Toward Scatterometer Winds Assimilation in the Mesoscale HARMONIE Model

Gert-Jan Marseille and Ad Stoffelen

***Abstract*—Data assimilation (DA) experiments have been conducted with the high-resolution limited-area model HirLAM Aladin Regional Mesoscale Operational NWP In Euromed (HARMONIE), which is operational at most weather centers, which are part of the European HirLAM consortium. Recently, the assimilation of scatterometer ocean surface winds was introduced, showing limited forecast skill improvement. Possible explanations are discussed. These include model bias and the time mismatch between observation and analysis time, which introduces nonnegligible correlated errors in a three-dimensional (3-D) variational assimilation system. Also, ignoring the time mismatch increases the innovation, i.e., the observation minus background (model short-term forecast), by about 20% for scatterometer winds. The use of observations as point observations in most DA systems needs reconsideration for mesoscale DA. The introduction of observation operators, taking into account the instrument footprint, would improve the innovation by about 5% for scatterometer winds. Additional directions for improved use of observations in HARMONIE are discussed based on the notice that DA is an inherent deterministic concept. Hence, the selection of the spatial scale for deterministic DA should depend primarily on the 4-D observation coverage rather than the effective model resolution.**

***Index Terms*—HirLAM Aladin Regional Mesoscale Operational NWP In Euromed (HARMONIE) model, mesoscale data assimilation (DA), representativeness error, scatterometer winds.**

I. Introduction

DATA assimilation (DA) has proven to be beneficial for numerical weather prediction (NWP) in global and limited-area hydrostatic models for many years [1]. However, demonstrating additional value from assimilating observations (both conventional and from satellites) in nonhydrostatic mesoscale models appears quite a challenge. For simplicity, we discriminate between mesoscale (nonhydrostatic) and global (hydrostatic) models in the remainder of this paper. There is a long history on how to exploit observational information for global models. Apparently, applying the same paradigms to mesoscale models is less effective, i.e., forecasts from mesoscale models profit less from adapting the forecast initial state through observations. Understanding the fundamental differences between mesoscale and global NWP is still in its infancy. Here, we aim to elaborate on growing ideas and possible strategies to better exploit observational information for mesoscale models.

Mesoscale models use much smaller grid sizes than global models and explicitly resolve atmospheric convective processes, which evolve more rapidly and on much smaller spatial scales than those resolved by global models. Correct initialization of small-scale weather phenomena requires a dense network of observations in all four dimensions. Data sparsity can induce phase errors, i.e., incorrect positioning of weather systems (akin to aliasing in the field of signal processing).

The main objective of this paper is to discuss strategies on how to optimally exploit observational information for convection-permitting mesoscale models, with focus on satellite winds from scatterometer instruments. An important aspect is that current DA systems treat all observations as point observations. This assumption may be valid for global models, which do not resolve scales typically below 100–150 km [2], [3]. However, when moving to smaller grid sizes and explicitly resolving convective processes on scales smaller than the observation footprint, this assumption needs reconsideration and more advanced observation operators may be needed to make optimal use of observational information.

Sections II and III provide the status of mesoscale DA with the preoperational version of the HirLAM Aladin Regional Mesoscale Operational NWP In Euromed (HARMONIE) model at KNMI. This includes the assimilation of satellite 10-m ocean surface winds from scatterometer instruments. The scatterometer impact on model forecasts appears marginal on lead times beyond 2 h, while winds are expected to effectively add mesoscale NWP skill [4] and good impacts are reported in global models [5]. This is a general trend observed for all observing systems and not specific for scatterometer winds. Aspects that may explain the limited observational impact are discussed. In Section IV, we discuss the mismatch between observation and analysis time and argue that the implementation of more advanced observation operators, taking into account instrument footprints, and effective deterministic model resolution improves the representativeness of observational information to the model state in DA. Section V discusses the possible directions to improve on mesoscale DA. Section VI concludes the paper.

II. Data

A. HARMONIE Model

European meteorological institutes as part of the high-resolution limited-area model (HirLAM) consortium are

Manuscript received June 29, 2016; revised October 4, 2016 and November 11, 2016; accepted December 13, 2016. *(Corresponding author: Gert-Jan Marseille.)*

The authors are with the Research Department of Satellite Observations, Royal Dutch Meteorological Institute (KNMI), De Bilt 3730, AE, The Netherlands (e-mail: Gert-Jan.Marseille@knmi.nl; stoffele@knmi.nl).

Digital Object Identifier 10.1109/JSTARS.2016.2640339

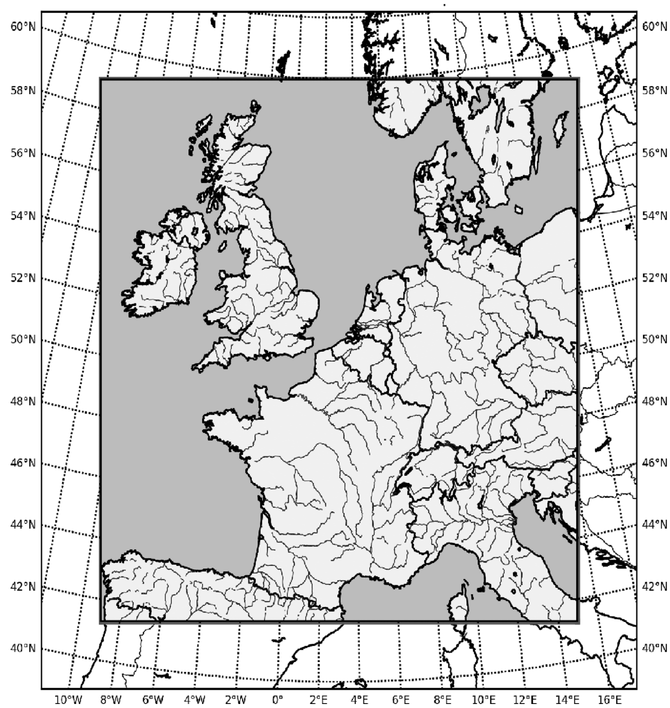

Fig. 1. HARMONIE mesoscale model domain, used operationally by KNMI. The domain is centered at 51° latitude, 3° longitude and is composed of 800 × 800 grid points covering a 2000 × 2000 km area, i.e., the model grid size is 2.5 km.

currently in the transition phase of moving from the operational hydrostatic model HirLAM to the nonhydrostatic convection-permitting HARMONIE model. HARMONIE is developed in cooperation with Météo-France and ALADIN,[1] and builds upon model components that have largely initially been developed in these two communities. At default horizontal grid size ≤2.5 km, the forecast model and analysis system are basically those of the AROME model from Météo-France [6]. In this paper, HARMONIE model cycle 38h1.2 was used with a grid size of 2.5 km for the model domain displayed in Fig. 1. The model top is at 10 hPa (∼26 km) and the number of model levels equals 65. The lateral boundaries are obtained from the global model of the European Centre for Medium-Range Weather Forecasts (ECMWF).

B. Scatterometer Ocean Surface Winds

A scatterometer is a satellite radar instrument, which provides a measure of wind speed and direction near the sea surface. Scatterometers measure the electromagnetic microwave backscatter by the wind-roughened ocean surface. Scatterometer wind information is organized on a grid of wind vector cells (WVCs) projected on the earth swath of the instrument. The number of across-swath WVCs determine the sampling resolution of the surface wind field and the wind information is considered to be Nyquist sampled, with modest correlation between neighboring WVCs. Each WVC contains between two and four ambiguous local wind vector solutions that are the result of the inversion of the wind geophysical model function (GMF), for a given set of backscatter values at a given scanning geometry [7], [8]. The ambiguity is mainly related to the double harmonic dependency of the GMF on wind direction. Each wind ambiguity is characterized by a solution probability that is determined based on the distance-to-GMF residual after the inversion.

The wind ambiguities, solution probabilities, and prior information from the ECMWF model 10-m background wind are used in a two-dimensional variational (2D-Var) ambiguity removal procedure [3] to produce an analyzed surface wind field, fitting one of the ambiguities at each WVC. This wind field is then used to select the wind vector ambiguity in each WVC that is closest to the analysis, based on vector difference, as the solution for the observed surface wind. A wind vector solution flag is set to the index of the selected wind ambiguity in each WVC. Finally, the backscatter measurements, wind ambiguities, scanning geometry, and wind vector solution flag, among others, are made available as a scatterometer wind product. A detailed overview of past and current operational scatterometers is provided in [9]. Here, we summarize the main characteristics of the scatterometers used in this study.

1) ASCAT: European C-band (rain insensitive [10]) scatterometers onboard the Metop-A and Metop-B satellites, which were launched into a sun-synchronous orbit on October 19, 2006 and September 17, 2012, respectively [11]. The satellite overpass time, expressed as local (equator crossing) time of ascending node (LTAN) is 10:30 UTC. The ASCAT coastal product that has 12.5-km sampling [9] is used in this study. Observations from these satellites are described in the remainder of this paper denoted as ASCAT-A and ASCAT-B. Both are still operational.

2) OSCAT: Indian Ku-band (rain sensitive) scatterometer on the OceanSat-2 satellite (en.wikipedia.org/wiki/Oceansat-2). Launched in June 2011 into a sun-synchronous orbit at 12:00 LTAN and operational until February 2014. The OSCAT product with 50 km sampling is used in this study [12].

3) HSCAT: Chinese Ku-band scatterometer on the Haiyang-2A satellite. Launched in August 2011 into a sun-synchronous orbit at 06:00 LTAN and still operational. The HSCAT product with 25 km sampling is used in this study, as processed by the Pencil-Beam Wind Data Processor (PenWP[2]).

It should be stressed that sampling is generally not the same as resolution. Sampling denotes the separation between adjacent observations, whereas resolution refers to the spatial scales resolved by the observations. In general, the resolution is chosen lower than the sampling distance. For instance, for ASCAT the cumulative spatial response function, which is a measure of the averaging in spatial domain, provides an estimate of the effective observation resolution, which is about 28 km for the ASCAT coastal product [13].

[1]ALADIN is the acronym for Aire Limitée Adaptation dynamique Développement InterNational, a collaboration of national meteorological services of Central and Eastern Europe on limited-area NWP.

[2]http://nwpsaf.eu/site/software/scatterometer/penwp/

C. 3D-Var DA

HARMONIE operates a 3D-Var DA system [14] with a 3-h cycling, i.e., eight analyses are performed each day. The observations operationally used by HARMONIE include surface pressure from SYNOP ground stations over land and sea (from ships) and buoys, wind observations from buoys, radiosonde and aircraft meteorological data relay (AMDAR), and temperature observations from radiosondes and AMDAR. In the remainder of this paper, these observing systems are called the conventional observing systems. The goal of DA is to find the best linear unbiased estimate (called analysis), which is a compromise 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 with the true atmospheric state. However, observations are imperfect too due to instrument imperfections and they are generally not fully representative of the model state variables. 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.

In general, the analysis equations in DA read, following [15] for the notational convention [16], [17]:

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}'[\mathbf{y}^o - \mathbf{H}\mathbf{x}^b] \tag{1}$$

$$\mathbf{K}' = \mathbf{B}'\mathbf{H}^{\mathbf{T}}[\mathbf{H}\mathbf{B}'\mathbf{H}^{\mathbf{T}} + \mathbf{R}']^{-1} \tag{2}$$

$$\mathbf{y}^o = \mathbf{H}\mathbf{x}^t + \boldsymbol{\epsilon} \tag{3}$$

where $\mathbf{x}^b$ is the n-dimensional background state vector, i.e., a short-term forecast initiated from the state analysis, $\mathbf{x}^a$ is the previous cycle, $\mathbf{y}^o$ is the m-dimensional vector with observations, $\mathbf{x}^t$ is the true (but unknown) state vector, and $\mathbf{H}$ is the $m \times n$ linearized observation operator matrix, which maps the model state space to the observed quantity. The latter may be a simple interpolation operator for some observing systems, but may be more complex, e.g., a linearized radiative transfer model for measured satellite radiances. 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 matrix transpose. The spatial spread of the observational information on the analysis state is determined by the prescribed background error covariance matrix, in the remainder of this paper also denoted as $\mathbf{B}'$, and the prescribed observation error covariance matrix, hereafter also denoted as $\mathbf{R}'$. If both describe well the true background and observation error covariances (denoted without primes, but which are generally unknown), then the resulting gain matrix, $\mathbf{K}'$ in (2), yields the best compromise, in statistical sense, of background and observations in the analysis, as in (1).

In practice, it is not trivial to correctly specify the $\mathbf{B}'$ and $\mathbf{R}'$ matrices. Ideally, $\mathbf{B}'$ should reflect the short-term model forecast errors. Currently, HARMONIE uses a climatological $\mathbf{B}'$ derived from downscaling (i.e., a 6-h forecast) four members of the ECMWF global ensemble over a six-week period and assuming homogeneity and isotropy. $\mathbf{R}'$ is taken as a diagonal matrix. The realism of this choice is further discussed in Section IV. For scatterometer winds, the prescribed observation error standard deviation of both wind components equals 1.47 m/s for ASCAT A and B and 1.45 m/s for OSCAT. It should be noted that further optimization of these values may be achieved following the procedure in [18].

In 3D-Var, it is assumed that all observations within the assimilation window have been measured at analysis time, i.e., typically the DA window center time. This is generally true for observations from radiosondes, SYNOP stations, and buoys. However, aircraft and satellite overpasses are asynoptic, introducing a time shift between observation and model background state. This timing issue can be partially resolved by using 3D-Var + first guess at appropriate time [19], which produces more accurate increments but still applied at the wrong time. This is resolved by more advanced assimilation schemes such as a 4D-Var DA scheme, which, however, is not available yet for HARMONIE. Alternatively, a 3D-Var-rapid update cycle [20] implementation is available, which employs a 1-h assimilation cycle.

D. Experimental Period

A number of observing system experiments (OSE) have been conducted with HARMONIE for the period November 15–December 31, 2013. This period was characterized by a predominant zonal flow and includes the December 5/6 "Mandela storm," which hit Northern Europe with at least ten casualties. Extreme winds in combination with spring tide caused extreme water levels at the Western European coast, the largest in the Netherlands since the 1953 flood disaster. The following experiments were conducted.

1) NO-OBS: No observations used in DA.
2) CONV-3h: Conventional observations used in 3D-Var with 3-h cycling.
3) CONV+SCAT-3h: Same as above but in addition the assimilation of all available ocean surface winds from ASCAT and OSCAT.
4) CONV+SCAT-THINN-3h: Same as above but including data thinning to 100 km observation spacing for ASCAT winds (currently, the default setting in HARMONIE).
5) CONV+SCAT-THINN-1h: Same as above but using a 1-h cycling.

III. Results

In 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. A widely adopted skill score is from statistics of observations minus model forecast, hereafter shortly denoted (o-f). Here, we focus on model scores for 10-m ocean surface wind. Ideally, observations (o) used for verification are from an observing system not used in DA. A potential candidate for 10-m model wind scores over oceans is 10-m wind observations from masts, e.g., on oil platforms. However, the coverage is too coarse for significant statistics over a six-week experimental period for the complete domain. Alternatively, one could use measurements from buoys, but these are too coarse. Scatterometer winds from HSCAT could be used, but these are available only twice per day in the HARMONIE domain. To assess the quality of 10-m model winds, it was therefore decided to use all observations from all the available scatterometers for verification, i.e., ASCAT-A, ASCAT-B, OSCAT, and HSCAT. The range of different overpass times of these satellites over the HARMONIE domain enables the verification of 10-m model winds for all forecast lead times although the number of available observations for verification may differ for different lead times. For 3-h cycling, analyses and subsequent forecasts are produced at 00, 03,..., 21 UTC. For the verification of scatterometer impact, only forecasts initiated from analyses in which scatterometer data were assimilated were used. Given the overpass times in Section II-B, it is inferred that observations from ASCAT can be used to verify forecasts with lead times: 0–4, 7–11, 15–17, 19–23 h; for OSCAT: 0–2, 4–5, 10–12, 17–19, 23–24 h, and for HSCAT: 0, 4–8, 12–14, 16–20, 22–24 h. As an example, ASCAT observations measured around 11 UTC have been used to verify the 2-h forecast initiated at 09 UTC. Clearly, verification scores differ for these other instruments, because of different instrument error characteristics and footprint size. To produce skill scores for all lead times with a single metric, one instrument was selected as reference. Matching lead times with other instruments were used to calibrate skill scores from different instruments. Next, missing lead times from the reference instrument were filled. ASCAT was selected as reference instrument because of the maximum availability of forecast lead times. The resulting skill scores are displayed in Fig. 2. Here, model forecast values were obtained through spatial and temporal interpolation of model fields to the observation location and measurement time. As such, observations are assumed point observations, in agreement with current treatment of observations in DA. This is further elaborated in Section IV.

The negative bias of the zonal wind component in the top left panel of Fig. 2 suggests an overestimation of the wind speed by the model relative to scatterometer winds given the predominant zonal flow during the experimental period. This is confirmed from Fig. 3 showing that HARMONIE ocean surface winds are too strong relative to scatterometer observations in particular for strong winds. Apparently, sea surface roughness increases not enough with friction velocity for high winds in HARMONIE. These results suggest that the formulation could be optimized. In addition, a coupled wave model could be used to improve the representation of high wind speeds above sea. This is part of further investigation and beyond the scope of this paper.

The center plots of Fig. 2 show that assimilation of observations improves the model simulations, i.e., the curve with crosses (no DA) is generally on top of the other curves. At analysis time (fc=0), model simulations using scatterometer in DA compare best to scatterometer observations. This is by design and is not very indicative on the performance of the DA system; a well-tuned DA system pulls the model state toward the true atmospheric state not only at observation locations, but also in nonobserved regions. Then, resulting model forecasts are also expected to be closer to the true atmospheric state. The CONV+SCAT-3h experiment shows the best scores at analysis time, which is not surprising since all available scatterometer

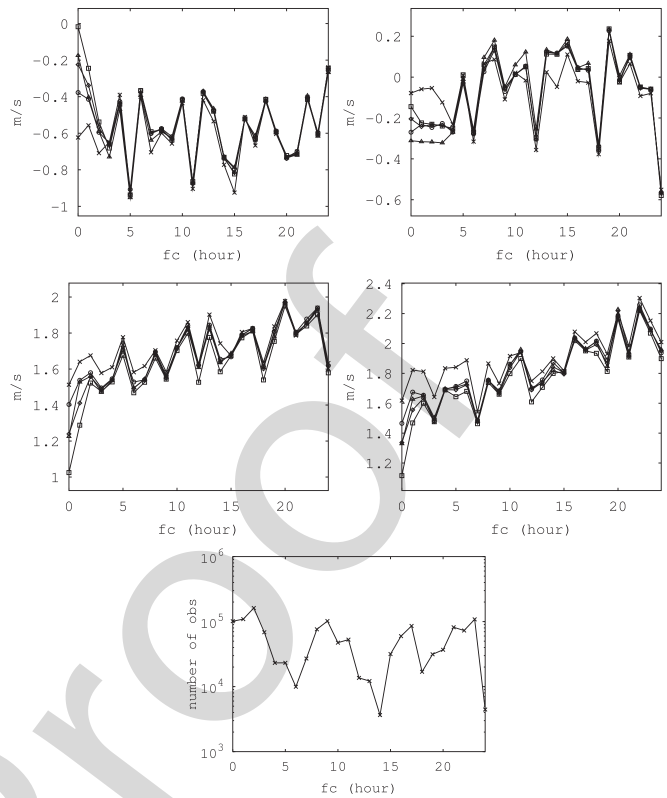

Fig. 2. Forecast skill scores for 10-m model wind over oceans; observation-minus-forecast (o-f) bias (top row) and standard deviation (center row) for the zonal (left panels) and meridional (right panels) wind components as a function of forecast range. Forecast (f) is obtained through spatial and temporal interpolation to the observation (o) location and measurement time. Observations include scatterometer ocean surface 10-meter winds from ASCAT on Metop-A and Metop-B, OSCAT, and HSCAT. Experiments included have been described in Section II-D: NO-OBS (crosses), CONV-3h (circles), CONV+SCAT-3h (squares), CONV+SCAT-THINN-3h (diamonds), and CONV+SCAT-THINN-1h (triangles). Scores are obtained from the six-week experimental period, see Section II-D, with the total number of used observations for verification in the bottom row.

observations were used both in the analysis and for verification. The thinning experiments (diamonds and triangles) only use about 1.6% of all available scatterometer observations in DA. Yet, the resulting analyses were drawn to the true state (represented by all scatterometer observations) substantially also some distance away from assimilated observations, i.e., the bias and standard deviation at analysis time have reduced substantially: the standard deviation of (o-f) at fc=0 for curves with diamonds and triangles is about halfway between the curves with circles (no scatterometer) and squares (all scatterometer). This shows the inherent redundancy in DA systems, that reducing the number of observations generally only marginally reduces the impact, due to the spatial filtering properties. For increasing forecast lead times, statistics of the experiments using scatterometer in DA (curves with squares, triangles, and diamonds) converge quickly to those of the experiment not using scatterometer (circles). At fc+3 and beyond the curves largely overlap, which is disappointing when comparing to global model skill scores. On the other hand, the experiment using all scatterometer observations tends to show the best scores, i.e., the squares curve is

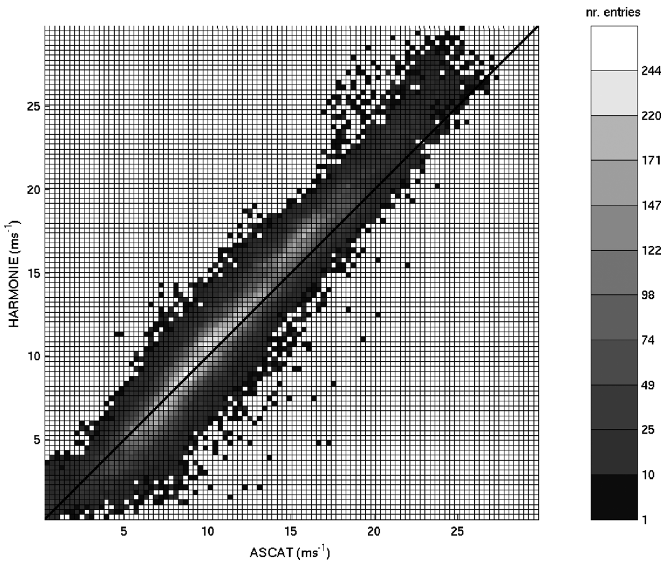

Fig. 3. Density scatter plots of 10-m wind speed (m/s) over sea, with scatterometer winds from ASCAT on Metop-B along the *x*-axis and HARMONIE model winds, averaged over the observation footprint, along the *y*-axis. Results are based on collocations over the complete six-week experimental period.

generally below the other lines over the complete forecast range, which is encouraging. Some additional remarks should be made. At fc+01, statistics of the experiment with 1-h sampling (triangles) are worse compared to the corresponding experiment with 3-h sampling (diamonds), despite the similar performance at fc+0. The better statistics for the 3-h cycling experiment (observations are used when measured within ± 1.5 h from analysis time) is because (some) observations used for verification of fc+01 have also been used in DA, which is not the case for the 1-h cycling experiment (where observations are used only when measured within ± 0.5 h from analysis time). It is thus encouraging that the 1-h cycling experiment scores better than the no scatterometer experiment at fc+01, indicating that the scatterometer has added value for the model on this term. This is further confirmed by the better scores at fc+02 for experiments using all scatterometer winds. Lack of skill for fc+03 and beyond is not typical only for scatterometer assimilation, but also for all observing systems (not shown). This seems to be fundamentally related to mesoscale DA and is currently under investigation. Possible explanations include the following.

1) The addition of nondeterministic small-scale model variance, which is weather dependent, but not well accounted for in $\mathbf{B}'$, $\mathbf{R}'$, and $\mathbf{H}$.
2) Model biases, of which Fig. 3 shows one example and the model tendency to quickly return to its (biased) climatology.
3) The time mismatch between analysis time and observation time for some observing systems, which is inherent for 3D-Var and further elaborated in Section IV.
4) Too much weight has been given to observations in the analysis, see [18].
5) Nonoptimal use of observations in mesoscale DA, e.g., (v-1) radiosondes drift from their launch location, which is ignored in DA, and (v-2) all observations are used as point observations, which may not be adequate for mesoscale DA.

Taking into account, the observation footprint and nondeterministic model variance (noise) in the observation operator are also discussed in Section IV.

IV. Observation Minus Background Diagnostic

Statistics of observation minus background, in the remainder shortly denoted as (o-b), is an important diagnostic for operational NWP centers, e.g., to check for model and/or observation biases. Here, we follow [17] who defines $\mathbf{x}^t$ as "the vector of coefficients obtained by projecting the true state of the atmosphere onto the model basis." Current NWP models yield a smooth simulation of the true atmospheric state and thus, from the above definition, $\mathbf{x}^t$ does only include spatial scales, which can be resolved by the NWP model.

Hence, when the NWP model state does not resolve all scales in the observation, the first term on the right-hand side of (3) cannot deliver the observation equivalent without error due to the missing spatial scales. This mismatch is called the observation representativeness error, denoted as $\boldsymbol{\epsilon}_r$. 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 spatial analyses and triple-collocation techniques [21], [22] provide estimates for the effective resolution of the model under investigation and the variance of the resulting observation representativeness error. As a rule of thumb, nowadays NWP models effective resolution equals seven to ten times the model grid size [2].

In addition, the instrument error, denoted as $\boldsymbol{\epsilon}_i$, accounts for instrument imperfections. Here, we ignore errors in the observation operator, i.e., from mapping from model to observation space. From (3) we may write for the background departures, also denoted innovation

$$\mathbf{y}^o - \mathbf{H}\mathbf{x}^b = \boldsymbol{\epsilon}_i + \boldsymbol{\epsilon}_r - \mathbf{H}\boldsymbol{\epsilon}^b \tag{4}$$

where $\mathbf{x}^b$ is the background state vector, i.e., a short-term forecast initiated with the previous analysis and the background error $\boldsymbol{\epsilon}^b = \mathbf{x}^b - \mathbf{x}^t$. The background departure covariance matrix, denoted in short $<(o-b)^2>$, with $<.>$ denoting the expectation operator equals

$$\begin{aligned} <(o-b)^2> \; &:= \; <(\mathbf{y}^o - \mathbf{H}\mathbf{x}^b)(\mathbf{y}^o - \mathbf{H}\mathbf{x}^b)^{\mathbf{T}}> \\ &= \; <[\boldsymbol{\epsilon}_i + \boldsymbol{\epsilon}_r - \mathbf{H}\boldsymbol{\epsilon}^b][\boldsymbol{\epsilon}_i + \boldsymbol{\epsilon}_r - \mathbf{H}\boldsymbol{\epsilon}^b]^{\mathbf{T}}> \\ &= \; \mathbf{R}_i + \mathbf{R}_r + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathbf{T}} \end{aligned} \tag{5}$$

where $\mathbf{B}$ is the background error covariance matrix defined as $\mathbf{B} =<\boldsymbol{\epsilon}^b{\boldsymbol{\epsilon}^b}^{\mathbf{T}}>$, $\mathbf{R}_i =<\boldsymbol{\epsilon}_i{\boldsymbol{\epsilon}_i}^{\mathbf{T}}>$ is the instrument error covariance matrix, and $\mathbf{R}_r =<\boldsymbol{\epsilon}_r{\boldsymbol{\epsilon}_r}^{\mathbf{T}}>$ is the observation representativeness error covariance matrix. The expression ":=" denotes "by definition." Here, it is as assumed that the cross correlations of the error terms vanish: $<\boldsymbol{\epsilon}_r{\boldsymbol{\epsilon}^b}^{\mathbf{T}}>=0$ by definition since the errors terms represent different spatial scales by construction; $<\boldsymbol{\epsilon}_i{\boldsymbol{\epsilon}^b}^{\mathbf{T}}>=0$ is plausible because correlations of uncertainties in the NWP model and instrument readings are often physically unreasonable; and $<\boldsymbol{\epsilon}_r \boldsymbol{\epsilon}_i{}^{\mathbf{T}}>=0$ because correlations of the representativeness error and instrument error are physically implausible as well.

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 [23]. Also, the representativeness error of observations separated by less than the model effective resolution are correlated and the corresponding covariance matrix is nondiagonal. Also, the true representativeness error is nonconstant and a function of local atmospheric turbulence [24]. It is thus likely that the total observation error covariance matrix $\mathbf{R} = \mathbf{R}_i + \mathbf{R}_r$ is a dense matrix for dense observations. However, for computational efficiency and also because of imperfect knowledge of these error sources, the total observation error covariance matrix is specified as a diagonal matrix in most operational DA systems, including HARMONIE, thus ignoring correlation of observation errors. In addition, the so-called superobbing or data thinning is employed to avoid overfitting due to misspecification of $\mathbf{R}'$.

Hence, observation preprocessing to reduce error correlations in $\mathbf{R}$ makes $\mathbf{R}'$ in better agreement with the exact matrix and the resulting gain matrix will be a better representation of the optimal Kalman gain. Reducing observation error correlations is therefore expected to yield an improved gain matrix and subsequent improved analyses and forecasts.

A. *Observation Representativeness Error*

Frehlich [25] introduces a spatial filter function g_j^m to simulate the projection of the continuous true state variable t_j onto the model basis. Then, the discrete representation of the true state variable in model space $\mathbf{t}_j^m$ can be written as a convolution of the true atmosphere and the filter function

$$\mathbf{t}_j^m(\mathbf{r}) = \int g_j^m(\mathbf{r}-\mathbf{s}) t_j(\mathbf{s}) \mathrm{d}\mathbf{s} \tag{6}$$

where $\mathbf{r}$ and $\mathbf{s}$ denote 3-D spatial coordinates, i.e., (6) denotes a 3-D spatial integration at a fixed time instant. This representation of the true state is in agreement with [17]. It is noted that the filter function is model dependent and an explicit expression is generally not known. Yet, some of its properties may be obtained from observation power density spectra or triple-collocation techniques [21]. For instance, in case that the model spectrum has a cutoff frequency with reduced energy on scales smaller than a certain threshold value, then g_j^m acts as a low-pass filter in spectral domain.

Similarly, the instrument footprint can be modeled through a spatial filter function g_j^o such that the discrete representation of the true atmosphere as observed by the instrument $\mathbf{t}_j^o$ may be written as a convolution of the continuous true atmosphere and the instrument filter function

$$\mathbf{t}_j^o(\mathbf{r}) = \int g_j^o(\mathbf{r}-\mathbf{s}) t_j(\mathbf{s}) \, \mathrm{d}\mathbf{s}. \tag{7}$$

For observing systems providing point measurements such as radiosondes or airplane reports, the instrument filter function is a delta-Dirac function and the mean atmospheric state within the observation sampling footprint simply equals the true atmospheric state at the observation location: $\mathbf{t}_j^o(\mathbf{r}_i) = t_j(\mathbf{r}_i)$, with $\mathbf{r}_i$ denoting the observation location. For ASCAT, measuring 10-m wind components, denoted by subscript $10\,\mathrm{m}$, over the ocean, $t_j \in \{u_{10\mathrm{m}}, v_{10\mathrm{m}}\}$ the footprint function is related to the 2-D cumulative spatial response function, mentioned in Section II-B.

Rewriting (7) in scalar form and selecting either $u_{10\,\mathrm{m}}$ or $v_{10\,\mathrm{m}}$ for the true state variable t_j, then scatterometer observation o, with instrument error ϵ_i, is related to the true atmospheric state averaged over the instrument footprint t^o through

$$o = t^o + \epsilon_i. \tag{8}$$

Similarly, the model background state b with background error ϵ^b is related to the true state in model space t^m through

$$b = t^m + \epsilon^b \tag{9}$$

and for the background departure

$$\begin{aligned} o - b = & \ (o - t^o) + (t^o - t^m) - (b - t^m) & (10) \\ = & \ \epsilon_i + \epsilon_r - \epsilon^b. & (11) \end{aligned}$$

Comparing with (5), the three error terms on the right-hand side of (11) are the instrument error, representativeness error, and background error in 1-D, respectively. For the unlikely situation that the instrument footprint filter function equals the NWP model filter function, i.e., $g_j^o = g_j^m$, and from (6) and (7), the second term on the right-hand side of (10) and (11) vanishes, i.e., the observation is fully representative of the model state variable. Otherwise, there are two remaining alternatives.

1) Observation Effective Resolution is Higher Than the Model Effective Resolution: For current global models, the width of the NWP model filter function (in spectral space) is generally smaller than of the instrument footprint filter [21], i.e., the representativeness error $(t^o - t^m)$ describes the atmosphere on scales between the observation footprint and model effective resolution. Averaging in observation space, also denoted superobbing, acts as a filter operation and thus reduces the representativeness error. The scatterometer sampling is very well suited for superobbing, through the application of a 2-D averaging window, such that the representativeness error of the resulting observation is close to the model effective resolution. To minimize the error correlations of neighboring observations, overlapping of adjacent nodes when performing superobbing should be avoided. Locations for superobbed observations are assigned to the center of the averaging window. Because locations are more remote after superobbing than of the standard product, both the correlation and variance of the representativeness error have reduced. In addition, averaging in observation space reduces the instrument noise, as given in (8). Overall, superobbing reduces (o-b) on average by reducing the first and second term on the right-hand side of (11). Alternatively, data thinning reduces the correlation of observation representativeness errors even more effectively when the footprint filter functions of selected observations have no or negligible overlap. Some data thinning thus seems effective for DA purposes because the assumption of uncorrelated observation error (diagonal $\mathbf{R}$) is better achieved, potentially yielding a closer to optimal gain matrix. However, data thinning does neither reduce the instrument error nor the representativeness error variance and hence does not reduce the total observation error variance and (o-b) variance. In addition, potential useful observational information is neglected, which makes data thinning potentially less attractive than superobbing.

2) Observation Effective Resolution is Lower Than the Model Effective Resolution: For mesoscale models with small grid sizes, the effective model resolution may be higher than the effective observation resolution or vice versa. For instance, HARMONIE with 2.5 km grid size has an effective resolution of 15–25 km, which is slightly higher than the 28 km effective resolution of the ASCAT coastal product, but substantially higher than for current Ku-band scatterometers (see Section II-B). Then, the model simulates scales not observed by the instrument. Currently, the observation footprint size is ignored in the observation operator, more precisely it is assumed to be a delta-Dirac function in (7). Similarly as described in Section IV-A1, averaging in model space will then reduce the correlation and variance of the representativeness error and in addition reduce the background error, as given in (9). Averaging in model space may be accomplished through the observation operator, using multiple grid points to simulate the observation, which will reduce (o-b) on average by reducing the second and third term in (11). Larger reductions of (o-b) may be expected for Ku-band scatterometers like QuikSCAT, OSCAT, and HSCAT because of their larger footprint size, hence increased filtering in model space, i.e., more effectively reducing the third term in (11) than for smaller footprints.

Observation minus background statistics were calculated for HARMONIE for the six-week experimental period. Observations of the various scatterometer instruments were compared with 10-m model winds from the CONV-3h experiment described in Section II-D, i.e., the experiment without assimilation of scatterometer winds. Three flavors of the model background were used that are as follows:

1) b, denoting spatial interpolation to the observation location, which reflects the calculation of (o-b) in 3D-Var;
2) (ii) b_t, the index t denoting time interpolation in addition, which would reflect the calculation of (o-b) in a 4D-Var DA system; and
3) (iii) $\bar{b}_t$, the overbar denoting averaging in model space, which reflects the application of an observation operator according to the instrument footprint, here eliminating the spatial representativeness error of observation and model.

Table I shows a clear negative bias for the zonal wind component for all instruments, confirming the earlier finding in Fig. 3 that HARMONIE overestimates the wind speed, in particular for strong winds. A strong reduction of the observation minus background standard deviation in the order of 20%, but a maximum reduction of 33% was observed, is obtained through temporal interpolation of model fields to the observation time, which confirms that the shift between observation time and analysis time is a substantial error source for 3D-Var. It is noted that for 3D-Var with a 3-h assimilation window observations measured ± 1.5 h from analysis time are used in the analysis. For conventional observing systems, the observation time is very close to analysis time with no or negligible mismatch between observation time and the background from the 3-h model forecast. However, scatterometer measurement times do not exactly coincide with analysis times. For each scatterometer instrument, the mean time difference between observation and analysis time, denoted as $\overline{\Delta t}$, and the mean absolute time difference, denoted as $\overline{|\Delta t|}$ are displayed in the table. The former has values in between -1.5 and 1.5, and the latter between 0 and 1.5 for a 3-h assimilation window. Ideally, both are close to zero. Worst case would be that all observations are obtained at one of the assimilation window boundaries, yielding $\overline{\Delta t} = \pm 1.5$ and $\overline{|\Delta t|} = 1.5$. When all observations are at the assimilation window boundaries with half the amount on both sides yields $\overline{\Delta t} = 0$, but $\overline{|\Delta t|} = 1.5$. From the numbers in the table, the time shift is largest for OSCAT and smallest for ASCAT.

TABLE I
SCATTEROMETER OBSERVATION MINUS HARMONIE BACKGROUND STATISTICS FOR THE 10-M ZONAL AND MERIDIONAL WIND COMPONENTS (IN M/S) AND VARIOUS OPTIONS FOR INTERPOLATION AND AVERAGING IN MODEL SPACE, SEE TEXT FOR DETAILS

(m/s)	Bias $u_{10\,\mathrm{m}}$	Stdev $u_{10\,\mathrm{m}}$	Bias $v_{10\,\mathrm{m}}$	Stdev $v_{10\,\mathrm{m}}$
ASCAT-A	(184.508 collocations); $\overline{\Delta t} = -0.28$; $\overline{\lvert\Delta t\rvert} = 0.74$			
$(o - b)$	−0.62	1.84	−0.16	2.04
$(o - b_t)$	−0.58	1.54	−0.18	1.65
$(o - \bar{b}_t)$	−0.57	1.46	−0.18	1.57
ASCAT-B	(193.917 collocations); $\overline{\Delta t} = -0.27$; $\overline{\lvert\Delta t\rvert} = 0.73$			
$(o - b)$	−0.60	1.77	−0.18	1.94
$(o - b_t)$	−0.51	1.56	−0.19	1.63
$(o - \bar{b}_t)$	−0.50	1.49	−0.19	1.56
OSCAT	(4.888 collocations); $\overline{\Delta t} = -0.81$; $\overline{\lvert\Delta t\rvert} = 0.99$			
$(o - b)$	−0.80	2.00	−0.09	2.03
$(o - b_t)$	−0.72	1.72	−0.16	1.52
$(o - \bar{b}_t)$	−0.72	1.68	−0.16	1.47
HSCAT	(23.961 collocations); $\overline{\Delta t} = -0.29$; $\overline{\lvert\Delta t\rvert} = 0.85$			
$(o - b)$	−0.46	1.61	−0.24	1.57
$(o - b_t)$	−0.46	1.36	−0.22	1.29
$(o - \bar{b}_t)$	−0.45	1.25	−0.22	1.18

In short: b spatial interpolation to observation location; b_t temporal interpolation in addition; $\bar{b}_t$ temporal interpolation plus averaging in model space according to the instrument footprint. $\overline{\Delta t}$ and $\overline{|\Delta t|}$ denote the mean time difference and the mean absolute time difference between observation and analysis time (in hours), respectively.

TABLE II
STATISTICS OF $(o - f_t)$ FOR 10-M ZONAL AND MERIDIONAL WIND COMPONENTS FROM ECMWF MODEL MODEL FIELDS AND FOUR SCATTEROMETER INSTRUMENTS

	t_f	Bias $u_{r\,10\,\mathrm{m}}$	Stdev $u_{10\,\mathrm{m}}$	Bias $v_{10\,\mathrm{m}}$	Stdev $v_{10\,\mathrm{m}}$
ASCAT-A	9.4	−0.22	1.52	0.34	1.67
ASCAT-B	9.5	−0.16	1.56	0.28	1.73
OSCAT	11.0	−0.31	1.41	0.41	1.41
HSCAT	5.6	−0.11	1.09	0.05	1.15

Collocated ECMWF data are obtained from the operational scatterometer products and include spatial and temporal interpolation to observation location and measurement time. The average lead time for collocation with ECMWF model forecasts is denoted as t_f (in hours). The number of collocations used in the statistics is given in Table I.

TABLE III
INFERRED MODEL ERROR (RIGHTMOST COLUMN) OF HARMONIE (HA) RELATIVE TO ECMWF (EC) FROM OBSERVATION MINUS FORECAST STATISTICS IN TABLES I AND II AND EFFECTIVE RESOLUTION OF THE VARIOUS SCATTEROMETER INSTRUMENTS

HA	$\sigma^2(o-b_t)$	$=$	$\sigma^2(\epsilon_i)$	$+$	$\sigma^2(\epsilon_r)$	$+$	$\sigma^2(\epsilon^{b_t})$
ASCAT	$\approx$		$=$		$<$		$>$
HSCAT	$>$		$=$		$\approx$		$>$
OSCAT	$>$		$=$		$>$		?
EC	$\sigma^2(o-f_t)$	$=$	$\sigma^2(\epsilon_i)$	$+$	$\sigma^2(\epsilon_r)$	$+$	$\sigma^2(\epsilon^{f_t})$

The question mark denotes inconclusive. $\sigma^2(.)$ denotes the variance of the variable between brackets. The top row is obtained from taking the variance of (11). Similar for the bottom row but now using forecasts from ECMWF rather than HARMONIE background.

Averaging in model space was done over a square window area, centered at the observation location. For ASCAT-A and ASCAT-B, the window size was chosen 28 km, i.e., equal to the effective resolution, which is about double the sampling distance. Studies on the effective resolution of OSCAT and HSCAT are not available and the window size was chosen double the sampling distance, i.e., 100 km for OSCAT and 50 km for HSCAT. Averaging in model space indeed yields a further reduction of the observation minus background standard deviation of 3–8%. This suggests that reducing "model noise," i.e., reducing model variability on scales, which the observation does not resolve, makes the model state more representative to the observation, hence reducing (o-b).

The operational scatterometer products disseminated by KNMI include winds from the ECMWF model for reference. This includes spatial and temporal interpolation from model winds to scatterometer wind locations and measurement times. The forecast fields available for time interpolation depend on the time schedule of the ECMWF forecasting system and dissemination. Forecasts initiated at 00 and 12 UTC are available to ECMWF member states about 6 h later. Then, from the satellite overpass times the closest ECMWF model fields are used for time interpolation. The average forecast lead time for collocation with scatterometer observations is denoted as t_f, which equals 9.4, 11.0, and 5.6 h for ASCAT, OSCAT, and HSCAT, respectively. Here, it is noted that HSCAT is currently not an operational product and processing can wait until the most recent forecasts from ECMWF are available. Table II shows the statistics of resulting observations minus ECMWF forecast including time interpolation, $(o - f_t)$, for the various scatterometers using the November/December 2013 ECMWF operational model version. ECMWF shows smaller biases for the zonal wind component but larger biases for the meridional wind component when compared to HARMONIE in Table I. The best statistics for HSCAT may be partly explained by the smallest value for t_f, but we note that the position of HSCAT, OSCAT, and ASCAT in the 4D-Var assimilation window is systematically different due to different LTAN. It is further noted that t_f is 2.2–2.7 h for HARMONIE against 5.6–11.0 h for ECMWF.

For ASCAT-A and ASCAT-B standard deviations of $(o - b_t)$ for HARMONIE are similar to $(o - f_t)$ for ECMWF. From (11), noting that the instrument error irrespective of the model and the representativeness error of ASCAT is smaller for HARMONIE than ECMWF, it is inferred that the model error must be larger for HARMONIE. This is summarized in Table III, which shows the inferred model error (rightmost column) of HARMONIE (HA) relative to ECMWF (EC) from observation minus forecast statistics in Tables I and II and the effective resolution of the various scatterometer instruments. For OSCAT and HSCAT $(o - f_t)$, statistics are generally better for ECMWF than $(o - b_t)$ for HARMONIE, despite the substantially larger forecast range needed for collocation with the ECMWF model. This may be explained partly because the effective resolution of these scatterometers is closer to the effective resolution of ECMWF, yielding a smaller representativeness error, in particular for OSCAT. More likely is that the HARMONIE model errors are substantially larger than for ECMWF. Assuming that both models perform equally well on the large scales, which has not been verified according to the authors knowledge, it is concluded that the additional energy of the HARMONIE model state on the turbulent scales contributes to the model error substantially, when quantified with standard root-mean-square-error (RMSE) metrics.

Exact quantification of the model errors requires estimates of the instrument and representativeness error, which may be obtained from triple-collocation techniques in principle. Yet, a rough estimate of HARMONIE model errors is obtained as follows. From Table III with focus on ASCAT: HARMONIE $\sigma^2(o - b_t) \approx$ ECMWF $\sigma^2(o - f_t)$. For HARMONIE, the background b_t is a 3-h forecast and for ECMWF f_t is a 9.4-h forecast. For HARMONIE, substituting in the top row of Table III yields: $\sigma^2_{\mathbf{HA}}(o - f_{3\mathrm{h}}) = \sigma^2(\epsilon_i) + \sigma^2_{\mathbf{HA}}(\epsilon_r) + \sigma^2_{\mathbf{HA}}(\epsilon^{f_{3\mathrm{h}}})$ and for ECMWF in the bottom row: $\sigma^2_{\mathbf{EC}}(o - f_{9.4\mathrm{h}}) = \sigma^2(\epsilon_i) + \sigma^2_{\mathbf{EC}}(\epsilon_r) + \sigma^2_{\mathbf{EC}}(\epsilon^{f_{9.4\mathrm{h}}})$. Subtracting both equations and noting that 1) $\sigma^2_{\mathbf{HA}}(o - f_{3\mathrm{h}}) \approx \sigma^2_{\mathbf{EC}}(o - f_{9.4\mathrm{h}})$, and 2) ASCAT and HARMONIE effective resolution are similar, i.e., $\sigma^2_{\mathbf{HA}}(\epsilon_r) \approx 0$ yields: $\sigma^2_{\mathbf{HA}}(\epsilon^{f_{3\mathrm{h}}}) \approx \sigma^2_{\mathbf{EC}}(\epsilon^{f_{9.4\mathrm{h}}}) + \sigma^2_{\mathbf{EC}}(\epsilon_r)$. If we further split the HARMONIE model error in two independent parts, representing the spatial scales resolved by the ECMWF model and remaining (small) spatial scales and assuming linear growth of ECMWF model errors over the first 10 h [26], and assuming that ECMWF and HARMONIE perform equally well on scales that ECMWF resolves (note that HARMONIE starts daily from ECMWF forecasts and uses ECMWF boundaries), yields: $\sigma^2_{\mathbf{HA(smallscales)}}(\epsilon^{f_{3\mathrm{h}}}) \approx \sigma^2_{\mathbf{EC}}(\epsilon^{f_{6.4\mathrm{h}}}) + \sigma^2_{\mathbf{EC}}(\epsilon_r)$, i.e., the HARMONIE model error on spatial scales not resolved by the ECMWF model can be approximated as the sum of the 6.4-h ECMWF model forecast error and the representativeness error of ASCAT in the ECMWF model. We note that for OSCAT on scales larger than 100 km, ECMWF has a better verification than HARMONIE. This implies that the variance added over the ocean by HARMONIE on these scales appears not to verify deterministically with scatterometer observations.

V. DISCUSSION

Mesoscale DA is still in its infancy and the additional value of observations for weather forecasting is currently limited to the first couple of hours. In general, forecast skill scores of HARMONIE are worse than for ECMWF (not shown here). Further progress in mesoscale DA may be achieved through the following notions, with focus on the HARMONIE model.

In general, one can only expect improvements on small scales when the large scales are correct. Therefore, one should verify that the mesoscale model performance on the large scales is compatible with the model in which it is ingested and which provides the lateral boundary conditions. The latter is generally a lower resolution version of the same model or a lower resolution global model. For HARMONIE, this means that its performance should be compatible or better to that of ECMWF on the scales that the latter resolves. To the authors' knowledge, this has not been yet verified for HARMONIE. A first target for HARMONIE could be to improve relative to ECMWF on the large scales, from faster cycling, hence ingesting recent observations more quickly than is done by ECMWF.

Correct specification of the background error covariance matrix $\mathbf{B}$ is a continuous challenge in DA, both for global and mesoscale models. For global models, $\mathbf{B}$ mainly imposes atmospheric balances at synoptic scales (>100–200 km) such as geostrophy. This specification may not be optimal on scales resolved by nonhydrostatic models where strong interactions between wind, temperature, and humidity dominate. As mentioned in Section II-C, as a pragmatic solution, HARMONIE uses a climatological $\mathbf{B}'$ derived from downscaled ECMWF ensemble members and imposes the assumption of homogeneity and isotropy of forecast error statistics. The resulting $\mathbf{B}'$ matrix structure functions produce large scale symmetric increments, which may be far from optimal for convective scale phenomena. The implementation of a flow-dependent $\mathbf{B}'$ matrix, better describing the atmospheric interactions on the mesoscale, is foreseen in the near future. However, there is a caveat here, related to the observation density as explained below.

Perfect knowledge of background and observation error characteristics would allow ingestion of all available good quality observations and contribute to an improved model state. Mesoscale models allow increasing variability on small spatial scales in their model state. As a consequence, the structure functions of the corresponding $\mathbf{B}$ matrix become smaller. In operational practice, $\mathbf{B}$ and $\mathbf{R}$ are not well known and pragmatic solutions applied to yield $\mathbf{B}'$ as mentioned above and a diagonal $\mathbf{R}'$. Marseille *et al.* [18] showed, using passive (independent not assimilated) observations, that increments degrade the model state further away from assimilated observations when using pragmatic $\mathbf{B}'$ and $\mathbf{R}'$. Introducing increasingly smaller, but imperfect, structure functions in $\mathbf{B}'$ is expected to degrade the model state increasingly closer to assimilated observations. A high density observation network is then required to prevent the model state to degrade away from assimilated observations.

However, the density of the observation network currently used in HARMONIE is too coarse, both spatial and temporal, to correctly initiate atmospheric processes on scales that the model can resolve. Hence, the 4-D observation density should be paramount on the model spatial scales to be adapted deterministically in DA. This may be achieved by constructing the structure functions of $\mathbf{B}'$ such that its filtering properties allow for a deterministic analysis on spatial scales prescribed by the density of the observation network. This prevents the introduction of variability in the model state on scales that the DA system cannot resolve deterministically and is therefore probably incorrect, which after evolution might also degrade the model state large scales. Increased (o-b) statistics of HARMONIE versus ECMWF, as discussed in Section IV, are an indication of increased variance in the HARMONIE model state, which does not verify deterministically. Spatial scales that cannot be analyzed deterministically should be treated in a probabilistic sense, hence removed (temporary) before DA. Running HARMONIE ensembles is foreseen in the near future. The ensemble mean is a smooth representation of the model state and as such a potential candidate to serve as input for DA. This needs further research. Potential alternative candidates are averaging in spatial domain or the application of a low-pass filter in spectral domain.

The DA concept is based on the assumption of an unbiased model and observations. HARMONIE shows an overestimate of the 10-m wind speed over the oceans, in particular for strong wind events. These biases should be removed for observations to have consistent dynamical impact on forecasts. Implementation of a new turbulence scheme in HARMONIE already shows better agreement of 10-m ocean winds with scatterometer observations however it is questionable if the turbulence scheme can be blamed for the overestimations during storm conditions. More likely, the main cause can be found in the applied surface stress roughness relationship. Nevertheless, reducing model biases is crucial for the effective use of observations in DA [27].

Finally, profile observations from radiosondes are assumed measured at launch location, ignoring drift from the launch location. This simplification introduces errors that can be corrected with relative ease.

VI. Conclusion

In a 3D-Var DA system, a shift between observation and analysis time contributes to the total error in the background departure, also denoted as innovation. This spatially correlated error propagates in the analysis increment hence reducing the quality of the resulting analysis. The current 3D-Var implementation of HARMONIE assumes that all observations are taken at analysis time, not correcting for any time shifts. For conventional (synoptic) observing systems such as radiosondes, synop stations, and buoys, this time shift is generally small. However, the time shift may be substantial for observations from asynoptic observing systems such as aircrafts and satellites. For satellite winds from scatterometer over the Atlantic, it was found that the time shift is in the order of 0.7–1 h on average, depending on the satellite orbit, for a 3-h assimilation window. Ignoring the time shift increases the innovation standard deviation by about 20% for scatterometer, but can be up to 33%.

Currently, all observations are used as point observations by HARMONIE. However, with continuously decreasing model grid sizes this assumption may no longer be adequate. A further reduction of innovation errors is obtained when taking into account the instrument footprint into the observation operator as part of the analysis equations. For scatterometer, a further reduction of 3–8% is achieved for the innovation standard deviation when introducing the observation operator in addition to resolving the time shift issue. These findings motivate the use of a 4D-Var assimilation system for HARMONIE and in addition the introduction of more advanced observation operators taking into account instrument footprints.

Comparing statistics of observation minus model between the nonhydrostatic convection-resolving HARMONIE model and the hydrostatic, relatively smooth, ECMWF model showed that adding variability to the HARMONIE model state on convective scales contributes to the model error substantially. In other words, mesoscale models look realistic but they are not real. This is largely explained by the incorrect positioning of small-scale rapidly evolving atmospheric phenomena, which are explicitly resolved by mesoscale models. Incorrect positioning of weather systems is strongly penalized when verified against observations using an RMSE-based metric. Nowadays DA in operational NWP is based on RMSE metrics and is inherently deterministic. Correct positioning of rapidly evolving small-scale phenomena then requires a high density observation network in all four dimensions if one aims to correctly initialize these phenomena deterministically. These observations may be provided by satellites, but for example also by ground-based radars, kilometer-spaced airplane readings [28], solar or wind production information from energy providers, etc.

Limitations of current mesoscale DA systems were discussed. Based on the notion that DA is inherently a deterministic approach, it was concluded that correct initialization of small-scale weather phenomena requires a dense network of observations in all four dimensions. Data sparsity then induces phase errors, i.e., incorrect positioning of weather systems through DA. This may be alleviated by constructing the structure functions of the prescribed background error covariance matrix such that its filtering properties allow for a deterministic analysis on model spatial scales prescribed by the density of the observation network.

References

[1] W. Lahoz, B. Khattatov, and R. Menard, Eds., *Data Assimilation: Making Sense of Observations*. Berlin, Germany: Springer-Verlag, 2010, doi: 10.1007/978-3-540-74703-1_1.

[2] W. C. Skamarock, “Evaluating mesoscale NWP models using kinetic energy spectra,” *Monthly Weather Rev.*, vol. 132, pp. 3019–3032, 2004.

[3] J. Vogelzang, A. Stoffelen, A. Verhoef, J. de Vries, and H. Bonekamp, “Validation of two-dimensional variational ambiguity removal on SeaWinds scatterometer data,” *J. Atmospheric Ocean. Technol.*, vol. 7, no. 26, pp. 1229–1245, 2009, doi: 10.1175/2008JTECHA1232.1.

[4] A. Stoffelen *et al.*, “The atmospheric dynamics mission for global wind measurement,” *Bull. Amer. Meteorol. Soc.*, vol. 86, pp. 73–87, 2005.

[5] P. Laloyaux, J. N. Thépaut, and D. Dee, “Impact of scatterometer surface wind data in the ECMWF coupled assimilation system,” *Monthly Weather Rev.*, vol. 144, pp. 1203–1217, 2016, doi: 10.1175/MWR-D-15-0084.1.

[6] Y. Seity *et al.*, “The AROME-France convective-scale operational model,” *Monthly Weather Rev.*, vol. 139, pp. 976–991, 2011.

[7] A. Stoffelen and D. Anderson, “Scatterometer data interpretation: Measurement space and inversion,” *J. Atmospheric Ocean. Technol.*, vol. 14, pp. 1298–1313, 1997.

[8] M. Portabella, and A. Stoffelen, “A probabilistic approach for SeaWinds data assimilation,” *Quart. J. Royal Meteorol. Soc.*, vol. 130, pp. 127–152, 2004, doi: 10.1256/qj.02.205.

[9] A. Verhoef, M. Portabella, and A. Stoffelen, “High-resolution ASCAT scatterometer winds near the coast,” *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2481–2487, Jul. 2012.

[10] W. Lin, M. Portabella, A. Stoffelen, A. Verhoef, and A. Turiel, “ASCAT wind quality control near rain,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4165–4177, Aug. 2015, doi: 10.1109/TGRS.2015.2392372.

[11] J. Figa-Saldaña, J. W. Wilson, E. Attema, R. Gelsthorpe, M. R. Drinkwater, and A. Stoffelen, “The advanced scatterometer (ASCAT) on the meteorological operational (MetOp) platform: A follow on for the European wind scatterometers,” *Can. J. Remote Sens.*, vol. 28, no. 3, pp. 404–412, 2002, doi: 10.5589/m02-035.

[12] A. Stoffelen *et al.*, “Research and development in Europe on global application of the OceanSat-2 scatterometer winds,” Final Rep. OceanSat-2 Cal/Val AO Project, Koninklijk Nederlands Meteorologisch Instituut, De Bilt, The Netherlands, Rep. NWPSAF-KN-TR-022, SAF/OSI/CDOP2/KNMI/TEC/RP/196, 2013.

[13] J. Vogelzang, A. Stoffelen, R. Lindsey, A. Verhoef, and J. Verspeek, “The ASCAT-6.25 wind product,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 2016.

[14] P. Courtier *et al.*, “The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation,” *Quart. J. Royal Meteorol. Soc.*, vol. 124, pp. 1783–1808, 1998.

[15] K. Ide, P. Courtier, M. Ghil, and A. C. Lorenc, “Unified notation for data assimilation: Operational, sequential and variational,” *J. Meteorol. Soc. Jpn.*, vol. 75, pp. 181–189, 1997.

[16] R. Daley, *Atmospheric Data Analysis* (Cambridge Atmospheric and Space Science Series). Cambridge, U.K.: Cambridge Univ. Press, 1999.

[17] A. Lorenc, “Analysis methods for numerical weather prediction,” *Quart. J. Royal Meteorol. Soc.*, vol. 112, pp. 1177–1194, 1986.

[18] G. J. Marseille, J. Barkmeijer, S. de Haan, and W. Verkley, “Data assimilation tuning and impact assessment using observations,” *Quart. J. Royal Meteorol. Soc.*, 2016, to be published, doi: 10.1002/qj.2882. Q4

[19] S. Massart, B. Pajot, A. Piacentini, and O. Pannekoucke, “On the merits of using a 3D-FGAT assimilation scheme with an outer loop for atmospheric situations governed by transport,” *Monthly Weather Rev.*, vol. 138, pp. 4509–4522, 2010, doi: 10.1175/2010MWR3237.1.

[20] S. de Haan and S. H. van der Veen, “Cloud initialization in the rapid update cycle of HIRLAM,” *Weather Forecasting*, vol. 29, pp. 1120–1133, 2014, doi: 10.1175/WAF-D-13-00071.1.

[21] J. Vogelzang, A. Stoffelen, A. Verhoef, and J. Figa-Saldaña, “On the quality of high-resolution scatterometer winds,” *J. Geophys. Res.*, vol. 116, 2011, Art. no. C10033, doi: 10.1029/2010JC006640.

[22] J. Vogelzang, G. P. King, and A. Stoffelen, “Spatial variances of wind fields: Their sensitivity to sampling strategy and their relation to second-order structure functions and spectra,” *J. Geophys. Res.*, vol. 120, 2015, Art. no. C10033, doi: 10.1002/2014JC010239.

[23] P. P. Weston, W. Bell, and J. R. Eyre, “Accounting for correlated error in the assimilation of high-resolution sounder data,” *Quart. J. Royal Meteorol. Soc.*, vol. 140, pp. 2420–2429, 2014.

[24] W. Lin, M. Portabella, A. Stoffelen, J. Vogelzang, and A. Verhoef, “ASCAT wind quality under high subcell wind variability conditions,” *J. Geophys. Res.*, vol. 120, no. 8, pp. 5804–5819, 2015, doi: 10.1002/2015JC010861.

[25] R. Frehlich, “The definition of ‘truth’ for numerical weather prediction error statistics,” *Quart. J. Royal Meteorol. Soc.*, vol. 137, pp. 84–98, 2011.

[26] L. Magnusson and E. Källén, “Factors influencing skill improvements in the ECMWF forecasting system,” *Monthly Weather Rev.*, vol. 141, pp. 3142–3153, 2013, doi: 10.1175/MWR-D-12-00318.1.

[27] W. de Rooy *et al.*, “Harmonie verification and evaluation,” HirLAM Aladin Report, to be published, 2016.

[28] S. de Haan, “High-resolution wind and temperature observations from aircraft tracked by Mode-S air traffic control radar,” *J. Geophys. Res.*, vol. 116, 2011, Art. no. D10111, doi: 10.1029/2010JD015264.

Gert-Jan Marseille was born in the Netherlands, in March 4, 1966. He received the M.Sc. degree in applied mathematics and the Ph.D. degree in MRI scan time reduction from Delft University of Technology, Delft, The Netherlands, in 1990 and 1997, respectively.

He is a Senior Scientist in the Royal Dutch Meteorological Institute (KNMI), De Bilt, The Netherlands, in the Research and Development Satellite Observations section, working on data assimilation and the European Space Agency ADM-Aeolus Doppler Wind Lidar mission.

Ad Stoffelen was born in the Netherlands, in February 25, 1962. He received the M.Sc. degree in physics from Eindhoven University of Technology, Eindhoven, The Netherlands, and the Ph.D. degree in meteorology on scatterometry from Utrecht University, Utrecht, The Netherlands, in 1987 and 1998, respectively.

He leads a group on satellite wind sensing in the Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands, and is responsible for scatterometer wind products on behalf of the European Organisation for the Exploitation of Meteorological Satellites. He is also much involved in the European Space Agency ADM-Aeolus Doppler Wind Lidar mission.

QUERIES

Q1. Authors: When accessing and uploading your corrections at the Author Gateway, please note we cannot accept new source files as corrections for your paper. Do not send new Latex, Word, or PDF files, as we cannot simply ”overwrite” your paper. Please submit your corrections as an annotated PDF or as clearly written list of corrections, with location in paper. You can also upload revised graphics to the Gateway.

Q2. Authors: Unless invited or otherwise informed, a mandatory Excessive Paper Length charge of $200.00 per page (beginning with page 7 and beyond) is required for papers in excess of six (6) printed pages for JSTARS. If you have any questions regarding overlength page charges, need an invoice, or have any other billing questions, please contact reprints@ieee.org as they handle these billing requests.

Q3. Author: Please check affiliations as set for correctness.

Q4. Author: Please update Refs. [13], [18], and [27].