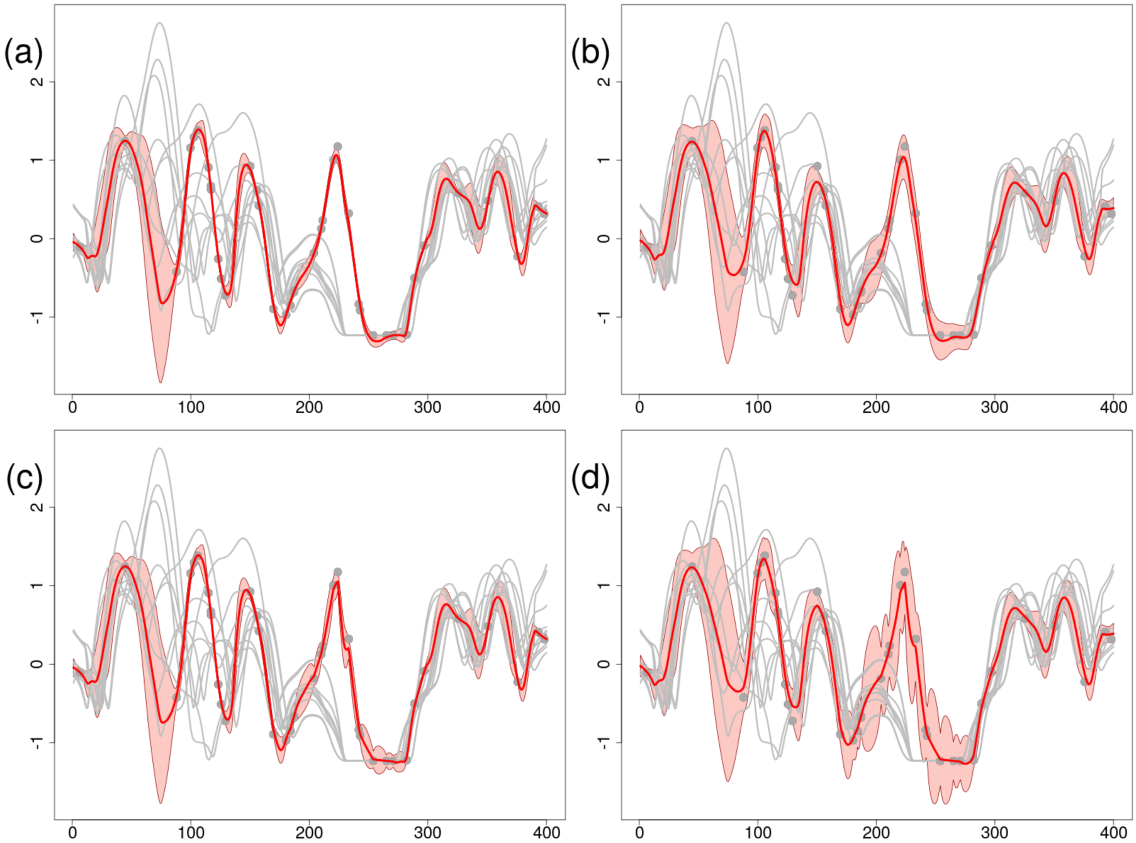

Figure 2. One-dimensional simulation in the transformed precipitation space. Analyses at grid points with different configurations. For all panels: $\varepsilon^2 = 0.1$, $L = 25\,u$. Panel *a*: $\overset{i}{\boldsymbol{\Gamma}}^{\mathrm{u}}$ with Gaussian function, $\alpha = 0.1$. Panel *b*: $\overset{i}{\boldsymbol{\Gamma}}^{\mathrm{u}}$ with Gaussian function, $\alpha = 1.0$. Panel *c*: $\overset{i}{\boldsymbol{\Gamma}}^{\mathrm{u}}$ with exponential function, $\alpha = 0.1$. Panel *d*: $\overset{i}{\boldsymbol{\Gamma}}^{\mathrm{u}}$ with exponential function, $\alpha = 1.0$. For each panel, the red line is the analysis (expected value), the pink area shows the interval between the 90th and the 10th percentiles, the gray dots and lines are the observations and backgrounds, respectively (see panel *b* in Fig. 1).

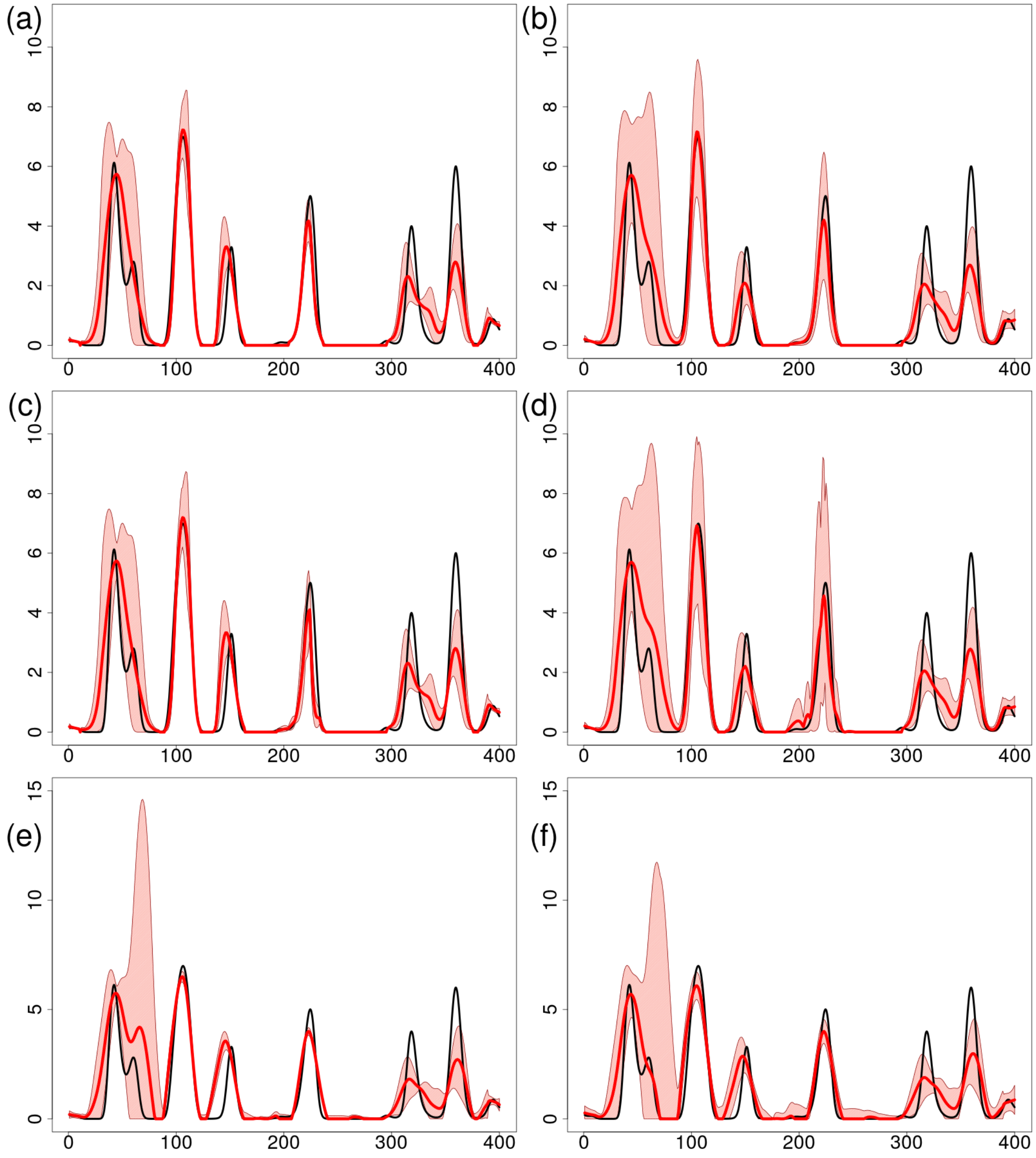

Figure 3. One-dimensional simulation in the original precipitation space (mm). Analyses at grid points with different configurations. For all panels: $\varepsilon^2 = 0.1$, $L = 25\,u$. The red line is the analysis (expected value), the pink area shows the 90th-10th percentile, the black line is the truth. Panels *a-d* as in Fig. 2. Panel *e*: no data transformation, $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{u}}$ with Gaussian function, $\alpha = 0.1$. Panel *f*: no data transformation, $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{u}}$ with Gaussian function, $\alpha = 1.0$. Note the different scale for precipitation between panels *e-f* and the others.

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

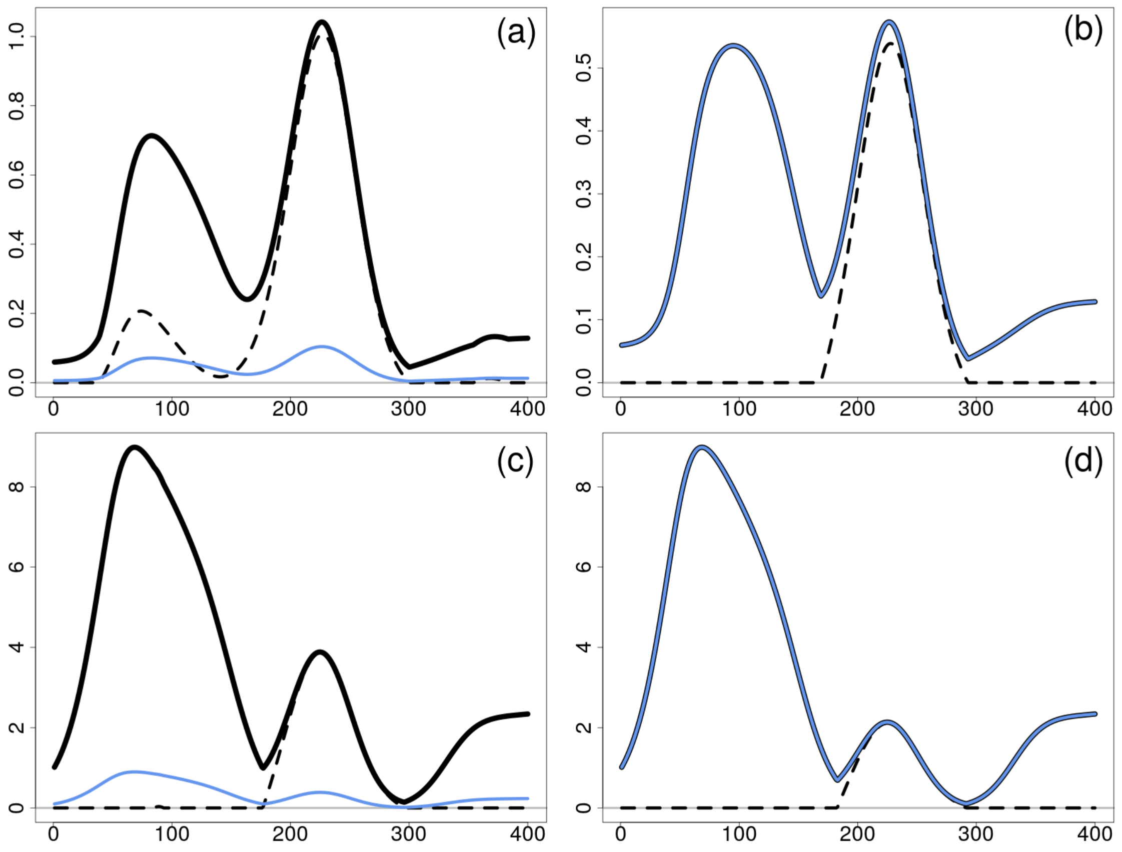

Figure 4. One-dimensional simulation. Error variances obtained with different configurations. The variances shown are: $\overset{i}{\sigma}{}_o^2(=\varepsilon^2\overset{i}{\sigma}{}_b^2)$, blue line; $\overset{i}{\sigma}{}_b^2$, black thick line; $\overset{i}{\sigma}{}_u^2$, black dashed line; $\overset{i}{\sigma}{}_f^2$ as the difference between $\overset{i}{\sigma}{}_b^2$ and $\overset{i}{\sigma}{}_u^2$. For all panels: $\alpha = 1$, $L = 25\,u$ and $\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{u}}$ with Gaussian function. Panel *a* (dimensionless quantities): Gaussian anamorphosis, $\varepsilon^2 = 0.1$. Panel *b* (dimensionless quantities): Gaussian anamorphosis, $\varepsilon^2 = 1.0$. Panel *c* (mm^2): no data transformation, $\varepsilon^2 = 0.1$. Panel *d* (mm^2): no data transformation, $\varepsilon^2 = 1.0$.

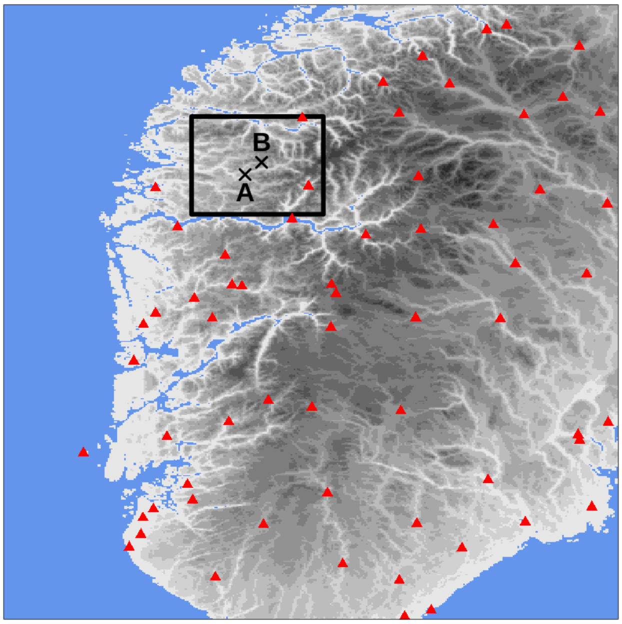

Figure 5. "South Norway" domain used in the simulations of Secs. 3.1.2- 3.2. The red triangles mark station locations used for cross-validation in Sec. 3.2. The gray shades indicate the altitude (from the lighter gray at 0 m to the darker gray at approximately 2400 m a.m.s.l.). The blue shade indicates the sea. The black box delimits the "Sogn og Fjordane" domain shown in Fig. 8, the crosses mark the two points A and B used in the following.

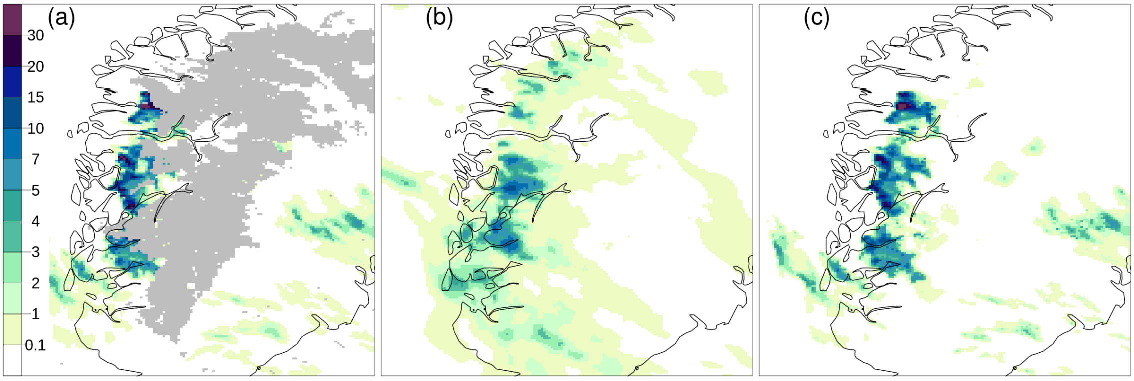

Figure 6. 2019-07-30 15:00 UTC, hourly precipitation totals over South Norway (see Fig. 5). Observations are shown in panel *a* over the same grid as the analysis. For each grid cell, the average of the observed values within the cell is shown. The background ensemble mean is shown in panel *b*. The analysis expected value is shown in panel *c*. The color scale is the same for all panels.

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

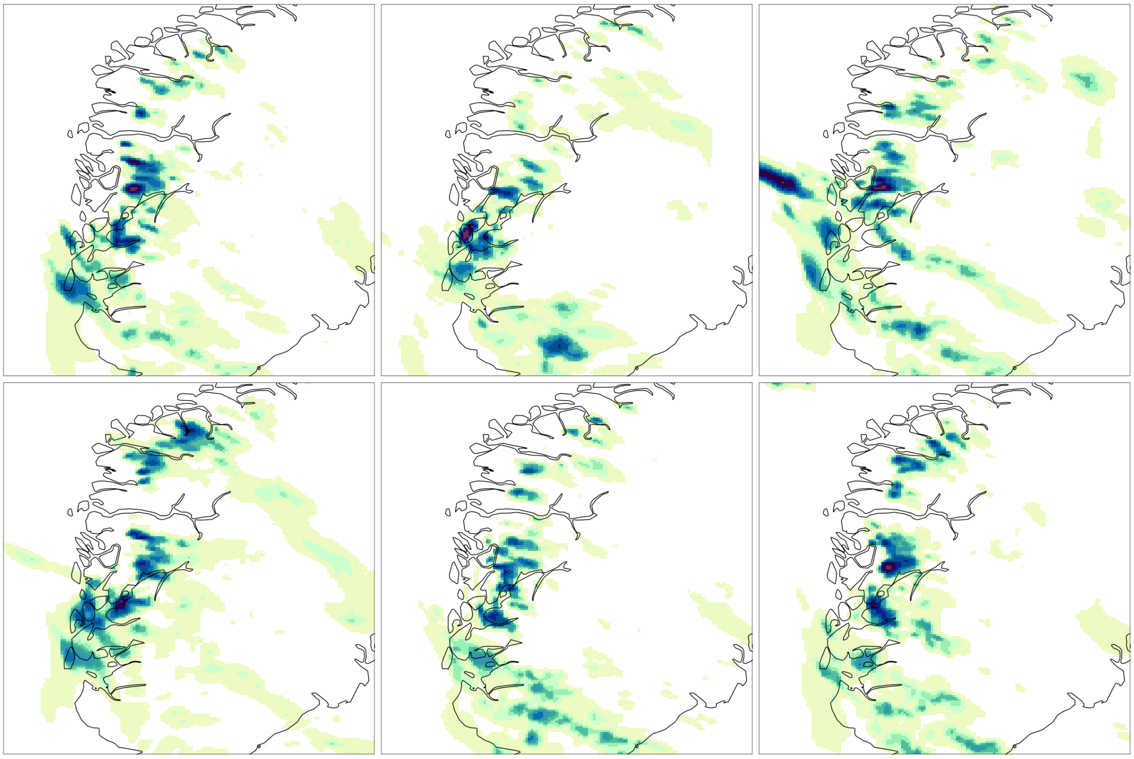

Figure 7. 2019-07-30 15:00 UTC, hourly precipitation totals over South Norway (see Fig. 5) for six of the ten background ensemble members. The color scale is the same as in Fig. 6.

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

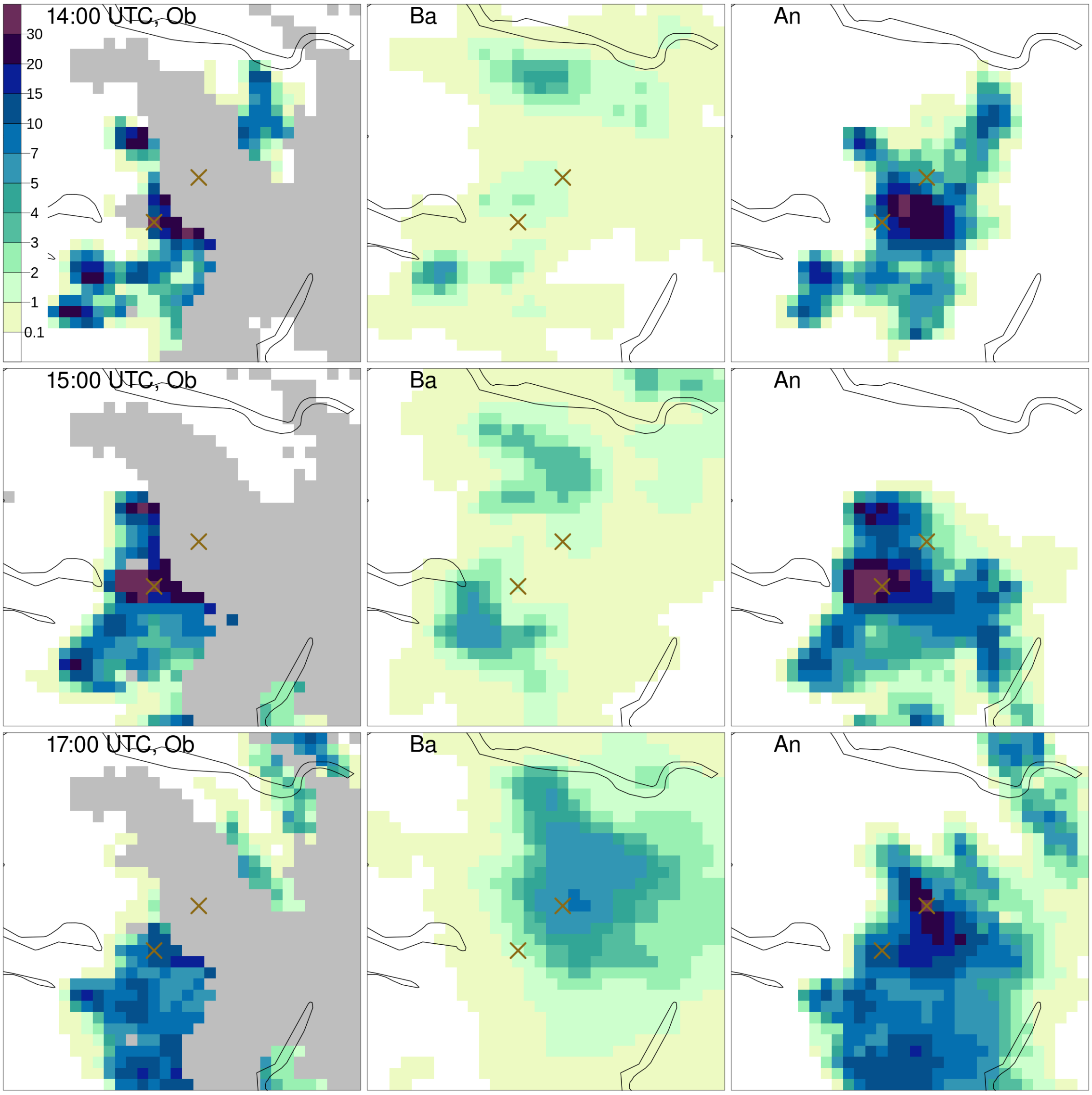

Figure 8. 2019-07-30 14:00 UTC (top row), 15:00 UTC (middle row), 17:00 UTC (bottom row) hourly precipitation totals over Sogn og Fjordane (see Fig. 5). The panels labeled with *Ob* (left column) show the aggregated observed values, as in Fig. 6. The panels with *Ba* (middle column) show the background ensemble mean. The panels with *An* (right column) show the analysis expected value. The crosses mark the A and B points of Fig. 5. The color scale is the same for all panels.

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

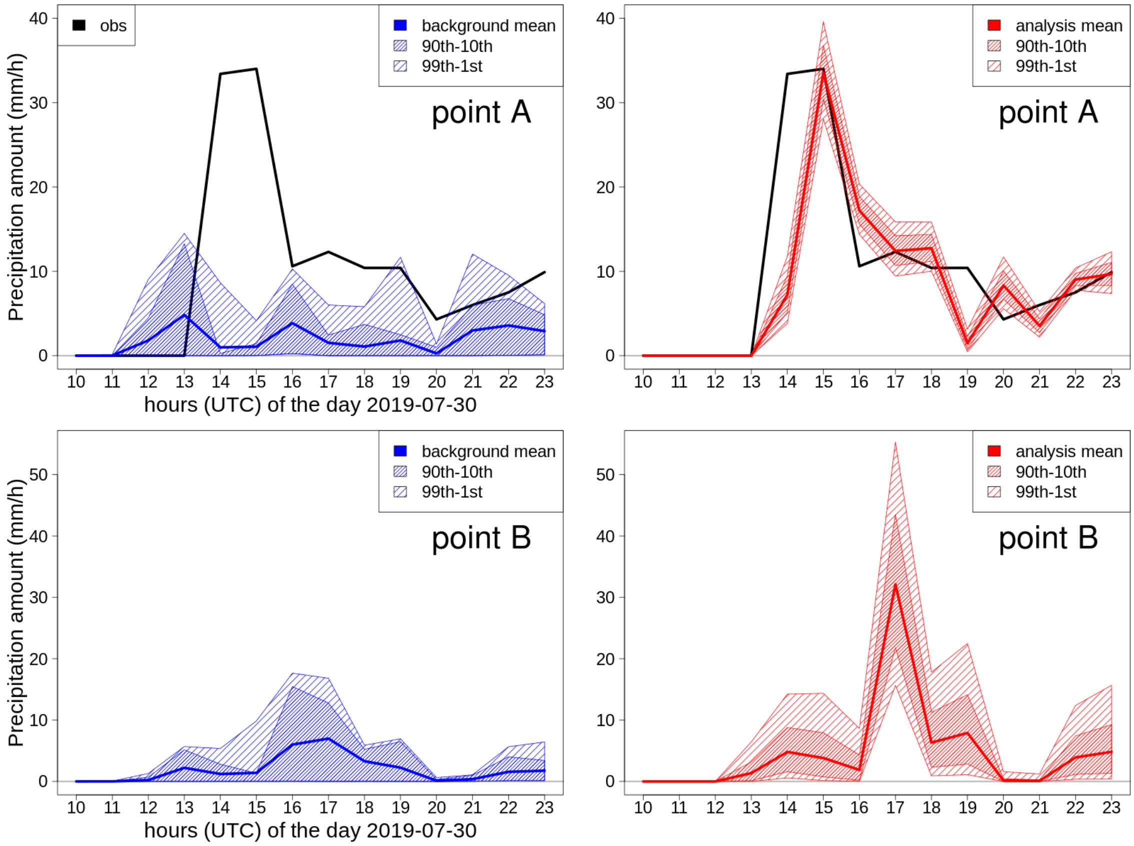

Figure 9. Time series of hourly precipitation totals for the period 2019-07-30 10:00 UTC to 23:00 UTC at points A (top row) and B (bottom row) of Fig. 5. The left panels show the background (blue). The right panels show the analysis (red). The blue (red) line shows the background (analysis) mean, the region with denser shading lines is the difference between the 90th and the 10th percentiles, the region with sparser shading lines is the difference between the 99th and the 1st percentiles. For point A, the closest observation, which is a radar-derived estimate, is shown (black line). Point B is in a region where observations are not available.

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

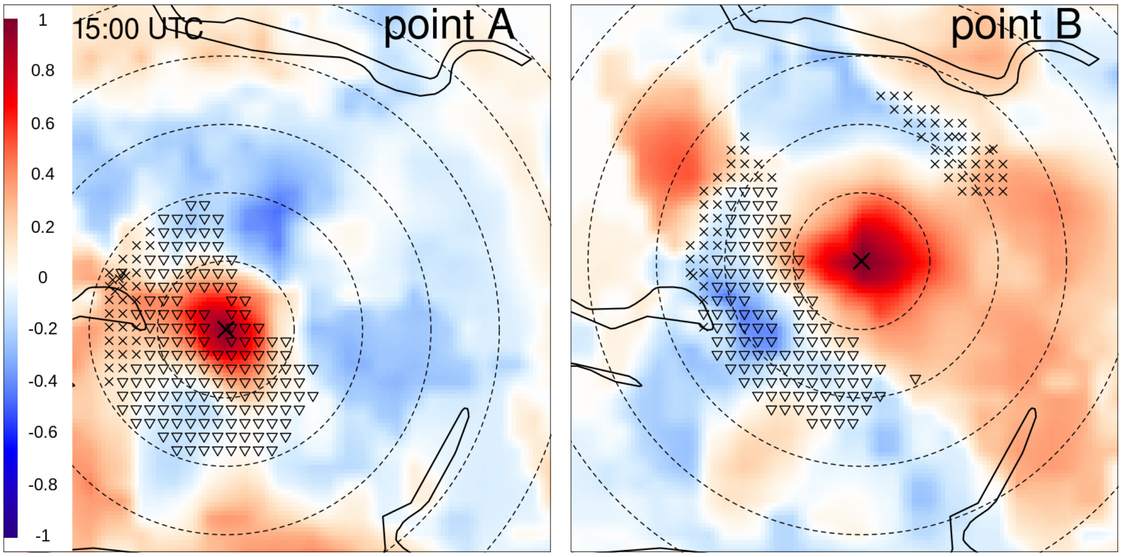

Figure 10. 2019-07-30 15:00 UTC, background error correlations $\overset{i}{\mathbf{\Gamma}}^{\mathrm{b}}{}_{i,:}$ of Eq. (16) used for spatial analysis of hourly precipitation totals over Sogn og Fjordane (see Fig. 5). The blue-red color shades show the background error correlations. With reference to Fig. 5, the left panel shows the background error correlations between point A and the grid points. For point B, the same quantity is shown in the right panel. The symbols show the closest 200 observations, the triangles are observations of precipitation, while the crosses are observations of no-precipitation. The concentric circles have their common center at either point A or B and they are distance isolines at: 10 km, 20 km, 30 km, 40 km and 50 km.

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

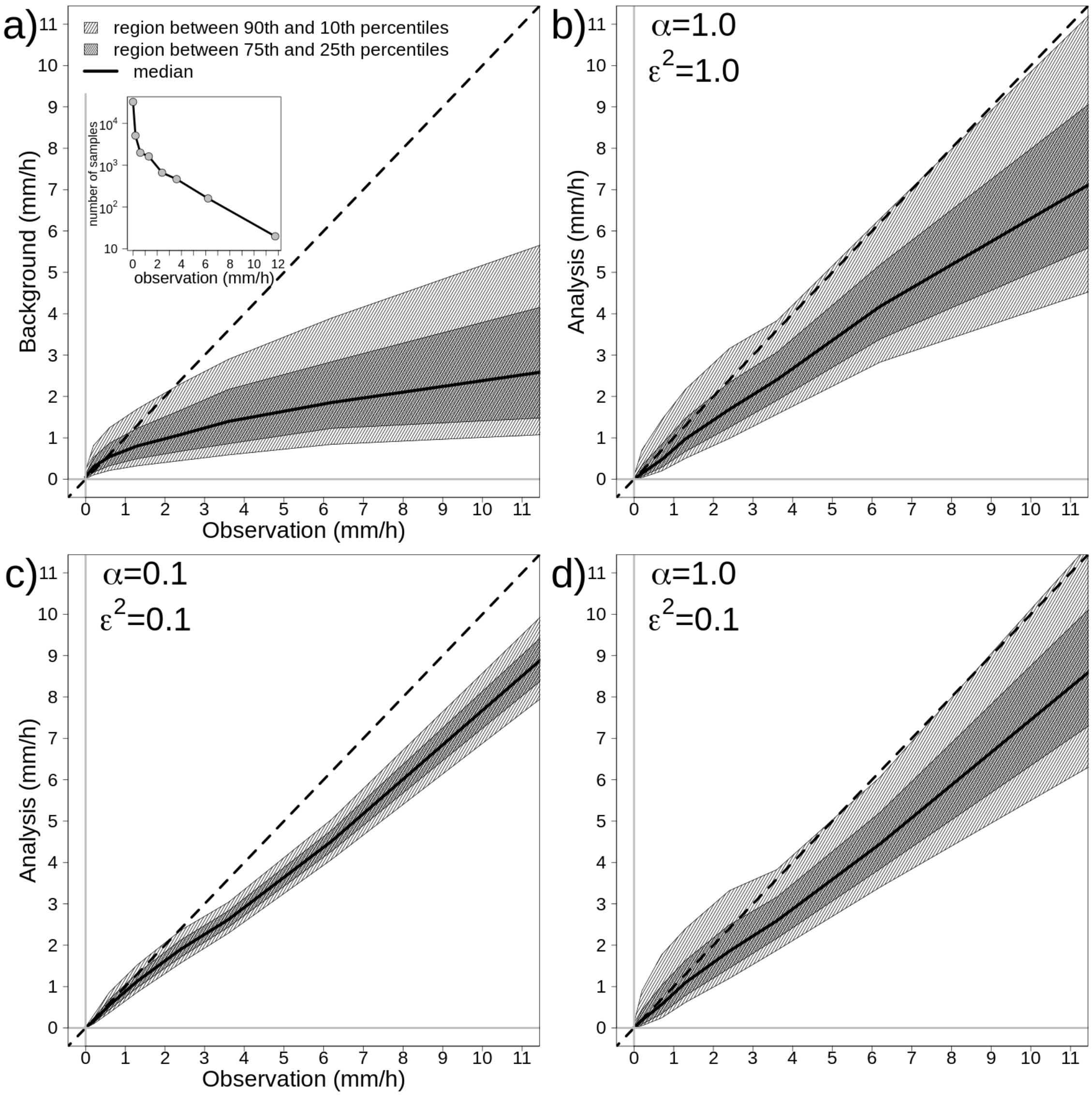

Figure 11. Summer 2019 hourly precipitation statistics for the cross-validation experiments. Panels: *a* background versus observations; *b* analysis $\varepsilon^2 = 1$ $\alpha = 1$ versus observations; *c* analysis $\varepsilon^2 = 0.1$ $\alpha = 0.1$ versus observations; *d* analysis $\varepsilon^2 = 1.0$ $\alpha = 0.1$ versus observations. The independent observations have been divided into classes, the number of samples within each class is shown in the inset of panel *a*. Within each class and for each probabilistic prediction, several percentiles have been computed. The regions between the average of the 90th and the 10th percentiles are shown by light gray shades. The regions between the average of the 75th and the 25th percentiles are shown by dark gray shades. The thick black line indicates the average of the medians. The dashed black line is the diagonal (1:1) line.

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

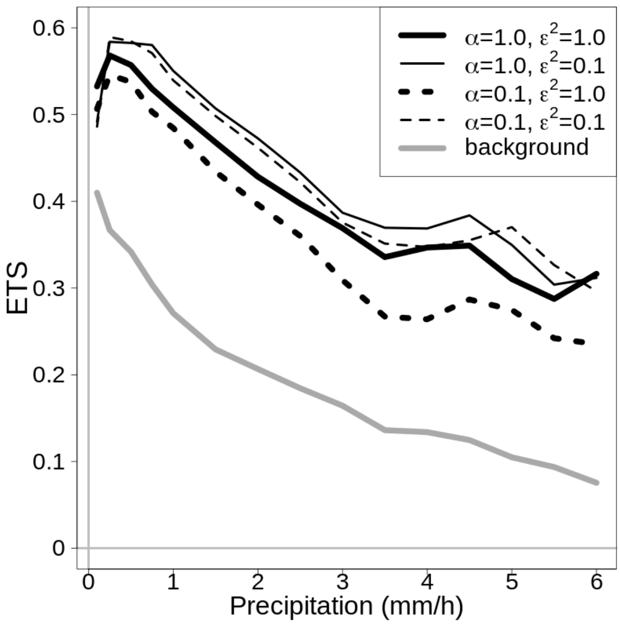

Figure 12. Equitable Threat Score (ETS) for summer 2019 hourly precipitation, as obtained through the cross-validation experiments. The black lines are the ETS curves for the analysis mean values, as indicated in the legend. The ETS curve for the background is the gray line. The precipitation thresholds defining the ”yes” events are reported on the x-axis.

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

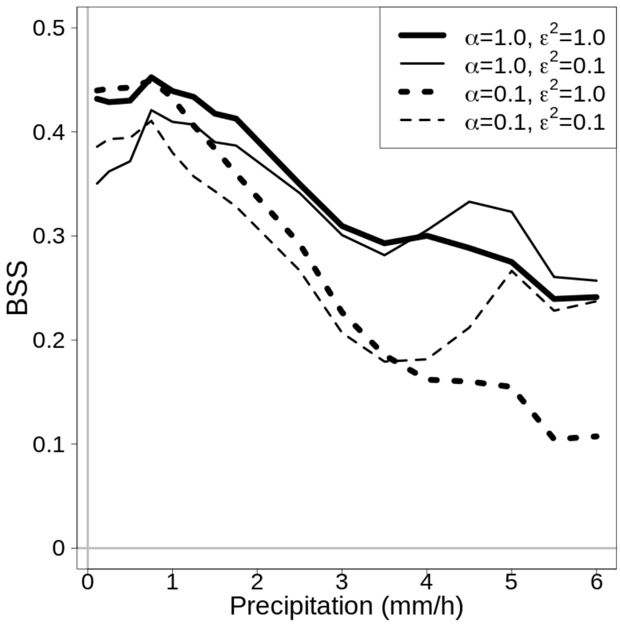

Figure 13. Brier Skill Score (BSS) for summer 2019 hourly precipitation, as obtained through the cross-validation experiments. Relative skill of the analysis probabilistic predictions over that of the background. Different analysis configurations are considered, as reported in the legend.

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

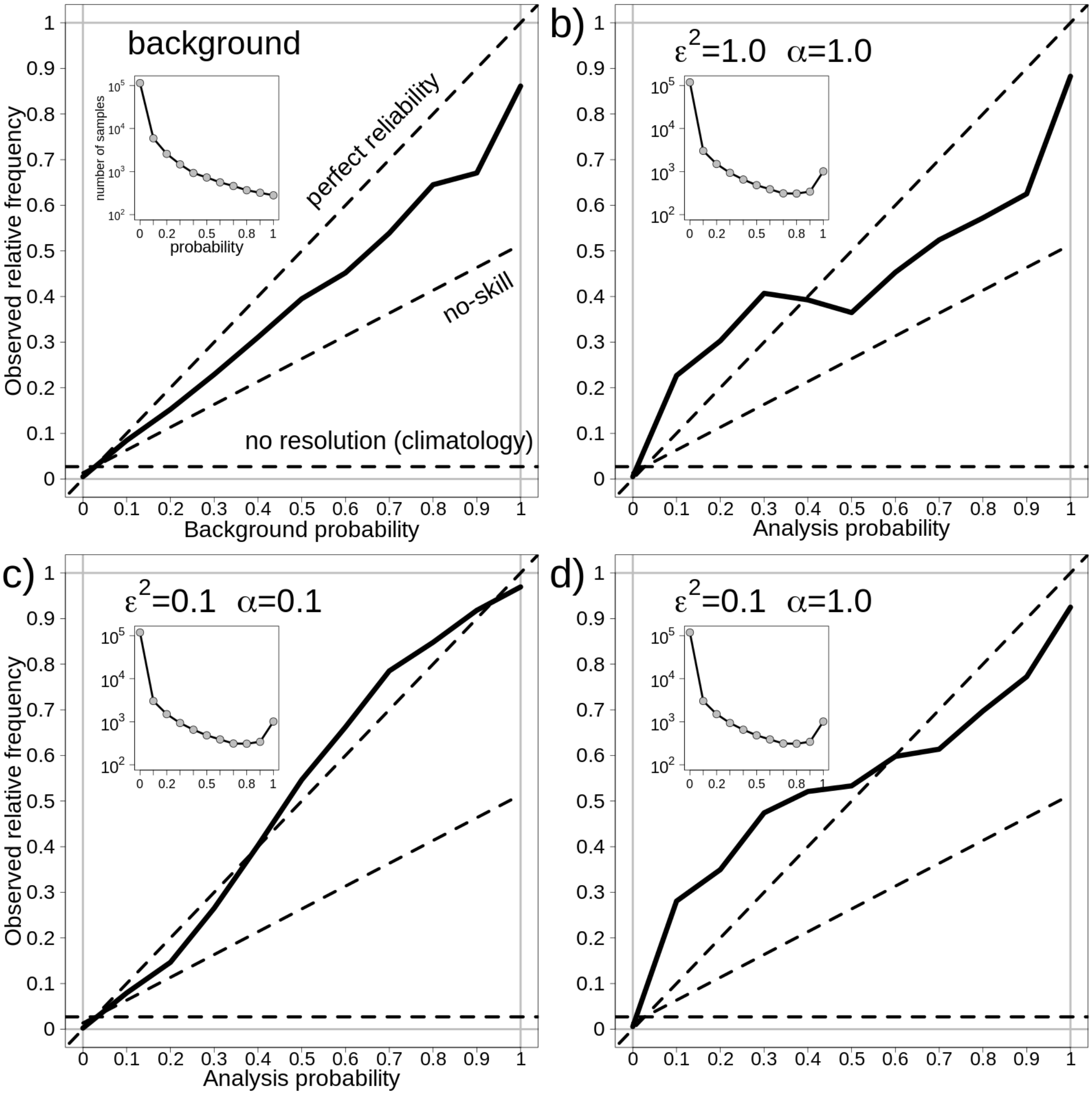

Figure 14. Reliability diagrams for summer 2019 hourly precipitation with a threshold of 1 mm/h, as obtained through the cross-validation experiments. The no-resolution and no-skill w.r.t. climatology lines are shown, as it is the diagonal (1:1) perfect reliability line. Panel *a* shows the background. Panels *b-d* show the analysis with the different configurations.

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

Table 1. Overview of variables and notation for global variables. All the vectors are column vectors if not otherwise specified. If $\mathbf{X}$ is a matrix, $\mathbf{X}_i$ is its ith column (column vector) and $\mathbf{X}_{i,:}$ is its ith row (row vector).

symbol	description	space	dimension
m	number of grid points	-	-
p	number of observations	-	-
k	number of forecast ensemble members	-	-
$\tilde{\mathbf{X}}^{\mathrm{f}}$	forecast ensemble	original	mxk matrix
$\mathbf{X}^{\mathrm{f}}$	forecast ensemble	transformed	mxk matrix
$\mathbf{x}^{\mathrm{f}}$	forecast ensemble mean	transformed	p vector
$\mathbf{A}^{\mathrm{f}}$	forecast perturbations	transformed	mxk matrix
$\mathbf{P}^{\mathrm{f}}$	forecast covariance matrix	transformed	mxm matrix
$\tilde{\mathbf{y}}^{\mathrm{o}}$	observations	original	p vector
$\mathbf{y}^{\mathrm{o}}$	observations	transformed	p vector
$\mathbf{x}^{\mathrm{t}}$	truth	transformed	m vector
$\tilde{\mathbf{x}}^{\mathrm{a}}$	analysis	original	m vector
$\mathbf{x}^{\mathrm{a}}$	analysis	transformed	m vector
$\boldsymbol{\eta}^{\mathrm{a}}$	analysis error	transformed	m vector
$\mathbf{P}^{\mathrm{a}}$	analysis error covariance matrix	transformed	mxm matrix
$\boldsymbol{\sigma}^{\mathrm{a}}$	analysis error standard deviation, $\sqrt{\mathrm{diag}\left(\mathbf{P}^{\mathrm{a}}\right)}$	transformed	m vector
$\mathbf{x}^{\mathrm{b}}$	background	transformed	m vector
$\boldsymbol{\eta}^{\mathrm{b}}$	background error	transformed	m vector
$\mathbf{P}^{\mathrm{b}}$	background error covariance matrix	transformed	mxm matrix
$\boldsymbol{\varepsilon}^{\mathrm{o}}$	observation error	transformed	p vector
$\mathbf{H}$	observation operator	transformed	pxm matrix
$\mathbf{L}$	reference length scales for localization	transformed	m vector
$\mathbf{D}$	reference length scales of the scale matrix	transformed	m vector
ε^2	relative quality of the background wrt observations	transformed	-
α	stabilization coefficient	transformed	-

https://doi.org/10.5194/npg-2020-20
Preprint. Discussion started: 19 June 2020

Table 2. Overview of variables and notation for local variables. All variables are specified in the transformed space. All the vectors are column vectors if not otherwise specified. If $\mathbf{X}$ is a matrix, $\mathbf{X}_i$ is its ith column (column vector) and $\mathbf{X}_{i,:}$ is its ith row (row vector).

symbol	description	dimension
p_i	number of observations in the surroundings of the ith grid point	-
$\overset{i}{\mathbf{H}}$	observation operator	$p_i \mathrm{x} m$ matrix
$\overset{i}{\mathbf{R}}$	observation error covariance matrix	$p_i \mathrm{x} p_i$ matrix
$\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{o}}$	observation error correlation matrix	$p_i \mathrm{x} p_i$ matrix
$\overset{i}{\mathbf{y}}{}^{\mathrm{b}}$	background at observation locations	p_i vector
$\overset{i}{\mathbf{P}}{}^{\mathrm{b}}$	background error covariance matrix	$m \mathrm{x} m$ matrix
$\overset{i}{\mathbf{\Gamma}}$	localization matrix	$m \mathrm{x} m$ matrix
$\overset{i}{\mathbf{V}}$	localization between grid points and observation locations	$m \mathrm{x} p_i$ matrix
$\overset{i}{\mathbf{Z}}$	localization between observation locations	$p_i \mathrm{x} p_i$ matrix
$\overset{i}{\mathbf{\Gamma}}{}^{\mathrm{u}}$	scale correlation matrix	$m \mathrm{x} m$ matrix
$\overset{i}{\mathbf{G}}{}^{\mathrm{b}}$	background error covariances between grid points and observation locations	$m \mathrm{x} p_i$ matrix
$\overset{i}{\mathbf{S}}{}^{\mathrm{b}}$	background error covariances between observation locations	$p_i \mathrm{x} p_i$ matrix
$\overset{i}{\mathbf{G}}{}^{\mathrm{f}}$	forecast error covariances between grid points and observation locations	$m \mathrm{x} p_i$ matrix
$\overset{i}{\mathbf{S}}{}^{\mathrm{f}}$	forecast error covariances between observation locations	$p_i \mathrm{x} p_i$ matrix
$\overset{i}{\sigma}{}_o^2$	observation error variance	-
$\overset{i}{\sigma}{}_b^2$	average background error variance	-
$\overset{i}{\sigma}{}_f^2$	average forecast error variance	-
$\overset{i}{\sigma}{}_u^2$	error variance for the scale matrix	-
$\overset{i}{\sigma}{}_{ob}^2$	sum of error variances (Eq. (10))	-