

**RESEARCH ARTICLE**

# Estimating the benefit of Doppler wind lidars for short-term low-level wind ensemble forecasts

Tatiana Nomokonova<sup>1,2</sup>  | Philipp J. Griewank<sup>3</sup>  | Ulrich Löhnert<sup>1,2</sup>  |  
Takemasa Miyoshi<sup>4</sup>  | Tobias Necker<sup>3</sup>  | Martin Weissmann<sup>3</sup> 

<sup>1</sup>Institute for Geophysics and Meteorology,  
University of Cologne, Cologne, Germany

<sup>2</sup>Hans-Ertel-Centre for Weather Research,  
Climate Monitoring and Diagnostics,  
Cologne/Bonn, Germany

<sup>3</sup>Department of Meteorology and  
Geophysics, University of Vienna, Vienna,  
Austria

<sup>4</sup>RIKEN Center for Computational  
Science, Kobe, Japan

**Correspondence**

U. Löhnert, Institute for Geophysics and  
Meteorology, University of Cologne,  
Albertus-Magnus-Platz, 50923 Cologne,  
Germany.

Email: [ulrich.loehnert@uni-koeln.de](mailto:ulrich.loehnert@uni-koeln.de)

**Funding information**

Hans-Ertel-Centre for Weather Research  
funded by the German Federal Ministry  
for Transportation and Digital  
Infrastructure, Grant/Award Number:  
BMVI/DWD4818DWD5B

**Abstract**

This work focuses on the potential of a network of Doppler lidars for the improvement of short-term forecasts of low-level wind. For the impact assessment, we developed a new methodology that is based on ensemble sensitivity analysis (ESA). In contrast to preceding network design studies using ESA, we calculate the explicit sensitivity including the inverse of the background covariance **B** matrix to account directly for the localization scale of the assimilation system. The new method is applied to a pre-existing convective-scale 1,000-member ensemble simulation to mitigate effects of spurious correlations. We evaluate relative changes in the variance of a forecast metric, that is, the low-level wind components averaged over the Rhein–Ruhr metropolitan area in Germany. This setup allows us to compare the relative variance change associated with the assimilation of hypothetical observations from a Doppler wind lidar with respect to the assimilation of surface-wind observations only. Furthermore, we assess sensitivities of derived variance changes to a number of settings, namely observation errors, localization length scale, regularization factor, number of instruments in the network, and their location, as well as data availability of the lidar measurements. Our results demonstrate that a network of 20–30 Doppler lidars leads to a considerable variance reduction of the forecast metric chosen. On average, an additional network of 25 Doppler lidars can reduce the 1–3 hr forecast error by a factor of 1.6–3.3 with respect to 10-m wind observations only. The results provide the basis for designing an operational network of Doppler lidars for the improvement of short-term low-level wind forecasts that could be especially valuable for the renewable energy sector.

**KEYWORDS**

covariance, data assimilation, ensemble sensitivity analysis, localization, low-level wind forecasts, network of Doppler lidars, observing system

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Quarterly Journal of the Royal Meteorological Society* published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

## 1 | INTRODUCTION

An accurate short-term wind forecast is crucial for the energy sector, industry, and civil services. For example, a reliable short-term wind forecast allows for a reduction of the economic impact of extreme weather, time schedule adjustment in transport networks, and enhanced population safety. The forecast also provides valuable information for electricity production controllers to manage conventional power sources in a timely fashion according to instantaneous changes in available renewable-source power (Sweeney *et al.*, 2020), and wind-based generation in particular. The changes in electricity production by wind farms are backed up by power plants of different, weather-independent types. However, a considerable change in electrical power output requires a lead time, which varies from 10 s to minutes for hydro reservoirs and simple-cycle gas turbines to a day for nuclear power plants (Gonzalez-Salazar *et al.*, 2018). In addition, some power plants, especially coal-fired ones, emit greenhouse gases and air pollutants, and therefore the allocation of electricity generation to such power plants needs to be minimized. Thus, a proper management of electricity generation requires beforehand knowledge of the low-level wind properties.

A reliable low-level wind forecast can help to estimate the available electricity generation by wind farms and therefore is beneficial for the distribution of energy resources. Unfortunately, low-level wind forecasts have a relatively large uncertainty, due to the high temporal and spatial variability of the wind field. One way to improve numerical weather prediction (NWP) forecasts substantially is the assimilation of additional, new observational data (Kalnay, 2002).

There are a number of studies showing the benefits of data assimilation (DA) for low-level wind forecasts using real observations from a single wind-profiling instrument (Sawada *et al.*, 2015; Pichugina *et al.*, 2017; Finn *et al.*, 2020; Li *et al.*, 2020; Hristova-Veleva *et al.*, 2021). For operational DA, however, it is beneficial to build a dense network of remote-sensing sites capable of continuous profiling of wind in the atmospheric boundary layer (ABL). Before investing in such a network, however, more investigation is required on what kind of instrumentation to acquire, how dense the network should be, and what level of forecast improvement to expect. On the one hand, utilization of real wind profiles for DA experiments is beneficial because it allows for a direct incorporation of observations and their measurement-error covariances and allows us to quantify the impact considering the actual atmospheric variability. On the other hand, this approach provides only limited information to estimate the quantitative benefit of a planned instrument

network and identify optimal locations of the instruments.

One approach to obtain this information is based on the ensemble transform Kalman filter (Bishop *et al.*, 2001; Majumdar *et al.*, 2001), where both the analysis variance reduction due to the assimilation of a hypothetical observation and the sensitivity of a forecast parameter to this analysis variance reduction can be estimated based on information from a precalculated forecast ensemble. Ancell and Hakim (2007) developed an efficient approach for this ensemble sensitivity analysis (ESA) to estimate the forecast variance reduction due to additional observations.

ESA has been used in a number of studies to provide guidance for targeted airborne dropsonde observations for tropical cyclones (Torn, 2014; Majumdar, 2016) and more recently in two network design studies for automatic weather stations in Antarctica (Hakim *et al.*, 2020; Tardif *et al.*, 2021). The efficiency of the ESA approach is based on the cancellation of the background covariance  $\mathbf{B}$  matrix determining the analysis influence and the inverse background covariance  $\mathbf{B}$  matrix in the ensemble sensitivity in the absence of localization. The analysis background covariance  $\mathbf{B}$  matrix is, however, usually localized, which can only be accounted for crudely in the efficient ESA approach of Ancell and Hakim (2007). In ESA, localization is usually applied to the forecast metric instead of the analysis influence, which introduces errors, given an unknown signal propagation speed in the atmosphere. Further uncertainty is introduced by different signal propagation for different variables and different perturbation scales in the atmosphere. Therefore most preceding ESA analyzed only the optimal location of one observation type, but not optimal networks of measurements of different variables. To overcome this limitation, we propose and apply a new approach for ESA that calculates both the analysis variance reduction and the forecast sensitivity explicitly, to allow for a correct representation of localization in the assimilation system.

The quality of ESA results furthermore depends strongly on the number of ensemble members. Many operational global NWP models only produce 14–50 ensemble forecasts due to computational restrictions (Leutbecher, 2019; Buizza and Richardson, 2017). Jacques and Zawadzki (2015), Leutbecher (2019), and Necker *et al.* (2020a) showed that a large number of ensemble members reduces covariance sampling errors considerably in comparison with smaller ensembles. Miyoshi *et al.* (2014) performed their first experiments with a 10,240-member global model ensemble and compared the results with those obtained from datasets with fewer ensemble members. They found that model ensembles with fewer than 100 members were affected considerably by spurious correlation and could not capture the non-Gaussian shape of

the probability density functions of certain atmospheric variables. The authors concluded that an ensemble with at least 1,000 members is not affected much by these two issues. Houtekamer and Zhang (2016) and Necker *et al.* (2020b) also showed that a high number of members allows for a reduction of statistical errors and spurious correlations and by this enables more accurate estimates of error covariances to be obtained. In this study, we therefore use the 1,000-member ensemble dataset of Necker *et al.* (2020b) to mitigate the effects of spurious correlations.

There are a number of instruments providing measurements of low-level wind potentially available or already being assimilated for short-term weather forecasts. One of the least expensive instruments that can provide reliable wind measurements is a surface weather station. This instrument typically measures wind at 2 or 10 m height and cannot characterize air motions throughout the complete ABL. A meteorological tower can deliver wind profiles up to a few 100 m, also including the typical hub height of wind turbines (around 80 m). However, a meteorological tower is expensive to build and maintain and it can only characterize one location. Wind sensors that are typically maintained on wind turbines provide wind observations at a certain height, but not wind profiles. Radiosondes can provide high-quality and high vertical resolution profiles of wind. Radiosondes, however, are typically launched only twice a day at distinct locations. Space-borne instruments have the advantage of global coverage, but typically have low temporal, angular, and spatial resolution. For instance, scatterometers with synthetic aperture on board *Sentinel-1A/B* satellites are capable of measuring near-surface winds over oceans with an accuracy of about  $2 \text{ m}\cdot\text{s}^{-1}$  (Ahsbabs *et al.*, 2018). The Doppler lidar Atmospheric LASer Doppler INSTRUMENT (ALADIN) on board the *Aeolus* satellite provides profiles of the zonal wind component with an accuracy varying from 1.9 to  $4.4 \text{ m}\cdot\text{s}^{-1}$  (Martin *et al.*, 2021). These instruments have a recurrence cycle of the order of a week and therefore cannot be used for an improvement of regional short-term wind forecasts. Commercial aircraft deliver wind profiles through the Aircraft Meteorological Data Relay (AMDAR) for limited areas close to airports, since they measure profiles only during their ascents and descents (de Haan, 2011; Petersen, 2016; Stone, 2018). Mode-Selective (Mode-S) is another aircraft-based source for upper-air wind observations, which can be obtained from appropriately equipped aircraft and provide upper-air wind observations giving local coverage of a flight area (de Haan, 2011; Stone, 2018).

Another type of instrument providing wind profiles in the ABL is a ground-based Doppler lidar (Lang and McKee, 2011; Frehlich, 2013; Gottschall *et al.*, 2017; Pichault

*et al.*, 2021). A Doppler lidar has a number of advantages relative to ground-based Sound Detecting And Ranging and radar wind profilers. First, this instrument delivers continuous profiles of low-level wind with high temporal and spatial resolution. Second, it can often profile wind during clear-sky conditions, which is not always possible by radar (Hirth *et al.*, 2017). Third, modern Doppler lidars are relatively cheap, compact, and low-maintenance instruments, making them suitable for building an operational network. Several studies have presented potential benefits of Doppler lidars for wind-power applications. For instance, Pichault *et al.* (2021) presented methodologies to use Doppler lidar observations for a prediction of onshore power generation at a small wind farm. Theuer *et al.* (2020) showed that a single Doppler lidar can be a more cost-effective solution for wind-power output prediction than a dual Doppler radar installation, especially in areas where an installation of relatively big radar systems is not feasible. Frehlich (2013) demonstrated that scanning Doppler lidars can provide information on the turbulence conditions, which is important for wind-turbine efficiency improvements. Thus, Doppler lidars are promising instruments for wind forecast and wind-power prediction. However, more effort and research needs to be directed towards the potential of a network of Doppler lidars for the improvement of short-term forecasts over the larger areas required, for example, by big wind farms.

Within the Hans-Ertel-Centre for Weather Research (HErZ: Weissmann *et al.*, 2014; Simmer *et al.*, 2016) we estimate the potential to improve the short-term forecast of low-level wind using a network of Doppler lidars. As a part of HErZ, this study addresses three important questions.

1. How much improvement in the short-term low-level wind forecast can be expected from Doppler lidars with respect to operationally assimilated surface in situ stations?
2. How does this improvement depend on the number of lidars in the network?
3. How does the improvement gained from lidars depend on the penetration of lidar signal through the ABL?

This study focuses on the urban Rhein–Ruhr Area (RRA) and surrounding regions. This area accommodates more than  $10^7$  people and energy-intensive industry. The RRA features  $\sim 5.1 \text{ GW}$  of installed wind farms, which was 7.4% of the total installed wind-turbine capacity of Germany in 2019. For our analysis, we apply a novel ESA approach to the first convective-scale 1,000-member ensemble simulation over Germany that was previously generated using the Scalable Computing for Advanced

Library and Environment regional model (SCALE-RM), as described in detail by Necker *et al.* (2020a). The simulation uses a full-physics nonhydrostatic regional model and consists of ensemble forecasts from 16 initial times. The period analyzed covers a period in May/June 2016.

The remainder of the article is outlined as follows. The new ESA methodology is described in Section 2. Section 3 provides details on the region and period of interest, as well as the 1,000-member ensemble used in this study. Sections 3.2, 3.3, and 4 discuss the experimental design in detail and present sensitivity studies on the choices of localization, observation errors, and regularization. Section 5 presents the main results and demonstrates the potential impact of a network of hypothetical Doppler lidars relative to operationally assimilated surface observations. First, Section 5 analyzes two case studies focused on the impact of Doppler lidars on low-level wind forecasts for different weather situations. Second, averaged results of our study are presented based on all available 16 forecasts in May/June 2016. Finally, we summarize and discuss the results and offer a short outlook in Section 6.

## 2 | METHOD

In this study we use a new ESA approach to estimate how an assimilation of Doppler lidar data affects the variance of a predicted quantity. The approach used in this study builds upon the ESA methods described in Ancell and Hakim (2007), Torn and Hakim (2008), Torn (2014), Coniglio *et al.* (2019), and Hakim *et al.* (2020), but extends preceding approaches by incorporating analysis localization while letting the impact of observations on the forecast propagate freely over time. A detailed comparison of the new approach with preceding ones is presented in a companion article by Griewank *et al.* (2022), along with an evaluation using a toy model.

### 2.1 | Ensemble sensitivity analysis

The foundation of ensemble sensitivity is to assume a linear relationship of a predicted quantity and the ensemble perturbations at the start of the forecast. This is in contrast to the adjoint sensitivity calculated via an adjoint model, which does not require an ensemble (Ancell and Hakim, 2007). For a given lead time, the predicted quantity  $j$  is represented by a  $1 \times n$  (row) vector  $\mathbf{j}$ :

$$\mathbf{j} = [j_1 \quad j_2 \quad \cdots \quad j_n], \quad (1)$$

where  $n$  is the number of ensemble members. The state vector  $\mathbf{x}$  consist of elements  $x_1, x_2, \dots, x_m$ . Ensembles of the state vector are given by a  $m \times n$  matrix  $\mathbf{X}$ :

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{bmatrix}. \quad (2)$$

It is generally assumed that the simulations have on average zero bias relative to true values, which is reasonable, since the main interest is in the second central moment of the predicted quantity. In this case, the forecast metric  $\delta\mathbf{j}$  represents deviations of the predicted quantity from its ensemble average. The forecast metric is a  $1 \times n$  vector. The variance of the forecast metric over ensemble members is calculated as follows:

$$\sigma^2 = \frac{1}{n-1} \delta\mathbf{j} \delta\mathbf{j}^T, \quad (3)$$

where the T superscript denotes transposition.

Similarly, deviations of state vector elements from corresponding ensemble means are elements of the  $m \times n$  matrix  $\delta\mathbf{X}$  (hereafter denoted as the state). The  $m \times m$  covariance matrix of the state (background covariance matrix hereafter) is denoted as

$$\mathbf{B} = \frac{1}{n-1} \delta\mathbf{X} \delta\mathbf{X}^T. \quad (4)$$

Using the first-order Taylor approximation,  $\delta\mathbf{j}$  can be approximated from  $\delta\mathbf{X}$ :

$$\delta\mathbf{j} \approx \mathbf{s}^T \delta\mathbf{X}, \quad (5)$$

via the sensitivity  $\mathbf{s}$ , which represents the  $1 \times m$  vector of partial derivatives in Equation (5):

$$\mathbf{s}^T = \left[ \frac{\partial j}{\partial x_1} \quad \frac{\partial j}{\partial x_2} \quad \frac{\partial j}{\partial x_3} \quad \cdots \quad \frac{\partial j}{\partial x_m} \right]. \quad (6)$$

The sensitivity can be expressed using  $\mathbf{B}$  by multiplying both sides of Equation (5) with  $(n-1)^{-1} \delta\mathbf{X}^T \mathbf{B}^{-1}$ :

$$\mathbf{s} \approx \frac{\delta\mathbf{j} \delta\mathbf{X}^T \mathbf{B}^{-1}}{n-1} = \text{cov}(\delta\mathbf{j}, \delta\mathbf{X}^T) \mathbf{B}^{-1}, \quad (7)$$

where cov is the cross-covariance matrix, elements of which are covariances between the predicted quantity  $j$  and elements of the state vector  $\mathbf{x}$ .

Taking into account the linear approximation of  $\delta\mathbf{j}$  in Equation (5), the variance of the forecast metric  $\sigma^2$  can be

approximated from the background covariance matrix  $\mathbf{B}$  using the sensitivity  $\mathbf{s}$ :

$$\begin{aligned}\sigma^2 &= \text{var}(\delta\mathbf{j}) = \frac{1}{n-1} \delta\mathbf{j}\delta\mathbf{j}^T \approx \text{var}(\mathbf{s}\delta\mathbf{X}) \\ &= \frac{1}{n-1} \mathbf{s}\delta\mathbf{X}\delta\mathbf{X}^T\mathbf{s}^T = \mathbf{s}\mathbf{B}\mathbf{s}^T,\end{aligned}\quad (8)$$

where  $\text{var}$  is the variance calculated over  $n$  elements.

## 2.2 | Estimate of variance change

When observations are incorporated, an update of state  $\delta\mathbf{X}$  is performed using a Kalman filter (Whitaker and Hamill, 2002; Hacker and Lei, 2015; Hakim *et al.*, 2020):

$$\delta\mathbf{X}_u = \delta\mathbf{X} - \mathbf{K}\mathbf{H}\delta\mathbf{X}, \quad (9)$$

where  $\mathbf{H}$  is the  $p \times m$  linear forward operator that maps the state space to the observation space,  $p$  is the number of observations included, and  $\mathbf{K}$  is the  $m \times p$  Kalman gain matrix.

Thus, the change in the variance of the forecast metric  $\Delta\sigma^2$  after assimilation of observations is as follows:

$$\Delta\sigma^2 \approx \text{var}(\mathbf{s}\delta\mathbf{X}_u) - \sigma^2 \approx \text{var}(\mathbf{s}(\delta\mathbf{X} - \mathbf{K}\mathbf{H}\delta\mathbf{X})) - \text{var}(\mathbf{s}\delta\mathbf{X}). \quad (10)$$

The relative variance change can then be found as a ratio of the variance change over the initial variance of the forecast metric:

$$\Delta\sigma_r^2 = \frac{\Delta\sigma^2}{\sigma^2}. \quad (11)$$

## 2.3 | Kalman gain and localization

The traditional Kalman gain matrix is defined as follows:

$$\mathbf{K}_t = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}, \quad (12)$$

where  $\mathbf{R}$  is the  $p \times p$  observation-error covariance matrix. The observation-error matrix  $\mathbf{R}$  used in this study is a diagonal matrix, due to the assumption of independent observations. In this study, we update the ensemble perturbations with the modified Kalman gain matrix used in the ensemble square-root filter designed by Whitaker and Hamill (2002):

$$\begin{aligned}\mathbf{K}_w &= \mathbf{B}\mathbf{H}^T \left[ \left( \sqrt{\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}} \right)^{-1} \right]^T \\ &\quad \left[ \sqrt{(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}) + \sqrt{\mathbf{R}}} \right]^{-1}.\end{aligned}\quad (13)$$

Whitaker and Hamill (2002) use the traditional Kalman gain matrix to update the ensemble mean and the modified

Kalman gain to update the ensemble perturbations. They showed that the modified Kalman gain matrix is more resistant to sampling errors related to a finite-size ensemble. The formulation of the Kalman gain matrix given in Equation (13) can process all available observations at once. Therefore, an iterative procedure, such as that used by Hakim *et al.* (2020), is not required.

Localization is a standard feature in ensemble assimilation systems needed to reduce ensemble sampling errors and increase computational efficiency. To estimate the benefit of an observation for a specific forecast system, the localization used by that forecast system needs to be taken into account. We use similar values and weighting functions to those used by the regional Deutscher Wetterdienst (DWD) assimilation system (Schraff *et al.*, 2016), as we aim to estimate the impact for this specific system.

In our ESA study, we apply two types of localization: (1) horizontal localization  $\mathbf{L}_h$ , that is, localization between different horizontal coordinates and (2) vertical localization  $\mathbf{L}_v$ , that is, localization between different altitude levels:

$$\mathbf{L} = \mathbf{L}_h \circ \mathbf{L}_v, \quad (14)$$

where  $\mathbf{L}$  is the  $m \times m$  localization matrix and  $\circ$  is a Schur product.

Both localization types are calculated using a Gaspari–Cohn function (Gaspari and Cohn, 1999). This function is Gaussian-shaped and the width of the function is defined by a free parameter, which is the localization scale. The application of the Gaspari–Cohn function reduces correlations between elements of the state vector to nearly zero after a certain distance or cut-off radius. The cut-off radius is derived from the localization length scale by multiplying the latter by  $2\sqrt{10/3} \approx 3.65$ . Elements of  $\mathbf{L}_h$  are calculated using radial distances between coordinates of the corresponding elements of the state vector. Elements of  $\mathbf{L}_v$  are calculated using differences in the natural logarithm of pressure between the altitude levels of the corresponding elements of the state vector.

The localization matrix  $\mathbf{L}$  is applied as a Schur product to the background covariance matrices  $\mathbf{B}$  in the modified Kalman gain matrix (Equation 13):

$$\begin{aligned}\mathbf{K} &= (\mathbf{L} \circ \mathbf{B}) \mathbf{H}^T \left[ \left( \sqrt{\mathbf{H}(\mathbf{L} \circ \mathbf{B})\mathbf{H}^T + \mathbf{R}} \right)^{-1} \right]^T \\ &\quad \left[ \sqrt{(\mathbf{H}(\mathbf{L} \circ \mathbf{B})\mathbf{H}^T + \mathbf{R}) + \sqrt{\mathbf{R}}} \right]^{-1}.\end{aligned}\quad (15)$$

Elements of the localization matrix range from 0 to 1 and depend on the radial distance between the corresponding locations of the state vector points. The precise



localization values used in this study are described in Section 3.2.5.

## 2.4 | Regularization of sensitivity

Ancell and Hakim (2007) discussed two issues which arise when attempting to calculate the sensitivity from the approximation in Equation (7). First, the background covariance matrix  $\mathbf{B}$  cannot be inverted if the ensemble size is smaller than the state size (Gombos and Hansen, 2008). Second, the computing time to inverse  $\mathbf{B}$  increases with the power of 3 in regards to state size  $m$ , making it prohibitively expensive to invert the background covariance  $\mathbf{B}$  matrix of the full model state.

A number of studies avoid this issue by neglecting off-diagonal terms of  $\mathbf{B}$  (Torn and Hakim, 2008). Such a diagonal approximation of the background covariance matrix (1) guarantees that the matrix is always invertible and (2) makes the inversion easy to calculate. Hacker and Lei (2015) discuss an implication of the diagonal approximation of  $\mathbf{B}$ . The authors show that neglected off-diagonal elements of  $\mathbf{B}$  in general lead to an overestimation of sensitivity to individual components of the state vector, which in turn leads to an inadequate quantitative characterization of predicted quantities. The authors report that the full background covariance matrix should be used for the sensitivity calculation, since this gives more accurate estimates of the predicted quantities, especially in the presence of fast dynamics, model errors, and large analysis ensemble spreads.

In this study, we solve the problem with singular matrices  $\mathbf{B}$  in Equation (6) by applying a regularization technique. We apply Tikhonov regularization (Tikhonov, 1965) to obtain the sensitivity:

$$\mathbf{s} \approx \text{cov}(\delta\mathbf{j}, \delta\mathbf{X}^T) \mathbf{B}^T (\mathbf{B}\mathbf{B}^T + \alpha^2\mathbf{I})^{-1}, \quad (16)$$

where  $\mathbf{I}$  is the  $m \times m$  identity matrix and  $\alpha$  is the regularization coefficient ( $\alpha \geq 0$ ). Hoerl and Kennard (1970) showed that Equation (16) is a least-squares estimator immune to the presence of highly correlated components in the state vector. Throughout this study, the sensitivity approximation given in Equation (16) is used, but a different form of regularization could have been applied, such as the singular value decomposition used by Ancell and Hakim (2007) and Hacker and Lei (2015). Section 4.1 explores the effects of the parameter  $\alpha$  and which values are suitable for this study.

A computationally very efficient way of avoiding  $\mathbf{B}^{-1}$  is to not compute the sensitivity at all. Instead, the change in variance as estimated through Equation (10) is reformulated so that the inverse  $\mathbf{B}$  of the sensitivity is balanced

by the  $\mathbf{B}$  from the Kalman gain shown in Equation (12) (Torn, 2014; Hakim *et al.*, 2020). However, this balancing cannot be achieved if the analysis influence of observations is localized, as the localized covariances of the Kalman gain ( $\mathbf{L} \circ \mathbf{B}$ ) cannot be balanced against the nonlocalized  $\mathbf{B}$  of the sensitivity. Instead, the variance change estimate can take analysis localization into account by splitting the forecast metric into local components and then applying distance-based localization weights to the covariances of the individual forecast metrics and the initial state (Hakim *et al.*, 2020; Tardif *et al.*, 2021). This approach is equivalent to having a nonlocalized sensitivity and a localized analysis if the observations only influence the forecast within the analysis localization, that is, the impact of the observations does not propagate over time. This propagation of the observation impact is also referred to as temporal localization evolution (Gasperoni and Wang, 2015).

In our companion study (Griewank *et al.*, 2022), we discuss in depth the errors that result from applying the approach of Tardif *et al.* (2021) when signal propagation is not accounted for appropriately, and estimating signal propagation is extremely difficult in a nonlinear three-dimensional atmospheric model. Different variables and perturbations at different scales propagate at differing speeds in the atmosphere, both horizontally and vertically. Tardif *et al.* (2021) can safely ignore signal propagation, as their metric is not calculated from the forecast but from the analysis. By applying Tikhonov regularization to calculate the sensitivity explicitly, we can take analysis localization into account without having to make any assumptions on signal propagation.

## 3 | DATA AND EXPERIMENTAL DESIGN

### 3.1 | Ensemble data and period

ESA requires an ensemble forecast that provides the model state and forecast metric. Our study utilizes an existing convective-scale 1,000-member ensemble simulation that has been computed on the K-computer at the RIKEN Center for Computational Science in Kobe, Japan (Miyoshi *et al.*, 2016a; 2016b). The 1,000-member ensemble simulation and its experimental setup are described in detail by Necker *et al.* (2020a). Below, we introduce the model data briefly, focusing on relevant aspects for the present study.

The 1,000-member ensemble simulation was done with the full-physics nonhydrostatic SCALE-RM and the SCALE Localized Ensemble Transform Kalman Filter (SCALE-LETKF) DA system (Hunt *et al.*, 2007; Lien *et al.*, 2017). Our study uses the high-resolution, convective-scale forecast output that is available from May 29–June 2 and

June 5–June 7, 2016. In this period, ensemble forecasts with 14-hr lead time have been initialized twice per day (at 0000 and 1200 UTC). The forecast domain ( $350 \times 250$  grid points) is centered over Germany. The model output has a 3-km horizontal resolution and 30 altitude levels ranging from the surface to the model top at 16.9 km.

The high-impact weather period analyzed is characterized by an atmospheric block over the Atlantic ocean (Piper *et al.*, 2016). A high-pressure system covered an area from Western France over Great Britain to Iceland. In the eastern part of France and over Germany, in contrast, there was a low-pressure system. Pressure gradients over Germany were weak, resulting in stationary conditions with relatively low wind speeds and heavy precipitation (Necker *et al.*, 2020a).

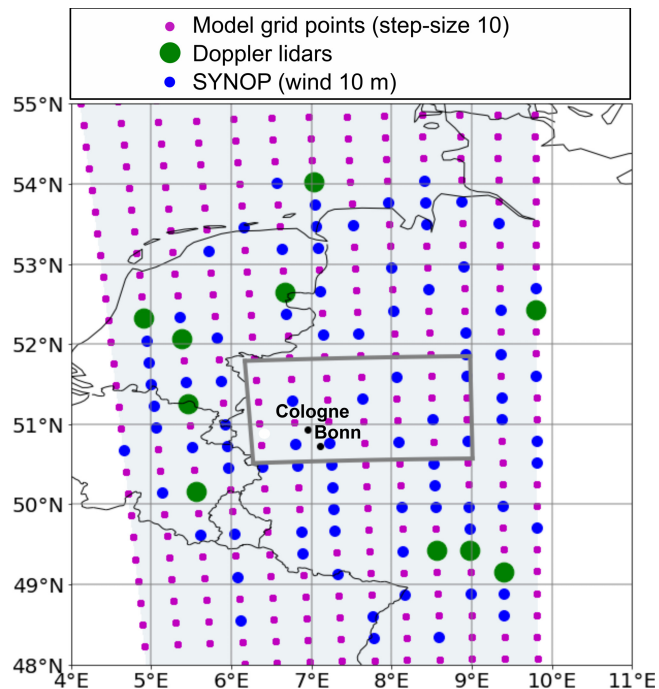
## 3.2 | Observational setup

### 3.2.1 | Surface synoptic stations

Surface synoptic observations (SYNOP hereafter) are typically taken by an automatic weather station and include measurements of temperature, relative humidity, pressure, and wind at the surface. Since our study focuses on wind, only 10-m wind components ( $u$  and  $v$ ) were used as SYNOP. Coordinates of SYNOP stations correspond to the locations of surface stations assimilated operationally by the German Weather Service (personal communication with Elisabeth Bauernschubert, DWD). We include only stations located within a radius of  $3^\circ$  from the center of the RRA domain. Since the reduced model resolution is 30 km, we placed surface stations at the closest model grid point. In total, there are 95 stations located within the circle area (Figure 1). SYNOP stations were included in all experiments in this study.

### 3.2.2 | Simulated Doppler lidar observations

A lidar is an active optical instrument that transmits a short laser pulse into the atmosphere and collects the light scattered back by atmospheric particles. In addition to backscatter signal intensity, Doppler lidars can measure a Doppler frequency shift between the transmitted and backscattered signals. The Doppler frequency shift is determined by radial motions of atmospheric scatterers (aerosols and clouds) due to their transport by horizontal wind and up/down drafts. Therefore, by using azimuthal scans at a fixed nonzenith elevation (Browning and Wexler, 1968) and assuming horizontal homogeneity of the wind field inside the cone defined by the scans, profiles of  $u$  and  $v$  components (or speed and direction)



**FIGURE 1** Locations of 95 SYNOP stations used in this study. The 95 SYNOP stations (blue dots) are located within a radius of  $3^\circ$  and from the center of the RRA shown by the gray rectangle. Magenta dots denote the model grid used for the state vector. Positions of the 10 Doppler lidars used for the experiment for sensitivity studies are shown as green dots. Note that coordinates of all shown SYNOP stations and Doppler lidars are mapped on the closest model grid points [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

of the horizontal wind can be retrieved from the radial velocity observed by Doppler lidars (Pichault *et al.*, 2021). The quality of the derived wind profiles depends strongly on the signal-to-noise ratio (SNR), which is determined by the backscattering efficiency of a volume of particles and the path integrated attenuation as well as the distance of the lidar to the backscattering volume. The former depends mainly on the size, shape, and number concentration of particles. Therefore, the maximum wind retrieval height is often limited to ABL height (up to a few km altitude: Schween *et al.*, 2014; Coniglio *et al.*, 2019; Pichault *et al.*, 2021). Aerosol concentrations in the free troposphere above are typically too low for a sufficiently high SNR. Also, the presence of liquid water in form of a cloud or a fog layer will lead to signal attenuation within a few 100 m of liquid base. Theuer *et al.* (2020) report that applicability of Doppler lidars for wind-farm applications is strongly limited in rain and fog conditions. Taking this into account, only the lowest five model output levels at 80, 429, 1,062, 1,853, and 2,845 m are considered for calculating simulated Doppler lidar observations. Note that in our study we refer to “simulated” observations, which are based on SCALE-RM model data. Applying different

maximum height ranges, we can thus investigate the impacts of these Doppler lidar limitations on DA.

### 3.2.3 | Assigned observation errors

The assimilation of observations relies on an appropriate estimate of the associated errors. Assigned observation errors in the context of data assimilation need to include the instrumental error as well as the error related to model representativeness (Hodyss and Nichols, 2015; Janjić *et al.*, 2018). In the case of the wind measurements with Doppler lidars, the instrumental error is associated with SNR, the type of scanning mode, and uncertainties in azimuth, elevation, and ranging of the instrument (Theuer *et al.*, 2020). In addition, the retrieval quality is also part of the instrumental error. For instance, Päsche *et al.* (2015) show that homogeneity of the airflow within the scanning volume and numerical stability of the retrievals with respect to small errors affect the accuracy of wind measurements. Päsche *et al.* (2015) evaluated wind profiles from a Doppler lidar with respect to ones from radiosondes and a radar wind profiler based on a one-year dataset. The authors showed that the root-mean-square (RMS) difference of Doppler lidar wind profiles does not change much up to 3 km height and does not exceed  $1 \text{ m}\cdot\text{s}^{-1}$  in speed and  $10^\circ$  in direction relative to the other two instrument types. In this study, we thus take a conservative estimate of the instrumental error of  $1 \text{ m}\cdot\text{s}^{-1}$  for both  $u$  and  $v$  components of wind measured with a Doppler lidar.

Representativeness error is associated with unresolved processes related to the differences in spatial and temporal scale between observations and model (Janjić *et al.*, 2018). For instance, due to complex surface terrain and topography, wind measurements can be influenced by small-scale circulations that may not be resolved by NWP models (WMO, 2018). In general, the representativeness error also includes errors of the forward model used (operator  $\mathbf{H}$  in Equation (9)): Janjić *et al.*, 2018). In this study, however, forward model errors are neglected because for grid points with Doppler lidars the state space is the same as the observation space.

We assume an observation error for Doppler lidars of  $2 \text{ m}\cdot\text{s}^{-1}$  for both  $u$  and  $v$  components of wind at all five altitude levels. Thus, in the case of uncorrelated instrumental and representativeness error, the latter is  $\approx 1.7 \text{ m}\cdot\text{s}^{-1}$ . Similar values of observation errors for wind observations in the lower troposphere (between 850 and 1,000 hPa) were estimated by Schraff *et al.* (2016) based on the Desroziers statistics (Desroziers *et al.*, 2005a; 2005b). The chosen observation error is also consistent with the one used in Coniglio *et al.* (2019) for assimilation of a Doppler lidar for short-term forecasts. The  $2 \text{ m}\cdot\text{s}^{-1}$  observation

error used meets the World Meteorological Organization's (WMO's) Observing Systems Capability Analysis and Review (OSCAR) requirement for an optimal (from a cost-benefit point of view) improvement of wind in global and high-resolution NWP application areas (Leuenberger *et al.*, 2020).

The wind observation errors of SYNOP surface stations are assumed to be  $3.5 \text{ m}\cdot\text{s}^{-1}$  following the values used in the operational data assimilation system of DWD. The higher value—with respect to the observation error of Doppler lidars—mainly originates from a lower representativeness of point-like surface observations. The DWD assimilation system for the limited-area mode currently assumes that observation errors are uncorrelated and we follow this approach in the present study. Accordingly, the observation-error covariance matrix  $\mathbf{R}$  is a diagonal matrix with a size of  $p \times p$ . Diagonal elements of  $\mathbf{R}$ —variances of observation errors—are the squares of the assigned observation errors. If observation errors are substantially correlated, we would overestimate the benefit of the observations. However, we do not expect SYNOP or Doppler lidar observations to have highly correlated errors.

### 3.2.4 | Observation space

The forward observation operator maps from model state space to observation space. For observations of model state variables, the observation operator is a linear function that can be written as a matrix  $\mathbf{H}$ . In our case,  $\mathbf{H}$  is a matrix with a size of  $p \times m$ . The number of observations included varies for different experiments. Note that, while we refer to our ESA estimates of the variance reduction for different Doppler lidar network designs as “experiments”, our experiments are only based on the existing ensemble simulations. We do not perform any data assimilation steps or additional model simulations. First, observations include  $u$  and  $v$  components of the 10-m wind, which represent a network of 95 operationally assimilated SYNOP stations (see Figure 1). For all experiments, the number of SYNOP observations is 190, which is two variables measured by 95 stations. Second, the number of Doppler lidar observations depends on the number of sites and the number of available ranges. The number of sites covers the range from 1 to 95 Doppler lidars. Note that the number of sites is fixed for an experiment, but may vary depending on the type of experiment setup described in Section 3.3. The number of available height ranges varies from 1 to 5. The number of ranges is the same for all Doppler lidars in the network within an experiment. Thus, the total number of Doppler lidar observations varies from 2 (one instrument, one range, and two measured quantities:  $u$  and  $v$  wind



components) to 950 (95 instruments, five ranges, and two measured quantities). Each observation is represented by a row in the  $\mathbf{H}$  matrix. For an observation, a single element in the corresponding row of  $\mathbf{H}$  is set to 1. This element corresponds to the type of observation, coordinates, and altitude of the observation. All remaining elements in the row are set to 0.

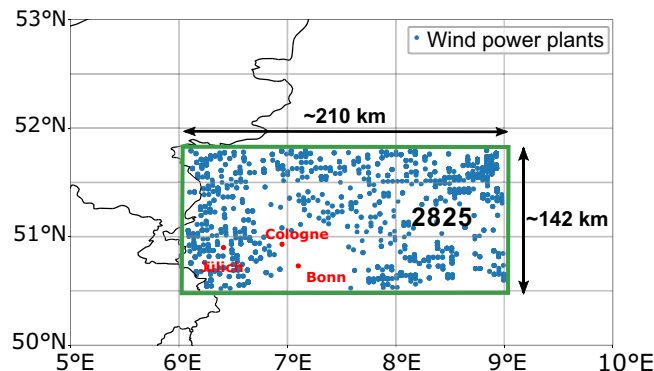
### 3.2.5 | Localization length scales

Our localization length scales are chosen from the range of values applied in the regional modelling system of DWD, as we are aiming to estimate the potential impact of observations for this particular modelling system. We decided to choose a static horizontal localization length scale of 50 km, which is the lowest value of the adaptive localization range (50–120 km) operationally applied by DWD (Schraff *et al.*, 2016). In the vertical, DWD applies a height-dependent localization scale that increases from 0.075 at the surface to  $0.5 \ln(h/\text{hPa})$  in the upper troposphere. We chose a static and narrow vertical localization length scale of  $0.1 \ln(h/\text{hPa})$ , since we focus on low-level observations up to 3 km.

## 3.3 | Experimental setups

Within this study, we designed a number of experiments to evaluate potential improvements in the short-term forecast of low-level wind using a network of Doppler lidars. Our experiments are based on three different setups depending on the number of Doppler lidars included in the network, as follows.

1. Setup 1 includes 10 randomly chosen Doppler lidars in the network and their random locations were chosen from the SYNOP station locations. This configuration with a fixed number and location of instruments in the network is used in Section 4 for sensitivity studies based on a single forecast.
2. In setup 2, we increased the number of Doppler lidars in the network from 5 to 95 with a step of 5 and chose their locations randomly from SYNOP station locations. For each combination of the number of Doppler lidars, the procedure is repeated 50 times. Setup 2 is applied in the experiments shown in Section 5.1.
3. Setup 3 is used in Section 5.2–5.4, when the network includes 25 Doppler lidars chosen randomly from the locations of SYNOP stations and a random choice is repeated 50 times for each combination of Doppler lidars.

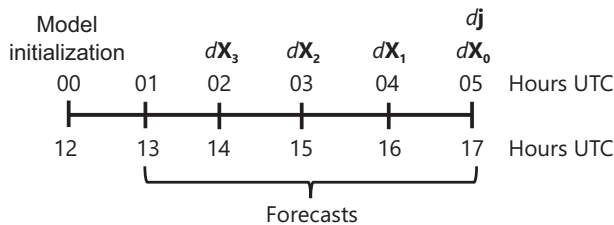


**FIGURE 2** Location of wind-power plants within the RRA. Data from 2019. The green rectangle depicts the borders of the RRA used in this study to derive the forecast metric. Blue dots indicate locations of individual windmills. The data were provided by Lukas Schmidt, who retrieved the data using the code available at <https://github.com/OpenEnergyPlatform/open-MaStR> [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

### 3.3.1 | Forecast metric

This study focuses on the Rhein–Ruhr Area (RRA), one of the biggest urban areas in western part of Germany. In 2019 there were 2,825 wind turbines registered in this area (Figure 2) with a total installed capacity of  $\sim 5.1$  GW. The RRA domain has a size of  $\sim 210 \times 142$  km<sup>2</sup> that corresponds to an area from  $50.5$ – $51.8^\circ$ N and  $6$ – $9^\circ$ E. We used averages of  $u$  and  $v$  components of the 80-m wind from SCALE-RM over the RRA domain as out forecast metric ( $\mathbf{j}_u$  and  $\mathbf{j}_v$ ). We chose the wind at 80-m height as that is a typical hub height of wind turbines.

In order to avoid spin-up effects in the SCALE-RM simulations (Necker *et al.*, 2020a), 00 and 01 hr forecasts are not used in this study (0000 and 0100 UTC for model initialization at 0000 UTC; 1200 and 1300 UTC for model initialization at 1200 UTC). Instead, for model runs at 0000 and 1200 UTC the predicted quantities are calculated for 0500 and 1700 UTC forecasts, respectively. In order to analyze different lead times, we use states at 0200, 0300, 0400, and 0500 UTC for the model initialization at 0000 UTC, as shown in Figure 3. These lead times correspond to lead times of 3, 2, 1, and 0 hr relative to the predicted quantities, respectively. For the initialization at 1200 UTC, states at 1400, 1500, 1600, and 1700 UTC are taken, respectively. Thus, the time of forecast metrics is fixed at 5 hr lead time, while different lead times are simulated by changing the time of the initial state. This required to compare the variance reduction in predicted quantities at different lead times properly. Shifting the forecast metric would result in a changing variance of the predicted quantity, which would make comparing different lead times difficult.



**FIGURE 3** A schematic diagram of times for the forecast metric  $\delta j$  and the state vector  $\delta X$ . The model initialization was performed twice a day at 0000 and 1200 UTC. The first two hours after model initialization are not used due to spin-up effects discussed in the text. The time of the forecast metric  $\delta j$  is fixed at 0500 and 1700 UTC for initializations at 0000 and 1200 UTC, respectively. For initialization at 0000 UTC, the time of state vector  $\delta X$  is 0200, 0300, 0400, and 0500 UTC for lead times of 3, 2, 1, and 0 hr, respectively. For initialization at 1200 UTC, the time of state vector  $\delta X$  is 1400, 1500, 1600, and 1700 UTC for lead times of 3, 2, 1, and 0 hr, respectively. Lead times of 0, 1, 2, and 3 hr are indicated by  $\delta X_0$ ,  $\delta X_1$ ,  $\delta X_2$ , and  $\delta X_3$ , respectively

### 3.3.2 | Model state space

Doppler lidars measure  $u$  and  $v$  components of wind. Since  $u$  and  $v$  wind components are usually prognostic variables of NWP models, we use these components in our state vector ( $X$  in Section 2.1). For the state we use a larger area than the RRA (model domain hereafter), which roughly ranges from 48–55°N and from 4–10°E (see Figure 1). Note that, due to computational restrictions, the present study uses a reduced horizontal and vertical grid of the available original model data. We only use every 10th grid in both the north–south and west–east directions. Such a subsampling results in a 30-km grid spacing (see Figure 1). Nevertheless, we compute the forecast response function using the original model output with 3-km resolution. In the vertical we use six of the original 30 levels, at 10, 80, 429, 1,062, 1,853, and 2,845 m height.

Each state vector (a column in  $X$ ) contains values of  $u$  and  $v$  components of wind at all grid points of the state domain at six different height levels. After the subsampling of the original model output (see Section 3.1), each level has 348 grid points, a single state vector has 4,176 elements, and the complete ensemble state  $X$  is a 4,176 × 1,000 matrix, where 1,000 is the number of ensemble members.

## 4 | SENSITIVITY STUDIES

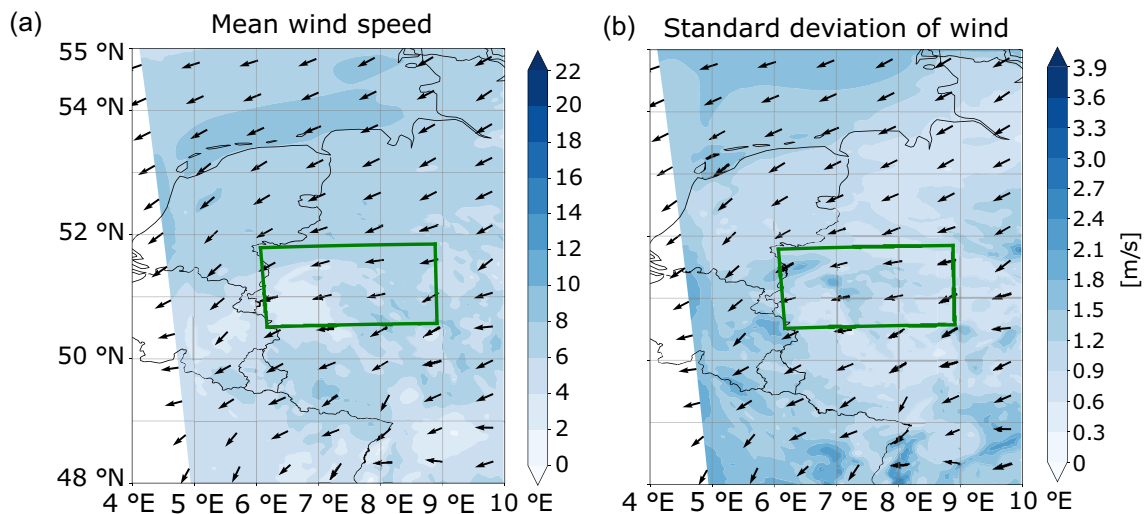
In order to demonstrate how the ESA algorithm performs, we selected one case study on May 29, 2016 that exhibited a wind pattern common for the period analyzed. In this case, the center of a low-pressure system was over the

northeastern part of France, producing easterly flow at 80-m height over the RRA Figure 4. The ensemble-averaged wind speed at 80 m for 3-hr lead time ranges from 4–10  $\text{m}\cdot\text{s}^{-1}$  with a standard deviation not exceeding 3  $\text{m}\cdot\text{s}^{-1}$ . At latitudes above 50°N, the wind speed and direction are nearly homogeneous, while at lower latitudes wind speed and direction have higher variability. Based on this single forecast case, we performed a set of sensitivity studies to test and evaluate how results change with different values of regularization coefficient, localization length scale, and observational error.

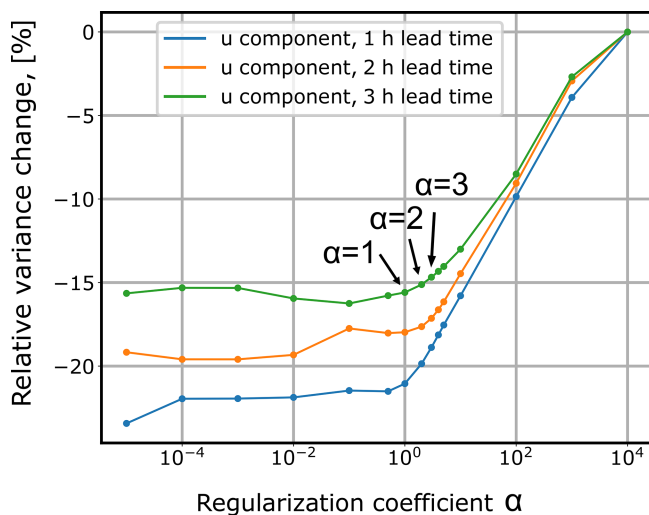
### 4.1 | Sensitivity of $\Delta\sigma_r^2$ to the choice of regularization coefficient

As shown in Section 2.4, we apply Tikhonov regularization to estimate the sensitivity (Equation 16). In order to find a suitable regularization coefficient  $\alpha$ , which controls the degree of regularization, we make a set of experiments with 95 SYNOP stations (10-m wind) and 10 randomly chosen Doppler lidar sites (setup 1 in Section 3.3), with only one—80 m—observed level. For these experiments, we use the model data for the 0000 UTC initialization on May 29, 2016. Figure 1 depicts coordinates of the 95 SYNOP stations (blue dots) and 10 Doppler lidar sites (green dots). As mentioned earlier, the SYNOP stations are always fixed and the Doppler lidar sites were chosen randomly.

We repeat the experiment with different values of  $\alpha$  ranging from  $10^{-6}$  to  $10^4$ . Figure 5 summarizes the results of these experiments and shows the relative variance changes  $\Delta\sigma_r^2$  for different lead times as functions of  $\alpha$ . Since DA, on average, leads to a reduction of forecast variance, values of  $\Delta\sigma_r^2$  change are mostly negative. When  $\alpha$  is too small (e.g.,  $< 10^{-2}$ ), the sensitivity estimate starts to become unstable and shows an oscillatory behavior (shown in Griewank *et al.*, 2022). The extreme effect occurs at  $\alpha = 10^{-6}$  when the variance change is positive (not shown). Positive values of variance reduction are unrealistic, since they imply that the assimilation increases the variance of the forecast metric. High values of  $\alpha$  (e.g.,  $> 10$ ) reduce the sensitivity too much, and the impact of observations is completely suppressed in the variance change estimates. For  $\alpha$  ranging from  $10^{-3}$  to 3, the effect of the assimilation shows minimal values for 2- and 3-hr lead times. For the 1-hr lead time, the minimum appears at a lower value of  $\alpha$ . At  $\alpha$  values of  $10^{-3}$  to 3, however, the estimated variance change at the 1-hr lead time is not much higher (a few per cent). For the following analysis we chose  $\alpha = 3$ . This value gives near-maximal estimates of  $\Delta\sigma_r^2$  for all lead times used and is high enough to suppress any unstable behavior. The chosen value  $\alpha = 3$



**FIGURE 4** (a) Mean and (b) standard deviation of wind at 80-m height on May 29, 2016 at 0500 UTC obtained from the model. Mean values and standard deviations are calculated over 1,000 ensemble members. Black arrows indicate wind direction at 80-m height. The RRA is shown by the green rectangle [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]



**FIGURE 5** Dependence of  $\Delta\sigma_r^2$  on  $\alpha$  for different lead times. Calculations are made for the 0000 UTC initialization on May 29, 2016. Simulated data from 10 Doppler lidars and 95 SYNOP stations are assimilated. The blue, orange and green lines correspond to 1-, 2-, and 3-hr lead times, respectively. The forecast metric is the  $u$ -component of the 80-m wind. Lidar observations are available at only one level (80-m height) [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

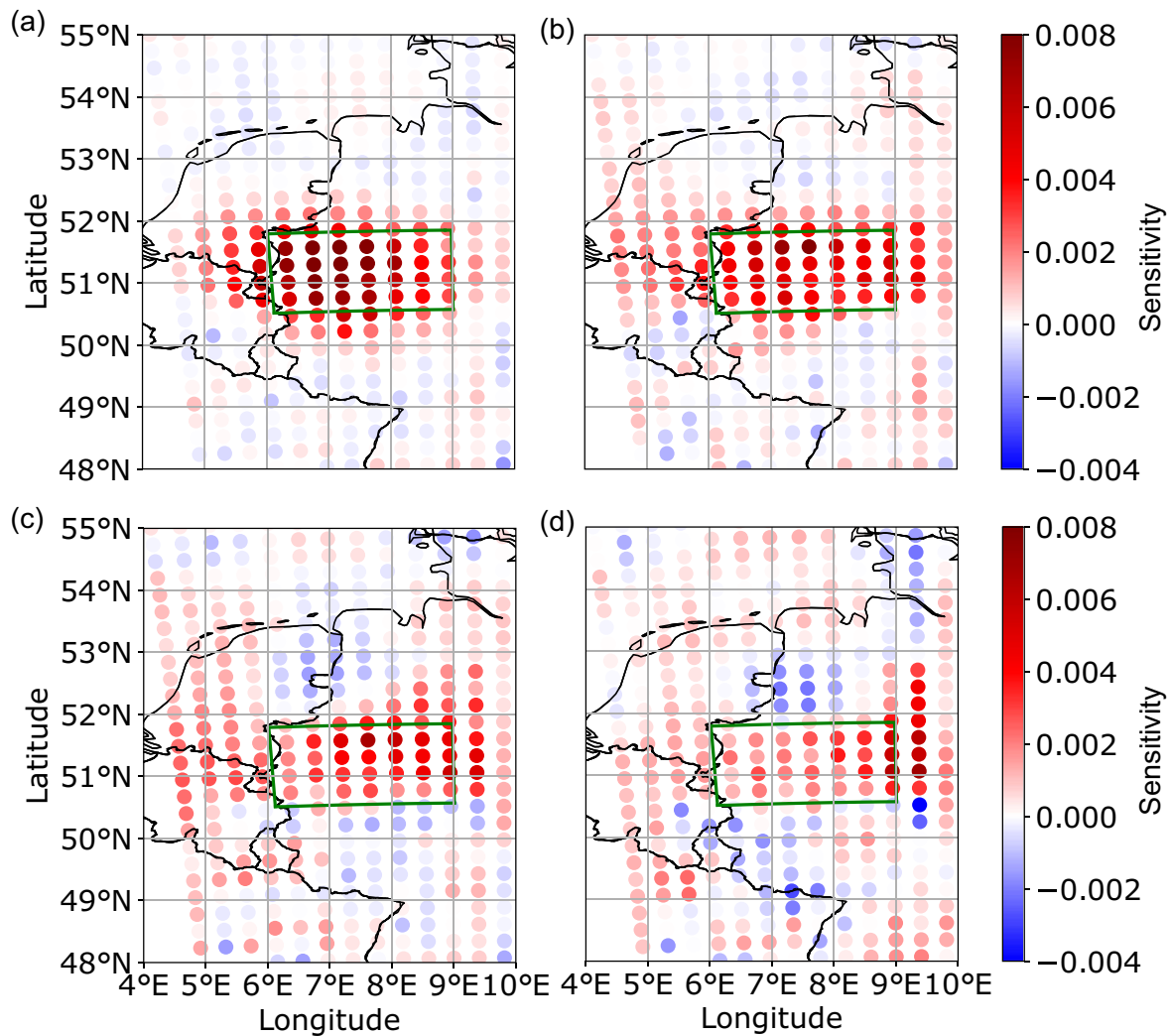
results in an uncertainty of about 2–3% in estimated  $\Delta\sigma_r^2$  relative to results obtained with  $\alpha$  values in the range from  $10^{-3}$  to 3.

Subsequently, we discuss the sensitivities regularized with  $\alpha = 3$  to determine if they match our theoretical expectations. Figure 6 shows the sensitivity of  $\delta j_{u,r}$  to the observations of the  $u$ -component of the 80-m wind for different lead times. We know that for the 0 hr lead-time (Figure 6a) the sensitivity should have a constant value

within the RRA, marked by the green rectangle, and be zero outside the RRA, and the computed ensemble sensitivity matches the expected pattern well, with only small values outside the RRA. The sensitivity for the 1- and 2-hr lead times (Figure 6b,c) also has maximum values in the center of the RRA, but as expected the maximal values decrease with lead time. The sensitivity outside the RRA, however, has increased. In the case of 3-hr lead time, the maximum sensitivity values are located to the east of the RRA. These sensitivity patterns are consistent with the wind conditions of that day (Figure 4). In the latitude range from 50–52°N, the wind was blowing homogeneously from east to west on that day. Apparent negative sensitivity values have considerably lower magnitude relative to positive sensitivities within the RRA.

## 4.2 | Sensitivity of $\Delta\sigma_r^2$ to the choice of observational error

In order to evaluate how ESA results change with localization scale and observational error, we performed a similar set of experiments to that in Section 4.1. Here, instead of changing  $\alpha$ , we fixed it to  $\alpha = 3$ , and the observation error of Doppler lidars was changed from 1 to 20 m·s<sup>-1</sup>. Three series of experiments were made: (1) with 95 SYNOP stations assimilated, (2) with 95 SYNOP stations and 10 Doppler lidars (setup 1 in Section 3.3) observing at 80 m only, and (3) with 95 SYNOP station and 10 Doppler lidars (setup 1 in Section 3.3) observing at all five altitude levels (or height ranges). Recall that the observation error of SYNOP stations is always fixed at 3.5 m·s<sup>-1</sup>. Figure 7 shows the resulting relative variance changes in the



**FIGURE 6** A demonstration of sensitivity of a forecast metric to selected elements of the state vector. The forecast metric is the  $u$ -component of 80-m wind forecast averaged over the RRA for 0500 UTC on May 29, 2016. Elements of the state vector are  $u$ -components of 80-m wind at all grid points of the model domain. The sensitivity is shown for (a) 0, (b) 1, (c) 2, and (d) 3 hr lead times. The sensitivity is calculated with Tikhonov regularization ( $\alpha = 3$ ) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

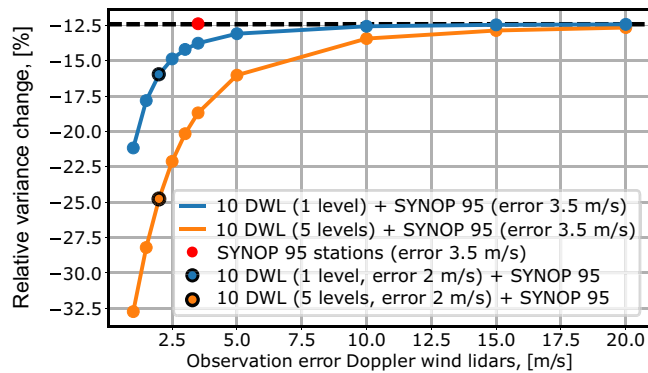
$u$ -component of the 80-m wind. As expected, the variance reduction becomes more negative at lower values of observation error. This is due to the fact that observations with a lower error have a higher weight compared with the model according to Equations (12) (13), and (15) used for Kalman gain calculations. It can be seen that, with the selected error assumptions, Doppler lidar observations reduce the variance considerably relative to SYNOP stations, even when only observations at 80 m are available. Information from more altitude levels increases the impact of Doppler lidar observations further. If the observation error was larger than  $3.5 \text{ m}\cdot\text{s}^{-1}$ , a network of 10 Doppler lidars would gain little (with respect to SYNOP) when only observations at 80 m are available. All five levels, however, would still give an improvement. It can also be seen that a further reduction of observation errors would be valuable for even stronger reduction of the forecast metric

variance. These results indicate that (1) our method estimates correctly that reducing observation errors increases their impact, (2) during cases with optically thick clouds, fog, and precipitation it is beneficial to use another wind profiling instrument less affected by presence of clouds (e.g., scanning cloud radars or radar wind profilers), and (3) a further improvement of short-term wind based on DA of Doppler lidars motivates the development of more powerful Doppler lidars, which should have lower observation errors.

### 4.3 | Sensitivity of $\Delta\sigma_r^2$ to the choice of localization length scale

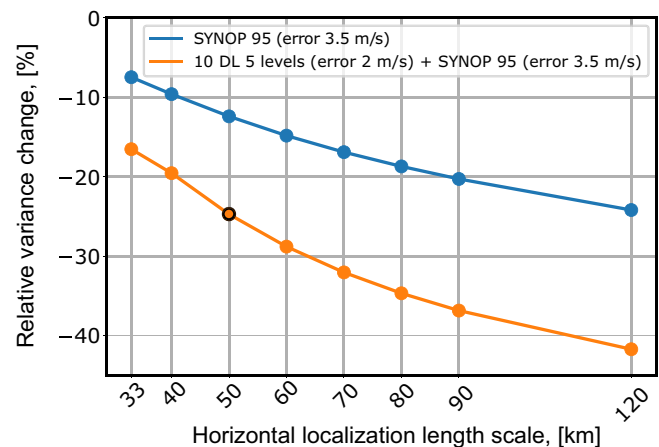
To understand the impact of a different horizontal localization scale better, we performed a set of experiments





**FIGURE 7** Dependence of  $\Delta\sigma_r^2$  on the observation error of Doppler lidars. Calculations are made for the 0000 UTC initialization on May 29, 2016. The forecast metric is the  $u$ -component of 80-m wind. A lead time of 3 hr is used. The blue and orange lines correspond to the assimilation of data from 10 Doppler lidars and 95 SYNOP observations. Lidar observations are available at 1 (blue line) and 5 (orange line) levels. The red dot indicates the observation error of SYNOP stations. The value of the red dot and the horizontal dashed line on the  $y$ -axis corresponds to  $\Delta\sigma_r^2$  when only the 95 SYNOP stations are assimilated. The black circles denote the  $2\text{ m}\cdot\text{s}^{-1}$  observation error of Doppler lidars assumed in this study. Regularization coefficient  $\alpha = 3$  [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

as in Section 4.1 with fixed  $\alpha = 3$  but for horizontal localization length scales varying from 33 to 120 km. Such an increase of the localization scale could, for example, be feasible in assimilation systems with a larger number of ensemble members that lead to fewer spurious correlations and a larger number of degrees of freedom. We repeated the experiments for (1) the 95 SYNOP stations only and also for (2) 95 SYNOP stations and 10 Doppler lidars (setup 1 in Section 3.3) observing all five altitude levels. The results are depicted in Figure 8. It can be seen that the estimate of  $\Delta\sigma_r^2$  depends strongly on the choice of horizontal localization length scale. In general, a larger horizontal localization scale increases the influence radius of observations, resulting in a more negative  $\Delta\sigma_r^2$ . These results underline that it is crucial to choose realistic localization scales that match the NWP system of interest. In the case of DA including both the 95 SYNOP stations and 10 Doppler lidars (orange solid line), the selected 50-km localization scale reveals differences of  $-40\%$  and  $+80\%$  relative to 33 and 120 km, respectively. If  $\Delta\sigma_r^2$  is evaluated against SYNOP stations only (i.e., ratio of the orange line over the blue one), the dependence on the horizontal localization scale is considerably less pronounced. In this case, the difference of the 50-km localization scale relative to 33 and 120 km is within  $\pm 10\%$ . In the following analysis, we evaluate  $\Delta\sigma_r^2$  when both SYNOP stations and Doppler lidars are assimilated with respect to the relative variance reduction when only SYNOP stations are assimilated.



**FIGURE 8** Dependence of  $\Delta\sigma_r^2$  on the horizontal localization length scale. Calculations are made for the 0000 UTC initialization on May 29, 2016. The forecast metric is the  $u$ -component of 80-m wind. A lead time of 3 hr is used. The blue line corresponds to the assimilation of 95 SYNOP stations. The orange line corresponds to the assimilation of data from 10 Doppler lidars and 95 SYNOP. Lidar observations are available at all five levels. The observation error is set to 3.5 and  $2\text{ m}\cdot\text{s}^{-1}$  for 10-m wind from SYNOP stations and wind profiles from Doppler lidars, respectively. The black circle denotes the choice of localization length scale applied in this study for our experiments. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

## 5 | NETWORK DESIGN STUDIES

Now that the sensitivity experiments have confirmed that our estimated variance changes fulfil our theoretical expectations, we now estimate the benefit of a potential Doppler wind lidar network for a short-term low-level wind forecast relative to conventionally assimilated SYNOP observations. For all available cases, we calculated  $\Delta\sigma_r^2$  for  $u$  and  $v$  components of the 80-m wind as predicted quantities. The calculation was performed separately for the  $u$  and  $v$  components. Recall, however, that both components are assimilated simultaneously. A number of experiments was performed. First, we derived values of  $\Delta\sigma_r^2$  for the assimilation of the 95 SYNOP stations only for 0, 1, 2, and 3 hr lead times. Then values of  $\Delta\sigma_r^2$  were obtained for the assimilation of 95 SYNOP stations and Doppler lidars.

For a given experiment, the number of Doppler lidar stations, the number of range altitude levels, and the lead time were fixed. For each experiment, locations of Doppler lidars were chosen randomly from the coordinates of the 95 SYNOP stations. An experiment for a combination of the number of Doppler lidar sites, available range altitude levels, and lead time was repeated 50 times to estimate how variable  $\Delta\sigma_r^2$  is depending on the exact lidar network configuration.

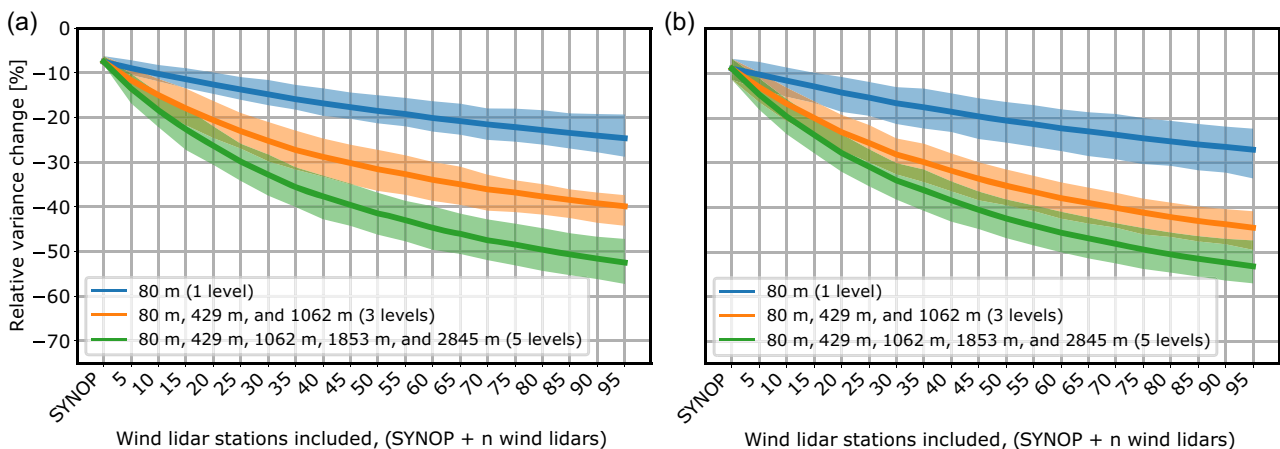
### 5.1 | Doppler lidar network density

In this section, we investigate the optimal size of the Doppler lidar network needed to achieve an improvement of low-level wind forecast in the RRA. We expect that the impact saturates for a certain size of Doppler lidar network, that is, with an increasing number of lidars the contribution of each following lidar to  $\Delta\sigma_r^2$  should become smaller than the contribution of previously added lidars. In order to check this, we run experiments for all 16 available cases. For these experiments, we increased the number of Doppler lidars in the network from 5 to 95 with a step of 5 (setup 2 in Section 3.3). For a given number of Doppler lidars in the network, the statistics of  $\Delta\sigma_r^2$  were calculated over 800 values of  $\Delta\sigma_r^2$  corresponding to all 16 available cases and 50 random lidar locations for each available case. The results of the experiment for the 3-hr lead time shown in Figure 9 depict an indication of the saturation effect. Note, however, that the divergence from a linear dependence between  $\Delta\sigma_r^2$  and the number of Doppler lidars starts at 20–30 instruments in the network when at least three levels are observed. The saturation effect is much less pronounced when only one level is available in Doppler lidar observations. The saturation effect starts at nearly the same number of Doppler lidars in the network for  $u$  and  $v$  components of the 80-m wind. Based on these results, we conclude that the most cost-efficient improvement of low-level wind in the RRA could maybe be achieved by a Doppler lidar network with about 25 instruments. For the analyzed cases, an assimilation of SYNOP stations and 25 Doppler lidars gives on average three times better  $\Delta\sigma_r^2$  relative to SYNOP stations only. In the case of SYNOP stations and 95 Doppler lidars, this improvement is only

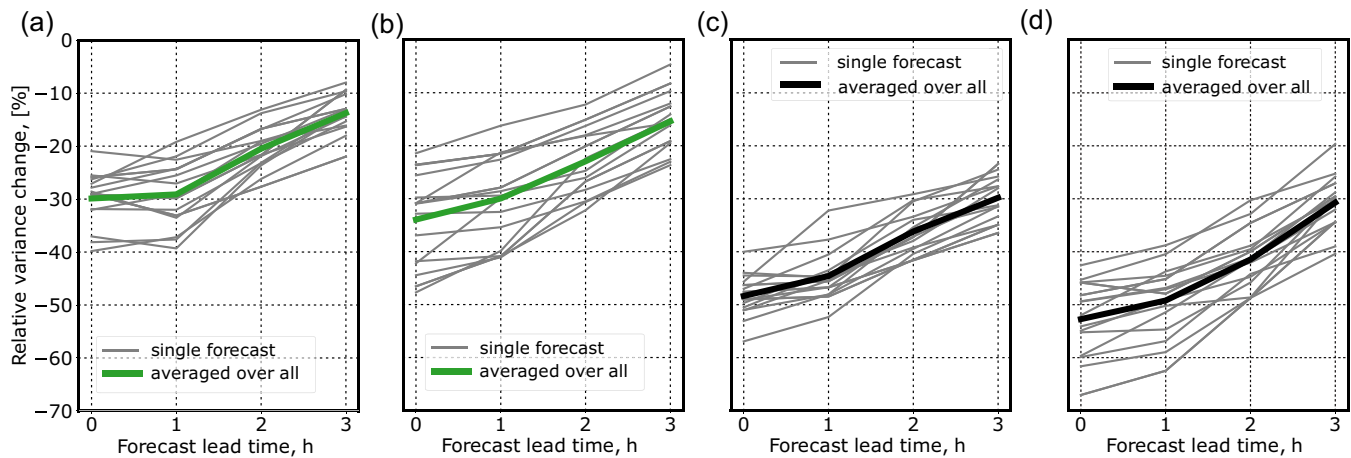
5.5 times. Note that this conclusion is valid only for a sufficiently deep ABL and when there is no cloud below 3-km height. An expected impact of a Doppler lidar network with 25 instruments is discussed in the next subsection. Such a network corresponds to one Doppler lidar per  $8 \times 10^3 \text{ km}^2$ . This is equivalent to installation of a Doppler lidar in every fourth SYNOP site in the area of interest.

### 5.2 | Situation-dependent forecast variability

In this subsection, we give a quantitative characterization of the improvement expected from a 25 Doppler lidar network (setup 3 in Section 3.3) in addition to the operationally assimilated SYNOP stations for the 16 different available forecasts. For each of 16 analyzed forecasts, derived values of  $\Delta\sigma_r^2$  were averaged over the 50 repetitions with randomly chosen Doppler lidar locations. These averaged values of  $\Delta\sigma_r^2$  are shown in Figure 10 as functions of the lead time. The results are shown for the case when only one level of lidar observations is available (Figure 10a,b) and when all five levels are available (Figure 10c,d). As expected, five observed levels give better results, that is, more negative  $\Delta\sigma_r^2$  in comparison with one level available. As was discussed in Section 3.1, the analyzed period is characterized by distinct weather patterns. In addition, the analyzed cases include both day and night-time forecasts. We see that the estimated benefit varies substantially depending on the atmospheric situation and whether the  $u$  or  $v$  component of the wind is chosen as the forecast metric. In the most beneficial case, the same number of lidars provides roughly twice the benefit as in



**FIGURE 9** Dependence of  $\Delta\sigma_r^2$  on the number of Doppler lidars in the network. Calculations are made for  $\Delta\sigma_r^2$  averaged over all 16 cases. The forecast metrics are the (a)  $u$ -component and (b)  $v$ -component of 80-m wind. The blue, orange, and green lines correspond to the assimilation of data from a different number of Doppler lidars and 95 SYNOP observations. The SYNOP value on the  $x$ -axis corresponds to only 95 SYNOP stations assimilated. Lidar observations are available at one (blue line), three (orange line), and five (green line) levels. The solid lines show the mean over 50 repetitions. The shaded areas depict 25th and 75th percentiles [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

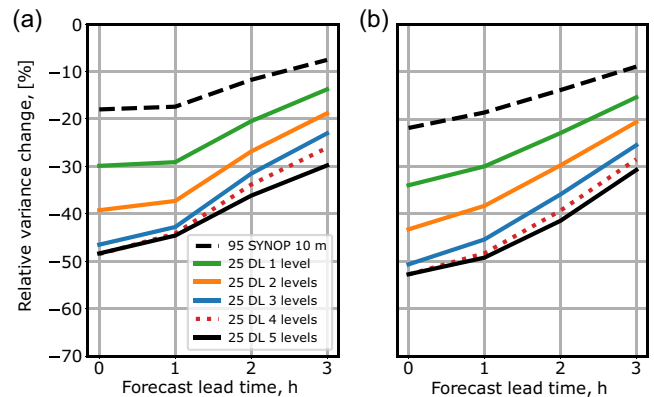


**FIGURE 10** Dependence of  $\Delta\sigma_r^2$  on the lead time. The forecast metrics are the (a,c)  $u$ -component and (b,d)  $v$ -component of 80-m wind. The assimilation is performed for simulated data from 25 Doppler lidars and 95 SYNOP stations. Lidar observations are available at (a,b) one and (c,d) five levels. Gray lines correspond to values of  $\Delta\sigma_r^2$  averaged over the 50 repetitions for a single case. The thick green and black lines show  $\Delta\sigma_r^2$  averaged over all 16 cases for 1 and 5 available levels, respectively [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

the weakest case. In all forecasts, the estimated benefit decreases as lead time increases from 1 to 3 hr, although the slope is again forecast-dependent. A deeper analysis on how the synoptic situation effects  $\Delta\sigma_r^2$  is outside the scope of this article.

### 5.3 | Vertical extent of the Doppler lidar measurements

This section shows results on the impact of a potential Doppler lidar network depending on the vertical extent of the Doppler lidar measurements. The vertical wind profiles from Doppler lidars can be influenced by the presence of hydrometeors—for example, fog, low-level liquid clouds, and precipitation—and therefore lead to different impacts on the low-level wind forecast. Figure 11 summarizes the impact of assimilation of SYNOP stations and 25 Doppler lidars with different vertical extents of Doppler lidar measurement on  $\Delta\sigma_r^2$  relative to assimilation of SYNOP stations only. The figure displays values of  $\Delta\sigma_r^2$  averaged over all 16 available cases. The results show that the assimilation of only 95 SYNOP stations (dashed black lines in Figure 11) yields on average 18–20% and 8–10%  $\Delta\sigma_r^2$  for 1- and 3-hr lead times, respectively. An additional network of 25 Doppler lidars improves these values, although the exact improvement depends on the number of observed levels. When only one level is observed (green solid lines in Figure 11) by all Doppler lidars, the improvement is a factor of 1.6 for 1-hr lead time. When three levels are available, that is, Doppler lidars profile the wind up to 1 km, the improvement is a factor of 2.3 for the 1-hr lead time. When wind profiles from 80 m to 1 km have been assimilated, the assimilation of wind profiles



**FIGURE 11** Dependence of  $\Delta\sigma_r^2$  averaged over all 16 available cases on the lead time. The forecast metrics are the (a)  $u$ -component and (b)  $v$ -component of 80-m wind. The assimilation is performed for simulated data from 95 SYNOP stations only (dashed black lines), and for 25 Doppler lidars in addition to the 95 SYNOP stations (other lines). Lidar observations are available at one (green lines), two (orange lines), three (blue lines), four (red dotted lines), and five (solid black lines) levels [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

from 1–3 km does not improve  $\Delta\sigma_r^2$  considerably on average. These results reveal that wind profiles up to 1 km are essential for the low-level wind forecast.

### 5.4 | Lead time

We investigated the impact of a Doppler lidar network for forecasts up to 3 hr. The absolute impact decreases with lead time and has approximately halved after a 3-hr forecast compared with analysis time, depending on the number of observed levels (Figure 11). Our results indicate that the impact of Doppler lidars is longer-lasting

compared with SYNOP observations, given the relative benefit of lidars with respect to SYNOP only. For instance, Figure 11 shows that for 1-hr lead time the improvement relative to the assimilation of only surface wind observations is a factor of 1.6 and 2.3 when the Doppler lidar data are assimilated with vertical measurement extent up to 80 m and 1 km, respectively. For 3-hr lead time, the improvement from assimilation of Doppler lidars increases to 2 and 2.7–3.3 for Doppler wind profiles up to 1 and 3 km, respectively. This result indicates that Doppler wind profiles are more beneficial than surface observations and the efficiency of assimilation of Doppler lidars is increasing with lead time. Especially noteworthy in this regard is that, while the fifth lidar level adds no additional benefit initially, its contribution grows with lead time. In contrast, while the benefit of the lower three lidar levels is greater, their contributions decrease with lead time.

## 6 | SUMMARY AND OUTLOOK

The main focus of our study is to estimate the impact of a hypothetical network of Doppler lidars on short-term forecasts of low-level wind in the RRA (Rhein–Ruhr area in the west of Germany). For this purpose, we developed and applied a new approach for ESA (Ensemble Sensitivity Analysis) that calculates the analysis variance reduction due to the assimilation of additional observations and its sensitivity to a forecast metric explicitly. In contrast to preceding methods, our approach accounts correctly for localization in the assimilation system without the need to prescribe the signal propagation of observation impact over time (in-depth analysis in Griewank *et al.*, 2022). This new method provides a very efficient approach for network design studies compared with computationally expensive Observing System Simulation Experiments (OSSEs) that require a complete model simulation for every potential configuration (Privé *et al.*, 2021).

Our study is conducted using output from an existing 1,000-member ensemble simulation (Necker *et al.*, 2020a). In total, we analyzed 16 forecasts of zonal and meridional wind on eight days during summer 2016. The benefit of assimilating hypothetical observations is characterized by the variance change  $\Delta\sigma_r^2$  relative to the original variance of the forecast metric before any assimilation. The benefit of additional Doppler lidar data is evaluated relative to the assimilation of surface-based wind sensor data assimilated operationally by the German weather service. The benefit is estimated for different potential networks of Doppler lidars randomly installed in the RRA. Results are discussed for  $u$  and  $v$  components of the 80-m wind averaged over the RRA. The main conclusions of our study are the following.

1. **Network density:** Our results indicate that, for the low-level wind in the RRA ( $\sim 210 \times 142 \text{ km}^2$ ), a network of 20–30 lidars is reasonable in the cost–benefit sense. Adding additional instruments to the network, in general, gives less improvement per instrument. For our study, we selected 25 instruments as a realistically achievable network. Such a network corresponds to installing a Doppler lidar at every fourth SYNOP site in our domain.
2. **Situation dependence:** The estimated value of the variance change  $\Delta\sigma_r^2$  varies by roughly a factor of two over the 16 separate forecasts in the summertime convective period analyzed.
3. **Vertical extent of the Doppler lidar measurements:** The exact benefit depends strongly on the penetration depth of the Doppler lidars, which is affected by the presence of hydrometeors. For example, low-level liquid clouds, fog, and precipitation can reduce the vertical measurement extent. Single-level lidar observations at 80-m height have an estimated benefit that is 1.6–2 times larger than that of the surface wind observations. The benefit of observing three levels up to 1-km height is 2.3–3.3 larger than that of the surface observations. Wind observations above 1-km height do not contribute considerably to further improvement, but their contribution grows with lead time. This result indicates that, for areas with frequent events of fog, low-level liquid clouds, and precipitation, scanning cloud radars and/or wind profilers could be beneficial to improve short-term low-level wind forecasts further.
4. **Lead time:** The estimated benefit of additionally assimilating wind profiles grows slightly with lead time compared with assimilating SYNOP only. For instance, for 1-hr lead time, the improvement relative to the assimilation of only surface wind observations is a factor of 2.3 when 1-km profiles from the Doppler lidars are assimilated. The improvement increases to 2.7–3.3 for 3-hr lead time.
5. **Sensitivity studies:** We conducted sensitivity studies to confirm that our approach behaves as expected in regard to observation errors, localization length scales, lead times, and number of observations. As results depend strongly on the chosen observation error and localization length scale, it is important to choose realistic values that mirror the operational modelling system of interest to achieve reliable impact estimates.
6. **Regularized ensemble sensitivity:** Our method relies on a regularized multivariate ensemble sensitivity, which requires us to choose the regularization coefficient  $\alpha$ . Overregularization suppresses sensitivities, while underregularization leads to extremely unrealistic and noisy sensitivities. We tested a range of regularization coefficients to find a suitable



value for our experimental setup. For a fixed value of  $\alpha = 3$ , the resulting sensitivities fulfil theoretical expectations for all lead times.

Our results show that Doppler lidars provide considerable potential to improve low-level wind forecasts at the hub heights of wind turbines (80 m), essential for sustainable energy applications. In addition to energy applications, as was shown in other studies, Doppler wind lidars can also improve predictions during the early evolution of severe weather (Coniglio *et al.*, 2019). While it is virtually impossible to predict precisely how much real observations would benefit an operational NWP system, due to biases, system complexity, and nonlinear interactions, our results emphasize further that Doppler lidars should be considered as an important component of future observational networks.

Future studies should extend similar investigations to longer periods to capture different weather situations and possible seasonal effects. In addition, it is possible to use the same approach for different forecast metrics or different ground-based instrumentation. An aspect that was beyond the scope of this study is the question of what minimum ensemble size would be required to guarantee reliable ensemble sensitivities. The 1,000-member ensemble applied mitigates the issue of sampling errors in ESA as shown by Necker *et al.* (2020a). If no large ensemble is available, a time-lagged ensemble or a climatology of forecasts (Hakim *et al.*, 2020; Tardif *et al.*, 2021) could allow a sufficient sample size to be obtained. A statistical sampling-error correction could help to mitigate sampling errors further when relying on smaller ensembles (Necker *et al.*, 2020b). Finally, this work has so far only analyzed the RRA located in the western part of Germany. The work can potentially be extended to more geographical regions, especially those having a high potential for wind-farm installations.

## AUTHOR CONTRIBUTIONS

**Tatiana Nomokonova:** conceptualization; formal analysis; methodology; software; visualization; writing – original draft; writing – review and editing. **Philipp J. Griewank:** conceptualization; methodology; writing – review and editing. **Ulrich Löhnert:** conceptualization; supervision; writing – review and editing. **Takemasa Miyoshi:** writing – review and editing. **Tobias Necker:** conceptualization; data curation; methodology; writing – review and editing. **Martin Weissmann:** conceptualization; supervision; writing – review and editing.

## ACKNOWLEDGEMENTS

This work has been carried out within and funded by the Hans-Ertel-Centre for Weather Research funded by the German Federal Ministry for Transportation and Digital Infrastructure (Grant number BMVI/DWD 4818DWD5B). The ideas of this research have been inspired by the WWRP Working group DAOS (“Data Assimilation and Observing Systems”) and CPEX-LAB (Cloud and Precipitation Exploration Laboratory within the Geoverbund ABC/J, <http://www.cplex-lab.de>). Collaborative efforts for this research have been supported through the EU COST Action CA18235 “PROBE” (European Cooperation in Science and Technology), funding agency for research and innovation networks (<http://www.cost.eu>). We acknowledge RIKEN for providing the SCALE-RM model data used in this study. We thank Elisabeth Bauernschubert from Deutscher Wetterdienst for providing us with the coordinates of the SYNOP stations used operationally for data assimilation. We appreciate Lukas Schmidt, who provided data of wind-power plants in Germany using the open code (<https://github.com/OpenEnergyPlatform/open-MaStR>) to scrap the data from the “Marktstammdatenregister” (MaStR). The ensemble data were postprocessed using the open-source project and Python package “xarray” (Hoyer and Hamman, 2017). We are grateful to the reviewers for their constructive suggestions, which helped to improve the article. Open Access funding enabled and organized by Projekt DEAL.

## CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.



## DATA AVAILABILITY STATEMENT

Code and output data are available on Zenodo (Nomokonova *et al.*, 2022, <https://doi.org/10.5281/zenodo.6331758>).

## ORCID

Tatiana Nomokonova  <https://orcid.org/0000-0002-8086-1403>

Philipp J. Griewank  <https://orcid.org/0000-0003-0906-3553>

Ulrich Löhnert  <https://orcid.org/0000-0002-9023-0269>  
Takemasa Miyoshi  <https://orcid.org/0000-0003-3160-2525>

Tobias Necker  <https://orcid.org/0000-0002-7484-3372>

Martin Weissmann  <https://orcid.org/0000-0003-4073-1791>

## REFERENCES

- Ahsbahs, T., Badger, M., Volker, P., Hansen, K.S. and Hasager, C.B. (2018) Applications of satellite winds for the offshore wind farm site anholt. *Wind Energy Science*, 3, 573–588.

- Ancell, B. and Hakim, G.J. (2007) Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting. *Monthly Weather Review*, 135, 4117.
- Bishop, C.H., Etherton, B.J. and Majumdar, S.J. (2001) Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Monthly Weather Review*, 129, 420–436.
- Browning, K.A. and Wexler, R. (1968) The determination of kinematic properties of a wind field using Doppler radar. *Journal of Applied Meteorology and Climatology*, 7, 105–113.
- Buizza, R. and Richardson, D. (2017) 25 years of ensemble forecasting at ECMWF. *ECMWF Newsletter*, 153, 20–31. Accessed January 5, 2022. <https://www.ecmwf.int/sites/default/files/elibrary/2017/18198-25-years-ensemble-forecasting-ecmwf.pdf>
- Coniglio, M.C., Romine, G.S., Turner, D.D. and Torn, R.D. (2019) Impacts of targeted AERI and Doppler lidar wind retrievals on short-term forecasts of the initiation and early evolution of thunderstorms. *Monthly Weather Review*, 147, 1149–1170.
- Desroziers, G., Berre, L., Chapnik, B. and Poli, P. (2005a) Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131, 3385–3396.
- Desroziers, G., Brousseau, P. and Chapnik, B. (2005b) Use of randomization to diagnose the impact of observations on analyses and forecasts. *Quarterly Journal of the Royal Meteorological Society*, 131, 2821–2837.
- Finn, T.S., Geppert, G. and Ament, F. (2020) Towards assimilation of wind profile observations in the atmospheric boundary layer with a sub-kilometre-scale ensemble data assimilation system. *Tellus A: Dynamic Meteorology and Oceanography*, 72, 1–14.
- Frehlich, R. (2013) Scanning Doppler lidar for input into short-term wind power forecasts. *Journal of Atmospheric and Oceanic Technology*, 30, 230–244.
- Gaspari, G. and Cohn, S.E. (1999) Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125, 723–757.
- Gasperoni, N.A. and Wang, X. (2015) Adaptive localization for the ensemble-based observation impact estimate using regression confidence factors. *Monthly Weather Review*, 143, 1981–2000.
- Gombos, D. and Hansen, J.A. (2008) Potential vorticity regression and its relationship to dynamical piecewise inversion. *Monthly Weather Review*, 136, 2668–2682.
- Gonzalez-Salazar, M.A., Kirsten, T. and Prchlik, L. (2018) Review of the operational flexibility and emissions of gas- and coal-fired power plants in a future with growing renewables. *Renewable and Sustainable Energy Reviews*, 82, 1497–1513.
- Gottschall, J., Gribben, B., Stein, D. and Würth, I. (2017) Floating lidar as an advanced offshore wind speed measurement technique: current technology status and gap analysis in regard to full maturity. *WIREs Energy and Environment*, 6, e250.
- Griewank, P., Weissmann, M., Necker, T., Nomokonova, T. and Löhnert, U. (2022) Ensemble-based estimates of the impact of potential observations. *Quarterly Journal of the Royal Meteorological Society*. Submitted.
- de Haan, S. (2011) High-resolution wind and temperature observations from aircraft tracked by mode-s air traffic control radar. *Journal of Geophysical Research: Atmospheres*, 116, D10111.
- Hacker, J.P. and Lei, L. (2015) Multivariate ensemble sensitivity with localization. *Monthly Weather Review*, 143, 2013–2027.
- Hakim, G.J., Bumbaco, K.A., Tardif, R. and Powers, J.G. (2020) Optimal network design applied to monitoring and forecasting surface temperature in Antarctica. *Monthly Weather Review*, 148, 857–873.
- Hirth, B.D., Schroeder, J.L. and Guynes, J.G. (2017) Diurnal evolution of wind structure and data availability measured by the DOE prototype radar system. *Journal of Physics: Conference Series*, 926, 012003.
- Hodyss, D. and Nichols, N. (2015) The error of representation: basic understanding. *Tellus A: Dynamic Meteorology and Oceanography*, 67, 24822.
- Hoerl, A.E. and Kennard, R.W. (1970) Ridge regression: applications to nonorthogonal problems. *Technometrics*, 12, 69–82.
- Houtekamer, P.L. and Zhang, F. (2016) Review of the ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 144, 4489–4532.
- Hoyer, S. and Hamman, J. (2017) xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, 5(1), 10.
- Hristova-Veleva, S., Zhang, S.Q., Turk, F.J., Haddad, Z.S. and Sawaya, R.C. (2021) Assimilation of DAWN Doppler wind lidar data during the 2017 convective processes experiment (CPEX): impact on precipitation and flow structure. *Atmospheric Measurement Techniques*, 14, 3333–3350.
- Hunt, B., Kostelich, E. and Szunyogh, I. (2007) Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, 230, 112–126.
- Jacques, D. and Zawadzki, I. (2015) The impacts of representing the correlation of errors in radar data assimilation. Part II: Model output as background estimates. *Monthly Weather Review*, 143, 2637–2656.
- Janjić, T., Bormann, N., Bocquet, M., Carton, J.A., Cohn, S.E., Dance, S.L., Losa, S.N., Nichols, N.K., Potthast, R., Waller, J.A. and Weston, P. (2018) On the representation error in data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 144, 1257–1278.
- Kalnay, E. (2002) *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge: Cambridge University Press.
- Lang, S. and McKeogh, E. (2011) Lidar and sodar measurements of wind speed and direction in upland terrain for wind energy purposes. *Remote Sensing*, 3, 1871–1901.
- Leuenberger, D., Haefele, A., Omanovic, N., Fengler, M., Martucci, G., Calpini, B., Fuhrer, O. and Rossa, A. (2020) Improving high-impact numerical weather prediction with lidar and drone observations. *Bulletin of the American Meteorological Society*, 101, E1036–E1051.
- Leutbecher, M. (2019) Ensemble size: how suboptimal is less than infinity?. *Quarterly Journal of the Royal Meteorological Society*, 145, 107–128.
- Li, L., Xie, N., Fu, L., Zhang, K., Shao, A., Yang, Y. and Ren, X. (2020) Impact of lidar data assimilation on low-level wind shear simulation at Lanzhou Zhongchuan International Airport, China: a case study. *Atmosphere*, 11, 1342.
- Lien, G.-Y., Miyoshi, T., Nishizawa, S., Yoshida, R., Yashiro, H., Adachi, S.A., Yamaura, T. and Tomita, H. (2017) The near-real-time SCALE-LETKF system: a case of the September 2015 Kanto-Tohoku heavy rainfall. *Scientific Online Letters on the Atmosphere*, 13, 1–6.
- Majumdar, S.J. (2016) A review of targeted observations. *Bulletin of the American Meteorological Society*, 97, 2287–2303.
- Majumdar, S.J., Bishop, C.H., Etherton, B.J. and Toth, Z. (2001) Adaptive sampling with the ensemble transform Kalman filter.

- Part II: Field program implementation. *Monthly Weather Review*, 130, 1356–1369.
- Martin, A., Weissmann, M., Reitebuch, O., Rennie, M., Geiß, A. and Cress, A. (2021) Validation of Aeolus winds using radiosonde observations and numerical weather prediction model equivalents. *Atmospheric Measurement Techniques*, 14, 2167–2183.
- Miyoshi, T., Kondo, K. and Imamura, T. (2014) The 10,240-member ensemble Kalman filtering with an intermediate AGCM. *Geophysical Research Letters*, 41, 5264–5271.
- Miyoshi, T., Kunii, M., Ruiz, J., Lien, G.-Y., Satoh, S., Ushio, T., Bessho, K., Seko, H., Tomita, H. and Ishikawa, Y. (2016a) “Big Data Assimilation” revolutionizing severe weather prediction. *Bulletin of the American Meteorological Society*, 97, 1347–1354.
- Miyoshi, T., Lien, G.-Y., Satoh, S., Ushio, T., Bessho, K., Tomita, H., Nishizawa, S., Yoshida, R., Adachi, S.A., Liao, J., Gerofi, B., Ishikawa, Y., Kunii, M., Ruiz, J., Maejima, Y., Otsuka, S., Otsuka, M., Okamoto, K. and Seko, H. (2016b) “Big Data Assimilation” toward post-petascale severe weather prediction: an overview and progress. *Proceedings of the IEEE*, 104, 2155–2179.
- Necker, T., Geiss, S., Weissmann, M., Ruiz, J., Miyoshi, T. and Lien, G.-Y. (2020a) A convective-scale 1,000-member ensemble simulation and potential applications. *Quarterly Journal of the Royal Meteorological Society*, 146, 1423–1442.
- Necker, T., Weissmann, M., Ruckstuhl, Y., Anderson, J. and Miyoshi, T. (2020b) Sampling error correction evaluated using a convective-scale 1000-member ensemble. *Monthly Weather Review*, 148, 1229–1249.
- Päschke, E., Leinweber, R. and Lehmann, V. (2015) An assessment of the performance of a 1.5  $\mu\text{m}$  Doppler lidar for operational vertical wind profiling based on a 1-year trial. *Atmospheric Measurement Techniques*, 8, 2251–2266.
- Petersen, R.A. (2016) On the impact and benefits of AMDAR observations in operational forecasting—Part I: A review of the impact of automated aircraft wind and temperature reports. *Bulletin of the American Meteorological Society*, 97, 585–602.
- Pichault, M., Vincent, C., Skidmore, G. and Monty, J. (2021) Short-term wind power forecasting at the wind farm scale using long-range Doppler lidar. *Energies*, 14, 2663.
- Pichugina, Y.L., Banta, R.M., Olson, J.B., Carley, J.R., Marquis, M.C., Brewer, W.A., Wilczak, J.M., Djalalova, I., Bianco, L., James, E.P., Benjamin, S.G. and Cline, J. (2017) Assessment of NWP forecast models in simulating offshore winds through the lower boundary layer by measurements from a ship-based scanning Doppler lidar. *Monthly Weather Review*, 145, 4277–4301.
- Piper, D., Kunz, M., Ehmele, F., Mohr, S., Mühr, B., Kron, A. and Daniell, J. (2016) Exceptional sequence of severe thunderstorms and related flash floods in May and June 2016 in Germany—Part 1: meteorological background. *Natural Hazards and Earth System Sciences*, 16, 2835–2850.
- Privé, N., Errico, R. and McCarty, W. (2021) The importance of simulated errors in observing system simulation experiments. *Tellus A: Dynamic Meteorology and Oceanography*, 73, 1–17.
- Sawada, M., Sakai, T., Iwasaki, T., Seko, H., Saito, K. and Miyoshi, T. (2015) Assimilating high-resolution winds from a Doppler lidar using an ensemble Kalman filter with lateral boundary adjustment. *Tellus A: Dynamic Meteorology and Oceanography*, 67, 23473.
- Schraff, C., Reich, H., Rhodin, A., Schomburg, A., Stephan, K., Perriñez, A. and Potthast, R. (2016) Kilometre-scale ensemble data assimilation for the COSMO model (KENDA). *Quarterly Journal of the Royal Meteorological Society*, 142, 1453–1472.
- Schween, J.H., Hirsikko, A., Löhnert, U. and Crewell, S. (2014) Mixing-layer height retrieval with ceilometer and Doppler lidar: from case studies to long-term assessment. *Atmospheric Measurement Techniques*, 7, 3685–3704.
- Simmer, C., Adrian, G., Jones, S., Wirth, V., Göber, M., Hohenegger, C., Janjic, T., Keller, J., Ohlwein, C., Seifert, A., Trömel, S., Ulbrich, T., Wapler, K., Weissmann, M., Keller, J., Masbou, M., Meilinger, S., Riß, N., Schomburg, A., Vormann, A. and Weingärtner, C. (2016) Herz: The German Hans-Ertel Centre for weather research. *Bulletin of the American Meteorological Society*, 97, 1057–1068.
- Stone, E.K. (2018) A comparison of mode-s enhanced surveillance observations with other in situ aircraft observations. *Quarterly Journal of the Royal Meteorological Society*, 144, 695–700.
- Sweeney, C., Bessa, R.J., Browell, J. and Pinson, P. (2020) The future of forecasting for renewable energy. *WIREs Energy and Environment*, 9, e365.
- Tardif, R., Hakim, G.J., Bumbaco, K.A., Lazzara, M.A., Manning, K.W., Mikolajczyk, D.E. and Powers, J.G. (2021) Assessing observation network design predictions for monitoring Antarctic surface temperature. *Quarterly Journal of the Royal Meteorological Society*, 148, 727–746.
- Theuer, F., van Dooren, M.F., von Bremen, L. and Kühn, M. (2020) Minute-scale power forecast of offshore wind turbines using long-range single-Doppler lidar measurements. *Wind Energy Science*, 5, 1449–1468.
- Tikhonov, A.N. (1965) Ill-conditioned problems in linear algebra and their robust method solution. *Doklady Akademii Nauk SSSR*, 163, 591–594.
- Torn, R.D. (2014) the impact of targeted dropwindsonde observations on tropical cyclone intensity forecasts of four weak systems during PREDICT. *Monthly Weather Review*, 142, 2860–2878.
- Torn, R.D. and Hakim, G.J. (2008) Ensemble-based sensitivity analysis. *Monthly Weather Review*, 136, 663.
- Weissmann, M., Göber, M., Hohenegger, C., Janjic, T., Keller, J., Ohlwein, C., Seifert, A., Trömel, S., Ulbrich, T., Wapler, K., Bollmeyer, C. and Deneke, H. (2014) Initial phase of the Hans-Ertel Centre for weather research—a virtual centre at the interface of basic and applied weather and climate research. *Meteorologische Zeitschrift*, 23, 193–208.
- Whitaker, J.S. and Hamill, T.M. (2002) Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130, 1913.
- WMO (2021). Guide to instruments and methods of observation – measurement of meteorological variables. Technical report WMO-N.8, Vol. I. Accessed on January 5, 2022. [https://library.wmo.int/doc\\_num.php?explnum\\_id=11386](https://library.wmo.int/doc_num.php?explnum_id=11386)

**How to cite this article:** Nomokonova, T., Griewank, P.J., Löhnert, U., Miyoshi, T., Necker, T. & Weissmann, M. (2023) Estimating the benefit of Doppler wind lidars for short-term low-level wind ensemble forecasts. *Quarterly Journal of the Royal Meteorological Society*, 149(750), 192–210. Available from: <https://doi.org/10.1002/qj.4402>