Assimilating atmospheric infrasound data to constrain atmospheric winds in a two-dimensional grid

Received: 23 October 2020 | Revised: 6 July 2021 | Accepted: 25 July 2021

DOI: 10.1002/qj.4141

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

Assimilating atmospheric infrasound data to constrain atmospheric winds in a two-dimensional grid

Javier Amezcua[1] | **Zak Barton**[2]

[1]University of Reading and UK National Centre for Earth Observation, Reading, UK

[2]University of Reading, Reading, UK

Correspondence
J. Amezcua, University of Reading and UK National Centre for Earth Observation, Reading, UK.
Email: j.amezcuaespinosa@reading.ac.uk

Funding information
European Commission FP7 and Horizon 2020 programmes, Grant/Award Numbers: 284387, 653980

Abstract

Infrasound waves travelling through atmospheric channels are affected by the conditions they encounter along their path. The shift in the back azimuth angle of a wavefront detected at the reception site depends on the cross-winds it encountered. Estimating the original field from this integrated measurement is an (ill-posed) inverse problem. By using a prior, this can be converted into a Bayesian estimation problem. In this work, we use the (ensemble) Kalman filter (EnKF) to tackle this problem. In particular, we provide an illustration of the setup and solution of the problem in a two-dimensional grid, depending on both across-track distance and height, which has not been done in previous works. We use a synthetic setup to discuss the details of the method. We show that one of the effects of along-track averaging (done in previous studies to simplify the problem) is to overestimate the magnitude of the analysed values, and propose that this will be a source of model error. We also illustrate the process with real data corresponding to nine controlled ammunition explosions that took place in the summer of 2018. In these cases, the real infrasound waves we study seldom reach higher than 40 km in height. However, the use of covariance-based methods (e.g., the EnKF) allows for updates in higher regions where the wave did not travel and where traditional observations are sparse. In fact, the larger impacts from observations in these cases are in the region of 40–60 km, in agreement with previous works. This study contributes to paving the way towards the ultimate goal of assimilating information derived from infrasound waves into operational numerical weather forecasting. More studies in quality control of the observations and proper validation of the results are urgently needed.

KEYWORDS

atmospheric infrasound acoustics, data assimilation, ensemble Kalman filter, infrasound observations, middle atmospheric dynamics, stratospheric winds

1 | INTRODUCTION

Infrasound waves can be generated by natural sources such as earthquakes, volcanoes and ocean microbaroms, and also from human activities such as mining and controlled detonations of old ammunition. Under certain atmospheric conditions, infrasound waves generated at the surface of the Earth can propagate through atmospheric wave-guides over large distances, potentially reaching up to high levels of the atmosphere before being reflected or refracted back towards the surface, where they can be detected by ground-based stations (Georges 1972; Garces *et al.*, 1998). The characteristics of a given wave are modified as it crosses the atmosphere, with variables such as wind, temperature and humidity affecting both the speed and orientation of the detected wavefront as measured at a receiving station. Hence, the received wave contains integrated information on the atmospheric regions it encountered along its path. This was recognised quite early (Groves, 1956; Donn and Rind, 1972; Rind *et al.*, 1973). However, trying to infer these atmospheric characteristics from an integrated observation such as the back azimuth angle of the incoming wavefront is an ill-posed inverse problem. Making use of a prior renders a Bayesian estimation problem, therefore providing an avenue to provide solutions (sometimes suboptimal yet still useful) to this problem (see e.g., Stuart, 2010; Ash *et al.*, 2016).

Information derived from detected infrasound waves can prove valuable in constraining winds and temperatures in the atmosphere, especially in poorly observed upper regions such as the stratosphere and mesosphere. The top of current numerical weather prediction models extends well into the mesosphere (Polavarapu *et al.*, 2005). It is recognised that processes in both the mesosphere and stratosphere, and their connection to the troposphere, are important for both numerical weather prediction and climate prediction (Orsolini *et al.*, 2011; Charlton-Perez *et al.*, 2013; Kidston *et al.*, 2015). However, the observational coverage of these two layers is nowhere near as dense as that of the troposphere. Some of the largest uncertainties in current reanalyses occur in these levels, leading the wind representation of mesospheric winds to lag behind other areas (Baker *et al.*, 2014; Korhonen *et al.*, 2019). For instance, Duruisseau *et al.* (2017) compared high-altitude (pressure less than 20 hPa) ERA-Interim values against wind measurements taken from a balloon radiosonde and noted that the standard deviations of the differences were greater than $5\,\mathrm{m\cdot s^{-1}}$ in some cases. Le Pichon *et al.* (2018) conducted a similar study comparing multiple reanalysis products against measurements from multiple ground-based instruments, finding that the standard deviation of the mean zonal wind difference at times could be larger than $20\,\mathrm{m\cdot s^{-1}}$ when looking specifically at heights of 40–60 km. It is therefore clear that any extra information that can be used to constrain such estimations better is of paramount value, as well as the methodology to extract this information in an adequate manner.

The present work is one of several studies paving the way to the assimilation of atmospheric infrasound data into numerical weather prediction models. Assink *et al.* (2019) provide a review on this ultimate goal. We follow on the steps of Blixt *et al.* (2019), Amezcua *et al.* (2020), and Vera-Rodriguez *et al.* (2020). These works make use of a data set resulting from controlled ammunition detonations at the Hukkakero site in Finland. These explosions generate infrasound waves that are detected by an array of powerful micro-barometers at the ARCES ground-based site in Karasjok, Norway. A series of daily detonations are performed over several days every summer, yielding observational over the course of the last two decades (Drob *et al.*, 2003; Gibbons *et al.*, 2007). These data have two desirable qualities. First, since waves are generated from controlled explosions, we know exactly the locations of the source and receiver of the wave, as well as the time of detonation and reception, the latter two within some error margin. Second, the event can be considered an individual pulse, which is not the case, for instance, with continuous sources in a general region (e.g., microbarom sources near Iceland), in which it is almost impossible to disentangle individual waves. The second is the proximity between the source at Hukkakero and the receiver at Karasjok, which is approximately 179 km, where the ARCES array of micro-barometers has been located since 1987 (Mykkeltveit *et al.*, 1990). This means that the infrasound waves travel for around 10 min, so the whole process can be considered instantaneous against synoptic-scale variability. This is not a feasible approximation, for instance, for waves generated by microbaroms near Iceland and detected in Norway, where the distance is an order of magnitude larger (around 2,000 km).

Blixt *et al.* (2019) used the ARCES data set to derive effective cross-wind velocities, which they compared with the ERA-Interim reanalysis data (Dee *et al.*, 2011). The authors tried different vertical averages of the along-track-averaged cross-wind to approximate the effective wind inferred from the observations. They also performed ray-tracing (e.g., Hedlin and Walker, 2013) for the different explosions to derive plausible tracks and maximum vertical penetration heights, and compute vertical sensitivities to the ray. Vera-Rodriguez *et al.* (2020) increased the scope of the estimation problem. While still working in the along-track average setting, they estimated three variables: tail- and cross-winds, as well as temperature, and used three observations: travel time, back azimuth angle and apparent velocity.

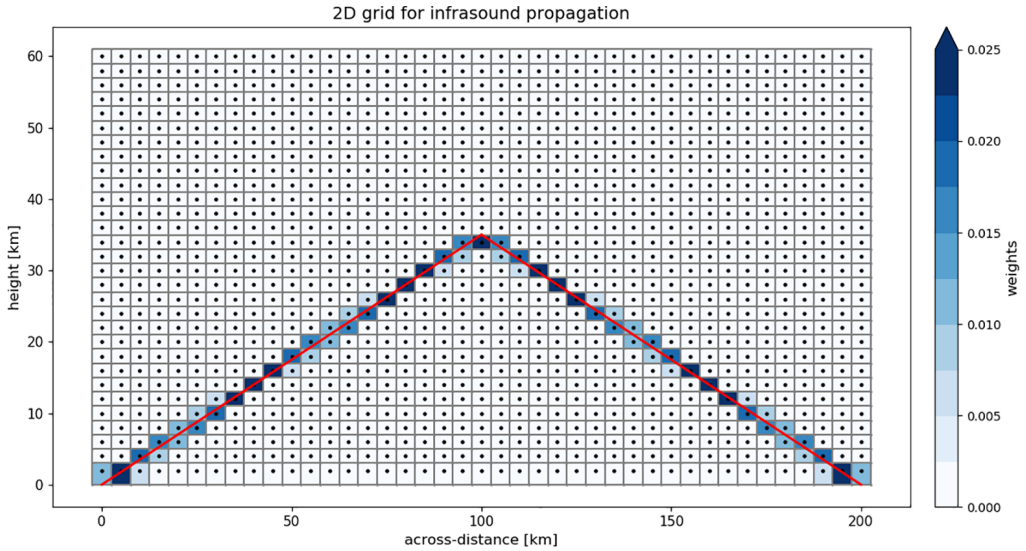

FIGURE 1 Sample two-dimensional atmospheric slab used in this work. The horizontal axis corresponds to the along-track distance, and the vertical axis to height. The red line shows the trajectory of an infrasound wave from emission, being reflected at some maximum height, and travelling back and being detected at the surface. The blue-shaded grid boxes are those touched by the ray, and the intensity of the colour is proportional to the distance travelled by the ray inside the grid boxes [Colour figure can be viewed at wileyonlinelibrary.com]

The authors performed an inversion using a heuristic algorithm developed by Vera-Rodriguez (2019). This algorithm relies on a smoothness constrain (as a regularisation strategy) to estimate these variables at 137 vertical levels. They used the ERA-5 EDA (ensemble data assimilation) ten-member reanalysis data to provide a departure point for the inversion algorithm and to identify ensemble members with values close to those estimated from observations (e.g., Simmonds *et al.*, 2020). One of the limitations they recognised is the fact that they could not estimate values at those levels above the maximum height reached by the infrasound waves, which is often around 40 km.

Amezcua *et al.* (2020) performed off-line data assimilation (DA) with the observed back azimuth angles from different explosions. They used the deterministic ensemble Kalman filter (DEnKF, Sakov and Oke, 2008) to estimate along-track-averaged cross-wind values at different levels of the atmosphere. As background values, they used wind values coming from the ten-member ERA-5 EDA reanalysis. Once more, the cross-wind values were considered constant along the whole 179-km track of the infrasound wave, but vertical variations were permitted. The estimation was not done, however, at all the 137 (unevenly spaced) vertical levels of the reanalysis, but instead on 6 (evenly spaced) layers of the atmosphere. The use of sample covariances coming from the background ensemble allowed the determination of upper-level winds, which were also the most uncertain in the reanalysis product. To make the DA successful, the authors used domain localisation and inflation in the EnKF. All the mentioned works considered a one-dimensional (1D) setting, which required along-track averaging. This is a major restriction that is overcome herein.

The main objective of this work is to provide a setup to perform DA with infrasound measurements on a two-dimensional (2D) grid, that is, allowing for along-track variations of the cross-wind for different numbers of along-track sections. As an illustration, consider Figure 1. A 2D atmospheric slab of $L_a = 200\,\text{km}$ (along-track distance) and $L_z = 60\,\text{km}$ (vertical distance) is discretised into N_a along-track sections, and N_z vertical levels, in this case with sizes $da = 5\,\text{km}$ and $dz = 2\,\text{km}$. An infrasound wave is emitted at the point $(0, 0)\,\text{km}$, travels to a maximum height attained in the middle of the horizontal domain at the point $(100, 35)\,\text{km}$, and is reflected back towards the surface, where it is detected at the point (2,000) km. The trajectory followed by the infrasound wave is shown by a red line, forming an isosceles triangle. A grid point is coloured white if the ray does not touch it, and in a shade of blue if the ray touches it. The shade of blue (labelled weight) is proportional to the distance the wave travels inside the grid box divided by the length of the total trajectory (more on this in the next section). The problem is then to use the integrated observation to update the cross-wind variable at the $N_a \times N_z$ grid points.

In Section 2, we briefly discuss the geometric limit for the propagation of infrasound waves in the atmosphere, which is the setup used in this and previous works. Perhaps the most important point of this section is to define the observation operator which acts in the DA problem. Section 3 discusses the Kalman filter and the ensemble implementation we use in this work. It also helps define

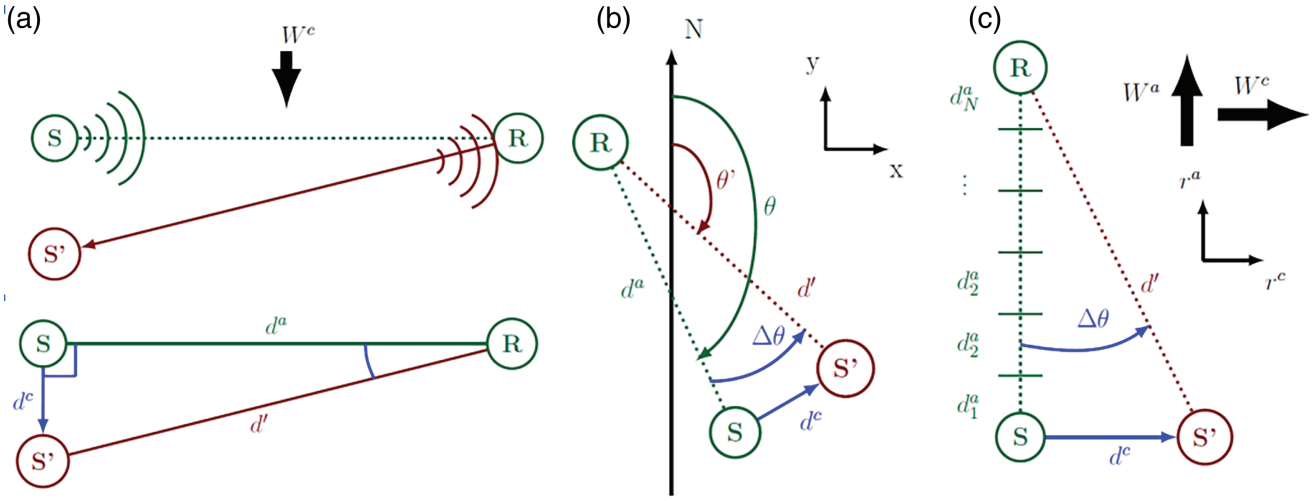

FIGURE 2 Simple schematic showing the geometric setup for infrasound transmission. (a) Effect of a cross-wind on a wave travelling from a source S to a receptor R. The cross-wind advects the wavefront and creates an apparent source S'. (b) Modification of the observed back azimuth angle of the wavefront at the reception site. (c) Discretisation of the process into N_a segments in the along-track direction of propagation [Colour figure can be viewed at wileyonlinelibrary.com]

the role of the elements defined in Section 2 in the DA process. Section 4 is, in our opinion, the most important of the work. Using a synthetic setup, we demonstrate the assimilation process under ideal conditions, in particular illustrating how the presence of correlations in the background wind field allow for updates beyond the wave path. In this section we also illustrate the effects that along-track averaging of the original fields (as used in previous works to simplify the problem) can lead to inaccurate results unless a source of representation error (e.g., Janjić *et al.*, 2018) is considered. In Section 5 we perform the DA process using data from nine explosions that took place in 2018 at Hukkakeiro. The sample covariances coming from a ten-member ensemble prove to be very noisy, and we propose a way to ameliorate this. Finally, Section 6 contains a summary and discussion about this work. Furthermore, there are many areas that still need to be tackled before proceeding to operational assimilation of infrasound waves for NWP. We discuss these in detail and provide ideas for the future.

2 | GEOMETRIC PERSPECTIVE OF THE PROPAGATION OF INFRASOUND WAVES IN THE ATMOSPHERE

The main phenomenon exploited in this study is that a background wind affects the propagation of an infrasound wave. More specifically, the wavefront of an infrasound wave can be advected by any cross-wind it encounters. The opposite is not true, meaning that an infrasound wave has a negligible effect on the background winds.

For the following explanation we rely on Figure 2. Panel (a) illustrates a wave originating at a source S and received at a receiver R some distance away. A cross-wind W^c (constant for now) is present along its path. If $W^c = 0$, the infrasound wave arrives at the receiver unaltered and can be traced back to its true source. If the wavefront, however, is translated by a cross-wind $W^c \neq 0$, then it can be traced back to an apparent source S'. A right-angled triangle is then formed between S, R and S'. The distance between S and R is called the along-track ground distance and is denoted by d^a, whereas the distance from S to S' is called the cross-track distance and is denoted by d^c.

When a wave is detected, one can measure the back azimuth angle θ, that is, the angle of the incoming ray with respect to north, as shown in panel (b). The bending caused by W^c leads to a measured angle back azimuth θ'. The effect of the cross-wind W^c is related to the difference between these angles (Diamond 1964):

$$\Delta\theta = \theta' - \theta, \tag{1}$$

where if $W^c > 0$ then $\Delta\theta < 0$, and if $W^c < 0$ then $\Delta\theta > 0$. To relate W^c and $\Delta\theta$, one uses the concept of celerity v, that is, the along-track distance d^a divided by the total propagation time T:

$$v = \frac{d^a}{T}, \tag{2}$$

where T is a known quantity, since we know the times of both explosion and detection. W^c can be directly related to d^c using the fact that, by construction, the cross-wind is parallel to d^c. Therefore,

$$d^c = W^c\, T. \tag{3}$$

From our geometric setup, the angle shift $\Delta\theta$ can be written as

$$\Delta\theta = -\arctan\left(\frac{W^c}{\nu}\right). \tag{4}$$

So far, W^c has been constant in time and space. When considering the explosions used in this study, we have that $d^a \approx 179\,\text{km}$, so assuming a constant W^c is not realistic. The next step is to consider W^c as a variable that changes with position along d^a, thus becoming $w^c(r^a, t)$, where r^a and r^c denote the directions that d^a and d^c face, respectively. This is shown in panel (c) of Figure 2. If we consider the waves generated at the source to be spherical, we always have a wavefront at d^a. Therefore, we always focus on the cross-wind along this line, and there is no reason to consider variation in the r^c direction. In the rest of this work, we continue to consider the wind fields to be time invariant. Since, in our setting, the whole infrasound transmission process lasts about 10 min, this is not a bad approximation. The general case with time-varying fields is more complicated and is not explored in this article.

Returning to Figure 2, in panel (c) we divide the distance d^a into N_a sections such that $d^a = \sum_{n_a=1}^{N_a} d_{n_a}$. In the n_a^{th} section, we can still apply Equation (3) to obtain the cross-wind shift $d^c_{n_a}$ as

$$d^c_{n_a} = w^c_{n_a} T_{n_a}, \tag{5}$$

where $w^c_{n_a}$ is the cross-wind value in the n_a^{th} section in the along-track direction, and T_{n_a} is the time the wavefront spends in this section. It should be clear that $T = \sum_{n_a=1}^{N_a} T_{n_a}$.

To calculate each T_{n_a}, we can relate the along-track distance ground segment length, $d^a_{n_a}$, with the along-track speed for each ground segment denoted by $w^a_{n_a}$ and the adiabatic speed of sound denoted by C, namely

$$T_{n_a} = \frac{d^a_{n_a}}{C + w^a_{n_a}}. \tag{6}$$

In reality, C is a function of temperature and atmospheric water content, but is considered constant here. Substituting Equation (6) into (5) and adding over all segments yields the apparent cross-wind displacement as

$$d^c = \sum_{n_a=1}^{N_a} \frac{d^a_n w^c_{n_a}}{C + w^a_{n_a}}. \tag{7}$$

To use (4) we can define an average cross-wind velocity W^c by dividing (7) by the total time T. We can write (4) in this more general case as

$$\Delta\theta = -\arctan\left(\frac{\sum_{n_a=1}^{N_a} \alpha_{n_a} w^c_{n_a}}{\nu}\right), \tag{8}$$

with the weights being

$$\alpha_{n_a} = \frac{d^a_{n_a}}{T\left(C + w^a_{n_a}\right)}. \tag{9}$$

These weights take into account the time spent within each section along $d^a_{n_a}$. They can also be written as

$$\alpha_{n_a} = \frac{T_{n_a}}{T}, \tag{10}$$

which implies that they are normalised.

The problem we are solving includes the wave travelling to a given height and being reflected back, whilst also travelling in the along-track direction. In Figure 1, the maximum vertical penetration is 35 km. Therefore, we introduce height as another dimension along which the wind fields vary. Note that, when projecting the trajectory onto the horizontal plane, our analysis for the relationship between the back azimuth and the effective cross-wind remains valid. We split the vertical distance into N_z sections, such that

$$z_{\max} = \sum_{n_z=1}^{N_z} z_{n_z}, \tag{11}$$

where $z_{\max}$ is the maximum height reached by a given wave. For every grid position $\{n_a, n_z\}$ in this two-dimensional setting, there is an associated cross-wind value given by $w^c_{n_a,n_z}$. Multiplying these winds by associated weights and summing over both N_a and N_z produces the weighted average cross-wind as

$$W^c = \sum_{n_a=1}^{N_a}\sum_{n_z=1}^{N_z} \alpha_{n_a,n_z}\, w^c_{n_a,n_z}. \tag{12}$$

The values α_{n_a,n_z} are the weights for each grid box. The weight is zero for grid boxes within which the wave does not propagate, and non-zero if it does. As done in the case of 1D propagation, Equation (12) can used to relate the cross-wind to the deviation in the back azimuth angle as

$$\Delta\theta = -\arctan\left(\frac{1}{\nu}\sum_{n_a=1}^{N_a}\sum_{n_z=1}^{N_z} \alpha_{n_a,n_z}\, w^c_{n_a,n_z}\right). \tag{13}$$

For our experiments, the weights α_{n_a,n_z} are obtained as a product of two factors. The first is the length of the wave trajectory in a given grid box, and the second factor is the vertical sensitivity to the wave transmission. This is obtained from a ray-tracing technique as done in Blixt *et al.* (2019).

3 | SOLVING THIS PROBLEM WITH THE (ENSEMBLE) KALMAN FILTER

3.1 | The setup of the problem in KF terms

DA techniques include two steps: forecast and analysis. As in Amezcua *et al.* (2020), in this study we perform off-line assimilation; that is, we only do the analysis step. The analysis is not fed back into the forecast model as improved initial conditions, hence there is no cycling in this study. In this section, we discuss the analysis step of the DA method that we use to approach the problem. We also relate the elements of the previous sections to those needed in the DA process. We choose the stochastic ensemble Kalman filter (SEnKF – Burgers *et al.*, 1998; van Leeuwen, 2020). The reason behind this choice will become evident when discussing covariance localisation.

The Kalman filter (KF, Kalman, 1960; Kalman and Bucy, 1961) relies on forecasting the first two moments of the state variables, and uses observations to update the values from background (or forecast) to analysis. The background value $\mathbf{x}^b \in R^{N_x}$ is considered to come from a multivariate Gaussian distribution (MGD) with mean $\boldsymbol{\mu}^b \in R^{N_x}$ and covariance matrix $\mathbf{B} \in R^{N_x \times N_x}$. An important fact that becomes useful in Section 4 is that a covariance matrix can be decomposed as

$$\mathbf{B} = \boldsymbol{\Sigma}^{1/2}\mathbf{C}\boldsymbol{\Sigma}^{1/2}, \tag{14}$$

where $\boldsymbol{\Sigma} \in R^{N_x \times N_x}$ is a diagonal matrix with the variance of the variables along the main diagonal, and $\mathbf{C} \in R^{N_x \times N_x}$ is a correlation matrix. A correlation matrix has 1's along the main diagonal, while all the off-diagonal elements are bounded between -1 and 1.

In the analysis step of the KF, the information coming from observations is incorporated. An observation $\mathbf{y} \in R^{N_y}$ is a linear transformation of the state variable via an observation matrix $\mathbf{H} \in R^{N_y \times N_x}$, and contaminated by observational error. Explicitly,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta}, \tag{15}$$

where $\boldsymbol{\eta} \in R^{N_y}$ is observational error coming from an MGD with zero mean (unbiased) and observational error covariance $\mathbf{R} \in R^{N_y \times N_y}$. In our problem, $N_y = 0$.

The difference between the observation and the mapping of the background mean into observation space is known as the innovation $\mathbf{d}^{ob} \in R^{N_y}$. Explicitly,

$$\mathbf{d}^{ob} = \mathbf{y} - \mathbf{H}\boldsymbol{\mu}^b. \tag{16}$$

The analysis step updates the $\boldsymbol{\mu}$ and $\mathbf{B}$ into their corresponding analysis values using the equations

$$\boldsymbol{\mu}^a = \boldsymbol{\mu}^b + \mathbf{K}\mathbf{d}^{ob} \tag{17a}$$

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\,\mathbf{B}. \tag{17b}$$

Note that the equation for the analysis mean is extremely simple to interpret: it is a linear combination of the background mean and the innovation. The weights in the linear combination are given by the rectangular matrix $\mathbf{K} \in R^{N_x \times N_y}$, the so-called Kalman gain, explicitly

$$\mathbf{K} = \mathbf{B}\mathbf{H}^{\mathbf{T}}\boldsymbol{\Gamma}^{-1} \tag{18a}$$

$$\boldsymbol{\Gamma} = \mathbf{H}\mathbf{B}\mathbf{H}^{\mathbf{T}} + \mathbf{R}. \tag{18b}$$

The Kalman gain can be regarded as the multivariate version of a ratio of variances: the background (co)variance in the numerator, and the total (co)variance $\boldsymbol{\Gamma}$ (background plus observation) in the denominator. The KF equation for the analysis covariance can be interpreted as a contraction of the background covariance.

There can be complications with respect to the framework we just described. For instance, it can be difficult to obtain robust estimators for the mean $\boldsymbol{\mu}^b$ and covariance $\mathbf{B}$ of the state variables. Instead, we have access to a sample of realisations of the state variable, called an ensemble, whose size is often considerably smaller than the number of state variables. Another complication is that observations can involve nonlinear transformations of the state variables. A Monte Carlo implementation of the KF known as the ensemble Kalman filter solves this problem. While the KF is optimal when the observation operator is linear, the ensemble alternative can provide suboptimal yet still useful estimators (Evensen, 2009). There are many implementations of the EnKF, from which we choose the stochastic ensemble Kalman filter (SEnKF; Burgers *et al.*, 1998; van Leeuwen, 2020) in this work.

Let us write our problem in terms of the notation used for the KF. The state vector $\mathbf{x} \in R^{N_x}$ is the collection of cross-wind values in the grid, concatenated in any order as long as one is consistent throughout all the steps of the method. We choose the following order (first vertical, then along-track):

$$\mathbf{x} = \left[w^c_{1,1}, \ldots, w^c_{1,N_z}, \ldots, \ldots, w^c_{N_a,1}, \ldots, w^c_{N_a,N_z}\right]^{\mathbf{T}}, \tag{19}$$

where $N_x = N_a N_z$. Consider that we have a background sample of N_e elements. This is labelled as background ensemble $\mathbf{X}^b \in \mathcal{R}^{N_x \times N_e}$ and can be written explicitly as the

matrix

$$\mathbf{X}^b = \left[\mathbf{x}_1^b,\ \mathbf{x}_2^b,\ \dots, \mathbf{x}_{N_e}^b\right]. \tag{20}$$

Now let us look at the observational side. Defining the sensitivity coefficients for each grid point following the same order as for (19), we have the vector $\boldsymbol{\alpha} \in R^{N_x}$,

$$\boldsymbol{\alpha} = \left[\alpha_{1,1},\ \dots, \alpha_{1,N_z},\ \dots,\ \dots, \alpha_{N_a,1},\ \dots, \alpha_{N_a,N_z}\right]^{\mathbf{T}}. \tag{21}$$

Many of the elements of this vector will be zero, since they correspond to grid points where the ray does not travel. We can finally write the (nonlinear) observation equation as

$$\mathbf{y} = -\arctan\left(\frac{\boldsymbol{\alpha}^{\mathbf{T}}\mathbf{x}}{\nu}\right) + \boldsymbol{\eta}. \tag{22}$$

In this case we have $N_y = 1$, that is, a single integrated observation from N_x state variables. Furthermore, $h : R^{N_x} \to R^{N_y}$ is a nonlinear function. Nonetheless, it is not too far from linear as long as the argument of the arc-tangent is relatively smaller than unity, since the Taylor expansion of this function is

$$\arctan(\xi) = \xi + \mathcal{O}(\xi)^3. \tag{23}$$

In our case, the argument of the arc-tangent is a fraction with an effective wind speed in the numerator and celerity in the denominator. The typical values of the denominator are 10^1 m·s^{-1}, whereas the celerity in the denominator is close to C, which is around 300 m·s^{-1}. Therefore the (nondimensional) value of the argument in the arc-tangent is around 10^{-1}. The linearisation of the observation operator around $\mathbf{x}_{ref} = \mathbf{0}$ is labelled as $\hat{\mathbf{H}} \in R^{N_y \times N_x}$. This becomes important later in this section, and is simple to write down as

$$\hat{\mathbf{H}} = -\frac{1}{\nu}\boldsymbol{\alpha}^{\mathbf{T}}. \tag{24}$$

The analysis step of the SEnKF is quite simple to implement since it performs the update for each one of the ensemble members separately, by repeatedly applying the KF analysis equation for the mean. The analysis value of the n_e^{th} ensemble member is

$$\mathbf{x}_{n_e}^a = \mathbf{x}_{n_e}^b + \mathbf{K}_e\mathbf{d}_{n_e}^{ob}. \tag{25}$$

The perturbed innovations for each ensemble member $\mathbf{d}_{n_e}^{ob} \in R^{N_y}$ are computed as follows (van Leeuwen, 2020):

$$\mathbf{d}_{n_e}^{ob} = \mathbf{y}^o - \left(h\left(\mathbf{x}_{n_e}^b\right) + \boldsymbol{\eta}_{n_e}\right). \tag{26}$$

The term $\boldsymbol{\eta}_{n_e} \in R^{N_y}$ is a realisation of the observational error, which is added to the mapping of the state variables into observation space (hence the word Şstochastic Ť in SEnKF). It is easy to ensure that the N_e perturbations are centred in zero. We have written $\mathbf{K}_e$ to denote that the gain is computed with the ensemble covariance $\mathbf{P}^b$, which is an estimator of $\mathbf{B}$. The ensemble-based gain is then

$$\mathbf{K}_e = \mathbf{P}^b\hat{\mathbf{H}}^{\mathbf{T}}\boldsymbol{\Gamma}_e^{-1} \tag{27a}$$

$$\boldsymbol{\Gamma}_e = \hat{\mathbf{H}}\mathbf{P}^b\hat{\mathbf{H}}^{\mathbf{T}} + \mathbf{R}. \tag{27b}$$

$\hat{\mathbf{H}} \in R^{N_y \times N_x}$ is the linearisation of the observation operator, which we found in Equation (24). There are ways to compute $\mathbf{K}_e$ that do not require the linearisation of the observation operator, but instead map all members into observation space, compute perturbation matrices in both state and observational space, and perform outer products (e.g., Hunt *et al.*, 2007; Sakov and Oke, 2008). We do not go down this path. The reason for our choice has to do with the use of model-space localisation and having an integrated observation operator, which is explained next.

3.2 | Implementing localisation

The Kalman gain has the important role of spreading the impact from observations to state variables. In our problem, this is particularly crucial since we have a single observation whose integrated information has to be distributed to N_x state variables. The quality of $\mathbf{P}^b$ is vital to the quality of $\mathbf{K}$ and therefore to the accuracy of the result of the DA process. Small samples are can lead to spurious covariances as a result of sampling noise. Localisation is a method to ameliorate this problem. It involves artificially modifying the sample covariances to reduce the magnitude of (and sometimes completely eliminate) undesired elements. There are several ways to implement localisation (see e.g., Sakov and Bertino, 2011). Here we only discuss the so-called B-localisation, which is adequate when dealing with observations which include integrated information; see for example Greybush *et al.* (2011) for the differences between B- and R-localisation.

B-localisation involves directly modifying $\mathbf{P}^b$ or its partial or total mapping into observation space, that is, $\mathbf{PH}^{\mathbf{T}}$ and $\mathbf{HPH}^{\mathbf{T}}$. The latter is useful when the computation of the full $\mathbf{P}^b$ is not feasible, but this is not an issue in our case. The simplest form of B-localisation occurs in model space (see e.g., Slyeva *et al.*, 2019). This requires directly replacing the sample covariance matrix for a modified version, that is, $\mathbf{P}^b \to \mathbf{P}_{loc}^b$. Model-space localisation requires the linearisation of the observation operator to perform the products $\hat{\mathbf{H}}\mathbf{P}^{\mathbf{b}}{}_{loc}\hat{\mathbf{H}}^{\mathbf{T}}$ and $\mathbf{P}^{\mathbf{b}}{}_{loc}\hat{\mathbf{H}}^{\mathbf{T}}$ required in the Kalman gain (Shlyeva *et al.*, 2019). Fortunately, in our case we have this readily available in Equation (24).

The localised background covariance matrix $\mathbf{P^b}_{loc} \in R^{N_x \times N_x}$ is often obtained by Schur (element-wise) multiplication of the original $\mathbf{P}^b$ with a tapering matrix $\mathbf{L} \in R^{N_x \times N_x}$, that is,

$$\mathbf{P}^b = \mathbf{P}^b \circ \mathbf{L}. \tag{28}$$

This tapering matrix is prescribed, and its design can come from our knowledge of the physical system, or using length scales coming from a tuning process to optimise (under some metric), the performance of the DA system.

Another option to modify $\mathbf{P}^b$ is to perform the decomposition indicated in Equation (14), trust the sample variances, and focus on modifying the sample correlation matrix $\mathbf{C}$. The ensemble correlation raised to a power (ECORAP; Bishop and Hodyss, 2007; ECORAP; Bishop and Hodyss, 2009) is a method which aims to "sharpen" the structures present in $\mathbf{C}$. In its simple implementation, it involves raising all and every element of the correlation matrix to a power p:

$$C_{i,j} \leftarrow \left(C_{i,j}\right)^p \text{sign}\left(C_{i,j}\right), \tag{29}$$

where the sign function ensures that the sign of the original element is not lost when k is even. The rationale behind this method is that correlations will become more Şpeakedť. If the correlation value is exactly equal to 1 it remains exactly so, values close to 1 remains close, while values close to 0 fade in magnitude after being raised to a power. The method we use in this work is a slight modification of ECORAP.

3.3 | Evaluating the effects of the DA process

To quantify the effects resulting from the DA process, consider some derived quantities. We write these in terms of the elements of the KF. When using the ensemble implementation, the corresponding sample estimators need to be substituted.

$$\mathbf{d}^{ab} = \boldsymbol{\mu}^a - \boldsymbol{\mu}^b \tag{30a}$$

$$\left(\mathbf{r}^{ab}\right)^2 = \text{diag}\left(\mathbf{A}\right) \div \text{diag}\left(\mathbf{B}\right). \tag{30b}$$

The first is the so-called analysis increment, that is, the difference of the analysis mean $\boldsymbol{\mu}^a$ minus the background ensemble mean $\boldsymbol{\mu}^b$. This can take both positive and negative values, with a value of zero meaning the observation does not have any effect. The second quantity is the ratio of the analysis variance of each state variable and its respective background variance. This ratio is bounded between zero and one, and the smaller the value, the more the observation reduces the uncertainty of the estimator of a given variable.

In our problem, we only have one (integrated) observation for each DA instant. As a consequence, both the observational departure d^{ob} and the total covariance γ^2 become scalars. Using these facts and Equation (24), the analysis increments are simply

$$\mathbf{d}^{ab} = -\frac{d^{ob}}{\nu\gamma^2}\mathbf{B}\boldsymbol{\alpha}. \tag{31}$$

Therefore, any spatial variation of the analysis increments is a result of either (a) spatial variations in the observational weights, or (b) spatial variations in the covariance matrix. This is illustrated in Section 4

It is also useful to look at quantities in observation space, in particular when in our problem everything is a scalar in that space, making comparisons simple. We denote the equivalent angle coming from the background mean, and its standard deviation, as

$$y^b = h\left(\boldsymbol{\mu}^b\right) \tag{32a}$$

$$\sigma_y^b = \sqrt{\hat{\mathbf{H}}\mathbf{B}\hat{\mathbf{H}}^{\mathbf{T}}}. \tag{32b}$$

The same can be done for the analysis:

$$y^a = h\left(\boldsymbol{\mu}^a\right) \tag{33a}$$

$$\sigma_y^a = \sqrt{\hat{\mathbf{H}}\mathbf{A}\hat{\mathbf{H}}^{\mathbf{T}}}. \tag{33b}$$

4 | SYNTHETIC EXPERIMENTS WITH THE KALMAN FILTER

4.1 | Setup

To illustrate the DA process, we perform idealized experiments with prescribed background settings and a synthetic observation. In this section we use the linearised version of the observation operator $\hat{\mathbf{H}}$ as the true observation operator $\mathbf{H}$. In this way we can use the exact KF and directly study the exact updates in the mean and covariance of the state variables, without the effects of nonlinearity and limited ensemble size. The experiments in this section use the 2D grid of Figure 1, which was first mentioned in Section 1. Recall that the domain is 200 km in the horizontal and 60 km in the vertical. It is discretised into $N_z = 30$ vertical levels centred at $z = \{2, 4, \ldots, 60\}$ km, and $N_a = 41$ along-track sections centred at $r^a = \{0, 5, \ldots, 200\}$ km.

TABLE 1 Background elements for the synthetic experiments

Background mean	Background variances	Background correlation matrix
Constant	Space dependent	Diagonal
Constant	Space dependent	Non-diagonal
Space dependent	Constant	Diagonal
Space dependent	Constant	Non-diagonal

This yields a total of 1,230 grid points, and the same number of state variables. The wave trajectory is shown by the red isosceles triangle with vertices at $(0, 0)$, $(100, 35)$ and $(200, 0)$ km, touching 65 out of 1,230 grid points, which is only around 5%. We consider uniform sensitivity in the vertical to the wave propagation. Hence, the weights in the observation operator are only proportional to the length travelled by the ray in different grid boxes.

We now need to prescribe background mean and covariance values. Table 1 presents a summary of the settings used. We study the separate effect of the spatial structure of the covariance and the magnitude of the variances. For this we use Equation (14) and prescribe two correlation matrices. The first one is simply the identity, which means that the background error is completely uncorrelated and innovations only occur at grid points through which the ray travels. The second correlation matrix has non-zero off-diagonal elements. This renders a case that is much more interesting, and also more realistic as we find in Section 5. The prescribed correlation matrix $\mathbf{C} \in R^{N_a N_z \times N_a N_z}$ for the non-diagonal case is

$$\mathbf{C} = \mathbf{C}_a \otimes \mathbf{C}_z, \tag{34}$$

i. e., a Kronecker product of two matrices: $\mathbf{C}_a \in R^{N_a \times N_a}$ for the along-track correlations and $\mathbf{C}_z \in R^{N_z \times N_z}$ for the vertical correlations. We choose these to be Toeplitz symmetric matrices. The first row of each one of these matrices uniquely defines each matrix. These rows are

$$\mathbf{C}_a[1, :] = \exp\left(-\frac{r^a}{25}\right) \tag{35a}$$

$$\mathbf{C}_z[1, :] = \exp\left(-\frac{(z-2)}{10}\right)\cos\left(\frac{\pi(z-2)}{15}\right). \tag{35b}$$

These matrices are illustrated in Figure 3. The left panel is for the vertical correlations, the middle panel for the along-track correlations, and the right panel is the resulting 2D correlation matrix. This last one participates in the DA process, after being scaled by the variances as in Equation (14). Note that we have allowed for both positive and negative vertical correlations but only positive horizontal ones. This is to mimic the real situations we found in Section 5.

For the observation side, we fix the celerity as $\nu = 300\text{m}\cdot\text{s}^{-1}$. This is in line with the values from real observations in Section 5. We also repeated the experiments with different values from $\nu = 260$ to $\nu = 340\text{m}\cdot\text{s}^{-1}$, and the results were qualitatively similar (not shown). We consider an observation with a given value and prescribed error standard deviation given as

$$\begin{aligned} \Delta\theta &= -0.5^\circ = -0.0087 \text{ rad} \\ \sigma^o &= \mathbf{R}^{1/2} = 0.1^\circ = 0.0017 \text{ rad}. \end{aligned} \tag{36}$$

The standard deviation of the error is large, about 10% of the observation value, but agrees with the literature.

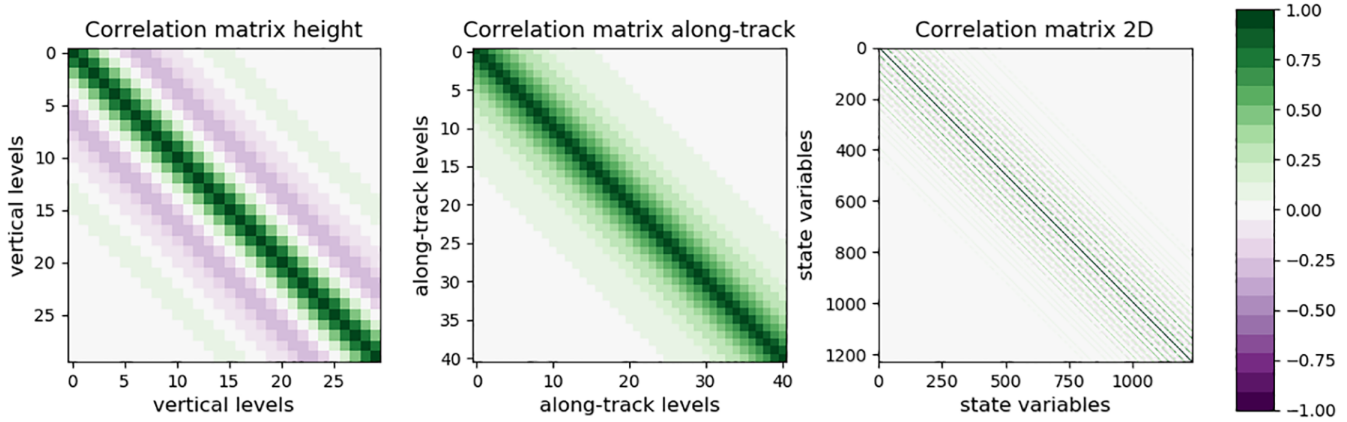

FIGURE 3 Prescribed correlation matrix used in the synthetic experiments. It is generated by the Kronecker product of a vertical correlation matrix (left) and an along-track correlation matrix (middle), to yield the 2D correlation matrix (right) [Colour figure can be viewed at wileyonlinelibrary.com]

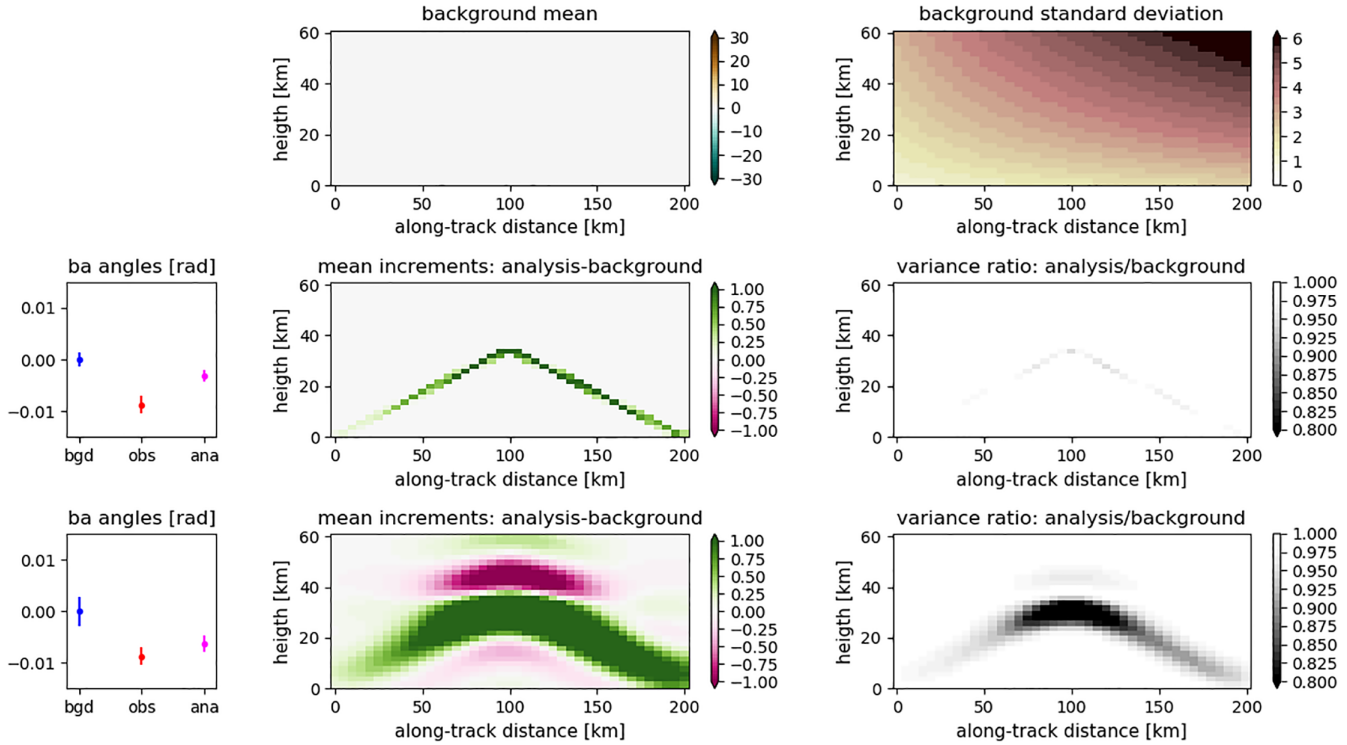

FIGURE 4 Summary of synthetic experiments using the KF to assimilate a single observation. The first row shows the background mean field (zero in this case) and the background standard deviation (variable in space). The rows show the results when using a diagonal **B** and a non-diagonal **B**. The columns show the results in observation space (left), and analysis increments (middle) and variance reduction (right) in the model space [Colour figure can be viewed at wileyonlinelibrary.com]

4.2 | Results under two settings

The setup and results of the first set of experiments (first two rows of Table 1) are shown in Figure 4. The first row in the figure shows the background mean (middle panel) and background standard deviation (right panel) for each grid point. These quantities are plotted as 2D maps. We set the background mean to be zero in all the domain. We allow, however, for spatial variations in the background wind variance. The top right panel illustrates the increase from the bottom left corner to the upper right corner. More precisely, the spatial variation is generated by the proposed expression

$$\sigma^b(r^a, z) = \left(1 + \frac{\sqrt{8}-1}{200} r^a\right)\left(1 + \frac{\sqrt{5}-1}{58}(z-2)\right), \tag{37}$$

that is, a product of a linear increase in the horizontal direction and a linear increase in the vertical direction. The results of assimilating this observation are shown in the middle and bottom rows of Figure 4. The left panels show comparisons of the actual observations (red) and the equivalent angles from background (blue) and analysis (magenta) in observation space, as defined in Equation (32) and (33), respectively. The middle panels show the analysis increments for each grid point, and the right panels show the variance ratios.

The results of using a diagonal **B** are shown in the middle row. We start with the elements in observation space. Since $\boldsymbol{\mu}^b$ is zero everywhere, the resulting background angle is $y^b = 0$. The observed angle is negative, which implies a positive equivalent cross-wind encountered by the wave, as discussed after Equation (1). In observation space, the filter is doing exactly what is expected: the analysis angle y^a lies between the background value and the observation. Furthermore, σ_y^a is smaller than both σ_y^b and σ^o. Now, we move to model space. Since **B** is diagonal, there are non-zero increments only at the grid points where the ray travelled, with the magnitudes of these increment being larger for those grid points that contain longer ray segments. Furthermore, the magnitude of the analysis increments is not symmetric with respect to the midpoint of the ray trajectory; in fact, it increases towards the top right, which is expected since $\sigma^b(r^a, z)$ in Equation (37) increases in the same direction. This agrees with the discussion on Equation (31). In this example, the largest analysis increments are around $1\,\mathrm{m{\cdot}s^{-1}}$. For the variance reduction from background to analysis, once more only the grid points where the ray travels are affected. The variances are reduced up to about 0.95 at most. As expected,

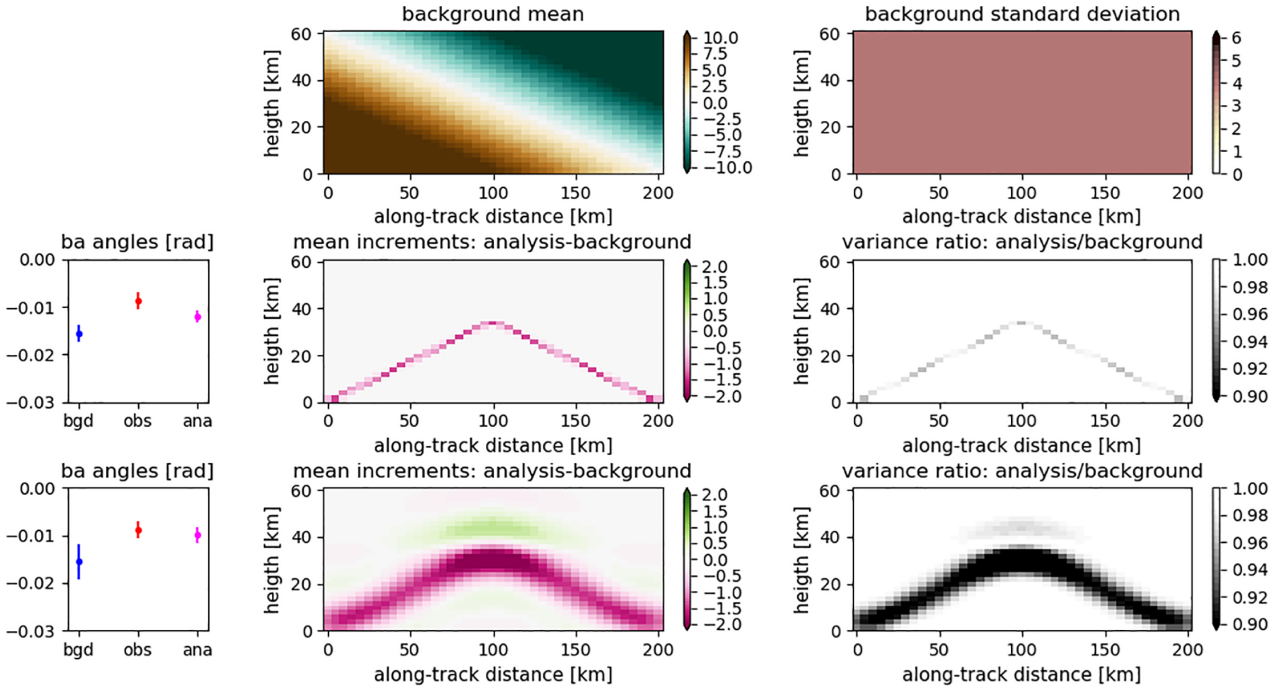

FIGURE 5 As in Figure 4, but for the case of variable background mean and constant background standard deviation [Colour figure can be viewed at wileyonlinelibrary.com]

the points along the trajectory closest to the top right area of the domain see the greatest reduction.

The bottom row of Figure 4 shows the results when using the non-diagonal **B** generated by Equations (34) and (37). In this case $\boldsymbol{\mu}^b$ is still zero, so we still have $y^b = 0$. However, σ_y^b is larger than that coming from the diagonal **B**. As a consequence, the analysis mean is closer to the observation value. The analysis increments are no longer limited to those grid points where the ray travels, since the non-zero covariances are able to communicate the effect of the observations to those grid points that the ray does not touch at all. There is a large area of positive increments (the expected sign) centred around the ray trajectory. These increments go up to around $4\,\mathrm{m{\cdot}s^{-1}}$. However, there are also regions of negative increments reaching around $-1.4\,\mathrm{m{\cdot}s^{-1}}$, which are just above and below the ray trajectory, and are a direct consequence of the negative correlations we prescribed in the vertical. The variance reduction is also not constrained to those grid points touched by the ray. The reduction of variances is larger than for the diagonal **B** case, with grid points reaching a ratio of around 0.8. As in the case with diagonal **B**, in both the analysis increments and variance ratio there is strong spatial dependence with larger effects towards the right and top of the domain.

The setup and results of the second set of experiments (last two rows of Table 1) are shown in Figure 5. In this case, $\sigma^b = 4\,\mathrm{m{\cdot}s^{-1}}$ in the whole domain. $\boldsymbol{\mu}^b$, however, changes over the 2D grid from -20 to $20\,\mathrm{m{\cdot}s^{-1}}$, according to

$$\mu^b(r_a, z) = 20\left(1 - \frac{z}{60} - \frac{r^a}{200}\right). \tag{38}$$

This variation was chosen such that the mean over the whole domain vanishes. This is not a realistic wind field, but it allows to prove some points in the following subsection.

Once more, we start with the results for the diagonal **B**. In observation space, $y^b \neq 0$; it is in fact negative, with a larger magnitude than that of y^o. Once more, y^a is between y^b and the y^a, so the DA is working. This time, the innovation d^{ob} is positive, so we expect a negative analysis increment. This is indeed what we find in the middle panel: the analysis increment is negative, with maximum increments of about $2.22\,\mathrm{m{\cdot}s^{-1}}$. The increments are restricted to the grid points touched by the ray, and their magnitude is just dependent on the associated weight in each grid box, since $\sigma^b(r^a, z)$ is constant at all grid points. The spatial variation of $\mu^b(r_a, z)$ coming from Equation (38) does not result in a spatial variation of either the analysis increments. The variance ratio reduction is limited to the grid points touched by the ray, any variation depends only on the weights of the grid points, and the minimum value of the ratio is about 0.96.

The bottom row shows the results of using a non-diagonal **B**. In observation space, y^b is exactly the same as with the diagonal **B** (as it should be since **B** plays no role in the computation of this y^b), but σ_y^b is larger. Hence, the observation has a larger impact in the assimilation, with y^a closer to y^o than in the diagonal **B** case. The analysis increments are larger in magnitude than with a

diagonal $\mathbf{B}$, this time reaching close to $4\,\mathrm{m\cdot s^{-1}}$ for some grid points. These are not restricted to the grid points touched by the ray; instead the non-zero covariance elements communicate the information to the grid points surrounding the ray. There are some negative increments in the area above the maximum height reached by the ray. Since the background variance is fixed in the whole domain, the spatial pattern of the increments depends only on the weights in the grid points and the prescribed correlations. The variance ratio can take values as low as 0.83, and it is not limited to the grid boxes in which the ray travels.

4.3 | The effects of spatial averaging

Consider that the grid in Figure 1 represents the true resolution of the atmospheric model in which we want to perform DA. The objective of the DA process is to find analysis values at the same resolution, which implies finding values for a total of $N_x = N_a N_z$ state variables (when we only have one physical variable per grid point). This is a difficult goal when using a single observation and imperfect information coming from a limited ensemble size. For this reason, one may be tempted (or plainly need) to reduce this problem to a 1D setting. One option is to perform along-track averaging. This was in fact done in Blixt *et al.* (2019), Amezcua *et al.* (2020) and Vera-Rodriguez *et al.* (2020). In these works the cross-wind was only allowed to vary with height, reducing the number of state variables to N_z (a different actual number for each work). The next discussion details along-track averaging, although the same steps apply for vertical averaging, and in fact we do both in experiments.

Averaging has two effects. The first effect is quite obvious: the size of the problem is reduced at the cost of losing information. We move from a 2D background field to a 1D background field. The resulting analysis is therefore also a 1D field. There is no way to (uniquely) transform this analysis back into a 2D field. The second effect is less obvious, and is particularly important given that our observation is an integrated quantity. To explain this, let us denote the along-track-averaged background mean as $\tilde{\boldsymbol{\mu}}^b \in R^{N_z}$, and the corresponding weights for the observation operator as $\tilde{\boldsymbol{\alpha}} \in R^{N_z}$. These elements are related to the original ones by

$$\tilde{\boldsymbol{\mu}}^b = \frac{1}{N_a}\boldsymbol{\Phi}\boldsymbol{\mu}^b \tag{39a}$$

$$\tilde{\boldsymbol{\alpha}} = \boldsymbol{\Phi}\boldsymbol{\alpha}, \tag{39b}$$

where the role of the matrix $\boldsymbol{\Phi} \in R^{N_a \times N_a N_z}$ is to perform sums. For the particular concatenation order we chose in Equations (19) and (21), $\boldsymbol{\Phi}$ is a block-row matrix with N_z repetitions of the $N_a \times N_a$ identity matrix:

$$\boldsymbol{\Phi} = \left[\mathbf{I}_{N_a}\ \mathbf{I}_{N_a}\ \ldots\ \mathbf{I}_{N_a}\right]. \tag{40}$$

The innovations in observation space are still scalars, and they can be written as

$$\tilde{\mathbf{d}}^{ob} = y^o - \frac{1}{N_a}\mathbf{H}\boldsymbol{\Phi}^{\mathbf{T}}\boldsymbol{\Phi}\boldsymbol{\mu}^b, \tag{41}$$

where $N_a^{-1}\boldsymbol{\Phi}^{\mathbf{T}}\boldsymbol{\Phi} \in R^{N_a N_z \times N_a N_z}$ is an idempotent matrix. In general,

$$\tilde{\mathbf{d}}^{ob} \neq \mathbf{d}^{ob}. \tag{42}$$

In fact, only under very limited circumstances do the innovations obtained from averaged inputs coincide with the innovations obtained from the full fields. One trivial case where they coincide is if the whole background field is equal to zero.

The analysis mean coming from the averaged data is

$$\tilde{\boldsymbol{\mu}}^a = \tilde{\boldsymbol{\mu}}^b + \tilde{\mathbf{K}}\tilde{\mathbf{d}}^{ob}, \tag{43}$$

with the gain

$$\tilde{\mathbf{K}} = \frac{1}{N_a^2}\boldsymbol{\Phi}\mathbf{B}\boldsymbol{\Phi}^{\mathbf{T}}\boldsymbol{\Phi}\mathbf{H}^{\mathbf{T}}\tilde{\boldsymbol{\Gamma}}^{-1} \tag{44a}$$

$$\tilde{\boldsymbol{\Gamma}} = \frac{1}{N_a^2}\mathbf{H}\boldsymbol{\Phi}^{\mathbf{T}}\boldsymbol{\Phi}\mathbf{B}\boldsymbol{\Phi}^{\mathbf{T}}\boldsymbol{\Phi}\mathbf{H}^{\mathbf{T}} + \mathbf{R}. \tag{44b}$$

In general,

$$\tilde{\boldsymbol{\mu}}^a \neq N_a^{-1}\boldsymbol{\Phi}\boldsymbol{\mu}^a. \tag{45}$$

Namely, doing DA using the original 2D field and then taking the along-track average of the result is not the same as doing DA using the 1D field as an input. The difference arises both in the innovations and in the way in which these innovations are spread back to the state space. Finally, the analysis increments resulting from $\tilde{\boldsymbol{\mu}}^a$ and $\tilde{\boldsymbol{\mu}}^b$ are

$$\tilde{\mathbf{d}}^{ab} = \tilde{\mathbf{K}}\tilde{\mathbf{d}}^{ob}, \tag{46}$$

which is not the same as the average of the analysis increments resulting from using the 2D fields; that is, in general,

$$\tilde{\mathbf{d}}^{ab} \neq N_a^{-1}\boldsymbol{\Phi}\mathbf{d}^{ab}. \tag{47}$$

We revisit the experiments from Table 1 to explore these effects. We start by performing a comparison in observation space in Figure 6. This figure has four panels. The top row corresponds to the case of fixed $\boldsymbol{\mu}$ and space-dependent σ^b, while the bottom row corresponds to the case of a space-dependent $\boldsymbol{\mu}$ and constant σ^b.

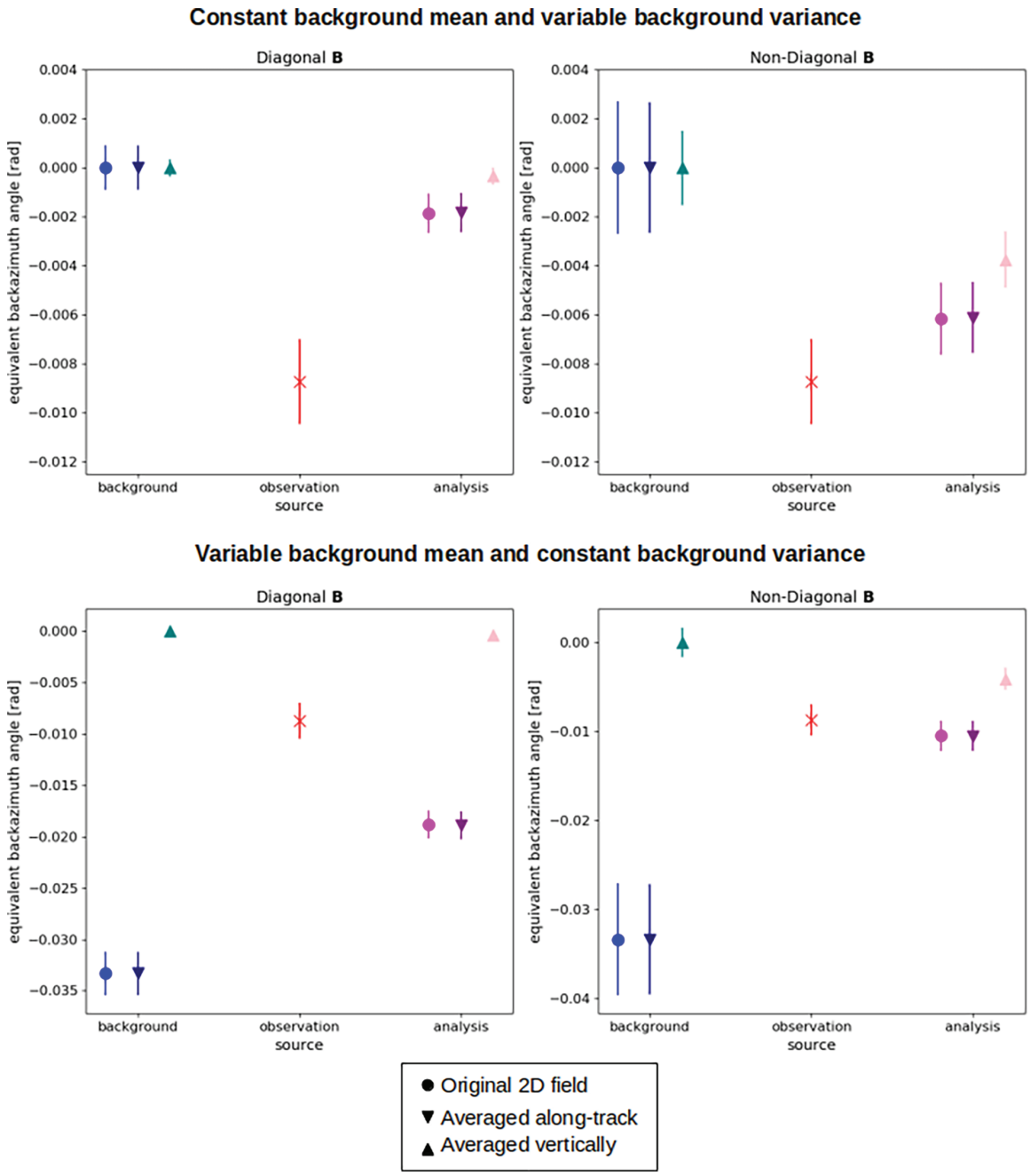

FIGURE 6 Effects of doing DA using summarised data. For different combinations of background mean and standard deviations (rows), and correlation matrices (columns), we show the equivalent angles coming from the background (blue lines), the observation (red) and the analysis (pink lines) [Colour figure can be viewed at wileyonlinelibrary.com]

The columns correspond to using a diagonal **B** (left) and a non-diagonal **B** (right). Each panel shows the equivalent values coming from the background (blue colours), the actual observation (in red) and the analysis (pink/purple colours). This time, for background and analysis we have three values: one coming from the full 2D field (circle), one coming from along-track-averaged data (down triangle) and one coming from vertically averaged data (up triangle). The value coming from the 2D field is the exact one, against which the other two should be compared.

Let us start with the case when $\boldsymbol{\mu} = \mathbf{0}$ and σ^b is space dependent. In this case all the background angles coincide, for diagonal and non-diagonal **B**. Also, the values for σ_y^b coming from the averaged data are smaller than those coming from the original 2D field. This is understandable if we think of the effect of averaging in the standard deviation of a random variable. The standard deviations are larger for the case of the non-diagonal **B**, as expected. Now we look at the analysis angles. In the case of a diagonal **B**, there is a large difference between the analysis coming from the 2D

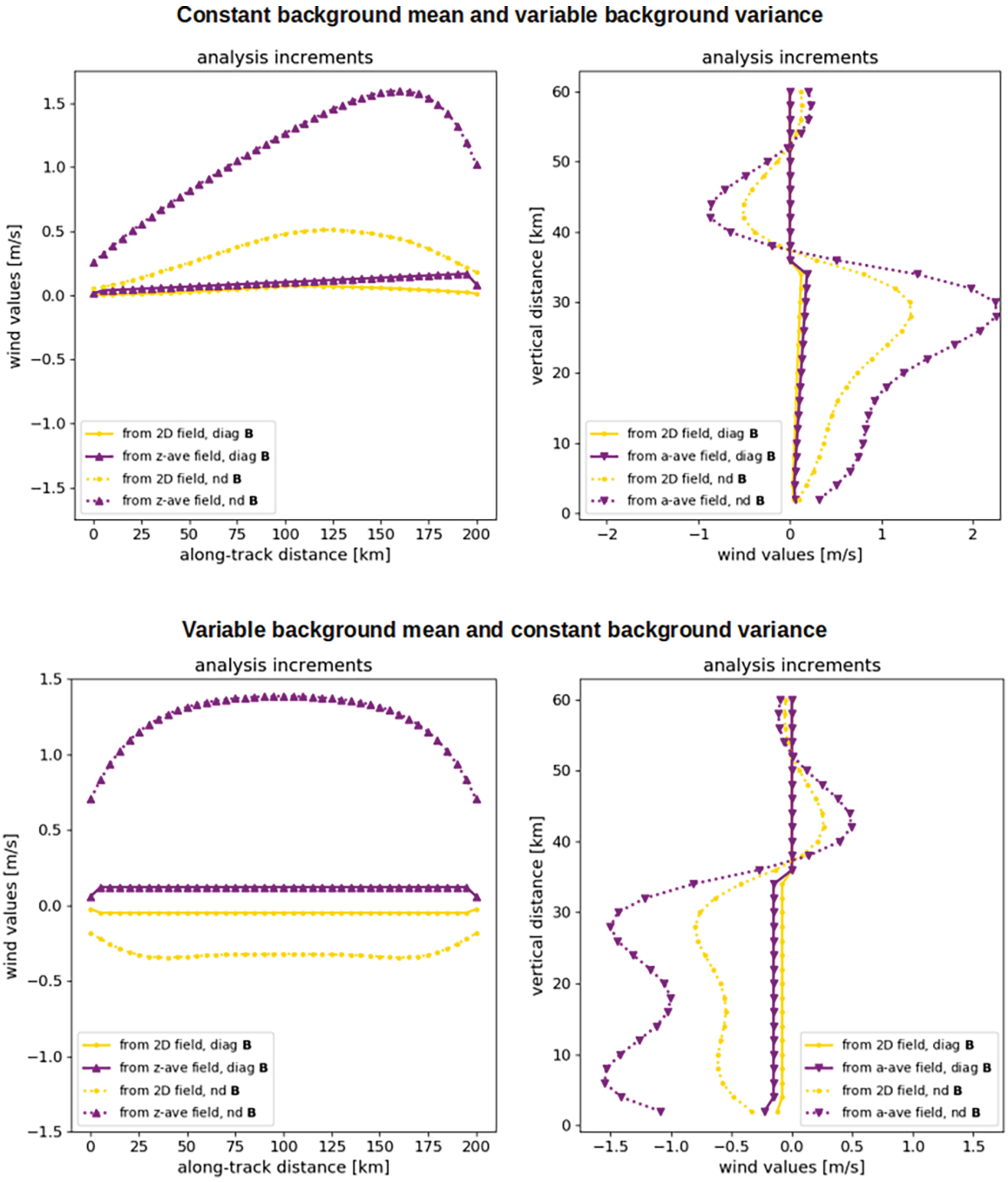

FIGURE 7 Effects of doing DA using summarised data, for different combinations of background mean and standard deviations (rows), in the case of vertical averages (left column) and along-track averages (right columns). In each panel we show analysis increments for two types of correlation matrices (line style), coming from using 2D fields in the DA (pink lines) and from using averaged data (purple lines) [Colour figure can be viewed at wileyonlinelibrary.com]

field and the analyses coming from the summarised data (either along-track or vertical averages). In fact, the impact of the observation in both averaged cases is very small. This impact is larger in the case of the non-diagonal $\mathbf{B}$, and the difference with respect to the analysis angle is smaller.

Now we move to the case of a space-dependent $\boldsymbol{\mu}^b$ and fixed σ^b. In this case the difference in the background angles coming from the 2D and the averaged fields can be quite large depending on the way the averaging is done. For the vertically averaged case, the equivalent angle is zero. This can be understood given the symmetry of the background mean wind field, and realising that the ray covers the whole along-track distance, but not all the vertical distance. The same differences are present in the angles corresponding to the analysis wind fields. The effect of the observation is larger in the case of the 2D field since the associated background standard deviations are larger. As in other cases, the values of σ_y^b are larger in

the case of non-diagonal **B**. After looking at observation space, we move to model space and look at the differences between $N_a^{-1}\mathbf{\Phi d}^{ab}$ () and $\tilde{\mathbf{d}}^{ab}$ (). These results are plotted in Figure 7. The rows have the same order as in Figure 6. The columns, however, show separately the cases for the vertical (left) and along-track (right) averages. In each panel we have four lines: the pink lines correspond to $N_a^{-1}\mathbf{\Phi d}^{ab}$, and the purple lines to $\tilde{\mathbf{d}}^{ab}$. The solid lines correspond to a diagonal **B**, and the dotted lines to a non-diagonal **B**. The largest analysis increments correspond to the case of having non-diagonal **B**. In fact they are close to zero when $N_a^{-1}\mathbf{\Phi d}^{ab}$ is computed either along-track or vertically. This is expected if we remember that only around 5% of the grid points are directly touched by the ray. The increments differ little between the two rows (i.e., with different combinations of background mean and variances), and also the error made when using averaged fields for the DA is small. This is not the case, however, for non-diagonal **B**. In this case the values of $N_a^{-1}\mathbf{\Phi d}^{ab}$ are large. In the vertical, they are not limited to the levels travelled by the ray, as is evident in the right column. In this case the difference between $N_a^{-1}\mathbf{\Phi d}^{ab}$ and $\tilde{\mathbf{d}}^{ab}$ is quite noticeable, with the latter overestimating the magnitude of the former in all cases. The mismatch is particularly bad for the case of variable $\mu^b(r^a, z)$ and constant σ^b, where the sign of $\tilde{\mathbf{d}}^{ab}$ is opposite to that of $N_a^{-1}\mathbf{\Phi d}^{ab}$ in the vertically averaged case. This is also the case for diagonal **B**, albeit barely noticeable due to the small magnitude of the increments.

The difference between $\tilde{\boldsymbol{\mu}}^a$ and $N_a^{-1}\mathbf{\Phi}\boldsymbol{\mu}^a$ is an example of representation error; see for example Janjić *et al.* (2018) for a nice discussion. This is something that has to be further studied and taken into consideration if the goal of assimilating infrasound wave angles into operational models involves large differences in the resolutions of the original grid and that used for the ray-tracing and ultimately the DA process.

5 | DATA ASSIMILATION EXPERIMENTS WITH REAL OBSERVATIONS

Now that we have illustrated the method in a synthetic setting, as well as the issues that may arise, we move to the assimilation of real observations, and using the ERA-5 EDA product as background.

5.1 | Observations

As outlined above, the data set used in this study contains observations related to contained explosions of ammunition at Hukkakero, Finland. While the infrasound waves generated by these explosions are received at multiple locations, we use data recorded by the ARCES array in Karasjok, Norway. Blixt *et al.* (2019) comment on the advantages that this site has over others, but a primary reason relates to the geographic setup that Hukkakero and Karasjok provide. The distance d^a is close to being parallel to the north pointing line, which implies that the cross-winds are similar to the eastward winds, which Blixt *et al.* (2019) note as being the dominant wind force for middle atmospheric winds (Drob *et al.*, 2003).

The infrasound waves caused by the explosions take approximately 10 min to arrive at the receiver. For each explosion, Blixt *et al.* (2019) performed ray-tracing simulations in an attempt to model the propagating waves. These simulations use data from ERA-Interim. The ray-tracing analysis yields vertical sensitivities for the computation of the effective cross-winds.

In this work we focus on the nine explosions which occurred in August 2018. Table 2 presents the dates and times of each explosion and three quantities. The first is the back azimuth angle (in radians), which is directly measured, where we consider an unbiased additive error following a normal distribution with a standard deviation of $\frac{1}{20}^\circ$ (about $\frac{1}{100}$ rad. Blixt *et al.* (2019) or Szuberla and Olson (2004) give details on the estimation of observational error in this case. The second quantity is the celerity (in $\mathrm{m{\cdot}s^{-1}}$), which is straightforwardly computed since the along-track distance is known and the travel time is measured. The travel time is considered to be perfectly measured, yielding exact celerity values, but this need not be the case. In fact, Vera-Rodriguez *et al.* (2020) considered errors in the travel time. Finally, we have the maximum height (in km) retrieved using the ray-tracing routine described before. This maximum height is also considered to be perfect, but this is not necessarily the case. Note that all the observed angles are negative, and all of them have magnitudes equal to or smaller than 10^{-2} rad, so we do not expect nonlinearity in the observation operator to be an issue. The sign indicates that the equivalent cross-wind, as seen by the observations, is positive, that is, with a positive westerly component.

5.2 | Background values from ERA-5 EDA

NORSAR has kindly provided the background cross-winds for this study. They have taken the zonal and meridional winds of the ten-member ERA-5 ensemble data assimilation (EDA) product and interpolated them to produce cross-winds along the great circle connecting Hukkakero and the ARCES site. These are provided at the 137 native

TABLE 2 Observation values used in our experiments for each of the explosions in August 2021

Time of explosion	Back azimuth angle (rad)	Celerity ($m \cdot s^{-1}$)	Maximum height (km)
August 18, 2018, 09:59	−0.0234049	290.474	36.8918
August 19, 2018, 10:59	−0.0356222	289.066	36.9371
August 20, 2018, 11:29	−0.0164235	285.834	37.4525
August 21, 2018, 11:00	−0.0076969	284.019	37.8886
August 22, 2018, 10:59	−0.0181689	280.019	41.6721
August 23, 2018, 12:59	−0.0094422	287.672	36.8179
August 24, 2018, 13:00	−0.0268955	280.019	46.8186
August 26, 2018, 14:00	−0.0391128	297.229	33.2536
August 27, 2018, 09:59	−0.0181689	291.896	35.3863

Note: The back azimuth angle is measured directly within some error, the celerity is computed from knowing the along-track distance and using the observed travel time, and the maximum height is retrieved using a ray-tracing routine.

vertical levels of ERA-5 (which are not equidistant), and at four equidistant points starting at the source and finishing at the receiver. The horizontal distance between points corresponds to about 60 km in the horizontal direction, which is very close to the native resolution of the ERA-5 EDA product of about 62 km.

In Amezcua *et al.* (2020) the vertical resolution was reduced by averaging native vertical levels into equidistant vertical layers. Following the discussion in Section 4 on the effects of averaging, we avoid averaging and perform DA at the native resolution. However, we eliminate levels above 60 km since the separation between levels is too large. We stop at the last level with thickness below 2 km, leaving a total of $N_z = 131$ vertical levels. With the $N_y = 4$ horizontal points, we have a total of 524 state variables.

When matching the background and observation times, the observations do not fall in "valid" reanalysis times. Figure 8 illustrates (in a rather qualitative way) this mismatch. The horizontal axis represents time. In the bottom row, the black dots show the times of four explosions on August 18, 19, 20 and 21, 2018, respectively, and in the upper row, the times for which the reanalysis values are valid. The time step of ERA-5 EDA is 3 hrs. We select and show the two closest times before and the two closest times after the explosion takes place. For all ensemble members, we perform linear interpolation to obtain background values at observation times. Figure 9 shows the background mean (left column) and the background standard deviations (middle columns) for each of the grid points in the 2D domain, for the first three explosions (rows). In each one of the panels we have sketched the trajectory of the wave up to the maximum height reached. Note that the situation is very similar for these days: lower-level westerlies and upper-level easterlies. The magnitudes of the cross wind, however, vary considerably, and on occasions can exceed 30 $m \cdot s^{-1}$. There is little variation along the horizontal direction. This is not surprising given that we are representing variations of a synoptic-scale variable over a rather short distance. In terms of the standard deviation, we see that it often reaches 1 $m \cdot s^{-1}$ in the highest regions of the domain, especially above 50 km. Nonetheless, there are other instances where high standard deviations are also present at lower levels. Once more, the variation with height is much more pronounced than the variation in the along-track direction. The right column of 9 shows the vertical sensitivity obtained by Blixt *et al.* (2019) using the ray-tracing routine, after multiplying by the length of the trajectory at each height, and normalising. This is considered to be dependent only on height, but in general it could depend on along-track direction and hence be different for the ascending and descending legs of the ray. The sensitivity is larger at lower levels, but the uncertainty is larger at higher levels. Therefore, it is not trivial to infer whcih regions will be more impacted by observations.

Having only $N_e = 10$ ensemble members for $N_x = 524$ state variables means that the sample covariance matrix likely suffers from considerable sampling errors. As explained in Section 3, the sample covariance matrix can be modified to eliminate spurious elements with localisation. We construct the tampered matrix $\mathbf{P}^b_{loc}$ using the decomposition in Equation (14), that is, treating variances and correlations separately. For the variances, we use those corresponding to the observation times. These values are shown in the middle column of Figure 9. The instantaneous correlation matrices for the first three explosions are shown in the left column of Figure 10. These are block matrices. The four blocks in the main block-diagonal correspond to the vertical correlation matrices for each one of the $N_a = 4$ across-track sections. The blocks outside the main block-diagonal are

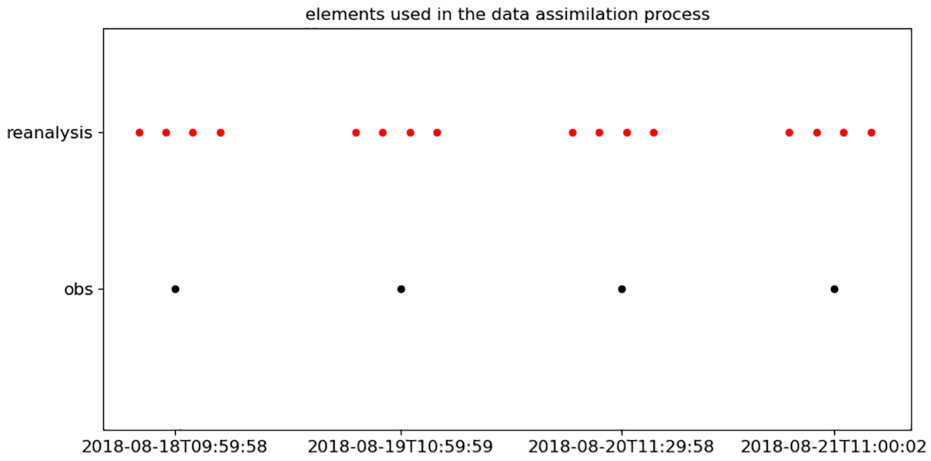

FIGURE 8 Timings of the ERA-5 EDA reanalysis values (red dots) and the explosions (black dots) for five detonations in the summer 2018. The reanalysis time series is interpolated to the explosion time to get the background mean and variance fields for the assimilation. The four closest times to a given observation are used to create a correlation matrix with low sampling noise [Colour figure can be viewed at wileyonlinelibrary.com]

correlations between along-track points at all different vertical levels. As expected, these matrices are noisy. To improve them, we do two things. First, we take the four times of ERA-5 EDA closest to the explosion time (as in Figure 8), and for each of them we compute the sample correlation matrix. Then, we take the average of these four correlation matrices, and consider this correlation matrix to be valid at the observation time. Basically we assume that the spatial structures quantified by the correlations evolve slowly, and the average structures are fixed within a window of 9 hrs centred roughly at the explosion time. The time-averaged correlation matrices are shown in the second column of Figure 9 for the first three explosions. These correlation matrices are considerably less noisy than the instantaneous one. Some features of the instantaneous matrices have been accentuated, like the positive elements on the main diagonals of each of the blocks. Being conservative and doubting the elements with small magnitude, we apply the ECORAP step described in Section 3. We choose the power $k = 2$. The matrices obtained in this way are shown in the third column of Figure 10. There are features that still resemble those appearing in the instantaneous matrices, but these tapered and averaged matrices are much cleaner. Also, note that there are features that definitely vary from one explosion to another, so these matrices do contain flow-dependent information. These matrices are used for the DA experiments.

As a final comparison, the panels in the right column of Figure 10 show histograms of the values of the elements in the different matrices: instantaneous correlations (black), averaged correlations (grey) and ECORAP averaged correlations (green). In all cases there are almost no negative values, pointing to the fact that in these experiments the majority of the increments will have the same sign. As different steps are performed, more correlations of intermediate value are transformed into values closer to zero. However, notice that the frequencies for correlation values of around 0.85 and larger do not change much, suggesting that we are not eliminating strong correlations (likely coming from dynamics) when trying to eliminate the sampling noise.

5.3 | Results

Due to the absence of a proper verification, these results should be considered illustrative and a step towards a more systematic and controlled assimilation. Figure 11 shows the observation space for the assimilation process corresponding to August 26, 2018. We have three groups of elements: blue dots corresponding to $y^b_{n_e}$, that is, the mapping into observation space of each one of the background ensemble members, a red dot and line representing the observed angle and its standard deviation, and pink dots corresponding to $y^a_{n_e}$, that is, the mapping into observation space of the analysis members. We have included two sets of pink points (different shades), corresponding to the analysis coming from instantaneous correlations and those corresponding to the analysis which uses the ECORAP time-averaged correlations. As we can see, the difference in observation space is barely noticeable. We chose this date because it is the one with the largest different between the background angles and the observed one, and even in this case the values are remarkably close, within 0.02 rad. The analysis ensemble members fall closer to the observation, and the spread is reduced.

Figure 12 shows the results for the first three explosions, on 18, 19 and 20 August. This figure has three columns. The first column shows, for each ensemble member, the difference between the given observation and the corresponding value coming from background and analysis; This is, it shows $y^o - y^b_{n_e}$ in blue and $y^o - y^a_{n_e}$ in red. From now on we only discuss the analysis coming from using the localised correlation matrices (using instantaneous correlations leads to very noise fields in model space). For each and every one of the ensemble

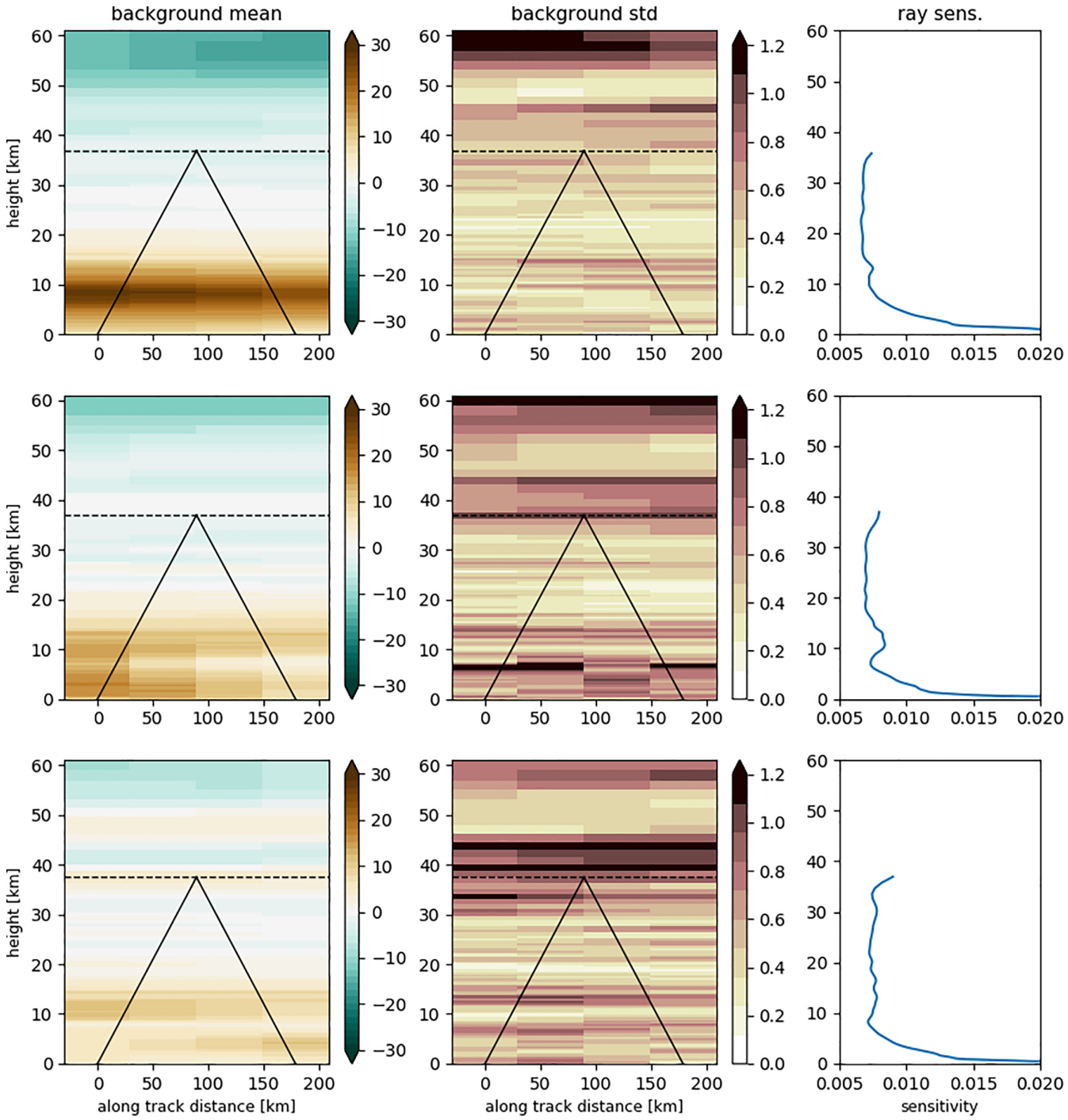

FIGURE 9 Cross-wind background mean (left panel) and background standard deviations (middle panel) for the explosions on August 18, 19 and 20, 2018. In each panel we show the guessed trajectory of the infrasound wave. The right column shows the vertical sensitivities needed for the observation operator. These depend on height [Colour figure can be viewed at wileyonlinelibrary.com]

members, $|y^o - y^a_{n_e}| < |y^o - y^b_{n_e}|$, which implies that the assimilation is achieving its purpose. Note that, for different rows, the vertical axis changes to show positive values (first row) and negative values (second and third row). In all our experiments, the maximum differences between $|y^o - y^b_{n_e}|$ are on the order of $10^{-3} - 10^{-2}$ rad.

The middle panel of Figure 12 shows the analysis increments for these explosions. There is strong horizontal correlation in the increments, meaning that similar increment often happen in all horizontal sections in a given vertical level, with small variations in magnitude. There are very few cases, however, where increments of different sign happen in the same vertical level. The top and bottom row show the typical behaviour we found in the DA process: analysis increments of around 10^{-1} m·s^{-1}, also coinciding with the magnitude found in

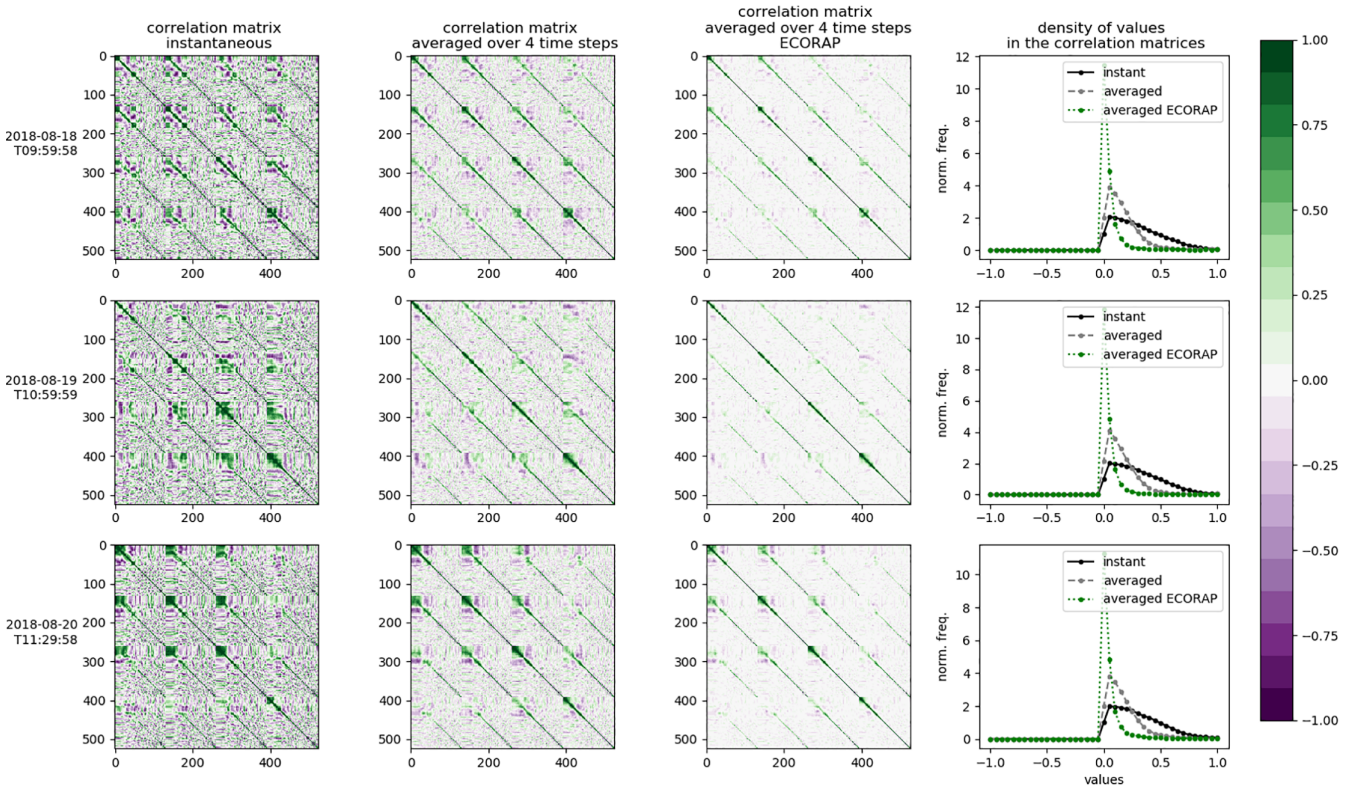

FIGURE 10 Sample correlation matrices computed from the ERA-5 EDA reanalysis for a grid of $N_a = 4$ horizontal points, and the lowest $N_z = 131$ vertical levels of the native resolution. For three times (rows), we show the instantaneous correlation matrix (left column), the averaged correlation matrix from the nearest four points in time (middle) and the averaged correlation matrix after being post-processed using ECORAP (right). The rightmost column shows the histogram for the values of the elements of the three correlation matrices [Colour figure can be viewed at wileyonlinelibrary.com]

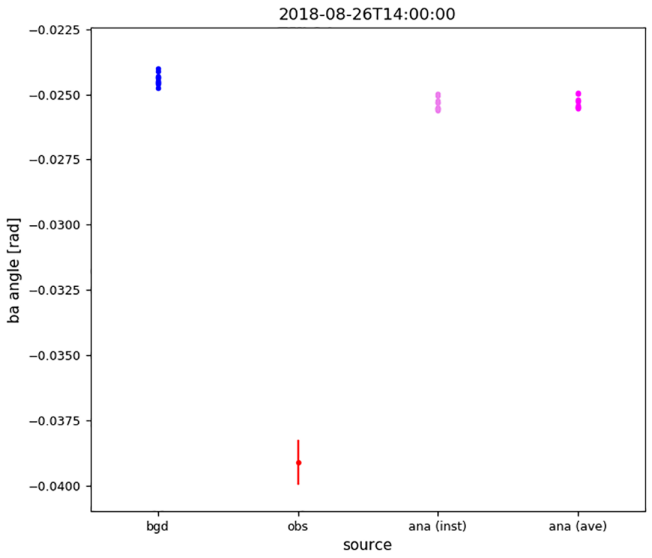

FIGURE 11 Equivalent angles in observation space for the explosion on 26 August. This date was chosen since it is one of the cases with largest difference between the background angles (blue dots, one for each ensemble member) and the observed angle (red). In two shades of pink we show the corresponding analysis angles, obtained when using the instantaneous correlation matrix and the post-processed one [Colour figure can be viewed at wileyonlinelibrary.com]

Amezcua *et al.* (2020), even when that work used averaged data. The middle row, however, shows a case in which we found unexpectedly large increments, with magnitudes frequently up to 1.2–1.5 $\mathrm{m \cdot s^{-1}}$. This was one of the cases with the largest differences between equivalent background angles and the observed angle. One requires increments verification to see if these increments are right. Nonetheless, cases like this stress the need for quality control of observations, which is routinely done for conventional observations in operational DA (see e.g., Chen and Kalnay, 2020).

Finally, the right panel of the figure shows the variance ratio. We see localised reductions in the variance, sometimes of 0.9 in the three explosions. There are large regions where the ratio is close to 1, which means the reduction in variance is negligible. This was also found in Amezcua *et al.* (2020), and considering what we observed in observation space (Figure 11), it is expected.

We end this section with a summary provided in Figure 13. We plot the analysis increments for the four along-track grid points, corresponding to each one of the panels from the leftmost (centred at the emission point) to rightmost (centred at the reception point). In each

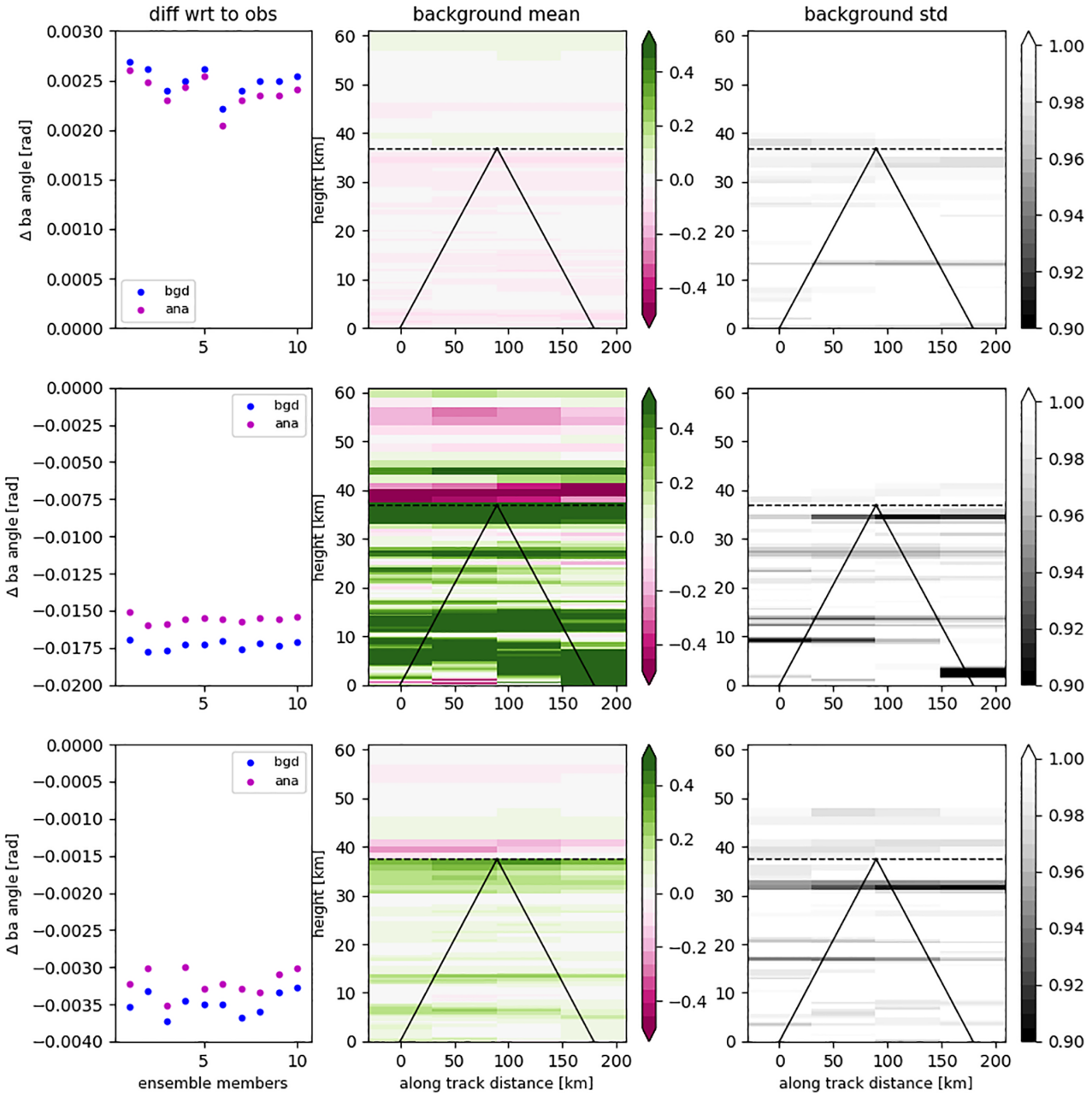

FIGURE 12 Results of the DA process for the explosions on August 18, 19 and 20, 2018. The left column shows the difference between the observation and background angles (blue) and the observation and analysis angles (magenta), for each of the ensemble members. The middle column shows the analysis increments for different grid points, and the right column shows the value of the variance ratio. In these two columns we show the guessed trajectory of the wave [Colour figure can be viewed at wileyonlinelibrary.com]

panel, we plot the analysis increments for the $N_z = 131$ vertical levels. There are several lines in each panel. The blue vertical line indicates the value of zero, and the grey dotted thin lines correspond to the profiles for each one of the nine explosions. The thick black line is the average analysis increment, and the red line is the median. These two summary statistics are computed over the nine explosions. We note a lot of variability over the nine explosions and height, although for the four horizontal sections the analysis increments tend to be positive. When taking the mean, the increments do tend to be positive as well, with especially large values around 5–15 km and 30–45 km in height. However, it seems that most of the contributions to these mean values come from outliers (maximum absolute values of around 1.5 $m \cdot s^{-1}$), including the explosions of 19 and 26 August. The median is therefore a more robust metric to asses the impact. The median is closer to zero for all vertical levels, but there are noticeable non-zero regions, especially in the 30–45 km region.

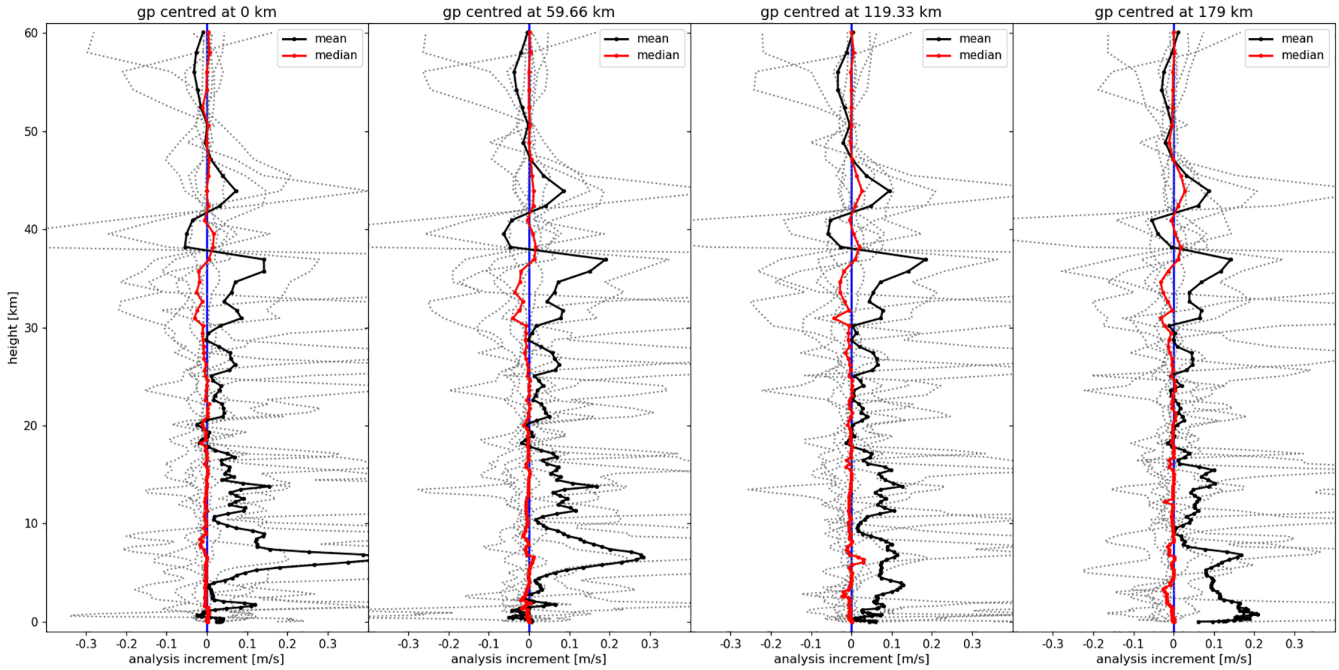

FIGURE 13 Summary of the analysis increments obtained for the nine explosions from 2018 for the $N_a = 4$ horizontal points (panels) and the $N_z = 131$ vertical levels (vertical axis in each panel). We show results for the individual explosions by grey dotted lines, the average increments over time in solid black and the median in solid red [Colour figure can be viewed at wileyonlinelibrary.com]

6 | SUMMARY AND FUTURE WORK

When an infrasound wave travels through the atmosphere, it is affected by the atmospheric conditions it encounters. The detection of these waves can provide useful information to help constrain values of atmospheric variables. This information, however, is not available in a straightforward manner but rather as an integrated quantity. A single change in back azimuth angle can be the result of non-unique combinations of cross-winds in different sections along the wave propagation. This makes inversion a practically impossible problem. Bayesian estimation, and in our case Kalman-filter-based data assimilation, can help solve this problem by using prior (or background) information. This problem is not too different, for instance, from that of retrieving satellite observations, which also measure integrated quantities (e.g., Lei *et al.*, 2015). The information coming from infrasound waves can be helpful for better estimating the winds in the stratosphere and mesosphere, where some of the largest uncertainties in models exist, and where observations are less frequent and less dense than in the troposphere.

In this article we have performed the following: First we explained and illustrated how observed back azimuth angles can be assimilated (using the Kalman filter and its ensemble implementation) to generate updates to a 2D cross-wind field. We illustrated this with synthetic experiments and different options for the background elements. Since it is sometimes necessary to summarize the input background fields, we explored how this averaging (either along-track or vertically) affects the results. We illustrated how this is a source of model error and can affect the results, particularly in the case of non-diagonal background correlations and for particular configurations of the background mean field.

We then moved to using real data for both background and observations. Expanding on the work of Amezcua *et al.* (2020), we used as background the ERA-5 EDA ten-member wind interpolated to the line joining the emission and reception sites of data collected from controlled explosions at Hukkaeiro. We use the native resolution of the reanalysis data in a 2D setting, about 60 km in the horizontal, and the lowest $N_z = 131$ (out of 137) vertical levels. Having only $N_e = 10$ ensemble members to represent covariance for $N_x = 524$ variables leads to very noisy estimators. To ameliorate this problem we did two things. We started by separating the covariance into correlation structures and standard deviations. The standard deviations come from the ERA-10 EDA members interpolated to the observation time (the same being true for the mean). For the correlations, we took the four reanalysis times closest to the observations, obtained the correlation matrix for each of those times, then took an average. Basically, we assume that the correlation structures evolve slowly over a 9-hr period centred roughly at the time of the

observation. To reduce the remaining small correlations, we used the method of ensemble correlation raised to a power. The results of the experiments with real data show that increments of the same sign often occur at all along-track points for a given vertical levels, with varying magnitude. In very few cases, we find analysis increments of different signs for a given vertical level. The variability over the vertical direction is much larger. The variance reduction is small and localised at some grid points, rarely yielding a ratio below 0.95. Finally, we provided a summary of the analysis increments for all horizontal points and all vertical levels. Analysis increments tend to be positive, and the mean analysis increment (over the nine explosions) is large in two regions: 5–15 km and 30–45 km. However, it seems that the magnitude of the mean is largely affected by two explosions (19 and 26 August) in which the observed angle differed from the analysis angles by around 10^{-2} rad, whereas for the other explosions this difference was close to 10^{-3}. We took the median analysis increment (again over the nine explosions), which resulted closer to zero, except for the region 30–45 km.

We identify four ways to improve this study. The first and obvious one is that we need to validate the DA results against independent observations, that is, those which have not been assimilated. Furthermore, in this study we have shown that quality control of the back azimuth angle may be necessary, since we found some cases with anomalously large analysis increments. A plausible source for independent observations for upper regions of the atmosphere are the wind lidars in the AEOLOUS satellite, launched in 2018 by the European Space Agency. The wind estimates from this instrument, based on work by Tan *et al.* (2008), have recently been validated by Witschas *et al.* (2020) and Khaykin *et al.* (2020). Johan Kero, Daniel Bowman and Joseph Eli Bird are working on infrasound transmission in the shadow zone observed on balloons in the lower stratosphere (Kero, personal communication). We are looking into using the winds detected in said balloons for validation.

The next two are methodological improvements which are applicable even in this situation where the distance travelled by the wave is relatively small. To begin with, we can expand our scope and use the three observations of Vera-Rodriguez *et al.* (2020), that is, back azimuth, apparent velocity and propagation time with their respective uncertainties. In order to use these observations, we need to relate them to the state variables; that is, we need to write down adequate observation operators. The aim of the problem would grow from only estimating cross-wind, to estimating tail-wind and temperatures as well for every grid point. In some cases these variables can be related in very nonlinear ways, as for instance in Equation (7). Though challenging, this is a feasible target for the type of method and setup we have used in this article. The third issue has to do with the way we have used the observation operator, in particular the weights in Equation (12) needed to compute the equivalent wind field. In a way, we are treating these weights as retrievals, since they are pre-processed; see for example Migliorini (2012) for the concept of assimilating retrievals versus raw observations and observation operators in the context of satellite DA. In our case, we are using a retrieved observation operator. Namely, the weights in the observation operator came from the vertical sensitivities generated by Blixt *et al.* (2019) using the ERA-Interim wind fields (a single-trajectory estimator) for the different explosions in the study. We used these weights for all the ensemble members coming from ERA-5. Being strict, the correct way to proceed with the problem is to perform the ray-tracing technique independently for each ensemble member. These would lead to different weights for each ensemble member.

The final challenge is one that will inevitably arise when studying transmission over longer length scales, on the order of thousands of kilometres. For instance, an infrasound wave travelling 2,000 km will roughly take almost 2 hrs to reach the detector from the source. This is over 3 hrs. Such an event can no longer be considered instantaneous with respect to synoptic variations, so the background wind field can no longer be considered static. How to introduce this time variation is the next goal.

AUTHOR CONTRIBUTIONS

Javier Amezcua: conceptualization; methodology; supervision; writing – review and editing. **Zak Barton:** investigation; validation; writing – original draft.

ACKNOWLEDGEMENTS

The thorough and diligent review of this work by two anonymous reviewers helped improve the manuscript substantially from its original form. Useful and interesting conversations with Peter Nasholm and his group in NORSAR also contributed to this article. The work was supported by the UK National Centre for Earth Observation (NCEO). This study was facilitated by previous research performed within the framework of the ARISE and ARISE2 projects (Blanc *et al.*, 2018; 2019) funded by the European Commission FP7 and Horizon 2020 programmes (grant agreements 284387 and 653980).

CONFLICT OF INTERESTS

The authors declare no conflicts of interest.

ORCID

Javier Amezcua https://orcid.org/0000-0002-4952-8354

REFERENCES

Amezcua, J., Nasholm, P., Blixt, M. and Charlton-Perez, A. (2020) Assimilation of atmospheric infrasound data to constrain tropospheric and stratospheric winds. *Quarterly Journal of the Royal Meteorological Society*. ISSN 1477-870X.

Asch, M., Bocquet, M. and Nodet, M. (2016) *Data Assimilation*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Assink, J., Smets, P., Marcillo, O., Weemstra, C., Lalande, J.-M., Waxler, R. and Evers, L. (2019). Advances in infrasonic remote sensing methods. In A. Le Pichon, E. Blanc, and A. Hauchecorne (Eds.), *Infrasound Monitoring for Atmospheric Studies*, pp. 605–632. Berlin: Springer.

Baker, W.E., Atlas, R., Cardinali, C., Clement, A., Emmitt, G.D., Gentry, B.M., Hardesty, R.M., Källén, E., Kavaya, M.J., Langland, R., Ma, Z., Masutani, M., McCarty, W., Pierce, R.B., Pu, Z., Riishojgaard, L.P., Ryan, J., Tucker, S., Weissmann, M. and Yoe, J.G. (2014) Lidar-measured wind profiles: the missing link in the global observing system. *Bulletin of the American Meteorological Society*, 95(4), 543–564.

Bishop, C.H. and Hodyss, D. (2007) Flow adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 133, 2029–2044.

Bishop, C. and Hodyss, D. (2009) Ensemble covariances adaptively localized with ECO-RAP. Part 1: tests on simple error models. *Tellus A: Dynamic Meteorology and Oceanography*, 61(1), 84–96. https://doi.org/10.1111/j.1600-0870.2007.00371.x

Blanc, E., Ceranna, L., Hauchecorne, A., Charlton-Perez, A., Marchetti, E., Evers, L.G., Kvaerna, T., Lastovicka, J., Eliasson, L., Crosby, N.B., Blanc-Benon, P., Le Pichon, A., Brachet, N., Pilger, C., Keckhut, P., Assink, J.D., Smets, P.S.M., Lee, C.F., Kero, J., Sindelarova, T., Kämpfer, N., Rüfenacht, R., Farges, T., Millet, C., Näsholm, S.P., Gibbons, S.J., Espy, P.J., Hibbins, R.E., Heinrich, P., Ripepe, M., Khaykin, S., Mze, N. and Chum, J. (2018) Toward an improved representation of middle atmospheric dynamics thanks to the ARISE project. *Surveys in Geophysics*, 39(2), 171–225.

Blanc, E., Pol, K., Le Pichon, A., Hauchecorne, A., Keckhut, P., Baumgarten, G., Hildebrand, J., Höffner, J., Stober, G., Hibbins, R., Espy, P., Rapp, M., Kaifler, B., Ceranna, L., Hupe, P., Hagen, J., Rüfenacht, R., Kämpfer, N. and Smets, P. (2019). Middle atmosphere variability and model uncertainties as investigated in the framework of the ARISE project. In A. Le Pichon, E. Blanc, and A. Hauchecorne (Eds.), *Infrasound Monitoring for Atmospheric Studies*, pp. 845–887. Berlin: Springer.

Blixt, E.M., Näsholm, S.P., Gibbons, S.J., Evers, L.G., Charlton-Perez, A.J., Orsolini, Y.J. and Kværna, T. (2019) Estimating tropospheric and stratospheric winds using infrasound from explosions. *Journal of the Acoustical Society of America*, 146(2), 973–982.

Burgers, G., Jan van Leeuwen, P. and Evensen, G. (1998) Analysis Scheme in the Ensemble Kalman Filter, Monthly Weather Review, 126, 1719–1724. Retrieved Aug 10, 2021, from https://journals.ametsoc.org/view/journals/mwre/126/6/1520-0493_1998_126_1719_asitek_2.0.co_2.xml.

Charlton-Perez, A.J., Baldwin, M.P., Birner, T., Black, R.X., Butler, A.H., Calvo, N., Davis, N.A., Gerber, E.P., Gillett, N., Hardiman, S., Kim, J., Kruuger, K., Lee, Y.-Y., Manzini, E., McDaniel, B.A., Polvani, L., Reichler, T., Shaw, T.A., Sigmond, M., Son, S.W., Toohey, M., Wilcox, K., Yoden, S., Christiansen, B., Lott, F., Shindell, D., Yukimoto, S. and Watanabe, S. (2013) On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models. *Journal of Geophysical Research: Atmospheres*, 118(6), 2494–2505. https://doi.org/10.1002/jgrd.50125.

Chen, T. and Kalnay, E. (2020) Proactive quality control: observing system experiments using the NCEP global forecast system. *Monthly Weather Review*, 148(9), 3911–3931. https://doi.org/10.1175/MWR-D-20-0001.1.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Ka·llberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597. https://doi.org/10.1002/qj.828.

Diamond, M. (1964) Cross wind effect on sound propagation. *Journal of Applied Meteorology and Climatology*, 3, 208–210.

Donn, W.L. and Rind, D. (1972) Microbaroms and the temperature and wind of the upper atmosphere. *Journal of the Atmospheric Sciences*, 29(1), 156–172.

Drob, D.P., Picone, J.M. and Garces, M.A. (2003) Global morphology of infrasound propagation. *Journal of Geophysical Research: Atmospheres*, 108(D21), 4680. https://doi.org/10.1029/2002JD003307.

Duruisseau, F., Huret, N., Andral, A. and Camy-Peyret, C. (2017) Assessment of the ERA-Interim winds using high-altitude stratospheric balloons. *Journal of the Atmospheric Sciences*, 74(6), 2065–2080. https://doi.org/10.1175/JAS-D-16-0137.1.

Evensen, G. (2009) *The Ensemble Kalman Filter*. Berlin: Springer-Verlag. isbn: 978-3-642-03711-5.

Garces, M.A., Hansen, R.A. and Lindquist, K.G. (1998) Travel times for infrasonic waves propagating in a stratified atmosphere. *Geophysical Journal International*, 135(1), 255–263.

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.

Georges, T.M. (1972) 3D ray tracing for acoustic-gravity waves. *Journal of the Acoustical Society of America*, 51, 147.

Gibbons, S.J., Ringdal, F. and Kværna, T. (2007) Joint seismic-infrasonic processing of recordings from a repeating source of atmospheric explosions. *The Journal of the Acoustical Society of America*, 122, EL158–EL164. https://doi.org/10.1121/1.2784533.

Greybush, S.J., Kalnay, E., Miyoshi, T., Ide, K. and Hunt, B.R. (2011) Balance and Ensemble Kalman Filter Localization Techniques. *Monthly Weather Review*, 139(2), 511–522.

Groves, G. (1956) Introductory theory for upper atmosphere wind and sonic velocity determination by sound propagation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 8(1), 24–38.

Hedlin, M.A.H. and Walker, K.T. (2013) A study of infrasonic anisotropy and multipathing in the atmosphere using seismic networks. *Philosophical Transactions of the Royal Society, A*, 371. https://doi.org/10.1098/rsta.2011.0542.

Hunt, B.R., Kostelich, E.J., Szunyogh, I. (2007) Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, 230, 112–126. https://doi.org/10.1016/j.physd.2006.11.008.

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. https://doi.org/10.1002/qj.3130.

Kalman, R.E. (1960) A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82, 35–45.

Kalman, R.E. and Bucy, R.S. (1961) New results in linear filtering and prediction theory. *Journal of Basic Engineering*, 83, 95–108.

Khaykin, S.M., Hauchecorne, A., Wing, R., Keckhut, P., Godin-Beekmann, S., Porteneuve, J., Mariscal, J.-F. and Schmitt, J. (2020) Doppler lidar at Observatoire de Haute-Provence for wind profiling up to 75km altitude: performance evaluation and observations. *Atmospheric Measurement Techniques*, 13, 1501–1516. https://doi.org/10.5194/amt-13-1501-2020.

Kidston, J., Scaife, A.A., Hardiman, S.C., Mitchell, D.M., Butchart, N., Baldwin, M.P. and Gray, L.J. (2015) Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. *Nature Geoscience*, 8, 433–440.

Korhonen, N., Hyvärinen, O., Kämäräinen, M., Richardson, D.S., Järvinen, H. and Gregow, H. (2019) Adding value to extended-range forecasts in Northern Europe by statistical post-processing using stratospheric observations. *Atmospheric Chemistry and Physics Discussions*, 2019, 1–18. https://doi.org/10.5194/acp-2019-679.

Le Pichon, A., Ceranna, L., Garcés, M., Drob, D. and Millet, C. (2006) On using infrasound from interacting ocean swells for global continuous measurements of winds and temperature in the stratosphere. *Journal of Geophysical Research*, D11. https://doi.org/10.1029/2005JD006690.

Le Pichon, A., Blanc, E. and Hauchecorne, A. (2018) *Infrasound Monitoring for Atmospheric Studies: Challenges in Middle Atmosphere Dynamics and Societal Benefits* (2nd edition). Berlin: Springer.

Lei, L. and Whitaker, J.S. (2015) Model space localization is not always better than observation space localization for assimilation of satellite radiances. *Monthly Weather Review*, 143(10), 3948–3955.

Migliorini, S. (2012) On the equivalence between radiance and retrieval assimilation. *Monthly Weather Review*, 140(1), 258–265. https://doi.org/10.1175/MWR-D-10-05047.1.

Mykkeltveit, S., Ringdal, F., Kvaerna, T. and Alewine, R.W. (1990) Application of regional arrays in seismic verification research. *Bulletin of the Seismological Society of America*, 80(6B), 1777–1800.

Orsolini, Y.J., Kindem, I.T. and Kvamstø, N.G. (2011) On the potential impact of the stratosphere upon seasonal dynamical hindcasts of the North Atlantic Oscillation: a pilot study. *Climate Dynamics*, 36(3-4), 579–588.

Polavarapu, S., Shepherd, T., Rochon, Y. and Ren, S. (2005) Some challenges of middle atmosphere data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 131, 3513–3527.

Rind, D., Donn, W.L. and Dede, E. (1973) Upper air wind speeds calculated from observations of natural infrasound. *Journal of the Atmospheric Sciences*, 30(8), 1726–1729.

Sakov, P. and Bertino, L. (2011) Relation between two common localisation methods for the EnKF. *Computers and Geosciences*, 15, 225–237. https://doi.org/10.1007/s10596-010-9202-6.

Sakov, P. and Oke, P.R. (2008) A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. *Tellus A*, 60(2), 361–371. https://doi.org/10.1111/j.1600-0870.2007.00299.x.

Shlyaeva, A., Whitaker, J.S. and Snyder, C. (2019) Model-space localization in serial ensemble filters. *Journal of Advances in Modeling Earth Systems*, 11, 1627–1636. https://doi.org/10.1029/2018MS001514.

Simmons, A., Soci, C., Nicolas, J., Bell, B., Berrisford, P., Dragani, R., Flemming, J., Haimberger, L., Healy, S., Hersbach, H., Horányi, A., Inness, A., Munoz-Sabater, J., Radu, R. and Schepers, D. (2020). Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5.1. Technical Memorandum 859.

Stuart, A.M. (2010) Inverse problems: a Bayesian perspective. *Acta Numerica*, 19, 451–559. https://doi.org/10.1017/S0962492910000061.

Szuberla, C.A., Olson, J.V. (2004) Uncertainties associated with parameter estimation in atmospheric infrasound arrays. *J Acoust Soc Am*. 115, 253–8. https://doi.org/10.1121/1.1635407.

Tan, D.G.H., Andersson, E., De Kloe, J., Marseille, G.-J., Stoffelen, A., Poli, P., Denneulin, M.-L., Dabas, A., Huber, D., Reitebuch, O., Flamant, P., Rille, L.E. and Nett, H. (2008) The ADM-Aeolus wind retrieval algorithms. *Tellus A*, 60, 191–205. https://doi.org/10.1111/j.1600-0870.2007.00285.x.

van Leeuwen, P.J. (2020) A consistent interpretation of the stochastic version of the Ensemble Kalman Filter. *QJR Meteorol Soc*, 146, 2815–2825. https://doi.org/10.1002/qj.3819.

Vera-Rodriguez, I. (2019) A heuristic-learning optimizer for elastodynamic waveform inversion in passive seismics. *IEEE Transactions on Geoscience and Remote Sensing*, 57, 2234–2248.

Vera-Rodriguez, I., Nasholm, S.P. and Le, P. (2020) A. Atmospheric wind and temperature profiles inversion using infrasound: en ensemble model context. *Journal of the Acoustic Society of America*, 148, 2423. https://doi.org/10.1121/10.0002482.

Witschas, B., Lemmerz, C., Geiß, A., Lux, O., Marksteiner, U., Rahm, S., Reitebuch, O. and Weiler, F. (2020) First validation of Aeolus wind observations by airborne Doppler wind lidar measurements. *Atmospheric Measurement Techniques*, 13, 2381–2396. https://doi.org/10.5194/amt-13-2381-2020.

How to cite this article: Amezcua, J. & Barton, Z. (2021) Assimilating atmospheric infrasound data to constrain atmospheric winds in a two-dimensional grid. *Quarterly Journal of the Royal Meteorological Society*, 1–25. Available from: https://doi.org/10.1002/qj.4141

APPENDIX A. GLOSSARY

This manuscript contains a myriad of terms, symbols and equations. We therefore present a glossary of the most important symbols to aid the reader.

Geometric setup for infrasound propagation	
Symbol	**Meaning**
S, S', R	Source, apparent source, receiver
r^c, r^a	Cross-direction, along-track direction
z_{max}, z_{n_z}	Maximum vertical penetration of the infrasound wave, height of the n_z^{th} vertical level (m)
$d^c, d^c_{n_a}, d^c_{n_z}, d^c_{n_a,n_z}$	Cross-distance: total, for the n_a^{th} along-track section, for the n_z^{th} vertical level, for the (n_a, n_z) box of a 2D grid (m)
$d^a, d^a_{n_a}, d^a_{n_z}, d^a_{n_a,n_z}$	Along-track distance: total, for the n_a^{th} along-track section, for the n_z^{th} vertical level, for the (n_a, n_z) box of a 2D grid (m)
$W^c, w^c_{n_a}, w^c_{n_z}, w^c_{n_a,n_z}$	Cross wind: constant, at the n_a^{th} along-track section, at the n_z^{th} vertical level, at a the (n_a, n_z) coordinate of a 2D grid ($\mathrm{m{\cdot}s^{-1}}$)
$W^a, w^a_{n_a}, w^a_{n_z}, w^a_{n_a,n_z}$	Along-track wind: constant, at the n_a^{th} along-track section, at the n_z^{th} vertical level, at the (n_a, n_z) coordinate of a 2D grid ($\mathrm{m{\cdot}s^{-1}}$)
T, T_{n_a}	Travel time: total, for the n_a^{th} along-track section, for the n_z^{th} vertical level, for the (n_a, n_z) box of a 2D grid (s).
$\alpha_{n_a}, \alpha_{n_z}, \alpha_{n_a,n_z}$	Weight: at the n_a^{th} along-track section, at the n_z^{th} vertical level, at the (n_a, n_z) coordinate of a 2D grid ($\mathrm{m{\cdot}s^{-1}}$)
$\boldsymbol{\alpha} \in R^{N_a N_z}$	Concatenation of the weights for the 2D grid. This becomes a vector.
$\theta, \theta', \Delta\theta$	Back azimuth angle in the absence (' presence) of cross-wind, change in back azimuth angle (rad)
v, C	Celerity, adiabatic speed of sound ($\mathrm{m{\cdot}s^{-1}}$)
Data assimilation	
Symbol	**Meaning**
$\mathbf{x} \in R^{N_x}$	Vector of N_x state variables. In our problem $N_x = N_a N_z$
$\mathbf{y} \in R^{N_y}$	Vector of N_y observations. In our problem, $N_y = 1$
$\boldsymbol{\eta} \in R^{N_y}, \mathbf{R} \in R^{N_y \times N_y}$	Additive observation error and observational error covariance matrix
$\mathbf{H} \in R^{N_x \times N_y}$	Linear observation operator (matrix)
$h : R^{N_x} \to R^{N_y}, \hat{\mathbf{H}} \in R^{N_x \times N_y}$	Nonlinear observation operator and its linearisation
$\boldsymbol{\mu}^b \in R^{N_x}, \overline{\mathbf{x}}^b \in R^{N_x}$	Background mean and its ensemble estimator
$\boldsymbol{\mu}^a \in R^{N_x}, \overline{\mathbf{x}}^a \in R^{N_x}$	Analysis mean and its ensemble estimator
$\mathbf{B} \in R^{N_x \times N_x}, \mathbf{P}^b \in R^{N_x \times N_x}$	Background error covariance matrix and its ensemble estimator
$\mathbf{A} \in R^{N_x \times N_x}, \mathbf{P}^a \in R^{N_x \times N_x}$	Analysis error covariance matrix and its ensemble estimator
$\mathbf{X}^b \in R^{N_x \times N_e}, \mathbf{X}^a \in R^{N_x \times N_e}$	Background and analysis ensemble matrices
$\mathbf{K} \in R^{N_x \times N_y}, \mathbf{K}_e \in R^{N_x \times N_y}$	Exact Kalman gain and its ensemble estimator
$\boldsymbol{\Gamma} \in R^{N_x \times N_y}, \boldsymbol{\Gamma}_e \in R^{N_x \times N_y}$	Total covariance and its ensemble estimator
$\boldsymbol{\Sigma} \in R^{N_x \times N_x} \in R^{N_x \times N_x}$	Diagonal matrix of variances, correlation matrix
$\mathbf{d}^{ob} \in R^{N_y}$	Observation innovations
$\mathbf{d}^{ab} \in R^{N_x}$	Analysis increments
$(r^{ab})^2 \in R^{N_x}$	Ratio of analysis and background variances