TECHNICAL MEMORANDUM

629

Assimilation of scatterometer data as equivalent-neutral wind

Hans Hersbach

Research Department

July 2010

Abstract

This document describes the assimilation of scatterometer data as equivalent-neutral 10m vector wind into the four-dimensional variational data assimilation (4D-Var) component of the integrated forecasting system (IFS) at the European Centre for Medium-Range Weather Forecasts (ECMWF). For given surface stress this quantity (also simply denoted by neutral wind) provides the wind at 10m height for which stability effects in the surface layer (SL) have been neglected. Sofar, at ECMWF, scatterometer data has been assimilated as 10m (non-neutral) wind, i.e., the actual wind at 10m height including stability effects. Since it is believed that scatterometer data relates more closely to surface stress than wind, the usage of an observation operator that is sensitive to neutral rather than to non-neutral wind should be more accurate.

Although it is straightforward to adapt the standard observation operator for surface wind, it is realized that the current assimilation system uses an estimation of exchange coefficients that is based on an old version of the ECMWF SL parametrization. Such coefficients are required to perform a proper vertical interpolation from the lowest model-level wind.

The reason for this is that only a limited set of model information is passed to the location where the observation operator is evaluated. In this document it is described how neutral wind can be included into this set, such that the correct information from the actual SL can be accessed. This also embraces the incorporation of the influence of the 4D-Var control vector in the minimization on this diagnostic surface field. Although not further explored here, this latter extension could be applied to other diagnostic surface fields as well.

Several assimilation experiments are performed at a resolution of T511 in early-delivery mode for 109 cases in the Autumn of 2009. From these it is verified that departures between scatterometer and model wind speed are slightly reduced when an observation operator for neutral wind is used. Impact on forecast skill is found to be relatively neutral, with some (not significant) positive impact over the Southern Hemisphere for the atmosphere and ocean surface waves. Most favourable results are obtained for an experiment in which neutral wind is fetched from the actual ECMWF SL.

The usage of scatterometer data as neutral wind has become the default configuration in IFS cycle 36r3, which represents the starting point of a future cycle (36r4) for the currently operational suite (36r2).

1   Introduction

Space-born scatterometer data provide accurate information on speed and direction of surface wind over the global oceans. Since the launch of the ERS-1 satellite in July 1991, global coverage of scatterometer data has been available without interruption. Applications vary from near-real time assimilation into numerical weather prediction models (NWP), the forcing of ocean models, to climate studies accessing the now 19-year data record.

At the European Centre for Medium-Range Weather Forecasts (ECMWF) scatterometer winds have been assimilated in the operational integrated forecasting system (IFS) from 30 January 1996 onwards. The four-dimensional variational assimilation system at ECMWF (4D-Var) allows for a dynamically consistent use of observations. In this way, information of scatterometer surface winds is propagated to the entire troposphere (Isaksen and Janssen, 2004). Currently (May 2010) data is used from the AMI scatterometer on-board the European Remote sensing Satellites ERS-2 (from June 1996 onwards), from the ASCAT instrument on the MetOp-A platform (from June 2007 onwards), and data from the SeaWinds instrument on-board QuikSCAT has been used from February 2002 until its failure in November 2009.

A scatterometer is a microwave radar that emits pulses at well-defined frequency and polarization to the Earth surface. A backscatter is recorded, from which over the global oceans information on the local surface wind conditions can be obtained. The main physical process is based on Bragg scattering where backscatter is related

to the intensity of surface water waves with wavelengths that are comparable to that of the emitted pulse. By choice of the scatterometer wavelength in the centimetre range, the strength of gravity-capillary surface waves is sensed. These in turn, are determined by the local surface stress $\vec{\tau}$, or in effect, the local surface wind condition. Since backscatter response also depends on the relative angle between the incident pulse and capillary wave direction, information on wind direction can be extracted as well. In practise, an empirical relation (called geophysical model function, GMF) is established between backscatter and 10m vector wind $\mathbf{u}_{10}$.

The connection between stress and 10m wind depends among other quantities on the stability of the surface layer (SL). If one assumes a constant stress layer at the surface and a form of Monin-Obukhov stability theory (Monin and Obukov, 1954), this relation can be estimated as:

$$\mathbf{u}(z) = \frac{\mathbf{u}_*}{\kappa} \left\{ \log\left(\frac{z+z_0}{z_0}\right) - \Psi_{\mathrm{M}}(\frac{z+z_0}{L}) + \Psi_{\mathrm{M}}(\frac{z_0}{L}) \right\} + \mathbf{u}_{\mathrm{oc}}, \tag{1}$$

where $\vec{\tau} = \rho_a u_* \mathbf{u}_*$, with ρ_a air density, $\mathbf{u}_*$ the friction velocity and u_* its magnitude, $z = 10\mathrm{m}$, and $\kappa = 0.4$ is the von Kármán constant. This relation in principle depends on variations in atmospheric stability (expressed by the stability function Ψ_{M} and Obukhov length L), air density, and ocean surface current $\mathbf{u}_{\mathrm{oc}}$. In the ECMWF formulation roughness length z_0 over the ocean depends for light wind on the kinematic viscosity ν ($1.5x10^{-5}\mathrm{m\,s^{-1}}$) and on a Charnock relation (Charnock, 1955) for stronger winds

$$z_0 = \alpha_{\mathrm{M}} \frac{\nu}{u_*} + \alpha_{\mathrm{ch}} \frac{u_*^2}{g}. \tag{2}$$

Here $\alpha_{\mathrm{M}} = 0.11$, $g = 9.80665\mathrm{m\,s^{-2}}$ is the gravitational acceleration, and α_{ch} depends on the (ocean-wave) sea state (Janssen, 1991). This introduces a sea-state dependency on the relation between stress and surface wind as well.

At ECMWF scatterometer wind is assimilated as 10m wind $\mathbf{u}_{10}$. Variations in stability, air density, ocean current and sea state (which affect the relation with stress) are, therefore, not accounted for. Although such fluctuations may be small on average, locally it may have a non negligible effect. For atmospheric stability, e.g., locally seasonally dependent differences appear. An estimation of such effects within the ECMWF framework can be found in Hersbach (2010a).

To address the issue of stability, the concept of equivalent-neutral wind is popular. It represents the relation between stress and wind in case stability effects are neglected. The neutral wind $\mathbf{u}_n(z)$ at height z is given by:

$$\mathbf{u}_n(z) = \frac{\mathbf{u}_*}{\kappa} \log(1 + z/z_0). \tag{3}$$

Such winds (from now on simply denoted by neutral), therefore, represent the wind (usually at 10-metre height) for given surface stress in case the marine boundary layer were neutrally stratified. On average, the marine boundary layer is weakly unstable, and the global average 10m neutral wind appears $\sim 0.2\mathrm{m\,s^{-1}}$ stronger than the non-neutral wind (see *e.g.* Brown *et al.* (2006)). For QuikSCAT, the empirical relation between scatterometer backscatter and wind has been trained on neutral wind (NSCAT-2, QSCAT-1 GMF, see Wentz and Smith (1999); Freilich and Dunbar (1999); Ebuchi *et al.* (2002)). For ERS-2 and ASCAT such relation has been based on (non-neutral) wind (CMOD5 GMF, see Hersbach *et al.* (2007)). Recently an extension for neutral wind has become available (CMOD5.N GMF, see Verhoef *et al.* (2008); Hersbach (2010a)).

In this manuscript it is described how scatterometer data can be assimilated as neutral wind in the ECMWF assimilation system. This can be achieved by an appropriate extension of the generic observation operator for surface wind. In Section 2 a brief overview of the ECMWF 4D-Var system is presented. Details on the extension of an observation operator for neutral wind are given in Section 3. While doing so, a few issues are addressed. Some technical complications are described in Section 4. All necessary modifications have

been compiled in a model branch from IFS cycle 36r1 (dal_CY36R1_neutral_full_dependencies), which was later merged into IFS cycle 36r3. Some new features are discussed in Section 5, while a more extensive list is provided in Appendix B. The results of a number of impact studies are described in Section 6, and the document ends in Section 7 with a discussion.

A list of acronyms as used and IFS subroutines as mentioned in this document may be found in Appendix A.

2   Four-dimensional data assimilation (4D-Var)

ECMWF uses the method of incremental four-dimensional data assimilation (denoted by 4D-Var, Courtier *et al.* (1994)). For a given assimilation window, data is collected and compared to the model state via a cost function J that is to be minimized with respect to an increment $\delta\mathbf{x}$ that corrects the background $\mathbf{x}^{\mathrm{b}}$ at the start of the assimilation window. Schematically, J is given by:

$$J(\delta\mathbf{x}) = J_{\mathrm{b}} + J_{\mathrm{o}} = \frac{1}{2}\delta\mathbf{x}^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{d})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}). \tag{4}$$

The background is a short forecast from the previous analysis cycle. The J_{b} term expresses the confidence in this field via the background error covariance matrix $\mathbf{B}$. At the end of the minimization, the final increment $\delta\mathbf{x}^{\mathrm{a}}$ is added to the background $\mathbf{x}^{\mathrm{b}}$ to provide the analysis $\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \delta\mathbf{x}^{\mathrm{a}}$.

The comparison between model and data is obtained via an observation operator H. It expresses what the value of observation should be according to the model. In 4D-Var the comparison between model and observation incorporates the timing of the observation as well. As a function of initial model state, H therefore includes a model integration from initial time to observation time. In (4), $\mathbf{H}$ is a suitable linearization of H around the background. The covariance matrix of observation errors $\mathbf{R}$ expresses the accuracy of the observational network. The innovation vector $\mathbf{d}$, also called the first-guess departure, is given by:

$$\mathbf{d} = \mathbf{y}^{\mathrm{o}} - H(\mathbf{x}^{\mathrm{b}}), \tag{5}$$

where the set of observations within the assimilation window is represented as a vector $\mathbf{y}^{\mathrm{o}}$. The model integration started from the background over the entire assimilation window, that is required to calculate H and innovation d for each observation in (5), is called the trajectory run (or outer loop). It is performed at the same resolution as the forecast model. Minimization of (4), though, is performed at a reduced resolution (inner loop). It requires tangent-linear (TL) and adjoint (AD) equations for the entire path from increment $\delta\mathbf{x}$ to observation cost function (which embraces a model integration to appropriate time plus the translation from the resulting model state to observation equivalent). The sequence of outer and inner loop is iterated a few times, in which the high resolution trajectory is readily updated (currently there are three inner loops). In a final trajectory run analysis departures are evaluated:

$$\mathbf{d}^{a} = \mathbf{y}^{\mathrm{o}} - H(\mathbf{x}^{\mathrm{a}}). \tag{6}$$

A more detailed description may be found in Part II of the IFS-documentation (2009).

2.1   Some relevant technical details

The evaluation of the non-linear observation operator H is handled by the routine `hop`. Here, required model information is fetched from a set of arrays in memory. These GOM arrays include values from model fields at the appropriate time step which have been horizontally interpolated to observation location as well (in a routine `obshor`). Besides wind, temperature, humidity and a number of other quantities on model levels the
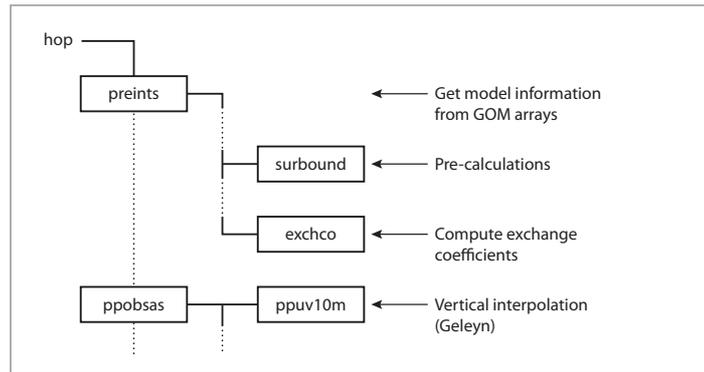
Figure 1: Snippet of the flow chart for the observation operator `hop`

GOM arrays contain information on a number of model surface fields. During the integration of the non-linear trajectory, the GOM arrays are filled step by step. This is managed by routines `cobs` and `cobslag`, which are part of a routine `scan2m`. This latter routine also contains routine `gp_model` that takes care of all model computations in grid-point space. This includes the ECMWF physics parametrization (`ec_phys/callpar`). The actual computation of H occurs at the end of the trajectory in one go for all observations, i.e., after all GOM arrays have been filled.

The evolution of TL and AD perturbations applied to linearization **H** occurs in routines `hoptl` and `hopad`. For model quantities on model levels the influx of linear perturbations and the passing back of AD dependencies are channelled by associated GOM arrays. For model surface fields such arrays are not used, which effectively means that all perturbations and dependencies in these quantities are neglected. Since surface fields are not part of the control vector **x**, this does make sense for 3D-Var (which was the operational environment when GOM arrays were introduced). In 4D-Var, however, where many surface fields are updated diagnosticly each time step (in `callpar`), such fields are affected by changes in the control vector at initial time, and strictly speaking, dependencies should be incorporated.

Surface fields reside in memory via a dedicated module (`surface_fields_mix`) where they are bundled into specific groups. Routines are available (in `su_surf_flds`) that greatly facilitate the inclusion of a new field, its initialization, and its inclusion into the interpolated trajectory that is used in the minimization. Other operations, such as archiving at specified intervals, updating in `ec_phys` and communication with GOM arrays, is more tedious.

3   Observation operator for equivalent neutral 10m wind

For surface data the non-linear observation operator H is handled in the routines `preints` and `ppobsas`, which are both called in `hop` (Cardinali *et al.*, 1994; Vasiljevic *et al.*, 1992). A flow diagram is presented in Fig. (1). Routine `preints` is responsible for the pre-calculation of a number of quantities, while the actual evaluation of H takes place in `ppobsas`. At the start of `preints` required model information is fetched from the GOM arrays. From these horizontally interpolated quantities, exchange coefficients for momentum and heat are determined, which are required in the vertical interpolation from lowest-model level to observation height. This latter step is performed in `ppobsas`.

3.1 Vertical interpolation

For operators that involve observations of wind or temperature below the lowest model level l (currently around 10m) the interpolation method of Geleyn (1988) is used. It is based on Monin-Obukhov theory in which simplified versions for stability functions are chosen. For the interpolation of wind, this method requires the relative wind $\mathbf{u}_l \equiv \mathbf{u}(z_l) - \mathbf{u}_c$ between the lowest model wind $\mathbf{u}(z_l)$ at height z_l and surface current $\mathbf{u}_c$, as well as the knowledge of the exchange coefficients b_N and b_D. These are defined as:

$$
\begin{aligned}
b_N &= \log(1 + z_l/z_0), &(7)\\
b_D &= \kappa\left(||\mathbf{u_l}||/||\mathbf{u_*}||\right), &(8)
\end{aligned}
$$

and are calculated at an earlier point of the code (see next subsection). Here z_0 is the surface roughness length for momentum and $\mathbf{u}_*$ the friction velocity. The interpolated model wind $\mathbf{u}(z)$ at a height z is estimated as:

$$
\mathbf{u}(z) - \mathbf{u}_c = \frac{\mathbf{u}_l}{b_D}\left[\log\left(1 + (z/z_l)(e^{b_N} - 1)\right) - S\right] \tag{9}
$$

where

$$
S = \begin{cases} (z/z_l)(b_N - b_D) & \text{stable case,}\\ -\log\left(1 + (z/z_l)(e^{(b_N - b_D)} - 1)\right) & \text{unstable case.} \end{cases} \tag{10}
$$

Although based on simplified stability functions $\phi(\eta) = 1 + \alpha_s\eta$ for stable and $\phi(\eta) = (1 - \alpha_u\eta)^{-1}$ for unstable stratification, where η is the ratio between height and Obukhov length, Geleyn (1988) shows that (9) provides a good estimate for the near-surface wind profile, and has correct values for $z = z_l$ and $z = 0$.

The neutral wind speed $\mathbf{u}_n(z)$ at height z is connected with the logarithmic profile (3). Since in (9), substitution $1/z_0 = \exp(b_N - 1)/z_l$ had been made, it directly follows that interpolation (9) can also be used for the exact evaluation of neutral wind, where the lhs is replaced by $\mathbf{u}_n(z)$ and S by:

$$
S = 0 \quad \text{for neutral wind.} \tag{11}
$$

For stable cases neutral wind $\mathbf{u}_n$ will be weaker ($S > 0$), while for unstable cases it will be stronger ($S < 0$) than the relative real wind $\mathbf{u} - \mathbf{u}_c$. Only for neutral stratification both coincide. Over the global oceans, the surface layer is typically weakly unstable, and 10m neutral wind speed is on average $0.2\,\mathrm{m\,s^{-1}}$ stronger than real wind.

In IFS, the method of Geleyn is coded in the routine `ppuv10m` which is called by `ppobsas`. The extension (11) for neutral wind is simple, and was incorporated in cycle 35r2. The incorporation of surface current, which was not part of the original method of Geleyn (1988) was included here as well. A study on its potential effect on the ECMWF forecast and assimilation system can be found at Hersbach and Bidlot (2009). Note that in the operational configuration, and for all applications discussed in this document, surface current is set to zero.

3.2 Exchange coefficients as used in the operational observation operator

The method of Geleyn relies on the availability of the exchange coefficients b_N and b_D. Evaluation of b_N (7) is straightforward, since it is directly related to surface roughness, which is in principle part of the GOM arrays. Quantity b_D depends in detail on stability, and requires knowledge of, e.g., friction velocity, surface stress or Obukhov length as it is calculated in the parametrization (within `callpar`). Neither of such quantities are available. As an alternative, b_D is based on an estimate from Louis *et al.* (1982), which was used in a previous physics package for IFS. In this method b_D is estimated from b_N and a stability-dependent correction factor f_M as:
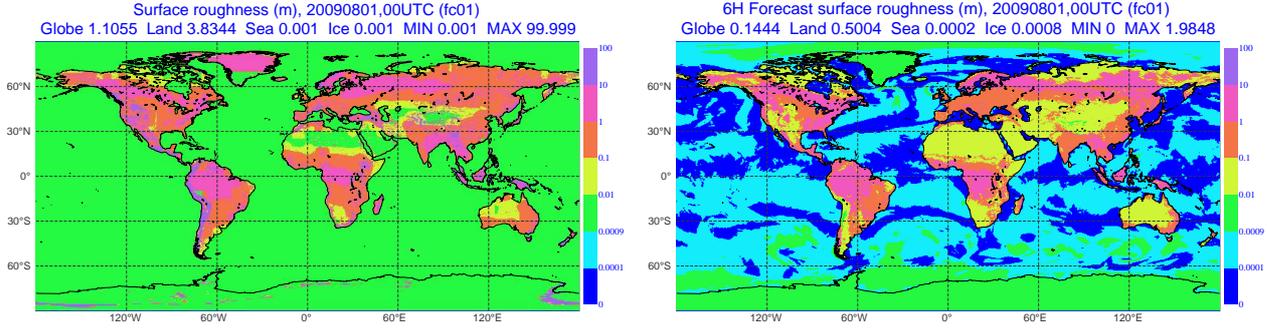
$$
b_D = b_N/\sqrt{f_M}, \tag{12}
$$

Figure 2: Global maps of surface roughness (SR, left) for the 00 UTC 20090901 DA (fc01) analysis, and the +6H forecast surface roughness (FSR, right panel).

where,

$$f_M = \begin{cases} 1 - 10R_i/(1 - 75R_i b_N \sqrt{1 + z/z_0}) & \text{unstable} \quad (R_i < 0), \\ (1 + 10R_i/\sqrt{1 + 5R_i})^{-1} & \text{stable} \qquad (R_i > 0). \end{cases} \tag{13}$$

Over the ocean the Richardson number R_i is estimated from temperature, wind and humidity at lowest model level, and skin temperature, which are all available in the GOM arrays. Other input quantities such as saturated humidity and wetness (set to unity over the ocean) are effectively estimated from wind, humidity and temperatures as well. In `preints` this latter is handled by routine `surbound`, while R_i, b_N, and f_M are evaluated in routine `exchco`.

In the TL and AD calculations, perturbations in (diagnostic) skin temperature and surface roughness are neglected.

3.3   Comparison with an offline calculation of the IFS surface-layer physics

The estimation of Louis *et al.* (1982) does not correspond to the latest version of the IFS surface layer (SL) parametrization that is used in the forecast model, where equations for the turbulent transfer of momentum, heat and moisture are addressed simultaneously. This latter occurs within the routine `vdfmain` (vertical diffusion), which is called under `callpar`. A concise description may be found in Part IV.3 of the IFS-documentation (2009). For the situation over water, a stand-alone version of the IFS SL physics is available, which can be used to test the similarity between the two schemes. Although this package, called OCFLX, addresses the same set of equations, these are solved in an iterative way, instead of the 'more than implicit' time-stepping method that is used in IFS. Input to OCFLX are the wind $\mathbf{u}_l$, temperature T_l, humidity Q_l, geopotential height Φ_l and pressure P_l at lowest model level, surface skin temperature and the ocean-wave Charnock parameter. Output are the surface fluxes for momentum (stress $\vec{\tau}$), heat and moisture, corresponding roughness lengths, and air density ρ_a.

Results from an offline extraction of Louis *et al.* (1982) from `preints` are compared to OCFLX on the basis of archived fields for the +6H T799 forecast component from the Daily-Archive (DA) analysis for 00UTC 20090801 of an analysis experiment with identifier fc01. For pressure, the exponent of the logarithmic surface pressure (LNSP) is used, and the geopotential height is determined from:

$$\Phi_l = \alpha R T_l (1 + R_{tv} Q_l), \quad \text{where} \quad \alpha = 1 - \log(\frac{1}{B})B(1 - B), \tag{14}$$

with $R = 287.0597\ldots$, $R_{tv} = 0.60777\ldots$ and $B = 0.997630119324$. For a typical case ($T_l = 289\mathrm{K}$, $Q_l = 0.01\mathrm{kg/kg}$) this corresponds to a height of $z_l = \Phi_l/g = 10.093\mathrm{m}$.

For OCFLX, exchange coefficients b_N and b_D are calculated from (7, 8), where friction velocity is determined as:

$$||\mathbf{u}_*|| = \sqrt{||\vec{\tau}||/||\rho_a}.\tag{15}$$

For `exchco`, the value for saturated humidity as calculated by OCFLX is used, while wetness is set to unity (as it is in `surbound`). An important point is that the GOM array for surface roughness does not contain the actual forecast surface roughness (archived as FSR) as calculated in the SL, but is filled with an analysis surface roughness field (SR), instead. This latter is based on climatology. Over the ocean, it contains a value of $1\mathrm{mm}$, which is typically an order of magnitude too high. A comparison between SR and FSR is presented in Fig. (2), which clearly demonstrates the difference between the two fields. For this reason, b_N (7) and f_M (13) are based on $z_0 = 1\mathrm{mm}$. For $z_l = 10\mathrm{m}$ this gives $b_N = 9.2$, while for the +6H forecast as indicated above, an average value of $b_N = 11.3$ is found from using FSR.

For both methods, the 10m neutral wind is calculated according to (9), with $S = 0$, and $z = 10\mathrm{m}$. As mentioned above, the lowest model level z_l is usually close to 10m, in which case (9) reduces to:

$$\mathbf{u}_n(z_{10}) \approx (b_N/b_D)\mathbf{u}_l.\tag{16}$$

For Louis *et al.* (1982) the ratio b_N/b_D is equal to the square root of correction factor f_M (13), and as a result, the effect on b_N by using SR rather than FSR is divided out for stable cases. For unstable cases there is a residual dependency.

The resulting 10m neutral wind is compared to the 10m wind as archived in the Meteorological Archival and Retrieval System (MARS). The difference is positive for unstable and negative for stable cases. Fig. (3) provides scatter plots for this stability correction for OCFLX (horizontal axis) versus `exchco` (vertical axis). The left panel shows that `exchco` as used in the operational model ($z_0 = 1\mathrm{mm}$) under estimates unstable corrections, and shows quite some scatter for stable conditions. On average, stability corrections are $0.06\mathrm{m\,s}^{-1}$ lower, i.e. $0.14\mathrm{m\,s}^{-1}$ versus $0.2\mathrm{m\,s}^{-1}$ for OCFLX.

Surprisingly, the usage of FSR (middle panel) deteriorates the comparison with OCFLX. Reason for this may be that it seems that the Louis *et al.* (1982) scheme had been calibrated for a fraction $z_l/z_0 = 5500$. At the time that this formulation was used in the ECMWF SL, the lowest model level was located at about $z_l = 30\mathrm{m}$, which translates to $z_0 \approx 5.5\mathrm{mm}$. The right-hand panel of Fig. (3) shows results for $z_0 = 5\mathrm{mm}$, from which it is seen that the agreement is indeed improved for unstable cases. For stable cases, where f_M (13) does not depend on b_N, all three choices give equivalent results.

The framework of Louis *et al.* (1982) may give reasonable results for 10m neutral wind because interpolation (16) mainly depends on the ratio b_N/b_D. For vertical interpolation to other heights, however, the too high value for z_0 will introduce biases. An example is buoy wind, which is typically observed at a height of 4 to 5m. The situation for the vertical interpolation to 4m is displayed in Fig. (4). Depending on 10m wind speed, corrections vary from $0\mathrm{m\,s}^{-1}$ to $-2.5\mathrm{m\,s}^{-1}$, with an average of $-0.6\mathrm{m\,s}^{-1}$. Interpolation based on Louis *et al.* (1982) with $z_0 = 1\mathrm{mm}$ overestimate the magnitude of the correction by on average $0.13\mathrm{m\,s}^{-1}$ (left panel). For moderate winds the mismatch is largest. For winds stronger than $14\mathrm{m\,s}^{-1}$ the sign changes (middle panel), which corresponds to the regime where the actual FSR becomes larger than 1mm. It is clear that a correct value for b_N is important for interpolation to heights other than 10m. Within the framework of Louis *et al.* (1982) best results are expected to set $z_0 = 5\mathrm{mm}$ in the correction factor (13), but to use $z_0 = \mathrm{FSR}$ in the definition (7) for b_N. This is confirmed by the right-hand panel of Fig. (4), which shows a large reduction in the wind-speed dependent interpolation error. This would require that FSR is passed to the GOM arrays, rather than SR.
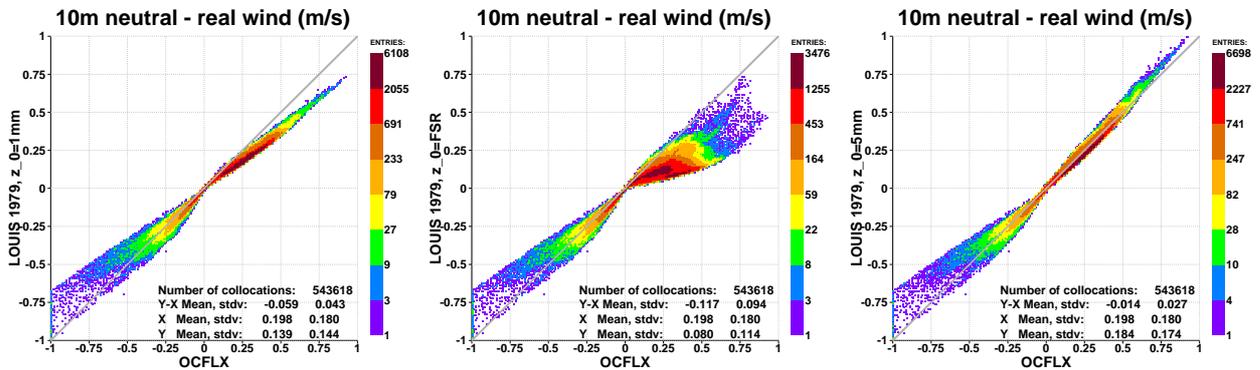
Figure 3: Scatter plots for the deviation of 10m neutral from 10m non-neutral wind over sea (T799 6-hour forecast from the 00 UTC 20090801 DA, fc01 analysis) between OCFLX (current SL formulation) and exchco *(Louis* et al. *(1982)), where the latter is based on a surface roughness of $z_0 = 1mm$ (left), $z_0 = FSR$ (middle), and $z_0 = 5mm$ (right panel).*
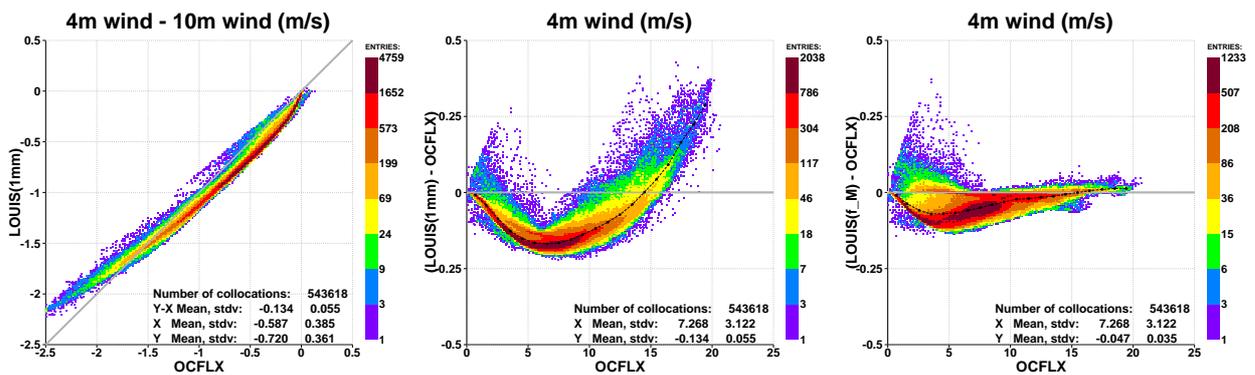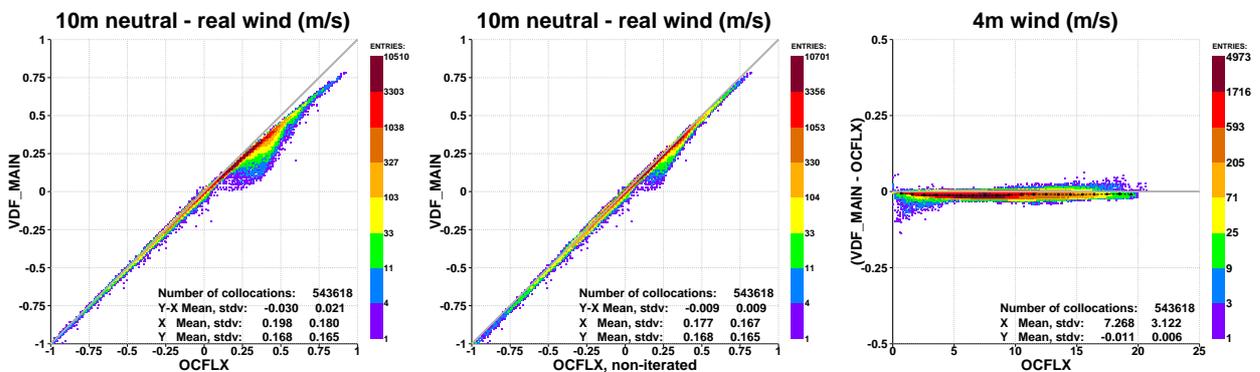

Figure 4: Scatter plots for the correction of 4m non-neutral wind from 10m non-neutral wind over sea (T799 6-hour forecast from the 00 UTC 20090801 DA, fc01 analysis) between OCFLX (current SL formulation) and exchco *(Louis* et al., *1982), using the standard choice of SR for surface roughness (left), the difference between Louis* et al. *(1982) and OCFLX (middle), and a similar difference for a blend between SR and FSR as explained in the text (right panel).*

Figure 5: Scatter plots for the deviation of 10m neutral wind from 10m non-neutral wind over sea (T799 6-hour forecast from the 00 UTC 20090801 DA, fc01 analysis) for OCFLX with iteration (left) and without iteration (middle panel) versus exchco_vdf, *and the difference in the estimation of non-neutral 4m wind speed (right panel).*

3.4 Extension of the non-linear observation operator

From the previous sub-section it emerges that it is desirable to have direct information on stability in the observation operator as it is calculated inside the SL. This can be established by an appropriate extension of the GOM arrays. One could pass on b_D directly, or use friction velocity u_*, from which b_D can be calculated from (8). Since the current goal is the construction of an observation operator for 10m neutral wind $\mathbf{u}_n(z_{10})$, it is proposed to add neutral wind itself. From this, quantity b_D can be reconstructed as:

$$b_D = \log(1 + z_{10}/z_0)\frac{||\mathbf{u_l}||}{||\mathbf{u}_n(z_{10})||}. \tag{17}$$

By definition (of the constant stress layer) no wind turning occurs in the SL, and therefore, substitution of (17) into vertical interpolation (9) with $S = 0$, directly provides neutral wind again. In other words,

$$H(\mathbf{x}) = \mathbf{u}_n(z_{10}). \tag{18}$$

Extension (17) together with (7), is implemented in a new routine called `exchco_vdf`.

Besides the inclusion of neutral wind in the GOM arrays, it is desirable that for surface roughness z_0, SR is replaced by FSR. It should be noted, however, that over the ocean FSR has a dynamic range of a few orders of magnitude. For this reason, it can locally vary substantially, which may lead to inaccurate results in the horizontal interpolation (`obshor`). Over water, there is an alternative to reconstruct z_0 from (3, 2). For given neutral wind and Charnock parameter α_{ch}, these equations can be solved for z_0 by iteration. In cycle 36r1, a routine is available (`z0sea`) that gives an accurate estimation of z_0 without the need of iteration (Hersbach, 2010c). The usage of this routine will result into more accurate estimates for z_0 over the ocean. Besides the neutral wind, this requires the inclusion of the wave Charnock parameter into the GOM arrays as well.

The Charnock parameter can be directly fetched from the ocean-wave model component of IFS, while the neutral wind is to be fetched from `vdfmain`. In the default version of cycle 36r1, neutral wind is not calculated, though. A logical place to include its computation is where the 10m (non-neutral) wind is determined (`sppcfl`). Originally, `sppcfl` was only called in the forecast model at time steps that required post-processing for archiving. Nowadays, it is called every time step inside the trajectory run as well, since the 10m wind is required for the assimilation of all-sky microwave radiances (Bauer *et al.*, 2006). Routine `sppcfl` is called under `vdfmain` after the computations in the SL have completed. It has access to the same stability functions Ψ_M and Obukhov length L. Over the ocean, the extension of `sppcfl` calculates neutral wind from lowest-model level wind $\mathbf{u}(z_l)$ as:

$$\mathbf{u}_n(z_{10}) = \frac{\log(1 + z_{10}/z_0)}{(\log(1 + z_l/z_0) - \Psi_M((z + z_l)/L) + \Psi_M(z_l/L))}(\mathbf{u}(z_l) - \mathbf{u}_c). \tag{19}$$

Over land, the computation is somewhat more complicated, since here a 10m wind is returned that is representative for open terrain, rather than over the average landscape within the model grid box. Details may be found in Hersbach (2010a).

Like for `exchco` results from the new routine `exchco_vdf` are compared to OCFLX offline, on the basis of the same +6-hour forecast as used in Section 3.3. For that forecast, the neutral wind (19) as calculated in `sppcfl` has been archived on the basis of the in MARS existing parameters (U10N,V10N). Together with archived FSR, b_N and b_B are calculated in `exchco_vdf`, and substituted in (9). For 10m neutral wind, the result is, again, compared to the 10m (non-neutral) wind. In the left panel of Fig. (5), such obtained stability corrections (y-axis) are plotted against results from OCFLX (x-axis). From this it is seen that both the relative bias and scatter is smaller than for exchco (left panel of Fig (3)). Especially for stable cases the comparison between `exchco_vdf` and OCFLX is excellent. For unstable cases, however (which are more common over

the ocean), `exchco_vdf` somewhat under-estimates OCFLX. Over all cases, the average stability correction is 0.03m s^{-1} weaker, i.e., 0.17m s^{-1} versus 0.20m s^{-1} for OCFLX. Most of the scatter is found to occur for light winds (see Hersbach (2010a)). As mentioned above, there is a difference in which the turbulent equations for momentum, heat and moisture are addressed. For OCFLX these are solved in an iterative way, while a 'more than implicit' time-stepping is used in `vdfmain`. If instead of the default of three iterations, no iterations are performed in OCFLX, the comparison with `exchco_vdf` appears to improve for unstable cases (middle panel of Fig (3)). The fact that scatter is reduced considerably indicates that for part, differences between OCFLX and `vdfmain` may indeed be connected to a difference in solution method.

Since routine `exchco_vdf` provides good estimates for b_N and b_D, it can also be used for vertical interpolation to other heights. The example for $z = 4\text{m}$ is given in the right-hand panel of Fig. (5). From this it is seen that the comparison with OCFLX is much better than what is found for `exchco` (see Fig. (4)).

For 10m neutral wind, Fig. (6) provides global maps for all stability corrections as regarded in this Section. From this it is seen, that the regional patterns for less than average unstable (blue) and more then average unstable (red) areas are very similar for most approaches. Only for the usage of FSR in Louis *et al.* (1982) (lower left), unstable patterns are largely missing. The similarity in patterns is not too surprising, since all are driven by the difference between air temperature and skin temperature. And this input is the same for all approaches. For the scheme that is used in the operational observation operator, i.e., Louis *et al.* (1982) with $z_0 = 1\text{mm}$ (top left), perturbations are clearly weaker than on the basis of the offline package OCFLX (middle right). Better amplitudes are obtained for $z_0 = 5\text{mm}$ (left middle). The usage of neutral wind from extension (19) in `vdfmain` (top right) produces unstable corrections that are slightly weaker than OCFLX. It does corresponds much closer to results for OCFLX without iteration in the solution method (bottom right).

3.5 Extension of the linearized observation operator

As mentioned above, perturbations in surface parameters are not incorporated in associated GOM arrays. For this reason any TL perturbation in $\mathbf{u}_n$ is cut to zero, while any AD dependency that is accumulated in `hopad` is nullified later on. The combination of (17) and (9) exactly delivers $\mathbf{u}_n$, as shown in (18). This means that the observation operator for neutral wind is independent on variations $\delta\mathbf{x}$. Therefore its contribution to the cost J_o is a constant, and effectively, the associated observations are not assimilated.

One way to avoid this problem is the assumption that a variation in neutral wind, due to a variation in model state is mainly driven by a change in the lowest model level wind, and not so much by a change in stratification. Since the effect of stability is largely contained in the ratio b_D/b_N (as e.g. according to the approach by Louis *et al.* (1982)), it is reasonable to keep this ratio constant. This means that for b_D the TL of (17) can be approximated by:

$$\delta b_D = \left(\frac{b_D}{b_N}\right) \delta b_N. \tag{20}$$

For the usual case that the lowest model level z_l is close to 10m, this leads to a perturbation in neutral wind (16):

$$\delta\mathbf{H} = \delta\mathbf{u}_n \approx \left(\frac{b_N}{b_D}\right) \delta\mathbf{u}(z_l). \tag{21}$$

For the AD, similar equations apply. Perturbations in $\mathbf{u}(z_l)$ are well propagated in the associated GOM arrays, so in effect approximation (20) and its AD analogue allow for a proper assimilation of neutral wind.

The more consistent way forward is to resolve the issue of the associated GOM arrays for diagnostic surface fields. For the TL this means that the calculation of perturbations are to be added to the TL of routines that calculate these parameters inside `callpar`, that these are horizontally interpolated to observation location,
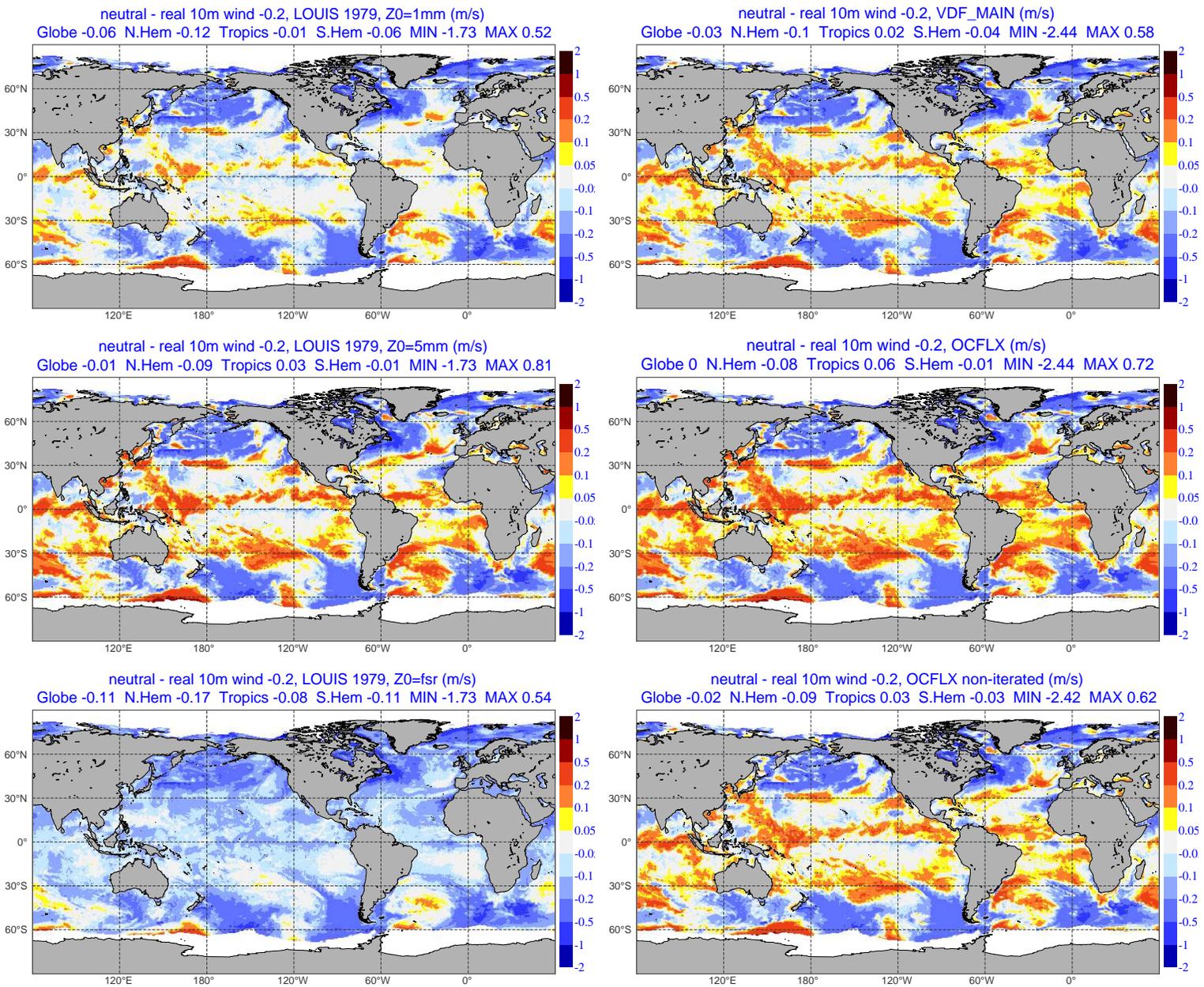
Figure 6: Global maps of the deviation of 10m neutral wind minus 0.2m s⁻¹ from real wind (00 UTC 20090901 DA favc, 3-hour forecast) as evaluated for several schemes.

and that they are stored into the responsible GOM array. In `hoptl` these perturbations are then read from memory, which then provide non-zero contributions to perturbations in H. For the AD, dependencies in surface parameters as built up in `hopad` are stored in the associated GOM array. These, in general non-zero values are read in the AD of the corresponding routines in `callpar`, and in this way contribute to the dependency of the initial control vector on the cost J.

For neutral wind, this extension appears to be possible. The necessary extensions in the TL (`sppcflstl`) and AD (`sppcflsad`) routines to `sppcfls` are made, which latter provides the simplified-physics version of `sppcfl`. For standard cycle 36r1, the associated GOM arrays are already allocated. They are just not used. Appropriate extensions in routines `cobstl/ad` provide the communication with the model physics, while extensions in the routines `preintstl/ad` provide the connection with the observation operator. The resulting proper flow of perturbations in neutral wind, now do allow for the exact TL and AD to (17). These are coded in `exchco_vdftl` and `exchco_vdfad`. These routines also contain an option for approximation (20), and depending on a logical in the argument list, one of the two methods is used. For z_0 dependencies are effectively handled by the TL (`z0seatl`) and AD (`z0seaad`) of the estimation from neutral wind and Charnock parameter in `z0sea`. Variations in the Charnock parameter are neglected, since that would require the TL and AD of the ECMWF wave model (WAM).

4 Technical complications

4.1 Issues regarding the trajectory run

In Section 3.4 it was shown how information on stability could be fetched from the physics in `callpar`, by the inclusion of neutral wind in the GOM arrays. In Fig. (7) some detail is provided on the flow in the trajectory run. The high-level routine `step0` takes care of all computations within one time step (NSTEP=0, 1, .., NSTOP). From this it is seen that the horizontal interpolation and storage of surface fields into GOM arrays occurs before diagnostic surface fields are updated. For that reason, at initial time step (NSTEP=0) surface fields contain their initialized values. These are based on the +3-hour forecast surface fields from the previous DCDA analysis that had been read in by `su_surf_flds`. Those surface fields that are not supplied by this first-guess will remain un-initialized. This is, e.g., the case for FSR (z_0) and neutral wind ($\mathbf{u}_n$). As a result, the observation operator as discussed in Section 3.4 will fail since it will find un-initialized values in the GOM arrays for neutral wind.

Prior to the execution of 4D-Var, the set of required first-guess fields are retrieved from MARS, or if still on disk from the ECMWF Field Data Base (FDB). At script level, any extra fields can simply be added to a list of requested parameters. For FSR, i.e., this would be straightforward. For 10m neutral wind, however, the situation is more complex because it is not archived in the DCDA forecast. This is only done for the operational DA forecast. Here neutral wind does not, like other parameters, result from a calculation inside IFS, but is derived from other fields from FDB, instead. In Section 3.4 it was shown how the computation of neutral wind can be included in `sppcfl`. This paves the way for archiving from inside IFS which requires the extension of a set of relevant routines that takes care of the post-processing (FULLPOS). Details may be found in Hersbach (2010a)). This then allows for the addition of neutral wind to the list of to be retrieved first-guess fields, and for the initialization of the relevant GOM arrays at the first time step.

Another consequence of the flow in the trajectory run is that the GOM arrays are filled with surface fields that have been updated the previous time step. As, e.g., can be seen from Fig. (7), the GOM arrays at NSTEP=1 rely on the results in the parametrization from NSTEP=0. This is not necessarily incorrect, since there are diagnostic fields that are strictly speaking valid for the next time step. For neutral wind and 10m wind, however, this is not the case, and for that reason, the values for neutral wind in the GOM arrays are out of sync with other variables
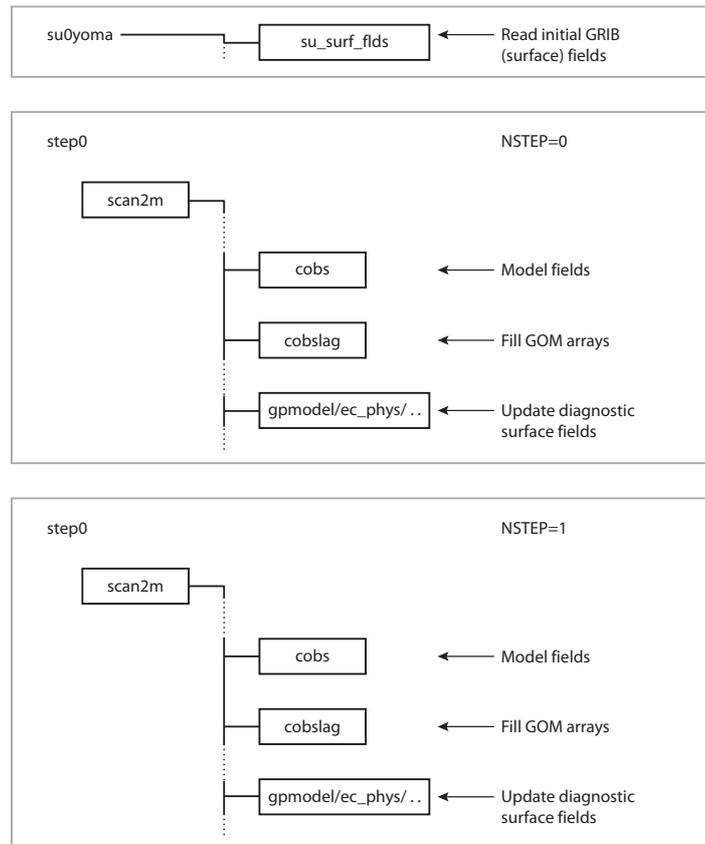
Figure 7: Snippet of the flow chart for the trajectory run

(such as fields at model levels). The update of neutral wind at final time step NSTEP=NSTOP is not used in the observation operator.

4.2   Issues regarding the minimization

For one specific iteration in the minimization, the TL is run from initial time to final time, which is followed by a run of the AD code from final time back to initial time. Similar flow diagrams as shown in Fig. (7) apply to `step0tl` and `step0ad`.

Since for the first time step neutral wind is taken from a first-guess field, its value will not be affected by any perturbation $\delta \mathbf{x}$ in the control vector at the start of the assimilation window. As a result, the cost J_o will be a constant for all observations for neutral wind that fall inside this first time slot. When the full dependencies of neutral wind are correctly handled by the associated GOM array, this will lead to a zero gradient with respect to such data. Assumption (20) assumes that perturbations in neutral wind are dominated by perturbations in the lowest model level wind. Since these are influenced by the control vector, this approach will lead to a non-zero dependency for the first time slot. Therefore, to take full advantage of the incorporation of perturbations in neutral wind, one should use the full TL and AD for `exchco_vdf` for all time steps, except for the first one, where assumption (20) is to be used. The choice for which approach to take can be conveniently provided by the inclusion of a logical in the GOM arrays that indicates whether surface fields have been updated sofar.

When a quantity is calculated in a non-linear code, its adjoint value is usually to be nullified in the corresponding AD code. When this rule is not obeyed, there is a risk that erroneous dependencies leak into upstream parts

of the code when the same variable is reused. A typical example is the calling of a certain routine for several time steps. Although variables are reused, they physically represent quantities at different time steps, and for that reason the adjoint of updated global variables need to be zero at the start of the call to the AD code. When the nullifying rule has been used consistently, such adjoint variables will be zero automatically. As a safeguard one could nullify all variables explicitly at the start of the adjoint routine, which eliminates any leaking from un-careful adjoint coding.

At the start of `ec_physad` such a reset is performed for all surface field perturbations. The risk is that this could overdo the job, since it also resets perturbations of variables that have not been updated yet. An example is the AD of the neutral wind that is read from the GOM arrays in the AD of `cobslag` and `cobs`. These quantities, which describe the influence of the control vector on the cost via the computation of neutral wind, must not be nullified at the start of `ec_physad`. The remedy is not to reset surface field perturbations anymore. It was carefully checked that this did not have side effects, which effectively means that for surface fields the adjoint has been properly coded.

In the minimization the same variables are used for TL and AD perturbations. To avoid interference, one should ensure that all perturbations other than those for the control vector are zero at the start of a TL run (from 0 to NSTOP), and similarly at the start of an AD run (from NSTOP to 0). This can be guaranteed by resetting variables at the very start and at the very end of a run. For surface fields this is performed in `ec_phystl` and `ec_physad`. With the incorporation of surface perturbations inside the observation operator, this has to be done at a higher level of the code. The proper location is `scan2mtl` and `scan2mad` which encapsulate all calculations for surface fields. For this reason the resetting of surface field perturbations is moved upwards to these routines.

5   Implementation for cycle 36r3

All extensions as discussed in Section 3 have been coded in a branch dal_CY36R1_neutral_full_dependencies , which was later merged into cycle 36r3 (for details, see Appendix B). Some had been prepared before and were already present in cycle 36r1. These include the adaptation of `ppuv10m`, the inclusion of `z0sea`, and their TL and AD versions. The decision on the choice on the various options is handled from the following logicals, which can be set at script level via the namelist `NAMOBS`:

- **LSCATT_NEUTRAL** (existing): If .true., for scatterometer wind vertical interpolation (9, 11) is used. Regarding the inversion from scatterometer observed backscatter to wind, $0.2\,\mathrm{m\,s^{-1}}$ is added to wind speed for QuikSCAT, while for ERS-2 and ASCAT geophysical model function CMOD5.N (Hersbach, 2010b), rather than CMOD5.4 (Abdalla and Hersbach, 2007) is used.
  If .false., scatterometer wind is assimilated as 10m non-neutral wind (currently operational).

- **LVDFTRAJ** (new): If true, for scatterometer data exchange coefficients are based on neutral wind from the ECMWF physics (`exchco_vdf`) and surface roughness is based on (`z0_sea`) Hersbach (2010c). In principle this new method could be used for other surface wind observations as well (such as buoys). It was decided, though, to concentrate on scatterometer data first. Also, if .true., the Charnock parameter, which is used to estimate z_0 is placed in the full time-dependent trajectory (as returned by the coupled wave model) rather than in the constant trajectory (using a value of $\alpha_{\mathrm{ch}} = 0.018$).
  If .false., the method of Geleyn (1988) is used (`exchco`, currently operational).

- **LVDFMIN** (new): If .true., for scatterometer data the full dependency of perturbations $\delta\mathbf{x}$ via neutral wind is incorporated. If .false., the ratio (u_N/u_D) is kept constant in the minimization.

- **LZ0FC** (new): If .true., FSR is placed in the GOM arrays rather than SR (currently operational). This FSR is also used in Louis *et al.* (1982), which is not optimal, as discussed in Section 3.3. Therefore, this switch should not be used at the moment.

At script level, the value of LSCATT_NEUTRAL (set in prep_ifs) is used for LVDFTRAJ as well. For that reason, the system remains bit-identical when LSCATT_NEUTRAL is .false. (which is the default in cycle 36r1). LVDFMIN is set to .true., LZ0FC is set to .false., and a script change is required in case this is not intended. Inside IFS, some quality control is performed (defrun). LVDFTRAJ is reset to .false. when LECMWF (configuration for ECMWF) or L3DFGAT (physics is run in trajectory) is .false..

5.1 Sanity checks

The AD and TL of the routines ppuv10m, z0sea, exchco_vdf and sppcfls were thoroughly tested. It was checked that the TL and AD code for individual routines match, and it was verified that finite differences convert to TL results in the limit towards zero perturbation size. In a reduced resolution of T42, a similar test on the match between the TL and AD was performed for the entire 4D-Var minimization. This option, which is provided in IFS by the setting of a variable NTESTVAR=3, was passed successfully for the configurations as described below, giving similar results to that for an experiment that was based on a clean 36R1 branch. This gives confidence on the move of nullifying statements for surface field perturbations at initial and final time from ec_phystl and ec_physad to scan2mtl and scan2mad, and the removal of such statements in ec_physad at all other time steps.

It was ensured that array boundaries were not violated and that non-initialized variables were not used in the rhs of any computation.

6 Impact studies

In order to test the various options as described in the previous Section, for the period from 15 July 2009 to 31 October 2009 a number of experiments are conducted at T511 in early-delivery mode. Details on configurations may be found in Table (6.1).

Except for a CONTROL run (experiment identifier f8tz), which used a clean version of cycle 36r1, all experiments are based on the branch dal_CY36R1_neutral_full_dependencies . The assimilation of scatterometer wind as neutral wind, using information from the ECMWF physics (both non-linear and linear) is explored in an experiment EXCHCO_VDF. The usage of a neutral wind observation operator on the basis of the estimation of b_D by Louis *et al.* (1982) and the usage of SR into b_N is investigated in an experiment EXCHCO.

The effect of the usage of FSR rather than SR in the standard configuration of cycle 36r1, i.e., assimilating scatterometer data as non-neutral wind, with Louis *et al.* (1982) is explored in an experiment FSR. Note that this change has virtually no effect on the assimilation of 10m wind data, since for that case vertical interpolation (9) is basically the identity operator. It does have an effect on the assimilation of wind at other observation heights, such as for buoy data. Prior to this experiment, two other experiments had been run that had used FSR rather than SR as well. These two experiments assimilated scatterometer wind in the new fashion (LVDFTRAJ=.true.), one, like EXCHCO_VDF taking account of neutral wind perturbations (LVDFMIN=.true.) and one neglecting them (LVDFMIN=.false.). Both appeared to give rise to very similar forecast skill, but, unfortunately performed worse (at a 95% significance level) than the CONTROL. It was only then realized that the usage of FSR rather than SR is not optimal for the Louis *et al.* (1982) scheme. By running experiment FSR, the effect of using forecast surface roughness can be isolated.

Name	Expid	Description	
CONTROL	f8tz	based on dah_CY36R1_backstitch	
		LSCATT_NEUTRAL=.false.	
EXCHCO_VDF	favr	LSCATT_NEUTRAL=.true.	LZ0FC=.false.
		LVDFTRAJ=.true.	LVDFMIN=.true.
EXCHCO	fblk	LSCATT_NEUTRAL=.true.	LZ0FC=.false.
		LVDFTRAJ=.false.	LVDFMIN=.false.
FSR	faxa	LSCATT_NEUTRAL=.false.	LZ0FC=.true.
		LVDFTRAJ=.false.	LVDFMIN=.false.

Table 1: List of experiments as discussed in this document. All are in a T511 early-delivery assimilation environment and have been run from 15 July 2009 to 31 October 2009. The second, third and fourth experiment are based on dal_CY36R1_neutral_full_dependencies.

The default configuration for dal_CY36R1_neutral_full_dependencies (i.e., LSCATT_NEUTRAL=.false.) should give equal results to a clean cycle 36r1. This was confirmed by an experiment (favq), which was run for a number of days. Indeed, bit-wise identical results were obtained with respect to the CONTROL.

6.1 Condition numbers

For all experiments displayed in Table (6.1) very similar condition numbers were found in the minimization. This indicates that the fetching of AD and TL information for surface field perturbations is handled well, and does not affect the performance of the minimization scheme.

6.2 Departure statistics

In Hersbach (2010a) it was observed that for two one-month periods scatterometer wind first-guess departures (5) are slightly reduced when based on neutral rather than non-neutral wind. This is confirmed in the present study. For both the EXCHCO_VDF and EXCHCO experiment, first-guess (and analysis departures) are somewhat reduced for ERS-2, ASCAT and QuikSCAT. For ASCAT, results are presented in Fig. (8), from which it emerges that the largest reduction is found in the Northern Hemisphere. Statistics are very similar for both experiments.

For ASCAT and QuikSCAT, maps of average first-guess departures are displayed in Fig (9) for the CONTROL (left) and EXCHCO_VDF experiment (right panels), respectively. For EXCHCO maps are similar to EXCHCO_VDF (not shown). The blue patterns around Newfoundland and the Hudson bay indicate areas of stable conditions, which are typical for late Summer. They are reduced in the EXCHCO_VDF, and this is probably the reason why for this season the reduction in departure statistics is largest in the Northern Hemisphere.

Analysis departures (6) are in general much lower than first-guess departures (5), since this expresses that the data is assimilated. A similar improvement is observed for both experiments as shown Fig. (8) as well. It emerges that this occurs for all time slots. Also for the first time slot (not shown), where the model neutral wind results from a first-guess field, and, therefore, is not influenced by the control vector, this appears the case. For the EXCHCO_VDF experiment this means that linearization (20), which is only used for the first time step, is working well. When, in contrast, for this step the full dependencies in neutral wind are retained it appears that the analysis departure is only marginally smaller than the first-guess departure. This, which was checked by an experiment that was run for a few days, indeed reflects that in such case scatterometer winds that fall in the first
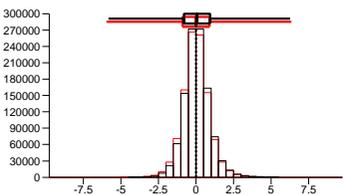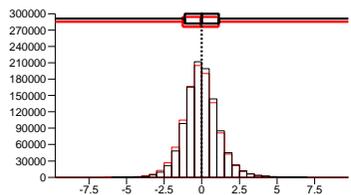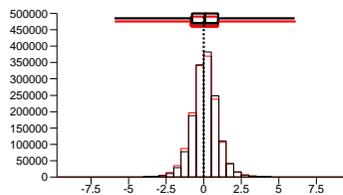
Figure 8: Statistics of first-guess (left) and analysis (right sub-panels) departures accumulated from 20090715 to 20091031, for ASCAT 10m wind as used in the observation operator H, i.e., neutral for EXCHCO_VDF (left) and EXCHCO (right panels), and non-neutral for the CONTROL (in parentheses). Statistics are stratified according to Northern Hemisphere (top) Tropics (middle) and Southern Hemisphere (lower panels). Smaller standard deviations indicate a better fit to the data.

Figure 9: Geographically time-averaged first-guess departures over the indicated period for ASCAT (top) and QuikSCAT (lower panels) 10m wind speed for the CONTROL (non-neutral, left) and the EXCHCO_VDF (neutral, right panels) experiment.
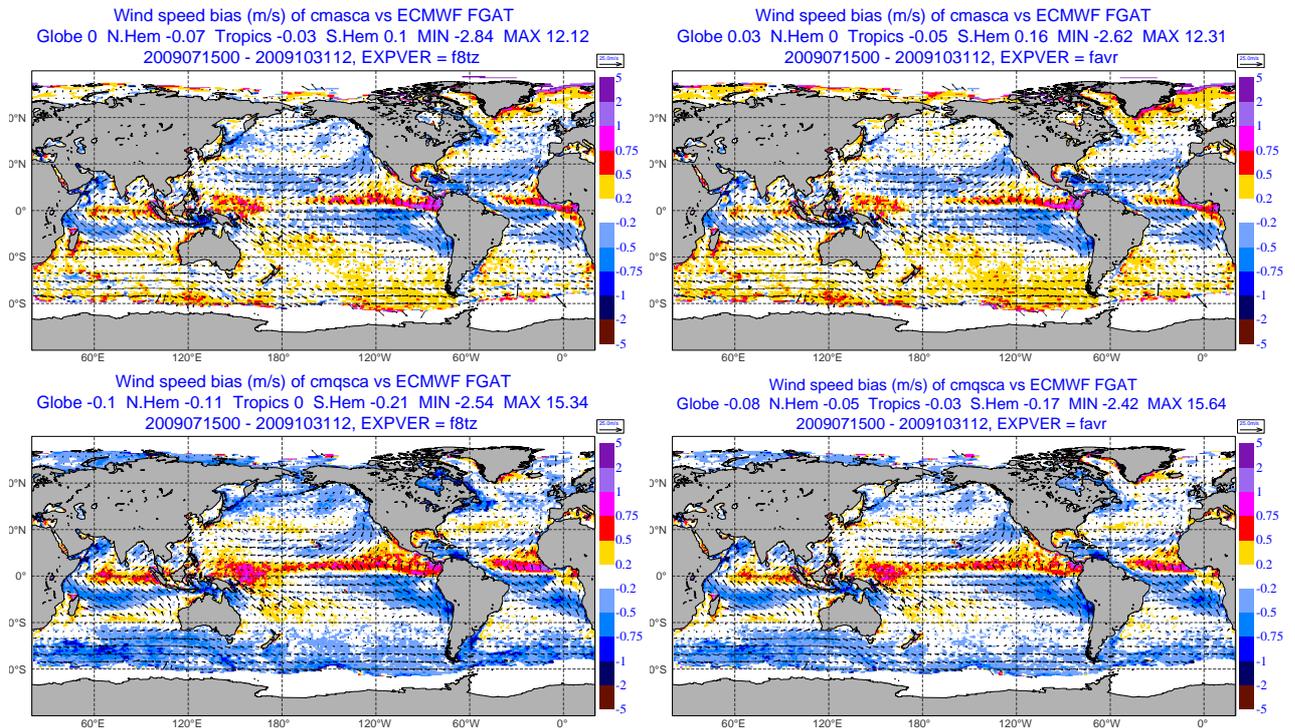
time slot are effectively not assimilated.

6.3 Forecast skill

The impact on average global forecast skill on geopotential at 500 hPa (Z500) in the atmosphere, and on significant wave height (SWH) at the ocean surface is presented in Fig. (10). For the EXCHCO_VDF and EXCHCO experiment, plots that indicate the statistical significance for Z1000 and Z500 are presented in Figs. (11, 12). In general the EXCHCO_VDF exhibits the slightly better performance. Regarding the Northern Hemisphere, all experiments under-perform the control around day 8 for geopotential, although EXCHCO_VDF is more comparable with the CONTROL. For EXCHCO and FSR, a negative impact which is significant on the 95% level is found for Z1000, while for EXCHCO_VDF this is on the brink of significance. For ocean waves, impact is more neutral, with EXCHCO_VDF giving a tiny (non-significant) improvement over the CONTROL. Over Europe (see Fig. (13)), similar results are found.

Over the Southern Hemisphere, a modest positive impact is found for EXCHCO_VDF and EXCHCO (not significantly, though) on geopotential. For ocean waves, EXCHCO_VDF clearly provides the best skill.

In general, the FSR experiment performs worst.

A more detailed view of the effect on forecast skill is presented in Fig (14), which shows the normalized average difference in day 3 forecast skill with respect to the CONTROL for Z1000. From this it emerges that no clear areas can be identified where forecast skill generally improves or deteriorates. The patchy patterns east of Japan are possibly the result of small scale and temporal fluctuations in atmospheric stability over the Kuroshio extension. At other levels, similar results are observed.
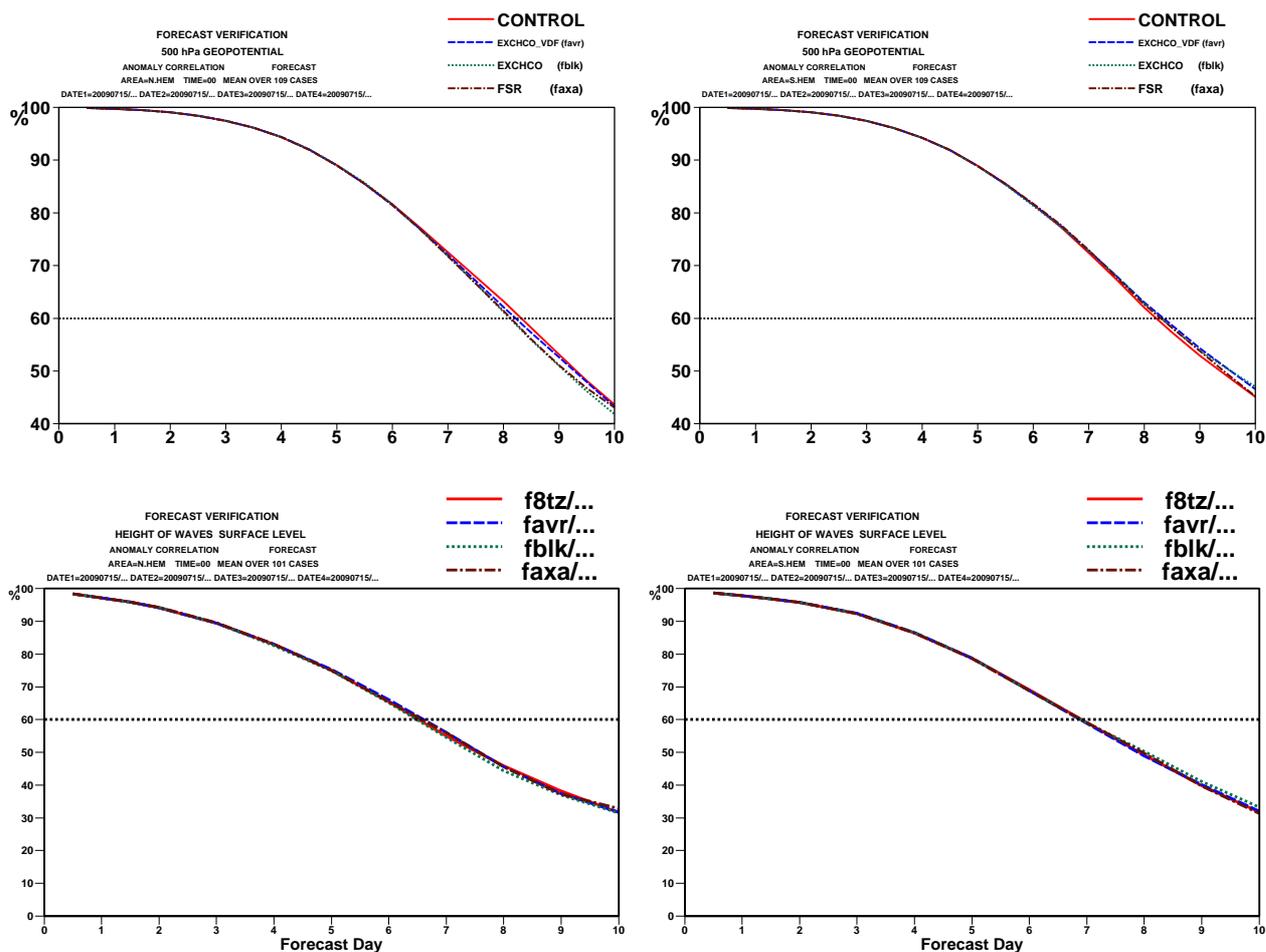
Figure 10: Anomaly correlation coefficient averaged over 109 cases over the Northern Hemisphere (left) and Southern Hemisphere (right panels) of the 500 hPa geopotential (top) and significant wave height (lower panels) for the CONTROL (red), EXCHCO_VDF (blue), EXCHCO (green) and FSR (brown) experiment. Higher values indicate a higher forecast skill.
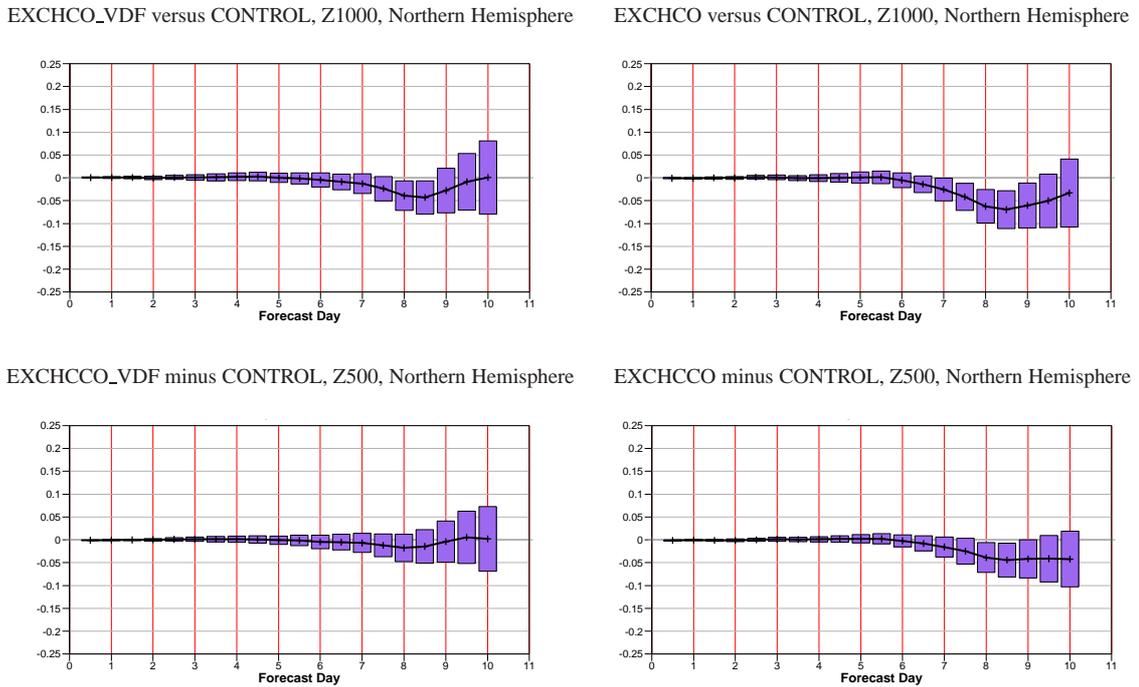
EXCHCO_VDF versus CONTROL, Z1000, Northern Hemisphere

EXCHCO versus CONTROL, Z1000, Northern Hemisphere

EXCHCCO_VDF minus CONTROL, Z500, Northern Hemisphere

EXCHCCO minus CONTROL, Z500, Northern Hemisphere

Figure 11: Normalized difference in anomaly correlation coefficient averaged over 109 cases over the Northern Hemisphere of the 1000 hPa (top) and 500 hPa (lower panels) geopotential for the EXCHCO_VDF (left) and EXCHCO (right panels) experiment. Positive values indicate an improvement, while the bars express the 95% confidence level.

EXCHCO_VDF versus CONTROL, Z1000, Southern Hemisphere

EXCHCO versus CONTROL, Z1000, Southern Hemisphere

EXCHCCO_VDF minus CONTROL, Z500, Southern Hemisphere

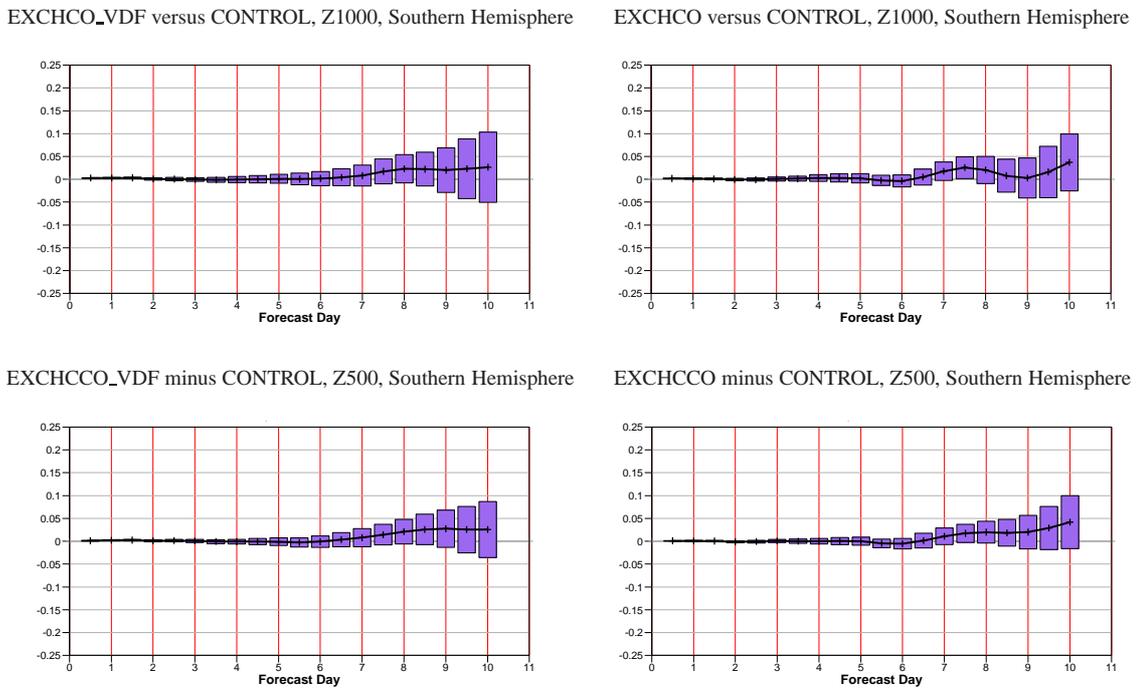EXCHCCO minus CONTROL, Z500, Southern Hemisphere

Figure 12: Normalized difference in anomaly correlation coefficient averaged over 109 cases over the Southern Hemisphere of the 1000 hPa (top) and 500 hPa (lower panels) geopotential for the EXCHCO_VDF (left) and EXCHCO (right panels) experiment. Positive values indicate an improvement, while the bars express the 95% confidence level.

Figure 13: Normalized difference in anomaly correlation coefficient averaged over 109 cases over Europe of the 1000 hPa (top) and 500 hPa (lower panels) geopotential for the EXCHCO_VDF (left) and EXCHCO (right panels) experiment. Positive values indicate an improvement, while the bars express the 95% confidence level.

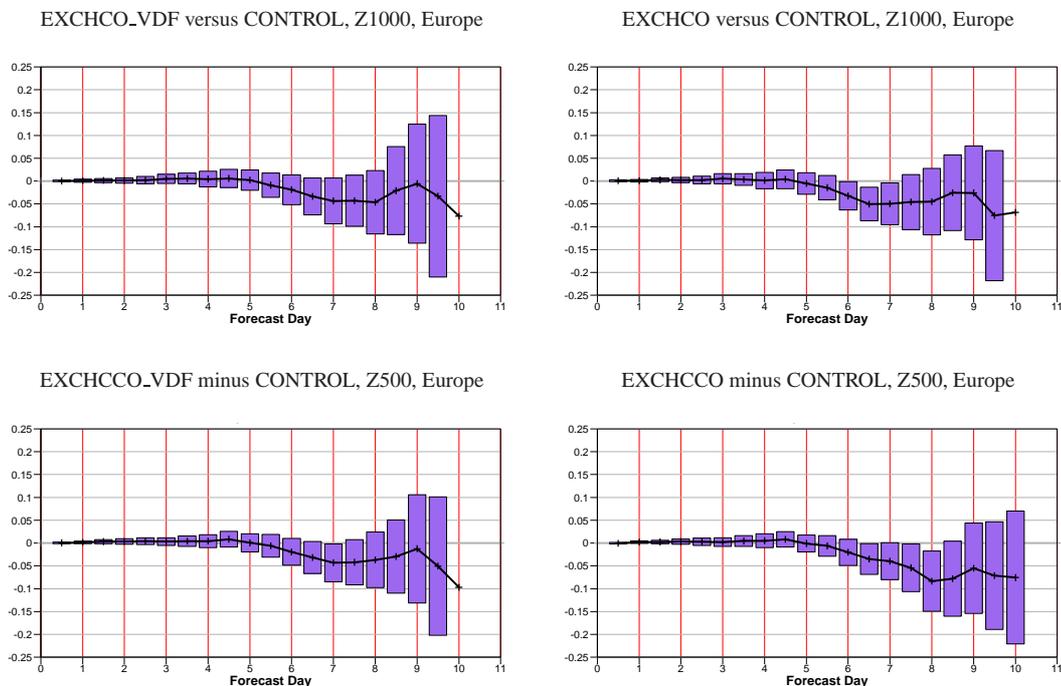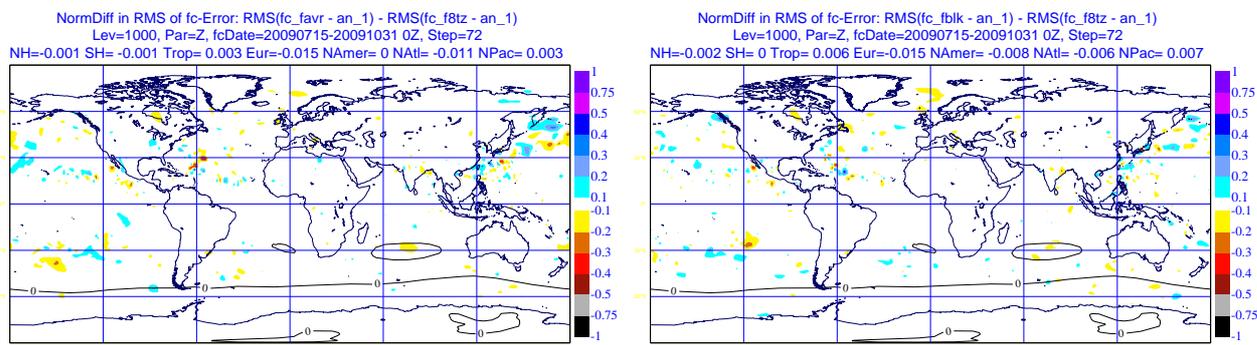*Figure 14: Normalized 3-day forecast error differences for the geopotential height at 1000 hPa compared to the CONTROL for EXCHCO_VDF (left) and EXCHCO (right), averaged over the 109-day period 20090715-20091031. Negative values indicate an improvement.*

6.4 Analysis differences

The effect on the average level of analysis fields was found to be very small (not shown). For surface wind, surface pressure, and significant wave height rather small-scale fluctuations appear over the Tropics. In general, no real coherent differences are found.

7 Discussion

In this document the assimilation of scatterometer data as equivalent-neutral 10m wind has been assessed in the ECMWF assimilation system. Rationale for the adaptation of the scatterometer observation operator is that scatterometer data is sensitive to surface stress, which is closer related to neutral than to non-neutral 10m wind. Indeed, it is found that the usage of an observation operator for neutral wind has a positive effect on departure statistics. Based on experiments performed at T511 in early-delivery mode for 109 cases in Autumn 2009, the impact on forecast skill appears rather neutral, although on the Southern Hemisphere some positive impact is found. The usage of scatterometer wind as neutral wind has become the default in cycle 36r3, and it is the objective to introduce this method in the operational suite for cycle 36r4.

The adaptation of the observation operator for surface (non-neutral) wind to neutral wind is in principle straight-forward. One complication is that the method that extrapolates the available lowest-level model wind to observation height Geleyn (1988) relies on the knowledge of neutral and non-neutral exchange coefficients b_N and b_D, respectively. These quantities are not directly available at the relevant part of the code, and have to be reconstructed. Currently (cycle 36r1), they are recomputed on the basis of Louis *et al.* (1982), which represents an old version of the ECMWF surface layer (SL) package. For both neutral 10m wind and non-neutral 4m wind (buoy data) differences are found with respect to a calculation based on the actual model SL. Although these differences are not large, they are not negligible with respect to the magnitude of the adaptations that are to be made to the observation operator for neutral wind. The neutral exchange coefficient is directly related to surface roughness. Rather than using of the actual model surface roughness (FSR) a climatological field (SR) is used, which is typically one order of magnitude too high over water. It appears, though that this constant climatological value (of 1mm) gives better results for Louis *et al.* (1982) than FSR does, probably because this SL scheme was developed for a fixed ratio between lowest-model level and roughness length, and FSR sensitively depends on surface wind speed.

To improve on this situation, a method was described that allows for a recalculation of exchange coefficients from the current physics parametrization. It was shown that this can be achieved by the inclusion of 10m neutral wind in the GOM arrays, which provide the communication between the model and the observation operator. Besides for the obvious case of an operator for 10m neutral wind, this quantity also provides the correct determination of (non-neutral) wind at other heights.

The fetching of neutral wind from the model physics assimilation system, requires that account is taken of tangent-linear (TL) and adjoint (AD) perturbations. If not, such observation operator would not depend on the control vector in 4D-Var, and as a result associated observations are not assimilated.

Sofar, perturbations in surface fields have been neglected in the ECMWF assimilation system. This is the correct approach in the frame work of 3D-Var, where no model integration is performed and unless the control vector contains surface fields itself, this will not influence information at the surface. For 4D-Var, though, the situation is different. Many surface quantities are updated diagnostically every time step, and in this way the control vector does affect those fields. Neutral wind is a good example.

It appears possible to extend IFS in such a way that full account is given for perturbations in diagnostic surface

fields. For the case of neutral wind it is found that the assimilation system performs very similarly to the case where 10m wind is estimated from lowest model wind. Also, when perturbations in 10m neutral wind are neglected, departures for scatterometer data appear relatively unchanged, i.e., the data is not assimilated. As an alternative a simpler option is provided as well. It is argued that perturbations in neutral wind could be estimated from lowest (non-neutral) wind under the assumption that stability does not significantly change in the minimization. It was found that both methods led to similar results in the minimization.

One complication is that the storage of model quantities into GOM arrays occurs before the call to the physics package which updates such quantities. For the first time step, therefore, values have to be fetched from first-guess fields. For neutral wind this requires the provision of a (archived) model field, which required an appropriate extension in the FULLPOS part of the IFS code. Since such initial surface fields are not influenced by the control vector, for this first time step perturbations have to be fetched from lowest model level wind to guarantee the assimilation of associated observations. The complication at the first time step could be remedied by storing model-field information into GOM arrays after the call to the model physics. Such a change would make the system more transparent, and it is suggested that this could be taken into consideration for the follow-up of IFS.

Rather than the inclusion of neutral wind in the GOM arrays, one could also consider the fetching of the exchange coefficients themselves. Although surface roughness is a fast changing quantity over the ocean, the neutral coefficient b_N, which is directly related to this behaves much more mildly. This facilitates horizontal interpolation. The omission of TL and AD perturbations in b_N and b_D effectively result in the neglect of changes in stability during the minimization such as expressed in (21). In other words, an observation operator for neutral wind will in this case still be able to incorporate essential information from the control vector. Such a replacement of neutral wind implies that the exchange coefficients are to be added in the list of surface fields. In addition, in the current order in which the model physics is called, initial first-guess fields are to be provided, i.e., archiving into MARS.

The incorporation of perturbations in neutral wind in the minimization as discussed in this document could be applied to other diagnostic surface fields as well. One candidate could be the skin temperature. Cardinali *et al.* (1994) found that it was difficult to assimilate 2m temperature over land, because a change in this quantity was not accompanied by a dynamically consistent change in surface temperature. This could lead to unphysical changes in the lapse rate. It would be worthwhile to investigate whether the account of perturbations in the diagnostic skin temperature could improve on this situation.

Acknowledgements

A Acronyms and abbreviations

A.1 General acronyms

4D-Var	Four-dimensional variational data assimilation
AD	Adjoint
AMI	Active Microwave Instrument
ASCAT	Advanced SCATterometer
CMOD	C-band Geophysical MODel function
DA	Daily Archive
DCDA	Delayed-cutoff Daily Archive
ECMWF	European Centre for Medium-range Weather Forecasts
ERS	European Remote sensing Satellite
ESA	European Space Agency
ESRIN	European Space Research INstitute
FDB	Fields Data Base
FSR	Forecast surface roughness
FULLPOS	Part of IFS that handles the post-processing for fields
GMF	Scatterometer Geophysical Model Function, relating wind to backscatter
GOM	Set of arrays that provide horizontally interpolated model information to the observation operator
IFS	ECMWF Integrated Forecasting System
L3DFGAT	IFS logical that indicates that the ECMWF model is run in the minimization
LECMWF	IFS logical that indicates the configuration for ECMWF
lhs	left-hand side
LSCATT_NEUTRAL	IFS logical that controls whether scatterometer data is assimilated as neutral wind
LVDFMIN	New IFS logical that controls whether perturbations in neutral wind are to be included
LVDFTRAJ	New IFS logical that controls whether neutral wind is to be fetched from the IFS SL
LZ0FC	New IFS logical that stores FSR rather than SR in the GOM arrays
MARS	Meteorological Archival and Retrieval System
NAMOBS	IFS namelist relating to observation operators
NSCAT-2	GMF for the NSCAT scatterometer
NSTEP	Time step in IFS, ranging from 0 (initial) to NSTOP (final time step)
NSTOP	Final time step
OCFLX	Offline version of the IFS SL physics over water
QSCAT-1	GMF for the QuikSCAT scatterometer
rhs	right-hand side
SL	Surface Layer
SR	Analysis surface roughness, which is based on climatology
T511	A triangular spectral grid with 511 waves along the equator
TL	Tangent linear
UTC	Coordinated Universal Time
U10N	u-component for 10m neutral wind as archived in MARS
V10N	v-component for 10m neutral wind as archived in MARS
WAM	third-generation ocean-WAve Model
Z1000	Geopotential height at 1000 hPa
Z500	Geopotential height at 500 hPa

A.2   Relevant IFS subroutines

`callpar`	Handles the ECMWF parametrization
`cobs`	Prepares model fields for storage into GOM arrays
`cobslag`	Fills GOM arrays
`defrun`	Defines and checks control parameters
`ec_phys`	Handles the ECMWF physics
`exchco`	Routine under `hop` that estimates exchange coefficients based on Louis (1979)
`exchco_vdf`	New routine under `hop` that estimates exchange coefficients from `vdfmain`
`gp_model`	Handles model calculations in grid point space
`hop`	Evaluates the observation operator
`obshor`	Provides the horizontal interpolation from model fields to observation location
`pppobsas`	Routine under `hop` that calculates the observation operator for surface observations
`ppuv10m`	Routine under `hop` that handles the vertical interpolation of wind
`preints`	Routine under `hop` that estimates exchange coefficients for surface observations
`scan2m`	High-level routine that encapsulates all computations in grid-point space
`sppcfl`	Routine under `vdfmain` that computes 2m temperature and 10m wind
`sppcfls`	Simplified-physics version of `sppcfl`
`step0`	High-level routine that handles all operations for one time step
`surbound`	Routine under `hop` that pre-calculates quantities that are required for `exchco`
`surface_fields_mix`	Module that defines groups of surface fields
`su_surf_flds`	Handles the initialization and setup for surface field
`vdfmain`	Routine in the ECMWF physics that includes the calculation of the SL
`z0sea`	Estimates surface roughness from neutral wind and Charnock parameter

Any associated tangent-linear and adjoint routines have the annex `tl` and `ad`, respectively.

B   Overview of code changes

The changes that were made in dal_CY36R1_neutral_full_dependencies , and which have been merged into cycle 36r3 are:

- The introduction of the diagnostic surface fields (group VDIAG), and their archiving into MARS of:

    - Neutral wind at 10 m x-component, y-component (131.228, 132.228)
    - Friction velocity (003.228)

- The calculation of 10-m neutral wind components and friction velocity in VDIAG (`sppcfl_mod`), For neutral wind, adaptations in simplified physics and its adjoint/tangent-linear code have been made as well (`sppcfls_mod`, `sppcflsad_mod`, `sppcflstl_mod`).

- Usage of first-guess fields of 10-m neutral wind components, forecast surface roughness, and forecast surface roughness for heat as input to the analysis.

- The introduction of 10-m neutral wind components in the GOM arrays. A logical LUPD informs whether surface fields have been updated.

- The introduction of the logicals LVDFTRAJ, LVDFMIN and LZ0FC in YOMOBS.

- Communication of perturbations in 10m-neutral wind components to corresponding GOM arrays (`cobsad/tl`). It would be straightforward to extend this new structure to other surface fields.

- Regarding the horizontal interpolation from model fields to observation location (`cobs`), the bookkeeping on which fields behave like scalars of vectors, is made more transparent.

- The resetting of surface perturbations at start and final steps in the adjoint and tangent-linear code has been moved to a higher level from `ec_physad/tl` to `scan2mdad/tl`. In addition, only perturbations are reset, rather than the complete (un-necessary) reset of associated non-linear surface fields.

- Information on the magnitude of the neutral-wind GOM arrays is printed for scatterometer data (`gathergom, prtgom`).

- Similarly to the requirement of the assimilation of all-sky microwave radiances (LEMWAVE, Bauer *et al.* (2006)), it is ensured that the post-processing of surface wind is being called under ec_phys when LSCATT_NEUTRAL=.true. (`suphli`).

- When LSCATT_NEUTRAL=.true., Charnock is saved in the high-res trajectory, which is then interpolated to the resolution of the minimization (KTRAJ=1 in `su_surf_flds`). Otherwise, a constant Charnock (value 0.018) is used in the minimization (KTRAJ=2).

References

Abdalla, S. and Hersbach, H. (2007). Final report for esa contract 18212/04/i-lg. *Esa contract report*, ECMWF, Shinfield Park, Reading, [available at http://www.ecmwf.int/publications/library/].

Bauer, P., Salmond, D., Benedetti, A., Saarinen, S. and Moreau, E. (2006). Implementation of 1d+4d-var assimilation of precipitation-affected microwave radiances at ecmwf. ii: 4d-var. *Quart. J. Roy. Meteor. Soc.*, **132**, 2307–2332.

Brown, A. R., Beljaars, A. C. M. and Hersbach, H. (2006). Errors in parametrizations of convective boundary layer turbulent moment mixing. *Quart. J. Roy. Meteor. Soc.*, **132**, 1,859–1,876.

Cardinali, C., Andersson, E., Viterbo, P., Thépaut, J.-N. and Vasiljević, D. (1994). Use of conventional surface observations in three dimensional variational assimilation. *Ecmwf technical memorandum*, ECMWF, Shinfield Park, Reading, [207].

Charnock, H. (1955). Wind stress on a water surface. *Quart. J. Roy. Meteor. Soc.*, **81**, 639–640.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. (1994). A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120**, 1367–1388.

Ebuchi, N., Graber, H. C. and Caruso, M. J. (2002). Evaluation of wind vectors observed by quikscat/seawinds using ocean buoy data. *J. Atmos. Oceanic Technol.*, **19**, 2,049–2,062.

Freilich, M. H. and Dunbar, R. S. (1999). The accuracy of the nscat 1 vector winds: comparisons with national data buoy center buoys. *J. Geophys. Res.*, **104 (C5)**, 11,231–11,246.

Geleyn, J. F. (1988). Interpolation of wind, temperature and humidity values from model levels of the height of measurement. *Tellus*, **40A**, 347–351.

Hersbach, H. (2010a). Calculation and archiving of friction velocity and 10m equivalent-neutral wind from ifs. *Ecmwf research department memorandum*, ECMWF, Shinfield Park, Reading, [R60.9/HH/1034].

Hersbach, H. (2010b). Comparison of c-band scatterometer cmod5.n equivalent neutral winds with ecmwf. *J. Atmos. Oceanic Technol.*, **27**, 721–736.

Hersbach, H. (2010c). A fit for sea-surface roughness and drag. *Ecmwf technical memorandum*, ECMWF, Shinfield Park, Reading, [DRAFT].

Hersbach, H. and Bidlot, J.-R. (2009). The relevance of ocean surface current in the ecmwf analysis and forecast system. In *Proc. ECMWF Workshop on Ocean-atmosphere interactions*, pp. 61–73, ECMWF, Reading, 10-12 November 2008.

Hersbach, H., Stoffelen, A. and de, S. H. (2007). An improved c-band scatterometer ocean geophysical model function: Cmod5. *J. Geophys. Res.*, **112 (C3)**, doi:10.1029/2006JC003743.

IFS-documentation (2009). edited by Bob Riddaway. *available from www.ecmwf.int/research/ifsdocs/*.

Janssen, P. A. E. M. (1991). Quasi-linear theory of wind wave generation applied to wave forecasting. *J. Phys. Oceanogr.*, **21**, 1,631–1,642.

Louis, J.-F. (1979). A parametric model of vertical eddy fluxes in the atmosphere. *Bound-Layer Meteorol.*, **17**, 187–202.

Louis, J.-F., Tiedtke, M. and Geleyn, J.-F. (1982). A short history of the pbl parametrization at ecmwf. In *Proc. ECMWF Workshop on Planetary Boundary Layer Parametrization*, pp. 59–80, ECMWF, Reading, 25-27 November, 1981.

Monin, A. S. and Obukov, A. M. (1954). Basic regularity in turbulent mixing in the surface layer of the atmosphere. *Geofiz. Inst., Akad. Nauk SSSR*, **24**, 163–187.

Vasiljevic, D., Cardinali, C. and Undén, P. (1992). Ecmwf 3d variational assimilation of conventional observations. In *Proc. ECMWF Workshop on Variational Assimilation with Emphasis on Three-dimensional Aspects*, pp. 389–436, ECMWF, Reading, 9-12 November, 1992.

Verhoef, A., Portabella, M., Stoffelen, A. and Hersbach, H. (2008). Cmod5.n - the cmod5 gmf for neutral winds. *Document external project, saf/osi/knmi/tec/tn/165*, EUMETSAT, [Available at http:www.knmi.nl/scatterometer/publications/].

Wentz, F. J. and Smith, D. K. (1999). A model function for the ocean-normalized radar cross section at 14 ghz derived from nscat observations. *J. Geophys. Res.*, **104 (C5)**, 11,499–11,514.