RMetS
Royal Meteorological Society

Assessment and tuning of data assimilation systems using passive observations

Gert-Jan Marseille,* Jan Barkmeijer, Siebren de Haan and Wim Verkley

Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands

*Correspondence to: G.-J. Marseille, KNMI, PO Box 201, 3730 AE De Bilt, The Netherlands. E-mail: Gert-Jan.Marseille@knmi.nl

Independent observations can be used to diagnose and tune a data assimilation (DA) system. Analysis increments generally improve the model state near assimilated observations but degrade it further away. High-resolution aircraft observations from Mode-S Enhanced Surveillance (Mode-S EHS) are used as an independent data source to verify increment degradation as a function of distance from assimilated observations. An adaptation of the inherently imperfect gain matrix in DA is proposed such that resulting analyses better fit the independent data source and as such draw model simulations closer to the true atmospheric state. It is found that the structure functions of the background-error covariance matrix of the experimental mesoscale HARMONIE model are appropriate but too much weight is given to observations relative to the model background. The ECMWF model is well tuned with a slight overestimation of temperature information in the upper troposphere.

Finally, a caveat is highlighted when comparing model forecasts from different experiments against observations. It is common practice to use the same observing system both in the analysis and for forecast verification. However, forecast verification is prone to sampling errors, yielding less favourable scores when using an independent data source. Avoidance of biased conclusions on the impact of observing systems, e.g. in observing system experiments (OSE), requires an independent data source (best practice) or a data source used in all experiments (best pragmatic practice) for verification of forecasts from different experiments.

1. Introduction

The goal of data assimilation (DA) in numerical weather prediction (NWP) is to find the best compromise (called analysis) between a model simulation (called background or first-guess) and observations. Model simulations tend to diverge from the true atmospheric state when evolving in time. Observations are used to keep the model on track of the true atmospheric state. However, observations are imperfect too. Tuning of a DA system, i.e. giving correct relative weight to the background and observations in the analysis and spreading the observational information in space, is a continuous challenge in NWP. The model analysis state vector, $\mathbf{x}^a$, is obtained from the analysis equations which in their general form read as (Lorenc, 1986; Daley, 1999)

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}'[\mathbf{y}^o - \mathcal{H}(\mathbf{x}^b)], \quad (1)$$

$$\mathbf{K}' = \mathbf{B}'\mathbf{H}^T(\mathbf{H}\mathbf{B}'\mathbf{H}^T + \mathbf{R}')^{-1}, \quad (2)$$

$$\mathbf{y}^o = \mathcal{H}(\mathbf{x}^t) + \boldsymbol{\epsilon}, \quad (3)$$

with $\mathbf{x}^b$ the background state vector, i.e. a short-term forecast initiated with the previous analysis, $\mathbf{y}^o$ is the vector with observations, $\mathbf{x}^t$ the true (but unknown) state vector and $\mathcal{H}$ the observation operator which maps the model state to the observed quantity. The observation operator may be a simple linear interpolation operator for some observing systems, but may be more complex, e.g. including a nonlinear radiative transfer model for measured satellite radiances. Matrix $\mathbf{H}$ is the partial derivative of $\mathcal{H}$ with respect to the state vector $\mathbf{x}$. The total observation error, $\boldsymbol{\epsilon}$, equals the sum of the instrument error and representativeness error which are discussed in detail below. Superscript T denotes the matrix transpose. The impact of observations on the analysis is determined by the prescribed background-error covariance matrix, denoted below by $\mathbf{B}'$, and the prescribed observation-error covariance matrix, denoted below by $\mathbf{R}'$. If both describe well the true background- and observation-error covariances (denoted without primes), then the resulting gain matrix, $\mathbf{K}'$ in Eq. (2), yields the best compromise, in a statistical sense, of background and observations in the analysis, Eq. (1).

In practice it is not trivial to correctly specify the $\mathbf{B}'$ and $\mathbf{R}'$ matrices. Ideally, $\mathbf{B}'$ should describe the short-term model forecast errors. Many DA systems rely on a climatological $\mathbf{B}'$ matrix as a pragmatic approach and which may be representative for background errors on average but is a poor estimate for each individual meteorological situation, because model forecast

errors are related to the actual weather situation and observation coverage, which both change with time.

The true **R** matrix is composed of errors from (i) instrument noise, (ii) observation representativeness and (iii) mapping from model to observation space, i.e. errors in the observation operator (Lorenc, 1986). For well-characterized instruments, the instrument error statistics are generally well-known with a diagonal covariance matrix for observations with uncorrelated errors but a dense matrix, e.g. for data from satellite sounders (Weston *et al.*, 2014). The representativeness error accounts for spatial scales resolved by observations but not by the NWP model. The smallest spatial scale resolved by a model is called the effective model resolution which differs for different models. Model and observation power density spectra or triple-collocation techniques (Vogelzang *et al.*, 2011) provide estimates for the effective resolution of the model under investigation and resulting observation representativeness error variance. As a rule of thumb, nowadays NWP models' effective resolution equals 7 to 10 times the model grid size (Skamarock, 2004; Vogelzang *et al.*, 2011). As a consequence, representativeness errors of observations separated by less than the model effective resolution are correlated and the corresponding covariance matrix is non-diagonal. Also, the true representativeness error is non-constant and a function of local atmospheric turbulence. Error covariances of the linearized observation operator strongly depend on the observing system and range from interpolation errors to more complex nonlinear errors, e.g. in radiative transfer functions for satellite radiance data. From the above, it is more likely that **R** is a dense matrix in general but, because of imperfect knowledge of these error sources and computational efficiency, **R'** is specified as a diagonal matrix in most operational DA systems, thus neglecting correlation of observation errors. This simplification is generally compensated by inflating the error variances, i.e. by increasing the matrix diagonal elements.

For a well-tuned DA system, **B'** and **R'** are such that the observational information draws the model state toward the true atmospheric state effectively. How can this be verified, given that the true atmospheric state is inherently unknown? Here a new methodology has been developed using independent observations, i.e. observations not used in the analysis (also denoted *passive* observations), in contrast to observations used in the analysis (denoted *active* observations). These passive observations represent the true atmospheric state. However, passive observations are imperfect and their error sources should be taken into account. Statistics of background and analysis departures, i.e. observation minus background or analysis are widely used to diagnose DA systems in general. Section 2 discusses the different characteristics of departure statistics for active and passive observations. Furthermore, analysis increments, i.e. analysis minus background, are expected to become worse with distance from assimilated observations. Passive observations can be used to test this and possibly find the radius of action where observations add positively to the analysis. Experimenting with a real NWP model is computationally expensive. Therefore, a simulation tool has been developed based on the analysis equations, which allows easy and cheap assessment of the sensitivity of the analysis to incorrectly specified **B'** and **R'** matrices. The tool is discussed in section 3 and it is shown that passive observations can be used to tune a DA system. In section 4, high-resolution aircraft observations from Mode-S Enhanced Surveillance (Mode-S EHS; de Haan, 2011) are used as passive observations in a six-week experiment with the non-hydrostatic mesoscale HARMONIE (HIRLAM Aladin Regional Mesoscale Operational NWP In Euromed) model. It is shown that one has to be careful with selecting passive observations for DA tuning. The same is true when assessing model forecast skill, e.g. in OSE, using observations. Naive selection of observations for forecast verification may lead to biased conclusions, as discussed in section 5. Finally, section 6 summarizes the conclusions from the study.

2. Departure statistics

Differences between observations and their model counterparts at the observation locations are called departures. The background departure, $\mathbf{d}^b = \mathbf{y}^o - \mathcal{H}(\mathbf{x}^b)$, denoted in short below as $(o - b)$, is used to correct the model simulation (background) and bring it closer to the true atmospheric state (Eq. (1)). Defining the analysis error $\boldsymbol{\epsilon}^a = \mathbf{x}^a - \mathbf{x}^t$, then from Eq. (1) and rearranging matrices, the analysis-error covariance matrix $\mathbf{A'}$, which is the result of using the gain matrix $\mathbf{K'}$ from Eq. (2), equals

$$\begin{aligned} \mathbf{A'} &:= \ <\boldsymbol{\epsilon}^a \boldsymbol{\epsilon}^{aT}> \\ &= \ [\mathbf{I} - \mathbf{K'H}]\mathbf{B}[\mathbf{I} - \mathbf{K'H}]^T + \mathbf{K'RK'}^T \end{aligned} \quad (4)$$

with $<.>$ denoting the expectation operator and $\mathbf{I}$ the identity matrix. In general it is assumed that the background error, defined as $\boldsymbol{\epsilon}^b = \mathbf{x}^b - \mathbf{x}^t$, and observation error are not correlated, i.e. $<\boldsymbol{\epsilon}^b \boldsymbol{\epsilon}^T> = 0$. $\mathbf{B} := <\boldsymbol{\epsilon}^b \boldsymbol{\epsilon}^{bT}>$ and $\mathbf{R} := <\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T>$ are the true, but generally unknown, background- and observation-error covariance matrices respectively. The expression ':=' denotes 'by definition'. For the unlikely situation of perfect knowledge of $\mathbf{B}$ and $\mathbf{R}$, substituting these in Eq. (2) yields the optimal gain matrix $\mathbf{K}$ which minimizes the analysis-error variance of the analysis-error covariance matrix (Bouttier and Courtier, 2002) and Eq. (4) reduces to

$$\mathbf{A} = [\mathbf{I} - \mathbf{KH}]\mathbf{B}. \quad (5)$$

From Eq. (3) the background departure covariance matrix of actively used observations, $\mathbf{y}^o$, denoted in short $< (o - b)^2 >$, equals

$$\begin{aligned} < (o - b)^2 > &:= <\mathbf{d}^b \mathbf{d}^{bT}> \\ &= <[\mathbf{y}^o - \mathcal{H}(\mathbf{x}^b)][\mathbf{y}^o - \mathcal{H}(\mathbf{x}^b)]^T> \\ &= <[\boldsymbol{\epsilon} - \mathbf{H}\boldsymbol{\epsilon}^b][\boldsymbol{\epsilon} - \mathbf{H}\boldsymbol{\epsilon}^b]^T> \\ &= \mathbf{HBH}^T + \mathbf{R}, \end{aligned} \quad (6)$$

with $\mathbf{H}$ the partial derivative of $\mathcal{H}$ with respect to model state $\mathbf{x}$ at $\mathbf{x}^t$ and assuming linearity: $\mathcal{H}(\mathbf{x}^b) = \mathcal{H}(\mathbf{x}^t) + \mathbf{H}\boldsymbol{\epsilon}^b$. From Eq. (1) we can write for the analysis departure of actively used observations, and assuming linearity,

$$\mathbf{d}^a = \mathbf{y}^o - \mathcal{H}(\mathbf{x}^a) = [\mathbf{I} - \mathbf{HK'}]\mathbf{d}^b. \quad (7)$$

Analysis departures are also denoted in short as $(o - a)$ below with covariance matrix

$$\begin{aligned} < (o - a)^2 > &:= <\mathbf{d}^a \mathbf{d}^{aT}> \\ &= <[\mathbf{y}^o - \mathcal{H}(\mathbf{x}^a)][\mathbf{y}^o - \mathcal{H}(\mathbf{x}^a)]^T> \\ &= <[-\mathbf{H}\boldsymbol{\epsilon}^a + \boldsymbol{\epsilon}][-\mathbf{H}\boldsymbol{\epsilon}^a + \boldsymbol{\epsilon}]^T> \quad (8) \\ &\overset{(7)}{=} [\mathbf{I} - \mathbf{HK'}] <\mathbf{d}^b \mathbf{d}^{bT}> [\mathbf{I} - \mathbf{HK'}]^T, \quad (9) \end{aligned}$$

where it is noted that $\boldsymbol{\epsilon}^a$ and $\boldsymbol{\epsilon}$ are not independent in Eq. (8), hence the covariance matrix of the analysis departures does *not* equal the sum of the analysis-error covariance and observation-error covariance for *actively* used observations in DA. For non-used (passive) observations in DA, denoted $\mathbf{y}_p^o$, similar relations hold, but now with subscript p: $\mathbf{y}_p^o = \mathcal{H}_p(\mathbf{x}^t) + \boldsymbol{\epsilon}_p$ with observation-error covariance $<\boldsymbol{\epsilon}_p \boldsymbol{\epsilon}_p^T> = \mathbf{R}_p$. The covariance of background departures, $\mathbf{d}_p^b = \mathbf{y}_p^o - \mathcal{H}_p(\mathbf{x}^b)$, is written

$$\begin{aligned} < (o_p - b)^2 > &:= <\mathbf{d}_p^b \mathbf{d}_p^{bT}> \\ &= <[\mathbf{y}_p^o - \mathcal{H}_p(\mathbf{x}^b)][\mathbf{y}_p^o - \mathcal{H}_p(\mathbf{x}^b)]^T> \\ &= <[-\mathbf{H}_p\boldsymbol{\epsilon}^b + \boldsymbol{\epsilon}_p][-\mathbf{H}_p\boldsymbol{\epsilon}^b + \boldsymbol{\epsilon}_p]^T> \\ &= \mathbf{H}_p\mathbf{B}\mathbf{H}_p^T + \mathbf{R}_p \end{aligned} \quad (10)$$

with $\mathbf{H}_\text{p}$ the partial derivative of $\mathcal{H}_\text{p}$ with respect to model state $\mathbf{x}$ at $\mathbf{x}^\text{t}$. Similarly the covariance matrix of analysis departures, $\mathbf{d}_\text{p}^\text{a} = \mathbf{y}_\text{p}^\text{o} - \mathcal{H}_\text{p}(\mathbf{x}^\text{a})$, is written

$$
\begin{aligned}
< (o_\text{p} - a)^2 > & := < \mathbf{d}_\text{p}^\text{a} \mathbf{d}_\text{p}^{\text{a}\,\text{T}} > \\
& = < [\mathbf{y}_\text{p}^\text{o} - \mathcal{H}_\text{p}(\mathbf{x}^\text{a})][\mathbf{y}_\text{p}^\text{o} - \mathcal{H}_\text{p}(\mathbf{x}^\text{a})]^\text{T} > \\
& = < [-\mathbf{H}_\text{p}\boldsymbol{\epsilon}^\text{a} + \boldsymbol{\epsilon}_\text{p}][-\mathbf{H}_\text{p}\boldsymbol{\epsilon}^\text{a} + \boldsymbol{\epsilon}_\text{p}]^\text{T} > .
\end{aligned}
\tag{11}
$$

Because passive observations have not been used in the analysis it is tempting to assume $< \boldsymbol{\epsilon}^\text{a}\boldsymbol{\epsilon}_\text{p}^\text{T} > = 0$ yielding $\mathbf{H}_\text{p}\mathbf{A}'\mathbf{H}_\text{p}^\text{T} + \mathbf{R}_\text{p}$ from Eq. (11) and using the definition in Eq. (4). However, the assumption of uncorrelated $\boldsymbol{\epsilon}^\text{a}$ and $\boldsymbol{\epsilon}_\text{p}$ is not true in general, because the representativeness error of active observations (as part of $\boldsymbol{\epsilon}$ and thus $\boldsymbol{\epsilon}^\text{a}$) and of passive observations (as part of $\boldsymbol{\epsilon}_\text{p}$) may be correlated depending on their separation in space and time. This is further explained in Appendix A. Here, only the final result of Eq. (11) is presented

$$
< (o_\text{p} - a)^2 > = \mathbf{H}_\text{p}\mathbf{A}'\mathbf{H}_\text{p}^\text{T} + \mathbf{R}_\text{p} - (\mathbf{X} + \mathbf{X}^\text{T}),
\tag{12}
$$

with $\mathbf{X}$ denoting the matrix, defined in Appendix A, including the cross-covariance of the representativeness errors of active and passive observations. In general, the analysis departure is larger for passive than active observations, i.e. the covariance from Eq. (12) is larger than from Eq. (9). This is further explored in the next section using a simulation tool, based on the analysis equations.

3. Simulation tool

A simulation tool has been developed to simulate the impact of observations in the analysis. The tool enables us to test the sensitivity to imperfectly specified background- and observation-error covariances on the analysis by calculating the analysis-error covariance from Eq. (4), given the true and prescribed background- and observation-error covariance matrices. For simplicity, the tool is one-dimensional. Observation sampling is equidistant along a track of given length L km and at a constant height level, thus simulating observations along (e.g.) an aircraft or satellite track. Observations are separated by the sampling distance, Δs, yielding $M = L/\Delta s$ observations along the track. The model grid size, Δx, determines the size of the state vector: $N = L/\Delta x$. The model grid definition and observation locations determine the observation operator $\mathbf{H}$ which is simply an $M \times N$ interpolation matrix in our simulations.

In the simulations discussed below, we fix L at 5000 km and Δx at 2.5 km, giving a state vector of dimension $N = 2000$ and thus background- and analysis-error covariance matrices of size 2000×2000. The choice for a 2.5 km grid is inspired by the HARMONIE model discussed in section 4. The sampling distance and thus the size of the observation vector and observation operator matrix is not fixed and part of the evaluation. The true background-error correlation, in the literature also denoted structure function, is assumed a Gaussian function with error standard deviation σ_B and length-scale (1 sigma) L_B making the true $\mathbf{B}$ a band-matrix. This definition for correlation length-scale is in agreement with Daley (1999). The true $\mathbf{R}$ matrix is the sum of a diagonal matrix, representing uncorrelated instrument errors with error standard deviation σ_I, and a non-diagonal matrix for the representativeness error with standard deviation σ_R. For the correlation length-scale of the latter, we take a Gaussian function with length-scale (1 sigma) $L_R = 25$ km, representing the smallest spatial scales the model can resolve. Calculations of spectra (Vogelzang et al., 2011) for the HARMONIE model show a strong drop relative to the theoretical $k^{-5/3}$ spectrum, from three-dimensional turbulence theory, starting at wavenumber, k, corresponding to 25 km spatial scales (not shown), i.e. 10 times the model grid size, slightly larger than the rule-of-thumb value of 7. The elements of the $N \times N$ true background- and $M \times M$

true observation-error covariance matrices, with indices i and j, are

$$
\mathbf{B}(i,j) = \sigma_B^2 \exp[-\{(i-j)\Delta x\}^2 / 2L_B^2],
\tag{13}
$$

$$
\mathbf{R}(i,j) = \sigma_I^2 \mathbf{I}(i,j) + \sigma_R^2 \exp[-\{(i-j)\Delta s\}^2 / 2L_R^2].
\tag{14}
$$

Both $\mathbf{B}$ and $\mathbf{R}$ are constructed such that they are positive definite matrices.

3.1. Atmospheric state and observation simulation

The background state is defined as an N-dimensional vector sum of the true atmospheric state and the background error: $\mathbf{x}^\text{b} = \mathbf{x}^\text{t} + \boldsymbol{\epsilon}^\text{b}$, with the covariance of $\boldsymbol{\epsilon}^\text{b}$ equal to $\mathbf{B}$ defined in Eq. (13). To simulate $\boldsymbol{\epsilon}^\text{b}$, a N-dimensional vector composed of random numbers from a Gaussian distribution with zero mean and standard deviation 1 is pre-multiplied by $\mathbf{B}^{1/2}$. From matrix algebra, the latter is obtained from an eigenvalue decomposition* $\mathbf{B} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\text{T}$, with $\mathbf{Q}$ the orthogonal matrix of singular vectors and $\boldsymbol{\Lambda}$ the diagonal matrix of strictly positive eigenvalues. Then $\mathbf{B}^{1/2} = \mathbf{Q}\boldsymbol{\Lambda}^{1/2}\mathbf{Q}^\text{T}$. Similarly observation errors are simulated from the $\mathbf{R}$ matrix Eq. (14). Table 1 shows the parameter values used to define the error matrices. These reflect the typical settings in the HARMONIE mesoscale model, further discussed in section 4.

$\mathbf{B}'$ for HARMONIE is constructed from downscaling members from the ECMWF Ensemble Prediction System (EPS) to generate a HARMONIE ensemble. Currently four members are thought to be sufficient with a downscaling performed over a period of typically one month. The ensemble spread is reflected in the resulting $\mathbf{B}'$ matrix. Some diagnostic tools are available to assess the overall correctness of $\mathbf{B}'$. These include plots of the background-error standard deviation, power density spectra and vertical correlation of temperature, humidity, divergence and vorticity as a function of pressure level (Brousseau et al., 2011). It is known that the ensemble spread is generally underdispersive, i.e. underestimating the real atmospheric spread. For each member, downscaling is typically performed over 6 h. It was found that this time span is too short for a mesoscale model to fully develop energy on the small scales. The resulting $\mathbf{B}'$ matrix will thus reflect better the model error large scales and its amplitude will underestimate the true error. This motivates the larger correlation length-scale used in $\mathbf{B}'$ and the smaller error standard deviation as compared to the truth.

In section 4.1, we consider AMDAR (Aircraft Meteorological DAta Relay) winds for active observations and aircraft Mode-S EHS winds for passive observations. Table 1 values reflect the corresponding error characteristics.

Setting $\mathbf{x}^\text{t} = 0$ for simplicity, but without loss of generality, 1000 realizations of $\mathbf{x}^\text{b}$ and $\mathbf{y}^\text{o}$ were simulated using the above equation for $\mathbf{x}^\text{b}$, the values in Table 1 and the analysis equations, Eqs (1)–(3). For each realization of $\mathbf{x}^\text{b}$ and $\mathbf{y}^\text{o}$, the analysis $\mathbf{x}^\text{a}$ is obtained from Eq. (1). This requires the gain matrix $\mathbf{K}'$ which is obtained from the prescribed background- and observation-error covariance matrices $\mathbf{B}'$ and $\mathbf{R}'$ respectively from Eq. (2). These latter two matrices deviate substantially from their true counterparts. $\mathbf{B}$ and $\mathbf{B}'$ follow from Eq. (13) with corresponding parameters values taken from Table 1. The true $\mathbf{R}$ matrix is calculated from Eq. (14) given the values in Table 1 for active observations. Following operational practice, $\mathbf{R}'$ is taken as a diagonal matrix with values representing the sum of the instrument error and representativeness error variances. In the case of expected error correlation for nearby observations, an inflation factor f_i is applied to the matrix diagonal elements. As a result $\mathbf{R}' = \sigma_{R'}^2 \mathbf{I}$ with $\sigma_{R'}^2 = f_\text{i}(\sigma_I^2 + \sigma_R^2)$. In our simulation tool, the inflation factor is determined from the ratio of the first off-diagonal element and diagonal element of matrix $\mathbf{R}$. For a ratio

---

*https://en.wikipedia.org/wiki/Eigendecomposition_of_a_matrix; accessed 22 July 2016.

Table 1. Parameter settings for a typical mesoscale model like HARMONIE (discussed in section 4): assumed true state and aircraft observations, including true and specified background- and observation-error standard deviations and length-scales.

Parameter		Unit	Setting
Δx		km	2.5
Δs	Active	km	200
Δs_p	Passive	km	2.5
σ_B	True	–	2.5
L_B	True	km	50
$\sigma_{B'}$	Prescribed	–	2.0
$L_{B'}$	Prescribed	km	250
σ_I	Active/passive	–	1.5
σ_R	Active/passive	–	1.0
L_R	Active/passive	km	25
$\sigma_{R'}$	Prescribed	–	See text
$L_{R'}$	Prescribed	km	0

The parameters are defined and discussed in the text. For σ values, no specific meteorological variable is considered, so the unit is taken as dimensionless.

value < 0.1, f_i is set to 1; between 0.1 and 0.25 it is set to 2.25, and for > 0.25 it is set to 4. From the values in Table 1, the inflation factor equals 2.25. In a simulation model setting, both the gain matrix based on prescribed errors and the optimal gain matrix can be calculated from Eq. (2). The corresponding suboptimal and optimal analysis-error covariance matrices are obtained from Eqs (4) and (5) respectively.

3.2. Spatial consistency of the gain matrix

The gain matrix in DA spreads the observational information in the model domain. The radius of action is largely determined by the correlation length-scale of $\mathbf{B}'$. One can ask whether the impact on the analysis away from locations of assimilated observations is as good as nearby. This was assessed by simulating passive observations with characteristics specified in Table 1, i.e. high-resolution sampling, Δs_p, with an observation on each model grid point and with identical error characteristics as the actively used observations. Given the high density of passive observations, departure statistics for passive observations $\mathbf{d}_p^b$ and $\mathbf{d}_p^a$ can be calculated as a function of distance d from assimilated (active) observations. Here $\mathbf{d}_p^b(d) = \mathbf{D}\mathbf{d}_p^b$ is defined with $\mathbf{D}$ a matrix of zeros and ones filtering those elements of vector $\mathbf{d}_p^b$ for which the location of the passive observation is at a distance d from the nearest active observation. The relative model state error reduction achieved by actively used observations a distance d away from the location of active observations, $\Delta_x(d)$, is defined as:

$$\Delta_x(d) := \frac{\text{Tr} < \mathbf{d}_p^b(d)\mathbf{d}_p^b(d)^T - \mathbf{d}_p^a(d)\mathbf{d}_p^a(d)^T >}{\text{Tr} < \mathbf{d}_p^b(d)\mathbf{d}_p^b(d)^T >} \quad (15)$$

$$= \frac{\text{Tr}\{\mathbf{D}(\mathbf{d}_p^b\mathbf{d}_p^{bT} - \mathbf{d}_p^a\mathbf{d}_p^{aT})\mathbf{D}^T\}}{\text{Tr}\{\mathbf{D}\mathbf{d}_p^b\mathbf{d}_p^{bT}\mathbf{D}^T\}}$$

$$\overset{(10,12)}{=} \frac{\text{Tr}[\mathbf{D}\{\mathbf{H}_p(\mathbf{B} - \mathbf{A}')\mathbf{H}_p^T + (\mathbf{X} + \mathbf{X}^T)\}\mathbf{D}^T]}{\text{Tr}[\mathbf{D}(\mathbf{H}_p\mathbf{B}\mathbf{H}_p^T + \mathbf{R}_p)\mathbf{D}^T]}, \quad (16)$$

with 'Tr' denoting trace, i.e. the sum of the matrix diagonal elements. Hence, each term in Eq. (15) equals the mean variance of all departures a distance d away from the nearest actively used observation. From Eq. (16) and ignoring the $\mathbf{X}$ terms (discussed below), it is found that a positive value for $\Delta_x(d)$ implies a smaller analysis than background error, while negative values imply a larger analysis than background error, i.e. the assimilated observations have degraded the model state relative to the background.

Figure 1 shows a steady decrease of observation impact on the analysis further away from the observation location irrespective
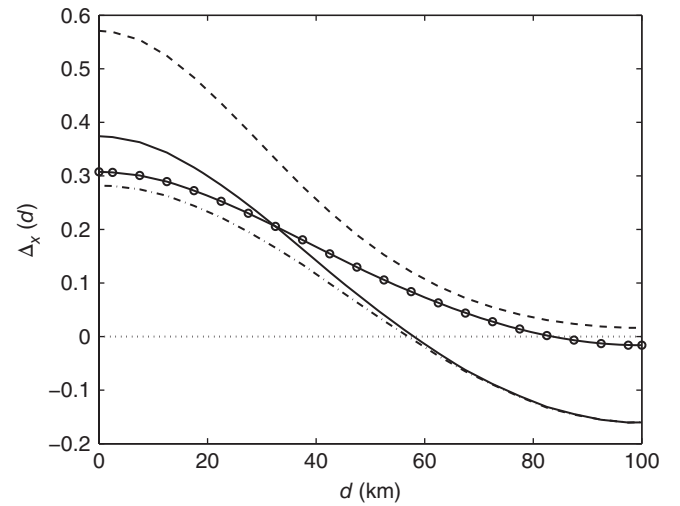

Figure 1. Relative model state error reduction, $\Delta_x(d)$, as a function of distance from assimilated observations. See the text for details. Results are for using the gain matrix $\mathbf{K}'$ from prescribed error covariance matrices (solid and dash-dot), the optimal gain matrix $\mathbf{K}$ (dashed) and the adapted gain matrix $\alpha\mathbf{K}'$ (solid with circles), discussed in section 3.3. Correlation of representativeness errors of passive and active observations has been taken into account, except for the dash-dotted line which is based on the assumption of no correlation. Negative observation impact is observed when curves drop below the dotted line.

of the used gain matrix, i.e. both from prescribed (solid and dash-dotted curves) and exact (dashed curve) error covariances matrices. The model state is improved close to assimilated observations. However the solid curve shows that the model state is degraded, on average, when the distance to the nearest assimilated observation exceeds 55 km when using the gain matrix from prescribed error-covariance matrices. The dashed curve shows that the model state has improved over the complete domain when using the optimal gain matrix, but however this is not available in practice. The dash-dotted curve is obtained from the same settings as the black curve but now ignoring the correlation of representativeness errors of active and passive observations. Hence, the difference between the black solid and dash-dotted curve is attributed to the $\mathbf{X}$ terms in Eq. (12); this difference is maximized for passive observations measured at the same location as active observations and vanishes for large separations, as expected.

3.3. Tuning of the gain matrix

The previous section showed that passive observations can be used to quantify the performance of an inherently imperfect gain matrix, i.e. whether increments from active observations have drawn the model state toward passive observations. Next, one can question whether passive observations can be used to adapt the gain matrix and improve subsequent analyses. The most simple adaptation to the gain matrix is through a scalar factor. Here, we aim to find a scalar adaptation of the gain matrix such that the resulting analyses are closer to passive observations. In other words, we aim to find scalar α which minimizes the norm $\|\mathbf{y}_p^o - \mathcal{H}\{\mathbf{x}^a(\alpha)\}\|$, given $\mathbf{x}^a = \mathbf{x}^b + \alpha\mathbf{K}'\mathbf{d}^b$. Substituting $\mathbf{x}^a$, and minimizing the norm by setting the derivative with respect to α equal to zero, yields

$$\alpha = \frac{\boldsymbol{\beta}^T\mathbf{d}_p^b}{\boldsymbol{\beta}^T\boldsymbol{\beta}}, \qquad \boldsymbol{\beta} = \mathbf{H}_p\mathbf{K}'\mathbf{d}^b = \mathbf{H}_p(\mathbf{x}^a - \mathbf{x}^b). \quad (17)$$

The latter equality is obtained by using Eq. (1), i.e. vector $\boldsymbol{\beta}$ simply is the projection of the analysis increment to passive observation space. For each of the 1000 realizations, α was computed from Eq. (17) and a new analysis calculated using the adapted gain matrix. Figure 2 shows the corresponding value of α for each realization. The mean value of about 0.6, which is smaller than 1,
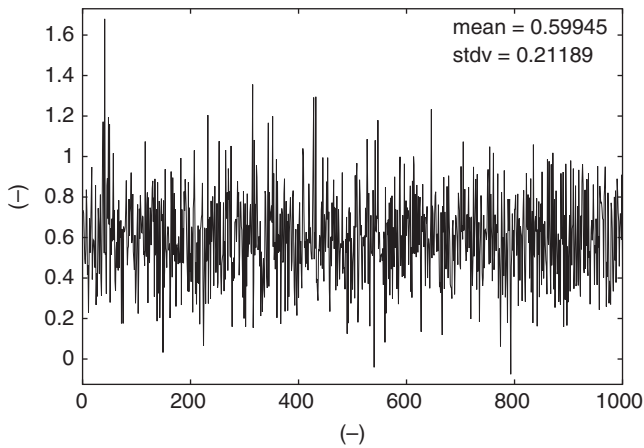
Figure 2. Gain factor adaptation α to be applied to $\mathbf{K}'$ for each of the 1000 realizations of simulated background state and observations.

indicates that too much weight has been assigned to observations relative to the background when using the prescribed error covariances. As such, passive observations can be used not only to diagnose but potentially also to tune the gain matrix. Interesting is the large spread for α with sometimes even negative values.

Following the same procedure as in the previous section, but now using analyses based on adapted gain matrices yields the solid curve with circles in Figure 1. Comparison with the black curve shows that the analysis has slightly degraded within about 30 km from actively used observations, but has improved substantially for larger distances, with larger area below the solid curve with circles than the solid curve, meaning that the model state is closer to the true state, on average, in particular away from assimilated observations through adapting the gain matrix. Clearly, the adapted gain matrix performs less than the optimal gain matrix (dashed curve).

3.4. Departure and model error statistics

Departure and model state error statistics have been calculated from the simulated observations and model states for the various options of the gain matrix. Figure 3 summarizes the results for active and passive observations. It should be kept in mind that statistics displayed in (a) are generally available in operational practice, however statistics in (b) are not available because the true atmospheric state is inherently unknown. On the other hand, (b) is of particular interest, i.e. how well have assimilated
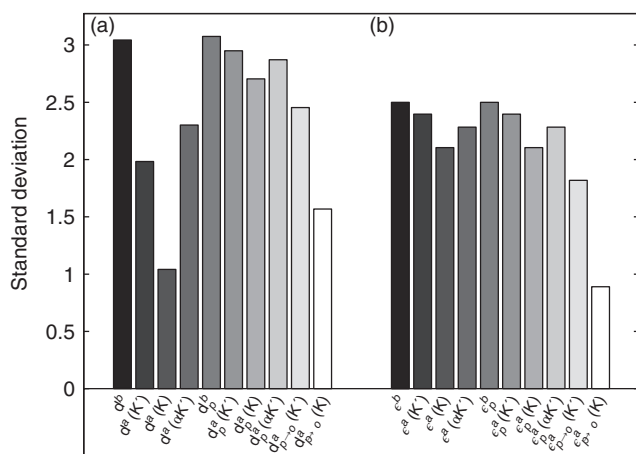
observations drawn the model state toward the truth, starting from the background? The discussion below aims to learn how departure statistics (a) relate to model error statistics (b), to achieve a similar translation in operational practice.

The leftmost bar of Figure 3(a) is the standard deviation of $\mathbf{d}^b$ for active observations which should be close to the square root of the sum of the true background error and observation error variance of active observations, i.e. from Table 1, $\sqrt{2.5^2 + 1.5^2 + 1^2} = 3.08$. The corresponding (same greyscale) bar in (b) is the standard deviation of the true background error, $\boldsymbol{\epsilon}^b$ with a value of 2.5 in agreement with Table 1. The second bar of both parts show the standard deviation of $\mathbf{d}^a$ and $\boldsymbol{\epsilon}^a$ based on the gain matrix from prescribed error covariance matrices. $\mathbf{d}^a$ is substantially smaller than $\mathbf{d}^b$, with a value of 1.99, however the analysis error standard deviation is only marginally smaller, with a value of 2.40, than the background error (compare first two bars in (b)). This is in agreement with the black curve of Figure 1 which shows an overall model state error reduction (area below the curve >0). The third bar of both parts is the result of $\mathbf{d}^a$ and $\boldsymbol{\epsilon}^a$ when using the optimal gain matrix in DA. This result is not available in practice, but shows a further reduction of $\mathbf{d}^a$, with a value of 1.06. It is noted that $\mathbf{d}^a$ is not necessarily minimized when using the optimal gain matrix. For instance, the unlikely situation of infinite weight given to observations (when the prescribed observation error equals zero or the background error is set to infinity) draws the analysis completely to the observations yielding $\mathbf{d}^a$ equal to zero. However, the resulting analysis does not minimize the analysis error. The third bar in (b) shows the minimum analysis error from using the optimal gain matrix, with a value of 2.10. Note that the minimum analysis error is only about 16% smaller than the background error (first bar in part 2) compared to the analysis departure in Figure 3 which is about 66% smaller than the background departure (bars 1 and 3 in (a)). The fourth bar of both parts is the result of $\mathbf{d}^a$ and $\boldsymbol{\epsilon}^a$ when using the adapted gain matrix. In (a), the result seems worse than using the non-adapted gain matrix (second bar), but as discussed above and confirmed below, the analysis departure of *active* observations cannot be used to draw any conclusion on the analysis error. In fact the larger value of the fourth bar is consistent with the finding above that less weight should be given to observations, through adaptation of the gain matrix with $\alpha < 1$, to better draw the model state to the true atmospheric state. The resulting analysis then draws closer to the background and less to observations, resulting in an increased value of 2.34 for $\mathbf{d}^a$, relative to the use of the non-adapted gain matrix. Indeed, in contrast to (a), the fourth bar in (b), i.e. the analysis error standard deviation after gain adaptation, has a value of 2.29, which is smaller than from the non-adapted gain (second bar) but larger than from the optimal gain (third bar of (b)), as expected.

The next four bars of both parts are similar to the first four bars of both parts but now for departure statistics from passive observations, o_p. For part 1, background departures in the fifth bar equal the first bar for identical error characteristics of active and passive observations (Table 1). The sixth bar, with a value of 2.95, has not reduced as much compared to the fifth bar as the second bar compared to the first bar. This is clear from the discussion in section 2. Using the optimal gain matrix further reduces $(o_p - a)$ for passive observations in the seventh bar, with a value of 2.76, but not as dramatic as the third bar. The eighth bar, with a value of 2.88, is the result of using the adapted gain matrix, which should result in a lower value than the sixth bar by construction. It can be easily shown that $< (o_p - a)^2 >$ is minimized for $\mathbf{K}' = \mathbf{K}$ (bar 7). For part 2, bars 1–4 equal the corresponding bars 5–8, simply because the calculation of the analysis error, $\boldsymbol{\epsilon}^a$, standard deviation is irrespective of observations. In fact from these results the total contribution of the $\mathbf{X}$ terms in Eq. (12) can be obtained as the sum of the analysis error variance (squared value of bar 2 in part 2) plus the observation error variance (from Table 1) minus the analysis departure variance of passive observations (squared value of bar 6 of part 1) which equals $2.40^2 + 1.5^2 + 1 - 2.95^2 = 0.31$.

Figure 3. Standard deviation of departure, $\mathbf{d}$ (part 1: 10 bars in (a)) and of model error, $\boldsymbol{\epsilon}$ (part 2: 10 bars in (b)). Statistics differ because of the used observations in the departures, either from active (no subscript) or passive (subscript p) observations and the used gain matrix in the analysis, $\mathbf{K}'$, $\mathbf{K}$ or $\alpha\mathbf{K}'$. The last two bars of each part are the result of active use of passive observations in data assimilation, the left(right) bar from analyses based on a gain matrix from prescribed(exact) error covariances.

Hence, for the settings in Table 1, the contribution of the representativeness error covariance to the analysis departure covariance of passive observations in Eq. (12) is small, which explains the similar (not identical because of the applied square root to go from variances to standard deviations) shape of bars 5–8 in part 1 and part 2; the analysis error variance is well estimated from passive observations by subtracting their total observation error variance from their analysis departure variance.

Finally, one may ask whether active use of passive observations would further improve analyses. Indeed the ninth and tenth bars are below bars 5–8 in both parts. From the last two bars in part 2 it is clear that actually assimilating the 2000 passive observations gives the best result, i.e. yielding a further reduction of the analysis error than when using only 25 observations. This seems evident, yet in actual global DA systems, including satellite observations, over 95% of the observations are neglected from assimilation, because the gain matrix is far from optimal, partly because actual observation-error correlations are not well known and/or not well accounted for and partly because $\mathbf{B}'$ is incorrect because it is based on climatology rather than the actual weather situation. Gain adaptation through a flow-dependent $\mathbf{B}'$ has already proved to be beneficial for global NWP (Bonavita *et al.*, 2012).

In summary, this section has shown, in a simulation environment, that analysis departures of active observations cannot be related directly to the analysis error. A direct relationship does exist for analysis departures of passive observations whose variance is close to the sum of analysis and observation error variances, depending on the correlation of the representativeness errors of active and passive observations. Furthermore, passive observations can be used to calculate an adaptation factor for an imperfect gain matrix used in DA. Use of the adapted gain matrix reduces the analysis error by construction. The sign of the adaptation indicates whether observations have been used too aggressively (adaptation < 1) or too conservatively (adaptation > 1) Direct assimilation of all observations, i.e. turning passive into active observations, would further reduce the analysis error. However, this is not common practice in DA nowadays, mainly because of imperfect knowledge of real background and observation errors and imposed simplifications for computational efficiency. As a consequence, nowadays practice is to neglect over 95% of the observations, at least in global DA. Rather than neglecting these data, the simulation tool has shown the potential of their use to verify and adapt the gain matrix. This idea is further explored in a real data assimilation system in the next section.

4. Use of passive observations in the HARMONIE mesoscale model

The non-hydrostatic convection-permitting HARMONIE model is developed in cooperation with the ALADIN project[†] and builds upon model components that have largely initially been developed in these two communities. At a default horizontal grid size of ≤2.5 km, the forecast model and analysis system are basically those of the AROME model from Météo-France (Seity *et al.*, 2011).

The construction of the prescribed background-error covariance matrix, $\mathbf{B}'$, for the current experimental version of the HARMONIE model, cycle cy38h1.2, was described in section 3.1. To test the performance of $\mathbf{B}'$, an experiment has been conducted with HARMONIE for the period 15 November to 31 December 2013. This period includes the 5/6 December 'Sinterklaas storm', also known as the 'Mandela storm', which hit Northern Europe with at least ten casualties. Extreme winds in combination with a spring tide caused extreme water levels at Western European coasts, the largest in the Netherlands since the 1953 flood disaster. HARMONIE boundary conditions are from the global ECMWF

model. For DA a 3D-Var (Courtier *et al.*, 1998) implementation is used with 3 h cycling, i.e. eight analyses are performed each day. Observing systems used include: SYNOP ground stations over land and sea (from ships), buoys, AMDAR (aircraft) and radiosondes. These are the *active* observing systems.

Figure 4 shows departure statistics for AMDAR aircraft data (black curves). Overall, the biases are much smaller for the analyses than the background and also the standard deviations have reduced substantially.

However, from the discussion in section 3 one cannot conclude on the correct tuning of the DA system in general and the correctness of $\mathbf{B}'$ in particular, based on this result only. The only conclusion is that DA draws the model state closer to assimilated observations.

4.1. DA tuning using passive Mode-S EHS observations

Recently, new observations have become available from mode selective (Mode-S) communication between aircraft and ground air traffic control radars (de Haan, 2011; Strajnar *et al.*, 2012). Communication takes place every 4 s. With an average aircraft speed of $250 \, \text{m s}^{-1}$, data every 1 km along the track are transferred to air traffic control from which local atmosphere wind components and temperature can be inferred (de Haan, 2011). Here, the so-called Enhanced Surveillance variant of Mode-S (Mode-S EHS) was used to derive atmospheric observations, denoted below as Mode-S. With on average 24 000 observations per hour (more during daytime, less at night), Mode-S provides a high-resolution dataset both in space and time, which has been assimilated successfully in NWP recently (de Haan and Stoffelen, 2012; Lange and Janjić, 2016). Figure 5 shows the typical daytime coverage of AMDAR and Mode-S observations.

Wind and temperature observations from Mode-S have not been used in the experiment (passive observations) and can thus be used as an independent referee to judge the quality of analysis increments obtained from the experiment. A total of 24 646 671 Mode-S observations for wind and temperature were obtained for the 6-week experimental period in the HARMONIE domain. The vertical coverage in Figure 6 shows a limited number of preferred flight levels (the spikes in the figure), with most of the observations taken in the upper troposphere above 500 hPa.

The departure statistics in Figure 4 show a slightly larger bias of background departures for Mode-S observations than for AMDAR, except in the lower troposphere for the zonal wind and in the upper troposphere for temperature, with larger biases for AMDAR background departures. As expected, biases of analysis departures are smaller for AMDAR because these observations were assimilated in contrast to Mode-S data. The standard deviation of background departures is smaller for Mode-S for both wind components, confirming the good quality of these data (de Haan, 2011). Also, the relative poor quality temperature observations, found in de Haan (2011), is apparent from the large background departures below 500 hPa in Figure 4(e). From the departure statistics of AMDAR and Mode-S, it is concluded that the quality of the latter is at least as good for winds, however a bias of ∼ $0.2 \, \text{m s}^{-1}$ is found for Mode-S meridional wind components. Mode-S temperature data are of poorer quality, but only in the lower troposphere.

Figure 4 also shows that the standard deviation of analysis departures is always smaller than background departures for both the passive (Mode-S) and active (AMDAR) observing systems. For active observing systems this is expected. Also, the standard deviation of background and analysis departures is much closer for passive than active observing systems, in agreement with the findings in section 3, in particular Figure 3(a). From the discussion in section 3.3, the analysis departure for passive observing systems is related closely to the analysis error. The smaller analysis than background departure for Mode-S then implies a smaller analysis than background error, on average. In other words, from the departure statistics of Mode-S observations it is inferred that the

---

[†]ALADIN: Aire Limitée Adaptation dynamique Développement InterNational, a collaboration by national meteorological services of Central and Eastern Europe on limited-area NWP.
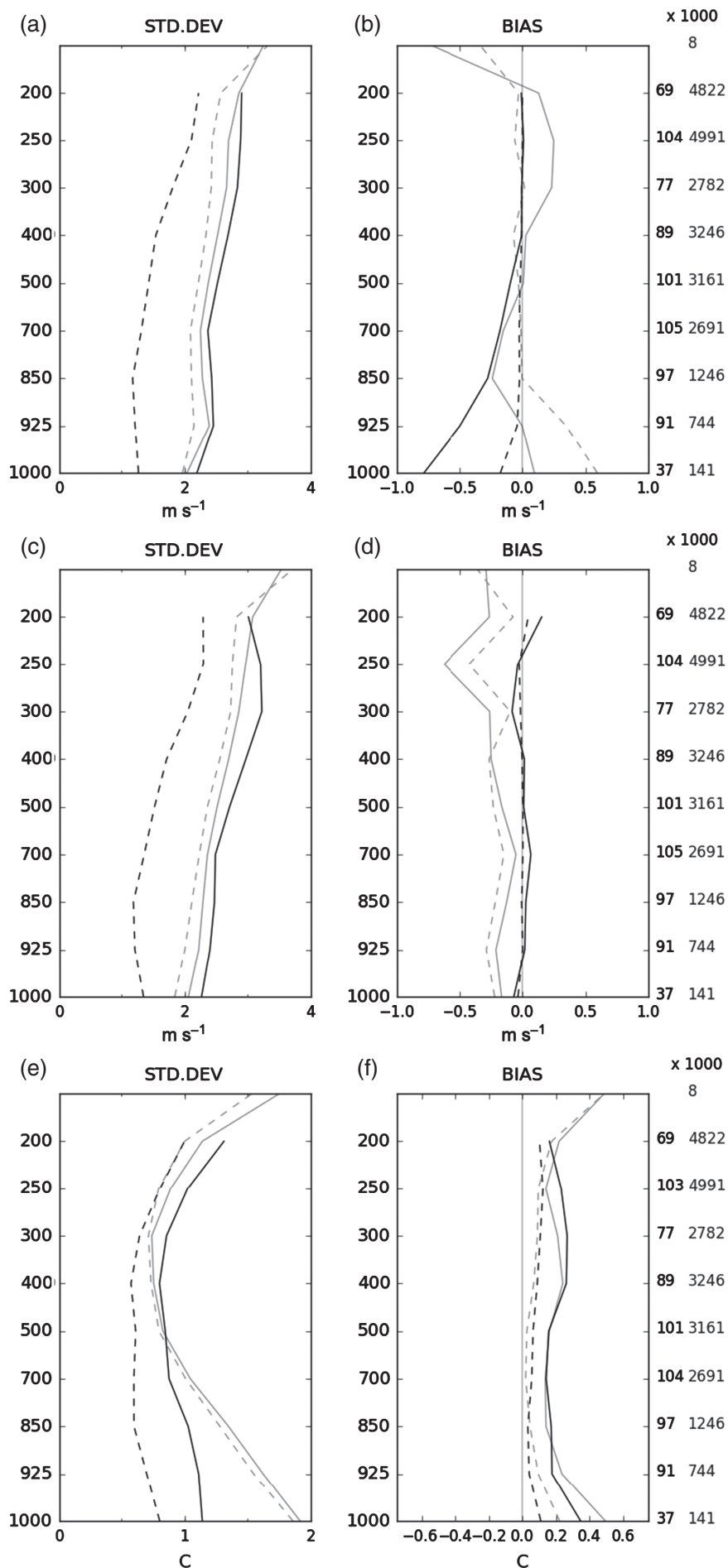
Figure 4. AMDAR (black) and Mode-S EHS (grey) departure statistics for (a, b) the zonal wind component, (c, d) the meridional wind component and (e, f) temperature. Panels (a, c, e) show standard deviation and (b, d, f) bias. The y-axis denotes pressure (hPa). Solid and dashed lines denote the background and analysis departures, i.e. $(o - b)$ and $(o - a)$, respectively. Statistics are based on the complete 6 week experimental period 15 November to 31 December 2013 in which AMDAR data were used actively and Mode-S EHS data passively. The numbers on the right-hand side give the total number of AMDAR (black) and Mode-S EHS (grey) data used in the statistics. These numbers should be multiplied by 1000.
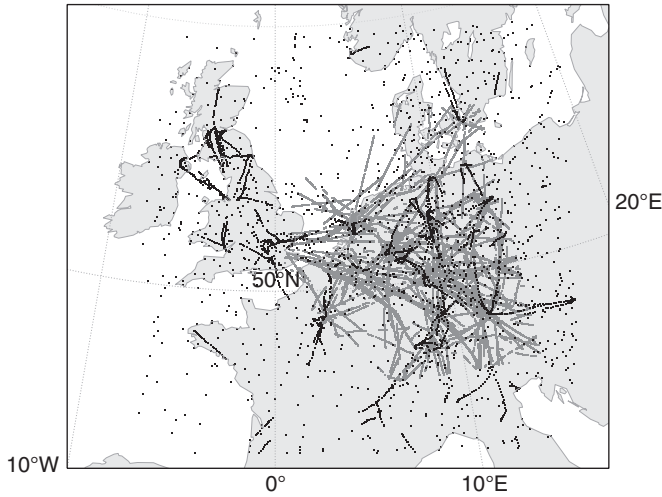
Figure 5. Typical daytime coverage of AMDAR (black) and Mode-S (grey) observations, here on 15 November 2013 near 1500 UTC. For AMDAR, locations of all observations used in the 1500 UTC analysis are displayed. For Mode-S, only locations of observations within 15 min from analysis time are displayed.
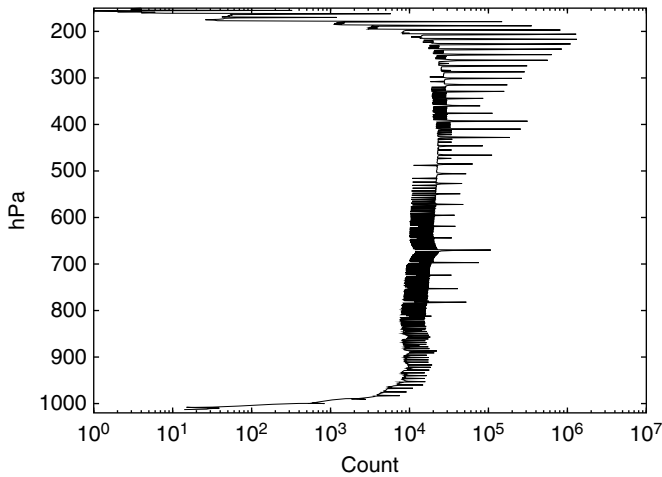

Figure 6. Mode-S data count as a function of pressure level over the 6 week experimental period. The *x*-axis is logarithmic.

active observations have drawn the model state closer toward the true atmospheric state, implying that the tuning of the HARMONIE DA system is at least adequate.

As in section 3.2, the model state error reduction from increments has been determined as a function of distance from active observations. Figure 7 shows that Mode-S observations are generally closely positioned to assimilated observations, i.e. within 50 km. The nearest active observation to each Mode-S observation is determined in three-dimensional space. From Figure 5 it is clear that, for the vast majority of Mode-S observations, the nearest active observation will be from AMDAR, considering that alternative upper-air measurements are from radiosondes only, with coarse spatial and temporal coverage.

From Figure 7 it appears that assimilated observations improve the analysis, on average, at least for distances up to 65 km from the observation location. This is an indication that the structure functions of the **B′** matrix, which control the spatial distribution of the increment, are appropriate, on average.

The increasing model state error reduction for the meridional wind and temperature at larger distances is in contrast to Figure 1. This is explained by noting that the denominator of Eq. (16) is constant in the simulation tool of section 3, i.e. the background departure variance is independent of the distance from assimilated observations, by construction. The consistent decrease of the curves in Figure 1 is explained by the consistent increase of analysis departures with distance from assimilated observations. However, Figure 8 shows that in real-world NWP both the analysis
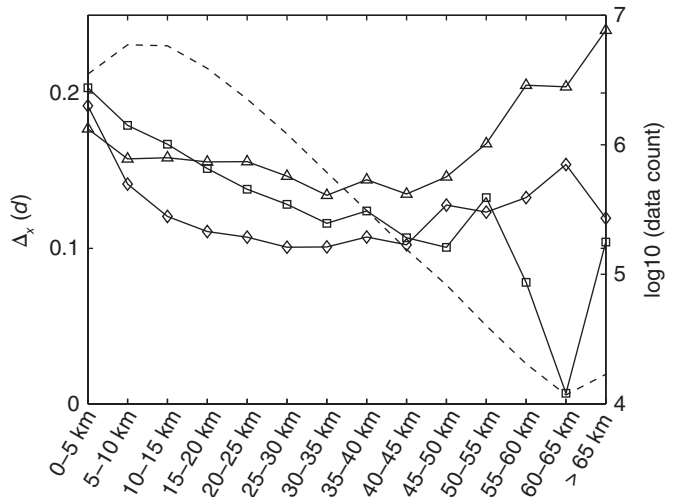
Figure 7. Relative model state error reduction as in Figure 1, but now for passive Mode-S zonal (squares) and meridional (diamonds) wind components and temperature (triangle) applied to HARMONIE model background and analysis fields. Matrices **D** (section 3.2) were designed such that background and analysis departures were assigned to 5 km bins ($d = 0-5$, $5-10$, $10-15$ km, etc.), depending on the distance d of passive observations to the nearest active observation. The dashed line denotes the number of Mode-S observations used over the complete 6 week period as a function of distance to the nearest assimilated observation. Note that the *y*-axis is logarithmic, i.e. the numbers along the axis are powers of 10. The data count is identical for all three parameters measured by Mode-S: u, v and t.
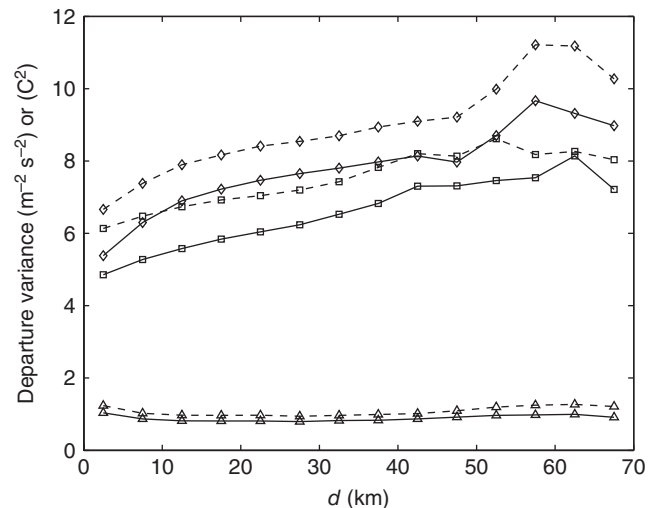
Figure 8. Mean model background departure variance (dashed) and analysis departure variance (solid) for passive Mode-S observations as a function of nearest distance between locations of Mode-S and assimilated observations for the zonal (square) and meridional (diamond) wind component and temperature (triangle).

and background departures increase away from assimilated observations. In fact their growth rate is quite similar, at least for distances up to 40 km. Then, the numerator of Eq. (16) is almost constant and the decreasing trend is explained by the increasing denominator with distance. Beyond 40 km distance, the growth rate of the background departure is stronger than the analysis departure for the meridional wind component and temperature (divergence of the curves with diamonds and triangles) explaining the corresponding growth of the corresponding bars in Figure 7. However, the growth rates beyond 40 km are less accurate because of a strong reduction of one or two orders of magnitude in the number of data (Figure 7).

Clearly, for constant background departures over the complete model domain, the trends of the bars in Figure 7 would be equal to the curves in Figure 1. The observation of increasing background departures away from assimilated observations is interesting in itself and is further elaborated in section 5.

Next, the procedure discussed in section 3.3 is used here to tune the gain matrix $\mathbf{K}'$ used in HARMONIE. For each analysis over the 6-week period adaptation factors are calculated for the two horizontal wind components and temperature, from Eq. (17), at pressure levels with at least 1000 Mode-S observations available within 15 min from analysis time. Figure 9 shows a typical example for the zonal wind component at 217 hPa, which corresponds to one of the spikes in Figure 6. At this pressure level almost 300 analyses in the 6-week period fulfilled the requirement of at least 1000 passive observations for the calculation of the gain matrix adaptation factor. The black line shows that the gain factor varies substantially for different analyses, similar to Figure 2. For Mode-S departures in the bottom panels, $(o_\text{p} - b)$ (light grey) is above $(o_\text{p} - a)$ (dark grey), on average, in agreement with the grey lines in the top panel of Figure 4 at 217 hPa. The black line is the result of $(o_\text{p} - a)$ after gain adaptation.

To arrive at this result, we did not rerun the experiment including gain adaptation, but instead made a highly simplified assumption of NWP as a univariate system, i.e. ignoring physical relations between model parameters. Then the analysis after gain adaptation, $\widehat{\mathbf{x}}_\text{a}$, can be calculated from the available analysis increment using Eq. (1) as follows. The adapted gain matrix yields

$$\widehat{\mathbf{x}}_\text{a} = \mathbf{x}^\text{b} + \alpha \mathbf{K}'[\mathbf{y}^\text{o} - \mathcal{H}(\mathbf{x}^\text{b})].$$

Following Eq. (17), and substituting the analysis increment into the above equation, yields

$$\widehat{\mathbf{x}}_\text{a} = \mathbf{x}^\text{b} + \alpha(\mathbf{x}^\text{a} - \mathbf{x}^\text{b}).$$

Resulting analysis departures have reduced by construction; the black line is below the dark grey line for all analyses. Clearly, the univariate assumption is not valid for NWP. The resulting gain matrix adaptations can therefore not be applied directly in operational NWP to improve analyses. A multivariate NWP system requires a more elaborate adaptation of the gain matrix which is outside the scope of this article. We come back to this in section 6.

The mean value of the black curve in Figure 9 gives an overall indication of whether the relative weight given to the model background and observations at this pressure level is correct, on average, for the 6-week period. The mean value of 0.89 indicates an overestimate of the relative weight given to the observations
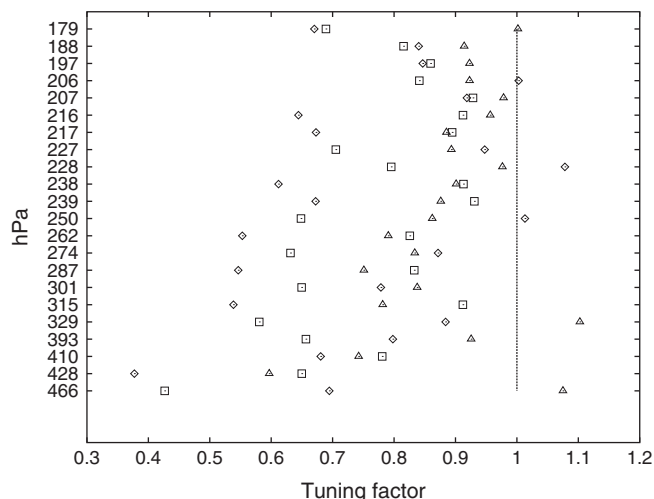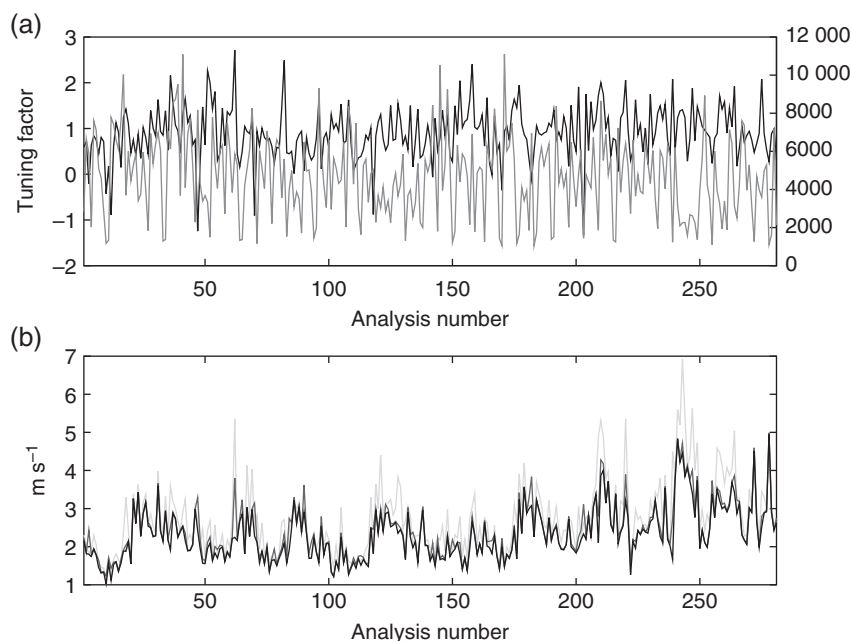

Figure 10. Gain matrix adaptation factor for the zonal (square) and meridional (diamond) wind component and temperature (triangle), averaged over the 6-week period for selected pressure levels with over 1000 Mode-S observations. The black dashed line denotes a gain factor of 1 for reference.

at this level. Similar calculations were performed for selected pressure levels and Mode-S measured parameters. Given the selection criteria and from Figure 6, pressure levels with sufficient Mode-S coverage are found in the upper troposphere only. The resulting Figure 10 shows that the gain matrix adaptation is generally below 1 for wind and temperature at all pressure levels. From this result we conclude that, on average, too much weight has been given to the observations in HARMONIE relative to the model background.

4.2. Active use of passive observations

In the previous section Mode-S observations were used passively to diagnose a DA system. One could ask whether instead active use of Mode-S would improve the model state, similar to section 3.3. Thus we ran another experiment with the HARMONIE model over the same 6-week period, now using Mode-S observations in addition. It was mentioned in section 4.1 that the distance between adjacent Mode-S observations is about 1 km, i.e. much smaller than the HARMONIE model effective resolution of about

Figure 9. (a) Gain matrix adaptation factor (black) for the zonal wind component at 217 hPa and the number of Mode-S observations (grey) used in the calculation of the adaptation factor. (b) Mode-S departure standard deviation, for $(o - b)$ (light grey), $(o - a)$ for the default gain matrix (dark grey) and $(o - a)$ for the adapted gain (black) used in the calculation of the analyses.
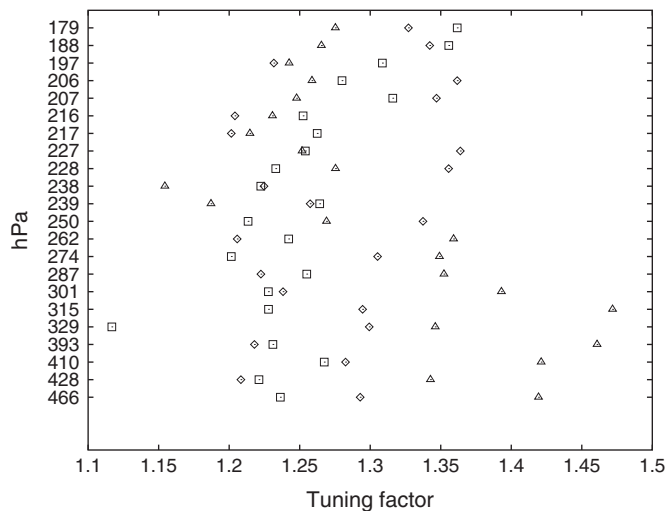
Figure 11. As Figure 10, but now with a limited number (through data thinning) of Mode-S data used in DA.

Figure 12. As Figure 10, but now for the ECMWF model, the version operational in December 2013.

25 km. To reduce error correlation of used Mode-S observations, we applied data thinning, such that the distance of actively used observations equals 15 km, i.e. less than 7% of the measured Mode-S data have been used in DA, leaving over 93% used passively.

Next, gain adaptation calculations were performed, similar to the previous section. In contrast to Figure 10, gain adaptation factors now show all values above 1 in Figure 11. This is explained by the fact that the maximum distance between active observations (including from Mode-S) and passive Mode-S observations is generally below 10 km, for the applied data thinning. Also, a smoothing operation has been applied to the raw Mode-S data (de Haan (2011)), which introduces some correlation in the error characteristics of Mode-S observations over distances up to 15 km. Clearly, a better fit to passive Mode-S observations is then obtained by giving more weight to Mode-S observations used actively in DA, which is achieved by enlarging the gain matrix. This example shows that the adaptive gain calculation suffers the same limitation as a DA system: use of observations with correlated errors should be avoided if not taken into account properly. In contrast, the correlation of errors for observations from AMDAR and Mode-S EHS, used in the previous section, is expected to be very small. First, because the number of aircrafts reporting both through AMDAR and Mode-S EHS is very limited (e.g. Figure 5). In addition, the measurement techniques and processing from both systems differ substantially. Finally, assessing the correlation of errors between Modes-EHS observations from different aircraft is currently under investigation at KNMI (de Haan, 2015). First results indicate that these error correlations are negligible, which is explained by the different on-board instrumentation of different aircraft companies and even within a single company for different plane types.

4.3. Application to the ECMWF model

Mode-S observations are not used by ECMWF and can thus be used to diagnose the ECMWF model. The situation is slightly different from HARMONIE because ECMWF operates a 4D-Var assimilation system with a 12 h assimilation window. Also from the ECMWF archive only forecasts initiated at 0000 and 1200 UTC are available. Because Mode-S observations at 0000 UTC are limited, we use for x^b in Eq. (17) the 12 h ECMWF forecast valid at 1200 UTC. The term $\mathbf{K}'[\mathbf{y}^o - \mathcal{H}(\mathbf{x}^b)]$ equals the analysis increment, which is obtained from archived analysis and corresponding background model fields. $\mathbf{H}_p$ maps the ECMWF model parameters to the location of passive observations, $\mathbf{y}_p^o$, here from Mode-S. This includes vertical interpolation from model to pressure levels and horizontal interpolation from the nearest model grid points to the observation location. No temporal
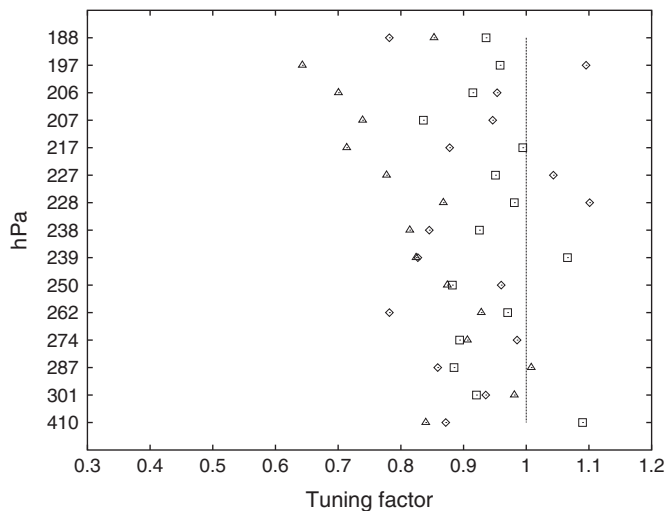
interpolation was needed because Mode-S observations measured within 15 min from verification time were selected.

Figure 12 shows the mean adaptation to the gain matrix needed to bring the ECMWF analysis closer to Mode-S observations. The same criteria as in Figure 10 have been adopted. This explains the reduced number of pressure levels in Figure 12, because of the limited number of analyses per day done by ECMWF (two) relative to HARMONIE (eight). From these figures, the gain adaptation factors are closer to 1, on average, for ECMWF, implying that the ECMWF DA system is well tuned in the sense that analyses are close to independent passive observations, here from Mode-S aircraft data. The values substantially smaller than 1 for temperature in the topmost pressure levels suggest that too much weight has been given to observations which impact the model temperature at these levels. We recall that the quality of Mode-S temperature is good in the upper troposphere (Figure 4(e, f)).

5. Observation impact assessment

Observation impact assessment is not limited to model analyses, as discussed in the previous sections, but includes model forecasts as well. In an OSE, one aims to assess the additional value of an observing system by comparing the skill of model forecasts from two different experiments, one denying and the other adding the new observing system under investigation. The former is in the remainder denoted the control (CTRL) experiment, the latter CTRL+NEWOBS. A widely adopted skill score is from statistics of observations minus forecast, in the remainder denoted in short as $(o - f)$. Observations used for verification (o) have not been used in the analysis and subsequent forecast (f) and thus one could naively think that o and f in $(o - f)$ are independent and the resulting skill score is independent of the used observing system for verification. The previous sections have learned that, even in the case of observations not used in DA, there is a caveat based on four points:

(i) increments are best close to actively used observations (both for a perfectly and imperfectly tuned DA system);
(ii) the correlation between increments and evolved increments in the short term is large on average, e.g. between 0.6 and 0.7 for HARMONIE 1 h evolved increments and slightly larger for ECMWF (not shown);
(iii) non-zero correlation of the analysis error (and subsequent forecast error) and the representativeness error of verifying observations (discussion at the end of section 2), with magnitude depending on the sampled air mass and spatial separation of observations used in the model forecast initial state and for verification;

(iv) distances between active observations and observations used for verification are generally smaller for observing systems used both in the analysis and for verification than when using observations from different observing systems in the analysis and for verification.

Given the large correlation between increments and evolved increments and maximum performance of increments near assimilated observations implies that the quality of short-term forecasts is likely to be best, on average, in areas near assimilated observations and thus model forecast verification in these areas is likely to give the best skill scores. This is confirmed by the curves with diamonds in Figure 8, showing increasing background departures and thus increasing background errors, from Eq. (10) and assuming non-varying observation errors with distance from locations of assimilated observations. As a further example, Mode-S observation density is largest and more or less fixed near airports and along air corridors. Hence, when using Mode-S both in DA and for verification, it is likely that the volume where forecast quality is best is better sampled by Mode-S than (e.g.) radiosondes, on average, thus yielding better skill scores when verified against Mode-S than against radiosondes. In addition, it is more likely that observations from Mode-S used in the forecast initial state and for verification sample the same airmass, implying non-zero correlation of their representativeness errors which reduces $(o - f)$ similar to $(o - a)$ in Eq. (12).

In general, verification of both CTRL and CTRL+NEWOBS with observations from the new observing system is likely to yield better forecast scores for the latter, at least in the short term. Skill scores are therefore likely to be biased toward experiments using observations for verification from the same observing system as used in the analysis. This might be the case even for relative poor-quality data from the new observing system, potentially leading to incorrect conclusions on the added value of the new observing system for NWP. Figures 13 and 14 show that the assimilation of observations has added value for HARMONIE wind and temperature; the dark grey (no DA) is always above the other curves (with DA). Comparing the light grey (CTRL) and black (CTRL+Mode-S) experiments, the skill score is largely overlapping when using observations from AMDAR for verification (a). Note that both experiments assimilate observations from AMDAR. When using observations from Mode-S for forecast verification, then (b) suggest a clear added value of using additional observations from Mode-S for forecasts up to $3-6$ h. From Figure 14(b), this is even true for the relatively poor-quality lower-troposphere temperature observations from Mode-S (Figure 4(e, f)). These biased conclusions are explained in the discussion above.

The results above explain the verification scores of Strajnar *et al.* (2015), who assimilated Mode-S observations and demonstrated positive scores when using Mode-S for verification, but neutral impact when using radiosondes for verification (their Figures 8

and 9). Also, their results for the summer period are less positive than the winter period, possibly because the latter contains a stable high pressure period with reduced advection, thus yielding higher correlations between increments and evolved increments. de Haan and Stoffelen (2012) show mainly neutral impact from assimilating Mode-S on 3 h forecasts at 400 and 875 hPa ('Ref' and 'RefM' experiments in their Figures 9 and 12). Clearly, for overall neutral impact, skill scores are irrespective of the observing system used for verification.

## 6. Summary, discussion, conclusions and outlook

Passive, i.e. not assimilated, observations are valuable to diagnose and tune a DA system. Neglecting observations from assimilation may have several reasons. For instance, imperfect knowledge of observation-error characteristics, in particular error correlations, is generally solved through data thinning strategies. In operational practice, about 95% of observations are not used for this reason. Also a DA system may not yet be ready to adopt observations from a relatively new observing system. Departure statistics from passive observations provide useful information on the tuning of the DA system, i.e. whether the correct weight has been given to observations in the analysis. This weight is based on prescribed, and inherently imperfect, background- and observation-error covariances. An adaptation factor for the gain matrix can be calculated such that the resulting analysis best fits the selected valuable passive observations. To be valuable, correlation of errors from passive observations and from active (assimilated) observations should be small, preferably zero. Denied observations from data thinning potentially suffer from this issue and should therefore be denied also from the gain adaptation calculation. As a consequence, passive observations used for DA diagnostic should ideally come from an observing system not used in assimilation.

In contrast to active observations, the covariance of analysis departures for passive observations is generally close to the sum of the analysis-error covariance and the observation-error covariance. The equality is not exact because the tempting assumption of zero correlation between analysis errors and errors from passive observations is violated when the representativeness error of actively used observations (as part of the analysis error) and passively used observations is non-zero. This may occur for observations closely separated in space and time and sampling the same air mass.

From the above property, the analysis-error characteristics (variance and correlation length-scale) can potentially be obtained from the analysis departures of passive observations, given well-known characteristics of the observation error for passive observations, used in the calculation of the analysis departure covariance. The resulting estimated analysis-error covariance
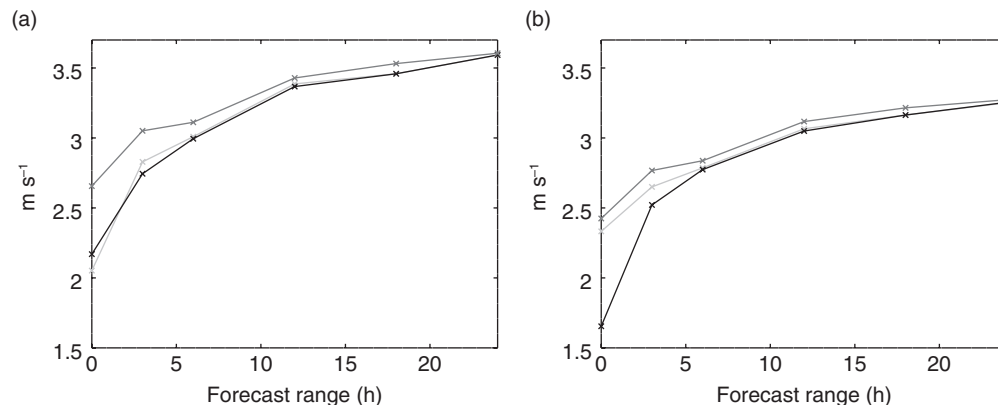
Figure 13. HARMONIE model forecast skill for the zonal wind component at 225 hPa. The skill score is defined here as the root mean square of $(o - f)$. Verification is done against data from (a) AMDAR and (b) Mode-S. Three 6-week experiments were conducted: no observations (dark grey), CTRL (light grey) and CTRL+Mode-S (black). The verification area is mainly Germany where Mode-S observation density is largest (Figure 5). For all forecast lead times and experiments, the same dataset is used for verification (8851 from AMDAR and 965 675 from Mode-S).
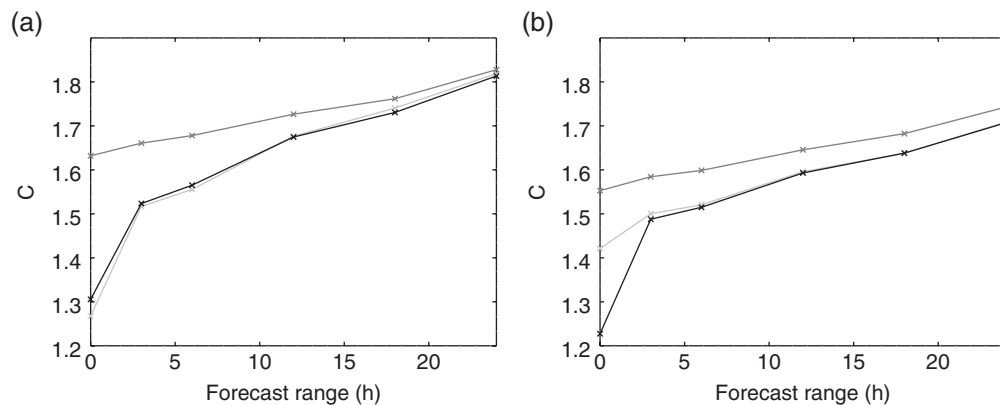
Figure 14. As Figure 13, but now for temperature at 875 hPa. The number of data used in the forecast verification was 14 699 (AMDAR) and 200 184 (Mode-S).

provides an alternative for the Desroziers diagnostic (Desroziers *et al.*, 2005) which is part of future research in our department.

Also, the difference of analysis and background departure statistics for passive observations identifies regions where increments (from assimilated observations) have improved or degraded the model state, on average. This has been done as a function of distance to assimilated observations and provides an indication of the correctness of prescribed structure functions of the background-error covariance matrix which spatially spreads the observational information in three-dimensional space. It was confirmed that analyses and subsequent backgrounds degrade with distance from assimilated observations.

These general findings were obtained in a simulation environment and confirmed when applied to a real DA system from the mesoscale HARMONIE model. Here, high-resolution aircraft observations from Mode-S EHS were used as passive observations. The high quality of these Mode-S EHS wind observations and temperature observations in the upper troposphere was confirmed by comparing with departure statistics from aircraft observations from AMDAR. It was found that analysis increments improve the model state, on average, for all distances away from assimilated observations, indicating that prescribed background- and observation-error covariance matrices are appropriate. On the other hand, from the calculated gain adaptation factors, it was concluded that the experimental HARMONIE version cy38h1.2 puts too much weight to observations relative to the model background. Similar calculations applied to the ECMWF model (note that ECMWF does not make use of observations from Mode-S EHS) indicated that their DA is generally well-tuned, but slightly overestimates temperature observation information in the upper troposphere.

Using Mode-S EHS as a passive observing system provides univariate adaptation factors for the wind components and temperature, which do not take into account physical relationships of these parameters, e.g. from geostrophic balance. Because NWP is a multivariate system, direct application of univariate adaptations to the gain matrix is not expected to be beneficial. Instead, the calculated gain adaptation factors reveal a general trend on the weight given to certain types of observations in the analysis. This information can be fed back to the gain matrix by adapting the variances of the corresponding parameters in the prescribed observation- and/or background-error covariance matrix.

The findings above culminate in a sketch of a potentially viable procedure to make maximum use of observations in a 3D-Var context, leaving untouched the structure functions from the prescribed climatological background-error covariance matrix (and diagonal observation-error covariance matrix):

1. Conduct a control experiment using conventional observations from radiosondes, AMDAR, SYNOP stations and buoys;
2. Calculate gain adaptation factors from Eq. (17) using passive observations from an independent (new), not

assimilated, observing system. The result is used to apply a scalar adaptation to the prescribed background- and/or observation-error covariances;
3. Repeat steps 1 and 2 with adapted background- and/or observation-error covariance matrices, until the mean value of the gain adaptation factors converges to 1;
4. The DA system is now well tuned to maximize the impact from observations from used observing systems and ready to adopt observations from the new observing system.

The procedure above progressively introduces new observing systems into the DA system while guaranteeing optimal use of accepted observing systems. Testing this procedure in an operational DA system is part of future work. New observing systems currently tested in the context of HARMONIE include: Mode-S EHS wind and temperature, scatterometer ocean surface winds, radar radial wind and reflectivity and humidity from the Global Navigation Satellite Systems (GNSS). Implementation of the procedure above is the subject of further research. Application of the tuning procedure is applicable too, in principle, in a 4D-Var context and using non-fixed but flow-dependent background-error covariances.

Finally, a caveat was revealed when assessing the impact of an observing system for NWP through verification of resulting forecasts with observations from the same observing system, which is a widely accepted approach in OSE. Forecast skill scores based on statistics of observations-minus-forecast $(o - f)$ may differ substantially, in particular for short-term forecasts, depending on the observing system used for verification. First, the tempting assumption of zero correlation of model forecast and observation errors at verification time may be violated in the calculation of $(o - f)$ statistics. Correlated representativeness errors of actively and passively used observations reduce the standard deviation of $(o - f)$. This is most likely when using an actively used observing system for verification. Second, analysis increments and their evolution in the short term are best near assimilated observations. That is, model forecast quality is not uniform over the verification area and forecast skill scores highly depend on the sampling of the verifying observing system. This explains the inconsistent skill scores found in Strajnar *et al.* (2015) and to a less extent in de Haan and Stoffelen (2012) who assimilated Mode-S EHS in their respective mesoscale models and showed better skill scores when verifying against Mode-S EHS rather than against radiosondes. The results demonstrate that one has to be careful when using observations for forecast verification when assessing the added value of new observing systems for NWP. The best way to avoid biased skill scores is to use observations from an independent observing system for verification, i.e. an observing system not being used in all experiments under investigation. In general, it is not easy to find such an observing system, because DA systems are eager to assimilate as many observations as possible. Alternatively, one could use for verification an observing system used in all experiments under investigation. Potential candidates are

observing systems more or less used as default for all assimilation experiments such as radiosondes and AMDAR for upper-air verification and SYNOP stations for near-surface verification.

Appendices

Appendix A: Analysis departure covariance of passive observations

Section 2 mentions that the analysis error, $\boldsymbol{\epsilon}^a$, and observation error from passive (i.e. not used in the analysis) observations, $\boldsymbol{\epsilon}_p$, are not necessarily uncorrelated because the representativeness error of active (i.e. used in the analysis) observations and passive observations may be correlated. Here, we discuss how this translates to the covariance matrix of analysis departures in Eq. (11). As mentioned in section 1, the total observation error, e.g. $\boldsymbol{\epsilon}$ in Eq. (3) equals the sum of the instrument error, below denoted with superscript i, and representativeness error, below denoted with superscript r. Active observations, denoted without subscript, and passive observations, denoted with subscript p, are then written

$$
\begin{align}
\mathbf{y}^o &= \mathbf{H}_o\mathbf{x}^t + \boldsymbol{\epsilon}^i + \boldsymbol{\epsilon}^r, \tag{A1}\\
\mathbf{y}_p^o &= \mathbf{H}_p\mathbf{x}^t + \boldsymbol{\epsilon}_p^i + \boldsymbol{\epsilon}_p^r. \tag{A2}
\end{align}
$$

For simplicity, but without loss of generality, we have assumed that mapping from model to observation space is accomplished through a linearized observation operator matrix, i.e. using $\mathbf{H}$ rather than $\mathcal{H}$. From these definitions the analysis departure of passive observations equals

$$
\begin{align}
\mathbf{y}_p^o - \mathbf{H}_p\mathbf{x}^a &\overset{\text{Eq. (A2)}}{=} -\mathbf{H}_p(\mathbf{x}^a - \mathbf{x}^t) + \boldsymbol{\epsilon}_p^i + \boldsymbol{\epsilon}_p^r\\
&\overset{\text{Eq. (1)}}{=} -\mathbf{H}_p\{\mathbf{x}^b - \mathbf{x}^t + \mathbf{K}'(\mathbf{y}^o - \mathbf{H}_o\mathbf{x}^b)\}\\
&\qquad + \boldsymbol{\epsilon}_p^i + \boldsymbol{\epsilon}_p^r\\
&\overset{\text{Eq. (A1)}}{=} -\mathbf{H}_p\{\mathbf{I} - \mathbf{K}'\mathbf{H}_o\}\boldsymbol{\epsilon}^b\\
&\quad - \mathbf{H}_p\mathbf{K}'\boldsymbol{\epsilon}^i - \mathbf{H}_p\mathbf{K}'\boldsymbol{\epsilon}^r + \boldsymbol{\epsilon}_p^i + \boldsymbol{\epsilon}_p^r. \tag{A3}
\end{align}
$$

Taking the covariance of Eq. (A3) and assuming no correlation between instrument errors, between instrument and representativeness errors and between the background error and all observation error sources, all cross-terms vanish except for the observation representativeness errors:

$$
\begin{align}
< (o_p - a)^2 > &:= < (\mathbf{y}_p^o - \mathbf{H}_p\mathbf{x}^a)(\mathbf{y}_p^o - \mathbf{H}_p\mathbf{x}^a)^T >\\
&= \mathbf{H}_p\mathbf{A}'\mathbf{H}_p^T + <\boldsymbol{\epsilon}_p^i\boldsymbol{\epsilon}_p^{iT}> + <\boldsymbol{\epsilon}_p^r\boldsymbol{\epsilon}_p^{rT}>\\
&\quad - (\mathbf{X} + \mathbf{X}^T) \tag{A4}
\end{align}
$$

with $\mathbf{X} = \mathbf{H}_p\mathbf{K}' <\boldsymbol{\epsilon}^r\boldsymbol{\epsilon}_p^{rT}>$. It is noted that the first three terms of Eq. (A4) correspond to the quadratic terms of the analysis departure covariance, by making use of Eq. (4). This result is in agreement with the tempting assumption discussed below Eq. (11) of uncorrelated analysis errors and errors of passive observations. Also, the observation-error covariance matrix for passive observations, defined above Eq. (10), equals the sum of the instrument-error and representativeness-error covariances:

$$
\mathbf{R}_p = <\boldsymbol{\epsilon}_p^i\boldsymbol{\epsilon}_p^{iT}> + <\boldsymbol{\epsilon}_p^r\boldsymbol{\epsilon}_p^{rT}> .
$$

When representativeness errors of active and passive observations are not correlated, $\mathbf{X}$ in Eq. (A4) vanishes such that the covariance matrix of analysis departures for *passively* used observations *does* equal the sum of the analysis-error covariance (mapped to observation space with the observation operator) and observation-error covariance because the observation and analysis error are truly independent. This is the case when passive and active observations are sufficiently separated, either in space or time.

The other extreme is when passive observations overlap with active observations, i.e. $\mathbf{H}_p = \mathbf{H}_o = \mathbf{H}$ at locations of passive observations. Then $\boldsymbol{\epsilon}_p^r = \boldsymbol{\epsilon}^r$ and $\mathbf{X} = \mathbf{H}\mathbf{K}' <\boldsymbol{\epsilon}^r\boldsymbol{\epsilon}^{rT}>$. Substituting these expressions in Eq. (A4) and then rearranging terms yields

$$
\begin{align}
< (o_p - a)^2 > &= [\mathbf{I} - \mathbf{H}\mathbf{K}'][\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}][\mathbf{I} - \mathbf{H}\mathbf{K}']^T\\
&\quad + (\mathbf{Y} + \mathbf{Y}^T) + <\boldsymbol{\epsilon}_p^i\boldsymbol{\epsilon}_p^{iT}> - <\boldsymbol{\epsilon}^i\boldsymbol{\epsilon}^{iT}>\\
&= < (o - a)^2 > + (\mathbf{Y} + \mathbf{Y}^T) + <\boldsymbol{\epsilon}_p^i\boldsymbol{\epsilon}_p^{iT}> - <\boldsymbol{\epsilon}^i\boldsymbol{\epsilon}^{iT}>, \tag{A5}
\end{align}
$$

with $\mathbf{Y} = \mathbf{H}\mathbf{K}' <\boldsymbol{\epsilon}^i\boldsymbol{\epsilon}^{iT}>$ and Eq. (9) was used in the last step. From Eq. (A5), the analysis departure is larger for passive than active observations in the case of positive definite $(\mathbf{Y} + \mathbf{Y}^T)$ and equal instrument error of passive and active observations. In general, the analysis departure is larger for passive than active observations, i.e. the covariance from Eq. (A4) is larger than from Eq. (9). This becomes more clear when considering the scalar case in Appendix B.

Appendix B: Numerical scalar example

For simplicity, we assume that active and passive observations are directly related to a model state variable and observed at an exact model grid point. Then $\mathbf{H}_o = \mathbf{H}_p = 1$. For active and passive observations, the analysis departure covariances are obtained from Eq. (9) and Eqs (A4, 4) respectively. For simplicity, perfect knowledge of all error sources is assumed, to deny the primes in the equations.

$$
\begin{align}
< (o - a)^2 > &= (1 - k)^2(\sigma_b^2 + \sigma_{oi}^2 + \sigma_{or}^2), \tag{B1}\\
< (o_p - a)^2 > &= (1 - k)^2\sigma_b^2 + k^2(\sigma_{oi}^2 + \sigma_{or}^2)\\
&\quad + \sigma_{pi}^2 + \sigma_{pr}^2 - 2k\sigma_{opr}, \tag{B2}
\end{align}
$$

with σ_b the true background-error standard deviation, σ_{oi} and σ_{pi} the instrument error standard deviation of active and passive observations respectively, σ_{or} and σ_{pr} the representativeness error standard deviation of active and passive observations respectively, and σ_{opr} is the cross-correlation representativeness error of the active and passive observation. From Eq. (2),

$$
k = \sigma_b^2/(\sigma_b^2 + \sigma_{oi}^2 + \sigma_{or}^2).
$$

Which of the two expressions is larger depends on the error properties. Assuming that the quality of active and passive observations are equal and their locations sufficiently far apart to ignore the representativeness cross-correlation, Eq. (B2) reads as

$$
< (o_p - a)^2 > = (1 - k)^2\sigma_b^2 + (1 + k^2)(\sigma_{oi}^2 + \sigma_{or}^2), \tag{B3}
$$

which always exceeds Eq. (B1) given positive k. For the extreme case of overlapping locations for passive and active observations and the above assumptions, Eq. (A5) reads as

$$
< (o_p - a)^2 > = < (o - a)^2 > + 2k\sigma_{oi}^2. \tag{B4}
$$

Taking $\sigma_b = 2, \sigma_{oi} = 2, \sigma_{or} = 1$, then $k = 4/9$, $< (o - a)^2 > = 2.78$ from Eq. (B1), $< (o_p - a)^2 > = 7.22$ from Eq. (B3) and

$< (o_{\mathrm{p}} - a)^2 >= 6.33$ from Eq. (B4). Also, the analysis error variance $\sigma_{\mathrm{a}}^2 = 2.22$ from Eq. (5). Substituting in the scalar variant of Eq. (A4), ignoring $\mathbf{X}$, yields

$$< (o_{\mathrm{p}} - a)^2 >= \sigma_{\mathrm{a}}^2 + \sigma_{\mathrm{oi}}^2 + \sigma_{\mathrm{or}}^2 = 7.22,$$

in agreement with the value above from Eq. (B3).

This example demonstrates the general properties of analysis departures which are smallest for actively used observations and largest for passive observations with locations sufficiently distant from active observations to ignore representativeness cross-correlations. For locations of passive observations approaching active observations, the analysis departure becomes smaller, but is still always larger than for active observations, given the assumptions above. For zero representativeness error, Eqs (B3) and (B4) are equal, as expected.

References

Bonavita M, Isaksen L, Hólm EV. 2012. On the use of EDA background-error variances in the ECMWF 4D-Var. *Q. J. R. Meteorol. Soc.* **138**: 1540–1559.

Bouttier F, Courtier P. 2002. *Data Assimilation Concepts and Methods*, Training Course Lecture Series. ECMWF: Reading, UK. http://www.ecmwf.int/sites/default/files/Ass_cons.pdf (accessed 21 July 2016).

Brousseau P, Berre L, Bouttier F, Desroziers G. 2011. Background-error covariances for a convective-scale data-assimilation system: AROME-France 3D-Var. *Q. J. R. Meteorol. Soc.* **137**: 409–422.

Courtier P, Andersson E, Heckley W, Pailleux J, Vasiljevic D, Hamrud M, Hollingsworth A. 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part I: Formulation. *Q. J. R. Meteorol. Soc.* **124**: 1783–1808.

Daley R. 1999. *Atmospheric Data Analysis*, Cambridge Atmospheric and Space Science Series. Cambridge University Press: Cambridge, UK.

de Haan S. 2011. High-resolution wind and temperature observations from aircraft tracked by Mode-S air traffic control radar. *J. Geophys. Res.* **116**: D10111, doi: 10.1029/2010JD015264.

de Haan S. 2015. Estimates of mode-S EHS aircraft derived wind observation errors using triple collocation. *Atmos. Meas. Tech. Discuss.* **18**: 12633–12661.

de Haan S, Stoffelen A. 2012. Assimilation of high-resolution mode-S wind and temperature observations in a regional NWP model for nowcasting applications. *Weather and Forecasting* **27**: 918–937, doi: 10.1175/WAF-D-11-00088.1.

Desroziers G, Berre L, Chapnik B, Poli P. 2005. Diagnosis of observation-, background- and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc.* **131**: 3385–3396.

Lange H, Janjić T. 2016. Assimilation of Mode-S EHS aircraft observations in COSMO-KENDA. *Mon. Weather Rev.* **144**: 1697–1711, doi: 10.1175/MWR-D-15-0112.1.

Lorenc AC. 1986. Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.* **112**: 1177–1194.

Seity Y, Brousseau P, Malardel S, Hello G, Bénard P, Bouttier F, Lac C, Masson V. 2011. The AROME-France convective-scale operational model. *Mon. Weather Rev.* **139**: 976–991.

Skamarock WC. 2004. Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Weather Rev.* **132**: 3019–3032.

Strajnar B. 2012. Validation of Mode-S Meteorological Routine Air Report aircraft observations. *J. Geophys. Res. Atmos.* **117**: D23110, doi: 10.1029/2012JD018315.

Strajnar B, Žagar N, Berre L. 2015. Impact of new aircraft observations Mode-S MRAR in a mesoscale NWP model. *J. Geophys. Res. Atmos.* **120**: 3920–3938, doi: 10.1002/2014JD022654.

Vogelzang J, Stoffelen A, Verhoef A, Figa-Saldaña J. 2011. On the quality of high-resolution scatterometer winds. *J. Geophys. Res.* **116**: C10033, doi: 10.1029/2010JC006640.

Weston PP, Bell W, Eyre JR. 2014. Accounting for correlated error in the assimilation of high-resolution sounder data. *Q. J. R. Meteorol. Soc.* **140**: 2420–2429.