Technical Memo

European Centre for Medium-Range Weather Forecasts

866

On the causes of systematic forecast biases in near-surface wind direction over the oceans

Irina Sandu, Peter Bechtold, Louise Nuijens[1], Anton Beljaars and Andy Brown

[1] Delft University of Technology, Netherlands

June 2020

Series: ECMWF technical memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/en/publications/

Contact: library@ecmwf.int

Abstract

The demand for more accurate forecasts of near-surface weather is growing at a rapid pace, especially in the context of increased use of renewable energy. Although improvements in numerical weather prediction systems have lead to improved forecasts of near-surface weather parameters over the years, systematic biases remain. Here we examine one of the most longstanding biases in ECMWF forecasts of near-surface weather, the bias in surface wind direction over the oceans. Since the 1990's the forecasted surface wind direction is rotated clockwise (anticlockwise) with respect to scatterometer (ASCAT) observations in the Northern (Southern) Hemisphere. Using a conditional analysis of the short-range forecast errors against ASCAT observations, we demonstrate that unstable boundary layers contribute the most to the remaining mean bias in surface wind direction over the oceans (3-5 degrees), although stable boundary layers that occur in the mid-latitude storm tracks during summer also contribute to some extent. Focusing on a typical trade wind region upstream of Barbados and using sensitivity experiments with the Integrated Forecasting System, we demonstrate that the surface wind direction bias in unstable boundary layers is related to the representation of turbulent and shallow convection momentum transport, and in particular to an apparent lack of friction in the lower part of the cumulus layer. We also show that both the forecast and analysis wind profiles are strongly sensitive to the momentum transport by shallow convection in the lowest 1.5 km in this region, which suggests that the analysis is only weakly constrained by wind observations in the trade wind boundary layer (at least prior to the assimilation of Aeolus observations). An evaluation of both forecast and analysis wind profiles with in-situ observations such as those provided by the EUREC4A field program, which recently took place around Barbados, would therefore be very valuable in the future.

1 Introduction

The increased importance of renewable energy and the need to forecast consequences of extreme events on near-surface weather well in advance call for better predictions of near-surface weather parameters such as winds, temperature or downwelling shortwave radiation. This requires a reduction of both systematic and random forecast biases, and an accurate representation of the sources of model and observation uncertainty in ensemble forecasting.

In this study we take a fresh look at one of the most systematic and longstanding biases in forecasts of near-surface weather in the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF), the bias in wind direction at the surface. Since the 1990's it has been found that the forecasted surface wind direction is generally veered (rotated clockwise) with respect to observations in the Northern Hemisphere (NH) both over land and over the oceans, while in the South Hemisphere (SH) it is backed (rotated anticlockwise) with respect to observations (*Hollingsworth* 1994). As schematically depicted in Figure 1, in the NH the surface wind is backed relative to the geostrophic wind, so that the wind veers with height through the boundary layer (the opposite being true in the SH). Assuming that the modelled geostrophic wind is similar to the observed geostrophic wind (*Brown et al.* 2005), the veering of the modelled surface wind with respect to observations translates into an underestimation of the wind turning throughout the boundary layer.

The systematic veering (backing) of the modelled surface wind direction in the IFS with respect to observations over the oceans in the NH (SH) is illustrated in Figure 2 by the negative (positive) differences in surface wind direction between scatterometer (ASCAT) observations and the short-range (12 hours) forecasts used in the data assimilation processs. Through the data assimilation process, short-range forecasts (or first-guesses) are optimally blended with millions of observations to create the initial conditions for the medium and extended range forecasts, or the analyses, which constitute the best possible recon-

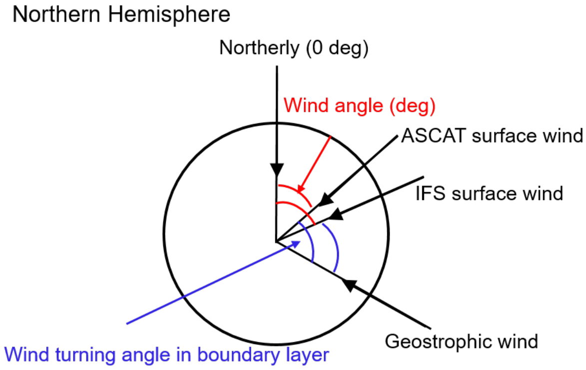

Figure 1: *Schematic representation of typical model errors in surface wind direction in the Northern Hemisphere, illustrated for north-easterly (trade-like) winds. In the Northern Hemisphere surface winds are backed with respect to the geostrophic wind, so that the wind vector veers (rotates clockwise) with height throughout the boundary layer. The wind direction at the surface (i.e. defined as the angle with respect to a northerly direction, red) has larger values in the IFS than in the ASCAT observations, which indicates that the model surface wind is veered with respect to the observations. Assuming the geostrophic wind direction is correct, this implies that the modelled wind turning throughout the boundary layer (blue) is underestimated.*

struction of the atmospheric state. The operational ECMWF analyses, and the deterministic ten days forecasts discussed in this study are nowadays produced at a horizontal resolution of 9 km, with 137 levels on the vertical. Although the errors in surface wind direction have somewhat decreased over time, in particular in the tropical convergence band due to successive changes to the forecasting system (see Figure 2, and Section 2 for a more in depth discussion), they remain qualitatively similar to those shown initially by *Hollingsworth* (1994) and then discussed in *Brown et al.* (2005; 2006). The forecasted surface wind direction is also veered (in the NH) with respect to observations over land (Figure 3). As over the oceans, the wind direction errors over land have decreased over time as the forecasting system components (model physics, numerics, resolution, data assimilation, use of observations) have improved (*Bauer et al.* 2015), but a mean bias of approximately 3 to 5 degrees remains. As discussed in Section 2 biases also remain in the near-surface wind speed. Further reducing these near-surface wind biases is important both for increasing the quality of wind forecasts at local scale and for improving the representation of the large-scale circulation. A better representation of winds in the boundary layer often translates into a better representation of surface drag (or stress); and surface drag plays a key role in the large-scale circulation (*Chen et al.* (2007), see also Section 2).

In a first detailed investigation of the causes of surface wind direction biases over oceans, *Brown et al.* (2005) concluded that although systematic these biases are more pronounced in regions where warm air is advected over cold sea surface temperatures (warm advection conditions hereafter). They showed that a large fraction of the cases with the larger errors are characterized by stable boundary layers, and concluded that a reduction of turbulent mixing in such conditions would be beneficial. As it will be discussed in more detail in the next section, turbulent mixing in stable conditions has been long known to be overestimated in global Numerical Weather Prediction (NWP) models, including the IFS (*Holtslag et al.* 2013, *Svensson and Holtslag* 2009). Moreover, *Brown et al.* (2005) suggested that the large surface wind direction errors in warm advection cases are also sensitive to the representation of mixing in unstable boundary layers, due to both convective and turbulent transport of momentum. *Brown et al.* (2006) went one step further and used Large-Eddy Simulations (LES) to explore the accuracy of the turbulence closure used for unstable situations in NWP models, for a variety of cases with different stability and baroclinicity. They concluded that a systematic error of about 5 degrees in surface wind direction is present for shallow boundary layers, when the modelled wind profiles are more well mixed than what is found in observations or LES. *Brown et al.* (2006) hypothesized that this could be due to a missing non-local counter-gradient term in the turbulence closure, such as that proposed by *Brown and Grant* (1997).

More recently, *Belmonte Rivas and Stoffelen* (2019) examined the differences in surface winds between the two most recent reanalysis data sets produced by ECMWF (ERA-INTERIM and ERA5) and scatterometer ASCAT observations. Reanalyses are the best possible reconstruction of the past atmospheric state, created with modern NWP forecasting systems such as the IFS, by optimally blending short-range forecasts and observations through powerful data assimilation methods such as the four-dimensional variational data assimilation used in the IFS (*Rabier et al.* 2000). Reanalyses of a past period (e.g. 1979 - present) are produced by using all the available past observations and a consistent recent version of the NWP system throughout the entire period. For example, ERA5, performed by ECMWF for the Copernicus Climate Change Service (*Hersbach et al.* 2019), is produced with cycle 41r2 of the IFS which was operational from March to November 2016, at a resolution of 30 km and with 137 levels on the vertical.

Comparing the first-guess winds from ERA5 and ERA-INTERIM with ASCAT winds, *Belmonte Rivas and Stoffelen* (2019) found that short-range near-surface wind biases over oceans are approximately 20 % lower in ERA5 than in ERA-INTERIM, due to the use of an improved NWP forecasting system and higher resolution, and are comparable to those of recent operational ECMWF forecasts (e.g. 2017 or 2019 shown in Figures 2 and 4). However, short-range forecast winds in ERA5, and ECMWF operational forecasts, are still characterized by excessive mean zonal winds (too westerly in the midlatitudes and too easterly in the tropics), defective mean meridional winds (not poleward enough in midlatitudes and not equarward enough in the trade regions) and veered (backed) surface winds in the NH (SH) compared to observations (Figures 5 and 7 of *Belmonte Rivas and Stoffelen* (2019) and Figures 2 and 4 here). This corroborates the analysis done in *Simpson et al.* (2018) which hypothesized that these errors could be due to some missing drag on low-level zonal flows. *Belmonte Rivas and Stoffelen* (2019) demonstrate that errors in the mean wind speed and direction in ERA-INTERIM and ERA5 are accompanied by errors in the transient component of the winds, and more precisely by an underestimation of transient wind activity compared to observations. They argue that some additional transient wind variability would induce a residual meridional circulation in the Ferrel cell, which would substract energy from the mean zonal wind, correct the meriodional mean wind biases and the stress curl (direction) biases. The underestimation of the transient wind variability could be due to a misrepresentation of mesoscale convective variability, and wind shear, which was also previously suggested by *Houchi et al.* (2010).

In this study we examine the systematic forecast errors in surface wind direction over the oceans in the IFS from a different angle. We combine an analysis of the short-range forecast errors against ASCAT observations, with an extensive set of sensitivity experiments performed with the IFS to identify the most likely causes of these longstanding errors. In Section 2 we relate the time evolution, and the gradual reduction of these biases, to known changes to the IFS. In Section 3 we highlight which are the regimes contributing most to the mean biases in surface wind direction shown in Figure 2, based on a stratification of these short-range forecast biases for different conditions. We then discuss the sensitivity experiments performed to assess how changes to the model physics affect the representation of boundary layer winds (Section 4). The only change that eliminates nearly entirely the surface wind direction bias is to turn off the momentum transport by shallow convection. Although this is not a desirable model change, it highlights very well the uncertainties in the representation of momentum transport in unstable boundary layers and in particular in the partition between turbulent and convective transport. Finally, in Section 5 we examine how forecast errors in wind profiles evolve with leadtime in an area typical of the trade

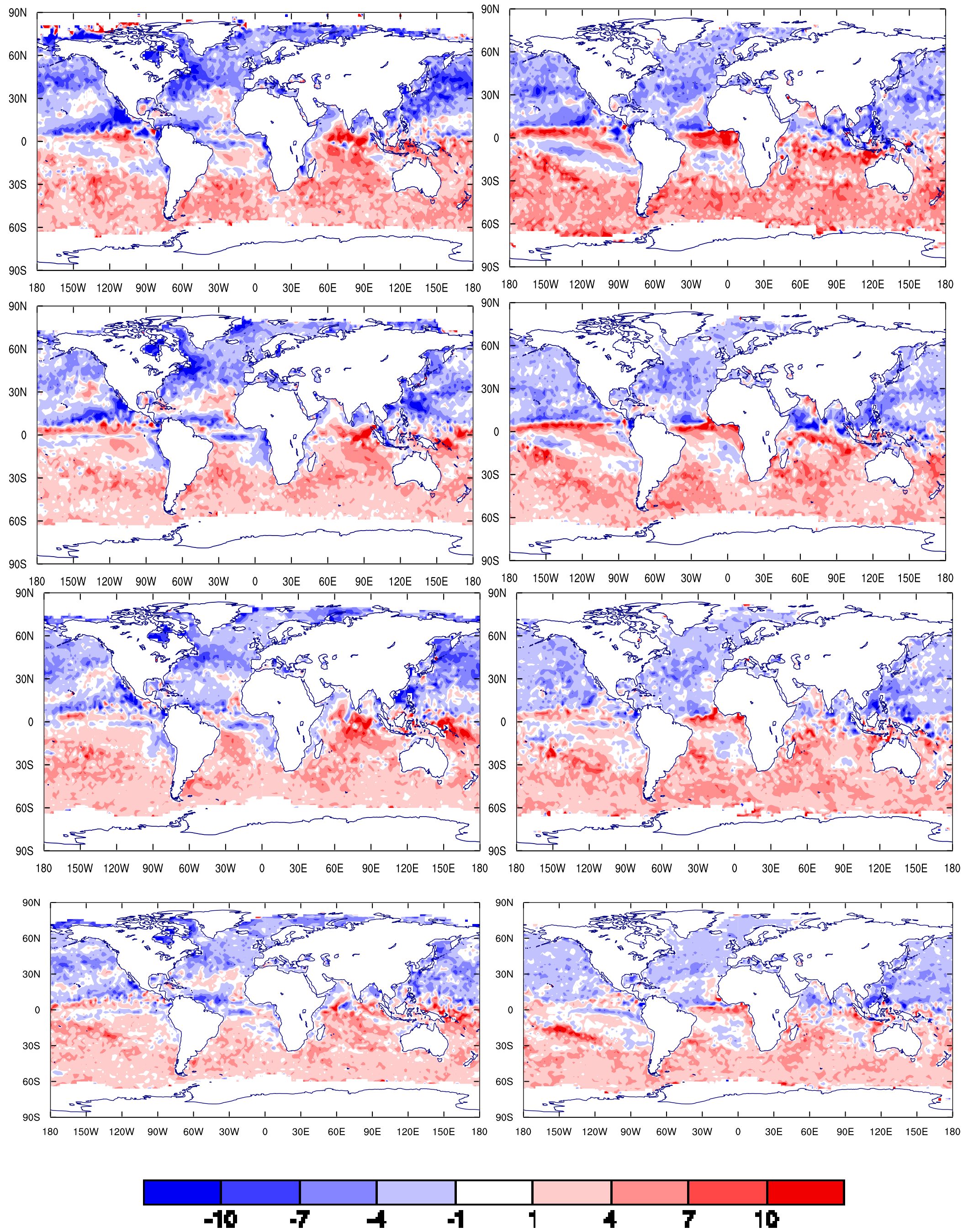

Figure 2: *Monthly mean first-guess departures (observations - short-range forecasts) of surface wind direction (degrees) with respect to ASCAT observations in the ECMWF operational deterministic forecasts for July (left) and December (right), for 2012, 2014, 2017 and 2019. These values represent averaged statistics from the 0000 and 1200 UTC analysis cycles, and observations and forecasts are binned on a 2 by 2 degrees grid.*

winds, upstream of Barbados, and attempt to relate the error growth to the processes at play. The aim is to understand whether forecast biases seen at short-range against ASCAT observations are representative of forecast errors in the medium-range, which are the main focus at ECMWF. The discussion in Section 5

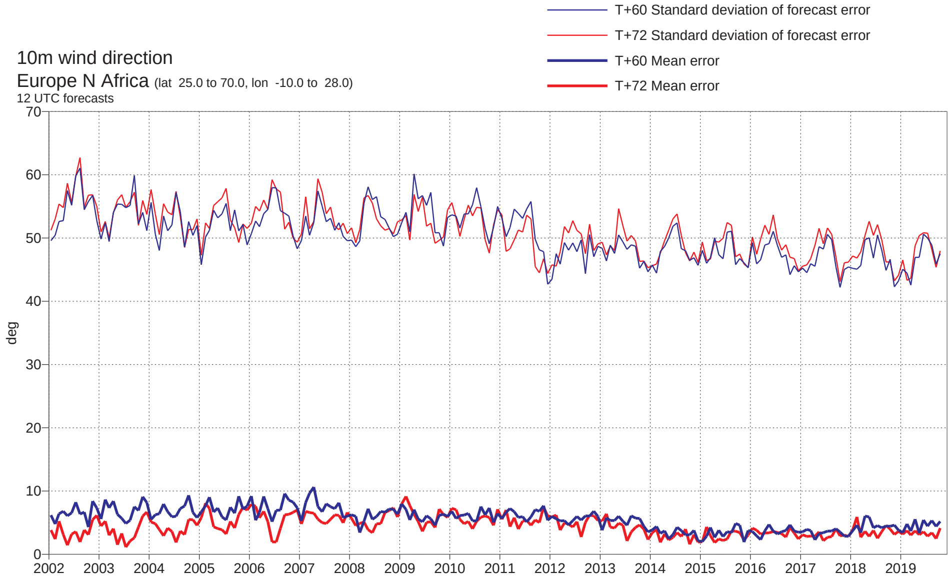

Figure 3: *Evolution of surface wind direction errors over Europe from 2002 to 2019 for the ECMWF operational deterministic forecasts. These are monthly values of mean errors (thick lines) and standard deviation (thin lines) at a lead time of 60 (red) and 72 hours (blue) of the daily forecasts initialized at 1200 UTC (verifying at 00 and 12 UTC, respectively). The verification is against observations from SYNOP stations over Europe (25-70N, 10W-28E).*

also highlights the fact that wind profiles in the boundary layer are not well constrained by observations, and that more observational constraints on the momentum budget, in particular in the cloud layer, are needed.

2 Time evolution of systematic surface wind direction biases in the IFS

As the NWP systems have tremendously improved over the past decades (*Bauer et al.* 2015), the forecast skill of upper air and near-surface weather parameters has continuously increased (*Haiden et al.* 2019). Surface wind direction biases in operational ECMWF forecasts have thus also gradually decreased both over ocean and land (Figures 2 and 3). While the verification against ASCAT observations (shown in Figures 2 and 4) can be easily done only in the framework of the data assimilation window, so for short-range forecasts, the verification over land using synop observations, suggests that surface wind direction biases are similar in the short and medium-range, albeit they slightly decrease with leadtime (not shown). We will come back to this aspect, of how wind biases evolve with leadtime, in Section 5. In this section we discuss the reduction of surface wind direction biases from the perspective of changes made to the IFS over time.

As many other global Numerical Weather Prediction (NWP) models, IFS has for a long time maintained too much diffusion in stable conditions compared to what would be expected based on the Monin-Obukhov similarity theory (*Holtslag et al.* 2013, *Svensson and Holtslag* 2009). In first order turbulence schemes using a K-diffusion closure such as that of the IFS, this is often done by prescribing so-called 'long-tail' stability functions which maintain considerably more mixing for Richardson numbers larger than 1 than the 'short-tail' functions prescribed by the Monin-Obukhov theory (see Fig. 1 of *Sandu*

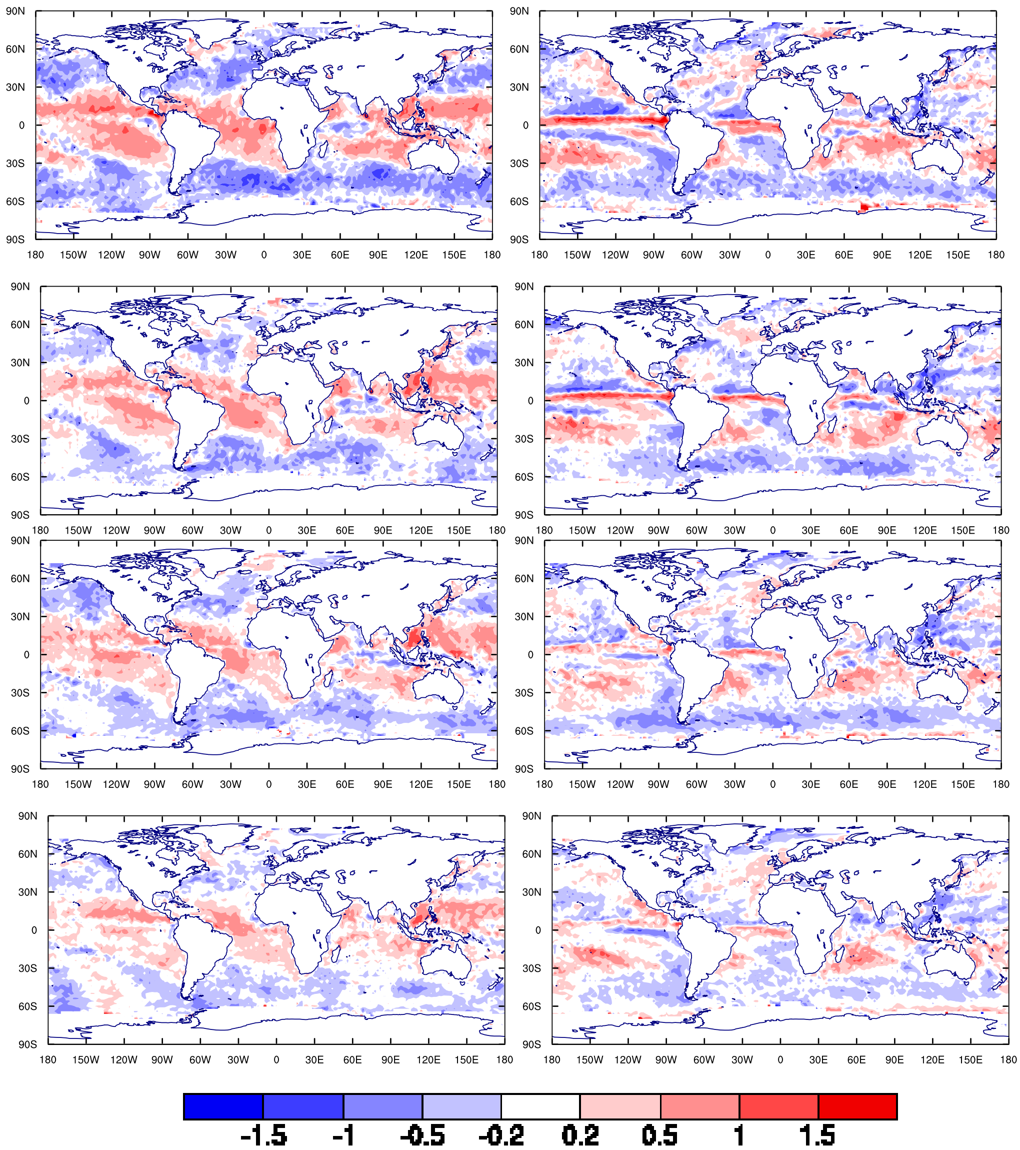

Figure 4: *Monthly mean first-guess departures (observations - short-range forecasts) of surface zonal (left) and meridional (right) wind (m/s) with respect to ASCAT observations in the ECMWF operational deterministic forecasts for December 2012, 2014, 2017, and 2019. These values represent averaged statistics from the 0000 and 1200 UTC analysis cycles, and observations and forecasts are binned on a 2 by 2 degrees grid.*

et al. (2013)). Additionally, the mixing can be (and is often) enhanced through the values chosen for the asymptotic mixing length entering the computation of the diffusion coefficients. Maintaining an artificially enhanced diffusion in stable boundary layers in global NWP models has been shown to be an effective way to reduce the cold near-surface temperature biases frequently encountered in stable boundary layers and to improve the representation of synoptic cyclones (*Beljaars and Viterbo* 1998, *Viterbo et al.* 1999). Maintaining more mixing in shallow stable boundary layers means that the radiative cooling at the surface is distributed over a deeper layer, allowing to maintain warmer near-surface temperatures. It thus avoids a runaway cooling regime in which the near-surface air becomes increasingly colder (due to radiative cooling at the surface) if very little mixing is maintained in an increasingly stably stratified boundary layer. Enhanced mixing in stable boundary layers also results in enhanced surface drag (or stress). Surface drag is very effective in controlling the large-scale circulation (*Chen et al.* 2007), in particularly in the North Hemisphere winter affecting both the synoptic systems and the planetary waves (*Sandu et al.* 2013). Until a few years ago, maintaining an enhanced surface drag by artificially enhancing the diffusion in stable boundary layers was seen as the only solution for maintaining a statisfactory large-scale performance of the IFS during the NH winter (*Sandu et al.* 2013).

Nonetheless, the enhanced diffusion in stable boundary layers also has detrimental effects for the quality of the forecasts, leading to an underestimation of the wind turning through the boundary layer, a veering (backing) of the surface winds in the NH (SH) and in too weak low-level jets (*Sandu et al.* 2013). Not surprisingly, the surface wind direction biases in the IFS were reduced on three occasions: in 1999, when the lowest model level was lowered from 30 to 10 m (not shown), in 2007 when the turbulent diffusion was reduced in free shear layers away from the surface, and in 2013 when the turbulent diffusion in stable conditions was finally reduced (Figure 3). The 2013 reduction in turbulent diffusion in stable layers was achieved through modifications of asymptotic mixing length (the 'long-tail' stability functions are still used both in stable boundary layers and in stable layers in the free-troposphere). As described in detail in *Sandu et al.* (2013/2014), the negative impact of this change on the large-scale circulation was counteracted by an increase in the strength of orographic drag. In practice this package of changes, which allowed a better representation of near-surface winds and of the large-scale circulation, consisted in a re-partioning of the parametrized surface drag between the different schemes that contribute to it: turbulence, orographic form drag (representing drag from topography with horizontal scales less than 5 km, *Beljaars et al.* (2004)) and low-level blocking and gravity wave drag (representing drag from topography with horizontal scales larger than 5km, *Lott and Miller* (1997)). It has since become obvious that the total parametrized surface drag and especially its partition between the various schemes varies widely among the main global NWP models, especially in regions with orography (*van Niekerk et al.* 2020, *Zadra* 2015). As surface drag is largerly unconstrained by observations, while it is very effective in controlling the large-scale circulation and thereby model performance, it has been widely used as tuning knob by the modelling community and the different schemes representing drag processes have been used interchangeably to achieve the desired amount of total surface drag and large-scale performance. This has led to a large inter-model spread, which consitutes one of the main sources of uncertainty for weather and climate models (*Sandu et al.* 2019).

Figure 2 shows however that over oceans, where stable conditions are not that frequent (see discussion in Section 3), the revision of turbulent diffusion in stable conditions implemented in the IFS in November 2013 lead to a fairly small reduction of the biases in surface wind direction (see also *Sandu et al.* (2013/2014)). While the surface wind direction errors for 2012 and 2014 look fairly similar (Figure 2), the errors in the zonal and meridional wind components were significantly reduced during this time frame (Figure 4). As IFS upgrades consist in a combination of changes to model physics, numerics, data assimilation methods and use of observations, it is hard to say what exactly caused the pronounced reduction in first-guess errors for the zonal and meridonal wind components. Other major changes to the operational ECMWF deterministic forecasts that occured in 2013 are an increase in the number of vertical levels from 91 to 137, and a revision of the convection scheme that resulted in an improved diurnal cycle of precipitation (*Bechtold et al.* 2014). Neither of these changes, nor other substantial changes to the ECMWF forecasting system (or forecast configuration) such as for example a major moist physics upgrade in 2015, the increase in horizontal resolution from 16 to 9 km for deterministic forecasts in May 2016, or the improved use of observations through continous data assimilation in July 2019 (*Lean et al.* 2019), completely eliminated the first-guess departures against ASCAT observations. It worth noting however that in 2019 these short-range errors are lower than ever (Figure 2 and 4).

3 Stratification of surface wind direction biases

A first step towards identifying the causes of the remaining short-range forecast biases in surface winds over the oceans, revealed by the comparison with ASCAT observations, consisted in stratifying these biases in a variety of ways. The stratification was done using the first-guess departures statistics from the ECMWF operational analysis for the months of July and December from 2012 to 2017 (a selection of which is included in Figures 2 and 4).

Similarly to *Brown et al.* (2005), the stratification aimed at disentangling which are the cases for which the largest errors are found by comparing the mean biases for different types of conditions, e.g. stable versus unstable boundary layers (defined as cases with negative versus positive surface buoyancy fluxes) or warm versus cold advection (defined as southerly versus northerly winds in the NH, and the opposite in the SH). The stratification was extended to include a multitude of other criteria: strong versus weak inversions at the top of the boundary layer (defined using either a lower tropospheric stability (*Klein and Norris* 1995) or an estimated inversion strength criterion (*Wood and Bretherton* 2006)), different degrees of cloudiness (defined using different classes of total cloud cover), backward or forward shear across the lower troposphere (defined as zonal wind increasing or decreasing with height), different degrees of boundary layer stability (defined using different ranges of surface buoyancy flux) or boundary layer depth (defined using the boundary layer height diagnostic), predominant wind direction (southerly/northerly/westerly/easterly), etc.

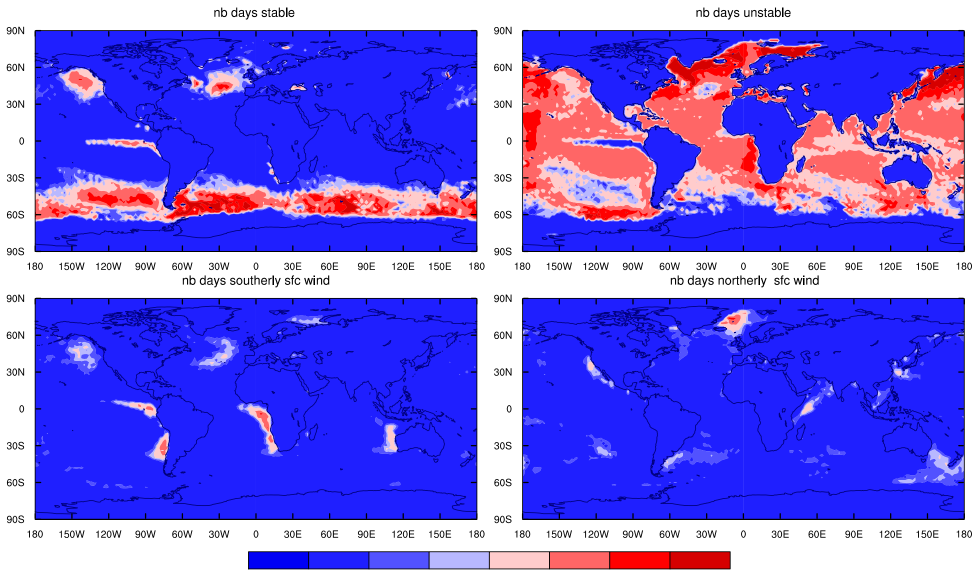

Figure 5: *Number of cases (for each 2x2 degrees box) characterized by (a) stable conditions, (b) unstable conditions, (c) southerly and (d) northerly winds in the short-range (12 hours) operational ECMWF deterministic forecasts for December 2017. As both the 00 and 12 UTC analysis cycles are considered, the maximum number of cases is 62. Stable/unstable cases are defined as cases for which the buoyancy flux is smaller/bigger than 5 W/m2 (the stable category thus also includes neutral cases); southerly/northerly winds are defined as cases for which the wind direction is between 150 and 210 degrees, and within 30 degrees from the northerly direction (as in Brown et al. (2005)).*

The novelty compared to the stratification done by *Brown et al.* (2005) was to multiply the mean bias for each of these different conditions by the frequency of occurence of those conditions within the respective month (see as an example in Figure 5 the number of cases characterized by stable/unstable boundary layers, and northerly/southerly winds for December 2017). This allows to highlight how much the biases in each category contribute to the mean bias in surface wind direction (or wind component) for the respective month shown in Figures 2 and 4. Figure 5 shows that, unsurprisingly, unstable conditions predominate over the oceans, while the number of cases with northerly or southerly flow is relatively small and limited to a few regions. Stable conditions are typically encountered in the mid-latitude storm tracks and are more frequent in the summer hemisphere storm tracks (i.e. December in the SH and July in the NH, not shown). For a NH summer month (July) the picture is similar (not shown), albeit with different regions where warm/cold advection cases are found and more frequent stable boundary layers in the NH storm tracks.

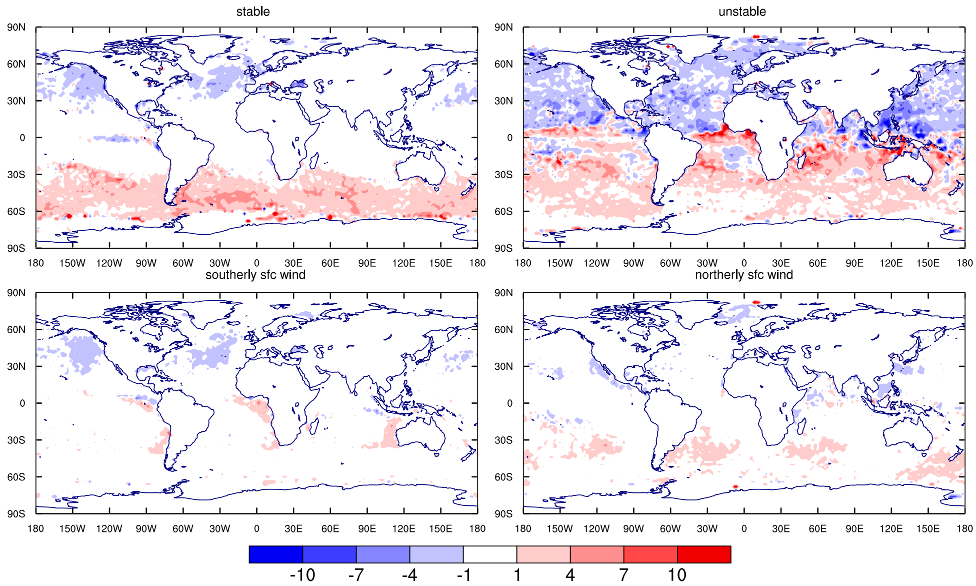

Figure 6: *Contributions to the monthly mean first-guess departures of surface wind direction (degrees) with respect to ASCAT observations for December 2017 due to days with (a) stable conditions, (b) unstable conditions, (c) southerly flow and (d) northerly flow. These values represent averaged statistics from the 0000 and 1200 UTC analysis cycles of the ECMWF operational deterministic forecasts, multiplied by the frequency of occurence of the respective category in December 2017 (see caption of Figure 5 for details on how the different categories were defined). For each grid box, the frequency of occurence for a certain category is defined as the ratio between the number of cases in that category and the total number of cases within the month. The sum of the errors in stable and unstable conditions equals the total error seen in Figure 2 for December 2017, while for the wind direction stratification the total error is equal to the sum of errors in northerly and southerly winds shown here, and westerly and easterly winds, which are not shown.*

The conclusions of our analysis are consistent with those of *Brown et al.* (2005) in the sense that the largest errors in surface wind direction are still found for warm advection cases, and more precisely for warm air intrusions in extra-tropical cyclones (not shown). The mean biases for southerly/northerly flows (indicative of warm advection in the NH and SH, respectively) are similar to those in Table 1 of *Brown et al.* (2005). However, when multiplying the surface wind direction errors in a certain type of conditions by the frequency of occurence of those cases, we find that warm advection cases (i.e. cases with southerly flow in the NH and northely flow in the SH) contribute little to the mean bias in surface wind direction within a particular month (Figure 6 bottom row). Stable conditions are also characterized by larger errors than unstable conditions, similarly to what was found by *Brown et al.* (2005). However, they contribute less to the overall bias than the unstable conditions, except for the summer hemisphere storm track where they contribute by a comparable amount (Figure 6 top row). Unstable boundary layers thus are found to contribute the most to the total error in surface wind direction both for December (Figure 6 top row) and July months (not shown), in all regions but the summer hemisphere storm tracks.

The further stratification of these biases, according to the additional criteria enumerated above, failed to lead to any conclusive results. This suggests that the remaining short-range forecast biases in surface wind direction (and speed) over oceans are mainly due to the representation of unstable boundary layers, irrespectively whether these are more or less convective, deeper or shallower, cloudy or clear.

4 Sensitivity of surface wind direction biases to shallow convection momentum transport

As a next step, we performed sensitivity experiments with the IFS to explore what aspects of the representation of turbulent and convective mixing affect the surface wind direction, and more generally as we will see in the next section, the representation of wind profiles in the boundary layer over the oceans. All the sensitivity experiments discussed here focused on December 2017 and were performed with recent IFS cycles (43r1 and 45r1 operational from November 2016 to July 2019). They were performed with the operational configuration used for ECMWF deterministic ten-day forecasts, albeit at a reduced horizontal resolution, i.e. 32 km instead of 9 km. Note however that the first-guess departures in surface wind direction against ASCAT observations in the control analysis experiment at 32 km resolution (CTL hereafter) are very similar (left panel of Figure 7) to those in the ECMWF operational forecasts at 9 km for December 2017 shown in Figure 2. This suggests that performing the sensitivity experiments at lower resolution does not affect the conclusions of our study.

A first comprehensive set of experiments consisted in forecast only experiments, i.e. experiments in which 24 hours forecasts are initialized each day (of December 2017 and July 2016) at 00UTC from the ECMWF operational analyses. These are quick to run and considerably cheaper (computationally) than so-called analysis experiments. In analysis experiments, all available observations are assimilated to create the initial conditions for the ten-days forecasts, which are then run from their own (cycling) analyses, like it is done for operational forecasts. While the impact of a model (e.g. physics) change can be fully estimated only through analysis experiments, forecast only experiments give a very good estimate of the impact of a certain model change and are always used as a first step in model development. This type of experiments was used here to explore the sensitivity of surface wind direction, and wind profiles, to a long list of changes to the physical parametrizations which can affect the wind representation in the boundary layer. To reduce the computational and archiving cost, the forecast only experiments were only 24 hours, instead of ten-days. This is sufficient to explore sensitivities in boundary layer winds because the response to changes in mixing takes place in a matter of hours and as we will see in Section 5 wind biases in the region of focus (upstream of Barbados) remain virtually constant at longer leadtimes.

A detailed description of the representation of physical processes in the IFS is beyond the scope of this study, and can be found in *IFS documentation* (2019). The aspects worth noting here are that the turbulence scheme is a first-order scheme using an eddy-diffusivity mass flux scheme (*Koehler et al.* 2011) for the convective boundary layer and 'long-tail' stability functions combined with an asymptotic mixing length depending on the boundary layer depth for stable boundary layers (*Sandu et al.* 2013/2014), while the shallow and deep convection scheme is the improved Tiedke mass-flux scheme (*Bechtold et al.* 2008; 2014). The model changes tested here include:

- reducing the mixing in stable boundary layers (by using 'short-tail' stability functions);
- reverting the modification made to the convective closure in 2013, in order to improve the diurnal cycle of convection (*Bechtold et al.* 2014);
- changes to the strength of convective entrainment/detrainment;
- changes to the shape of the *Troen and Mahrt* (1986) K-profile used to define the eddy-diffusivity component of the turbulence scheme in unstable boundary layers;
- switching off the mass-flux component of the turbulence scheme (note that this is not applied to momentum at present, so supressing it only affects winds indirectly through the impact on temperature and humidity);
- adding a counter-gradient term in unstable boundary layers as suggested in *Brown and Grant* (1997);
- switching off the entrainment at the top of unstable boundary layers;
- switching off the wave model, which is used to compute surface roughness over the ocean through the Charnock formulation, and hence has the potential of changing surface drag;
- switching off the convective momentum transport by shallow or deep convection.

The only change that resulted in a systematic and quasi-uniform backing (veering) of the forecasted surface winds over oceans in the NH (SH), and hence in a better alignment with ASCAT observations, was switching off the momentum transport by shallow convection. The reduction of mixing in stable boundary layers also reduces the surface wind direction biases (by 1 to 4 degrees) in the areas where stable boundary layers are encountered in the mid-latitude storm tracks (Figure 5). As stable boundary layers over oceans are more frequent in summer, the effect was most pronounced in the NH storm track in July and in the SH storm track in December (not shown). As the sensitivity to the degree of mixing in stable boundary layers is well-known and has been discussed in previous work (*Sandu et al.* 2013), the further experimentation and analysis therefore focused on the role of momentum transport by shallow convection. Its role and the uncertainties in its representation are less studied, and recent studies suggest that the mass flux approach may be less appropriate for transport of momentum than it is for the transport of heat and moisture (*Saggiorato et al.* 2020, *Schlemmer et al.* 2017, *Zhu* 2015).

To fully estimate the impact of momentum transport by shallow convection on circulation and assess the change in first-guess departures of surface winds against ASCAT observations, an analysis experiment in which the momentum transport by shallow convection was switched off was performed for December 2017 (NOCMT hereafter). In this experiment ten-days forecasts were started each day at 00UTC from their own analysis. Figure 7 (right) illustrates that indeed this change results in a systematic and nearly uniform change in surface wind direction. The opposite signs, and commensurate magnitudes, seen in the right and left panels of Figure 7 suggest that switching off shallow convection momentum transport goes a long way in reducing these systematic short-range biases in surface wind direction. This improvement in short-range surface wind direction is accompagnied by an improvement in surface wind speed in the Tropics against ASCAT observations, which is mainly associated with a strengthening of the easterlies by 0.2 to 0.5 m/s (not shown, but discussed in the next section).

However as expected, switching off shallow convection momentum transport leads to a considerable and significant deterioration of the forecast performance in the lower-troposphere (950 to 700 hPa), and more particularly in the Tropics, throughout the entire ten-day forecast range (not shown). A deterioration of

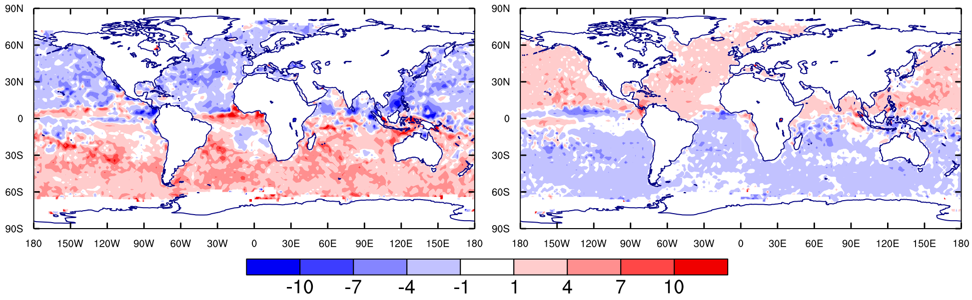

Figure 7: *Monthly mean first-guess departures of surface wind direction (degrees) with respect to ASCAT observations in the CTL analysis experiment for December 2017 (left) and the mean change in these first-guess departures when the shallow convection momentum transport is switched off in the NOCMT experiment (right). These values represent averaged statistics from the 0000 and 1200 UTC analysis cycles.*

6 to 8 % in the root mean square error of the wind vector is seen in the medium-range (3 to 6 days ahead) at 850 hPa when verified against own analysis, which is the standard practice for NWP verification. The representation of temperature is also significantly deteriorated in particular in the Tropics (a 3 to 5 % increase in the root mean square error at 500 hPa in the Tropical band being the most severe deterioration).

Given the conclusions reached by *Simpson et al.* (2018), that the too strong zonal winds in ECMWF short-range forecasts could be due to a lack of low level drag over the oceans, two additional analysis experiments were performed for December 2017. In these experiments, the surface drag coefficients over oceans for heat and respectively momentum, were increased by 10 %. The heat drag coefficients change had virtually no effect on the first-guess departures against ASCAT. The increase of the surface drag coefficients for momentum resulted as expected in a small decrease in the zonal component of the surface winds, and a small improvement (about 1 degree) of the short-range forecast surface wind direction with respect to ASCAT in the extratropics (not shown).

The improvement in (short-range) surface winds when switching off the momentum transport by shallow convection does not necessarily indicate that there is something wrong with the way this process is parametrized in the IFS. What shallow convection momentum transport effectively does is to mix winds across a deep layer, thus allowing the surface to "feel" the influence of the winds in the overlying cloud layer/free troposphere (and the other way around). So the reduction of the surface wind error in the NOCMT experiment could be simply due to the fact that convective momentum transport mixes erroneous winds from other levels down to the surface. In the next section we explore more in detail what happens with the wind profiles in the forecast, and in the analysis, when the shallow convection momentum transport is turned off in the NOCMT experiment. We also seek to understand whether the short-range biases in boundary layer winds discussed so far are consistent with biases seen at longer leadtimes (in the medium-range).

5 Boundary layer wind profiles and their sensitivity to changes in the forecasting system in a region upstream of Barbados

To better understand how the winds near the surface and in the lower troposphere change when the shallow convection momentum transport is switched off, we examine the wind profiles in a certain region upstream of Barbados. We focus on this region for two reasons: first, it is a typical undisturbed trade-wind region, and more importantly, this region has been better sampled with in-situ wind profile observations than any other region on Earth. Two recent observational efforts in this area were NARVAL (*Stevens et al.* 2019), and more recently (in Jan/Feb 2020) EUREC4A, the largest observational campaign ever of the coupled atmosphere-ocean system (*Bony et al.* 2017). Although comparison with in-situ wind observations from these observational efforts will be the subject of future studies, it is useful to document the wind errors in the IFS, their evolution and sensitivity to different aspects of the forecasting system in the same region.

5.1 Sensitivity of analysis and short-range forecast wind profiles to changes in the forecasting system

We start by looking at the mean wind profiles of the forecasts (at a leadtime of 12 hours) and of the verifying analyses from the CTL (black) and NOCMT (red) analysis experiments (Figure 8). Near the surface, the differences beetwen the 12 hours forecasts and the verifying analyses (i.e. the differences between the full and dashed pairs of black and red lines), corroborate what we have seen from the first-guess departures against ASCAT observations: the zonal winds forecasted in CTL are too strong with respect to the verifying (CTL) analysis and they considerably slow down in the NOCMT experiment. The near-surface zonal winds forecasted in NOCMT thus become very close to the near-surface winds in the two analyses, which are very similar at 12 UTC. The forecasted meridional component of the wind does not change much near the surface in NOCMT, which suggests that, at least in this region, the change in surface wind speed and direction is mainly due to the effect of the shallow convection momentum transport on the zonal component of the wind.

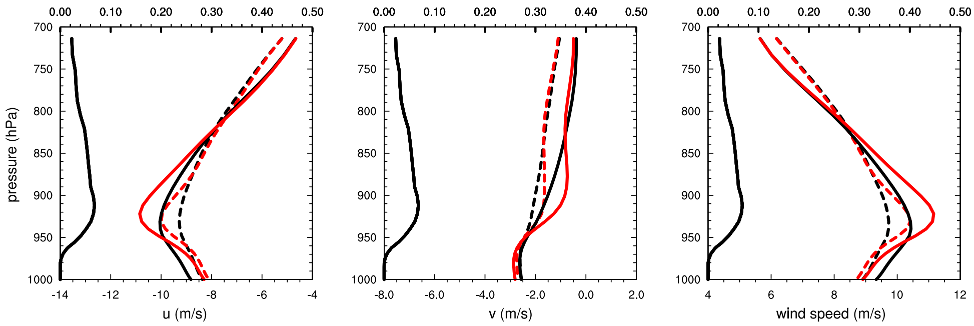

Figure 8: *Mean profiles of zonal (left), meridional wind (middle) and wind speed (right) from the 12 hours forecasts (full lines) and the analyses (dashes) of the CTL (black) and NOCMT (red) experiments for December 2017. These profiles represent averages over the 31 daily forecasts (verifying analyses) over a region east of Barbados (10-15N, 49-59W). In these experiments the forecasts are initialized daily at 00 UTC, so the mean analyses profiles shown here are for 12 UTC when the 12 hours forecasts verify. The cloud fraction from the 12 hours forecasts of the CTL experiment is also shown (black line on the left of each plot with values shown by the top x-axis).*

Higher-up, in the lowest half of the cloud layer, the zonal winds accelerate in the NOCMT experiment, both in the short-range forecasts and in the analysis. The acceleration in the forecasts is stronger than that in the analysis, and hence the forecast error is larger, corroborating the above discussion on the deteriorated forecast performance in the NOCMT experiment. The meridional component of the wind also changes in this region, contributing to the changes in wind speed seen in the right panel of Figure 8.

The changes in the forecasted wind profiles of the NOCMT experiment are not surprising: switching off shallow convection momentum transport, in this region where the zonal wind increases with height in the lower boundary layer, prevents stronger winds from near-cloud base being transported to the surface. Hence the near-cloud base winds remain stronger, while the near-surface winds decelerate compared to the CTL experiment. What is more surprising at first sight is the change seen in analysis in the lowest part of the cloud layer at 12 UTC (red dashed compared to black dashed line in Figure 8). At other analysis times (in particular 6 and 18 UTC) the analysis also changes in the sub-cloud layer (not shown). If the analysis was well constrained by observations, changes to the model physics should only lead to small changes in the analysis.

It is worth baring in mind however that the only direct information on the vertical structure of winds in the data assimilation system used to create the initial conditions for the forecasts (the analyses) is provided by radiosondes, and recently by the Aeolus satellite (which only started to be assimilated in the ECMWF operational analysis in January 2020, so after the period of focus here). All other wind measurements constraining the analyses are indirect: surface winds derived from scaterometter data and winds at certain atmospheric levels derived by tracking the motion of clouds from satellites (the so-called Atmospheric Motion Vectors or AMV). Winds are of course also indirectly constrained by assimilating temperature and humidity (conventional and satelitte) observations, which constrain the large-scale circulation (i.e. the pressure gradients). More information on the different observing systems assimilated in the IFS can be found in *Bormann et al.* (2019), *Lawrence et al.* (2019), who performed a large set of so-called Observing System Experiments (OSE) for the winter 2017/2018 and summer 2016. In these experiments, different observing systems are withdrawn (or denied) from the data assimilation when creating the initial conditions for the forecasts, in order to assess their impact on forecast skill.

Given that the OSE experiments performed by *Bormann et al.* (2019) cover December 2017 (and are performed in a similar configuration to the analysis experiments performed here), we used these experiments to explore how the wind profiles in the analysis and in the 12 hours forecasts change in our region of focus when certain observations (conventional observations, microwave temperature and humidity observations, infrared temperature observations or AMVs) are denied. We also performed additional analysis experiments in which we denied the AMVs only in the lower troposphere (below 600hPa), denied the ASCAT observations, and turned off the balance operators in the data assimilation. The balance operators are applied to 'transform' temperature increments into winds increments.

None of the changes to the use of observations, or the balance computation, had a systematic impact on the ASCAT first-guess departures (not shown). The first-guess departures against ASCAT could not be examined for the experiment in which the ASCAT observations were withdrawn. However, for all the data denial experiments, including the one for ASCAT observations, the 12 forecast and analysis wind profiles for the Barbados regions look very similar to those for the CTL experiment shown in Figure 8. The only noticeable, albeit very small changes to the analysis, are in the ASCAT and AMV denial experiments, for which the zonal winds in the analysis become slighlty more easterly (by 0.1 m/s on average) near the surface (not shown). This suggests that ASCAT and the AMVs act to slightly slow down the zonal component of the wind, and thereby the wind speed, near the surface in this region. The AMVs also affect the analysed meridional wind and zonal wind components above the cloud layer (above 800 hPa). The wind profiles in the 12 hours forecasts in both experiments are identical to those in the CTL experiment. The fact that the analysed wind profile does not change much in any of the denial experiments, does not mean that those observations do not play a role in constraining the wind profiles. What often happens when one observing system is withdrawn from the data assimilation system is that other observing systems compensate for its loss and play a bigger role in constraining the analysis.

5.2 Temporal evolution of wind profiles

As a final step of our investigation we explored how the wind profiles evolve during the forecast in the selected region. Figure 9 shows that in the first hours of the CTL forecasts (1 to 12), the forecasted zonal winds gradually accelerate, and they accelerate faster in the lower part of the cloud layer than near the surface. This behaviour explains most of the bias in wind speed at 12UTC with respect to the analysis (red versus dashed black line in Figure 9, or full versus dashed black in Figure 8). The meridional winds don't vary much in the first 12 hours of the CTL forecasts, except between 9 and 12 UTC, which is likely due to a diurnal cycle in the meridional winds (see also Figure 11). In the NOCMT experiment, the forecasted zonal winds also gradually (and even more pronouncedly) accelerate near cloud base, but these fastening winds near cloud base are not transported to the surface as it is the case in the CTL experiment (not shown). As discussed above, this results in better forecasted winds at the surface but worse winds (with respect to the analysis) at cloud base at 12 UTC (Figure 8).

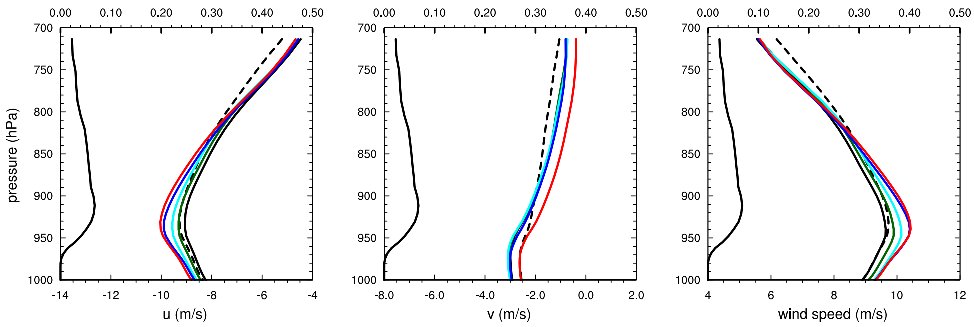

Figure 9: *Mean profiles of zonal (left), meridional wind (middle) and wind speed (right) for the forecasts of the CTL experiment for December 2017, at a leadtime of: 1 hours (dark green), 3 hours (cyan), 9 hours (blue) and 12 hours (red). The wind profiles from the CTL analyses at 00UTC (when the forecasts start) and 12 UTC (when the 12 hours forecasts verify) are shown in black full and dashed, respectively. As in Fig. 8 the profiles represent averages over the 31 daily forecasts (verifying analyses) over a region east of Barbados (10-15N, 49-59W). The cloud fraction from the 12 hours forecasts of the CTL experiment is also shown (black line on the left of each plot with values shown by the top x-axis).*

Figure 10 illustrates the processes that contribute to the changes in zonal wind in the boundary layer in the first hours of the CTL and NOCMT forecasts. Below cloud base (below 950hPa), turbulent mixing slows down the winds, while dynamics, and in the CTL run convection, accelerate them. In the lower part of the cloud layer where the largest errors in zonal winds are seen (between 950 and 900 hPa), turbulent diffusion, and in the CTL run convection, partially compensate the acceleration due to the dynamics. In the CTL, the balance between dynamics, diffusion and convection is not perfect, resulting in a net accelerating total tendency (full red line) throughout the lower troposphere. This accelerating total tendency which is slightly more pronounced around 950hPa, explains the forecast drift during these first 3 hours seen in Figure 9, and the corresponding forecast errors (against analysis) during these first hours. In the NOCMT experiment, the convective tendencies are nearly zero, the remaining small tendencies being due to momentum transport by deep convection. In the lack of deceleration from the convection around cloud base, the imbalance between the dynamics and diffusion tendencies is even larger in the NOCMT experiment, explaining well the larger errors in that layer (Figure 8). Near the surface, without an acceleration due to convective momentum transport, the total tendency is nearly zero, so the zonal winds remain quasi-constant.

The errors in wind profiles (with respect to the analysis) developed in the first hours do not evolve much

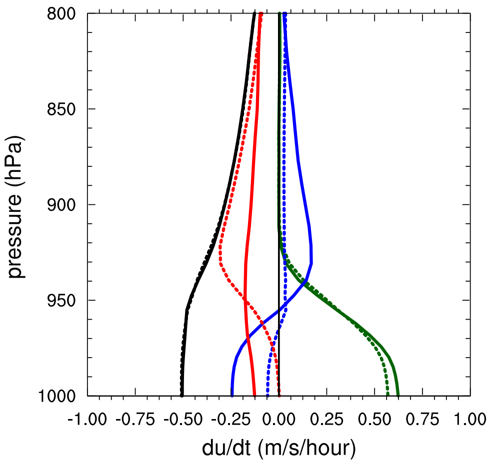

Figure 10: *Tendencies for the zonal wind component from the dynamics (black), turbulent diffusion (green) and convection (blue) schemes for the forecasts from the CTL (full) and NOCMT (dashed) experiments for December 2017. The total tendency, indicating the model drift is shown in red. These tendencies represent averages over the same region east of Barbados, and over the first 3 hours of the 31 forecasts of each experiment.*

throughout the medium-range either in the CTL (Figure 11, top) nor in the NOCMT experiment (not shown). The full blue lines in the top panels of Figure 11 illustrate that the forecasts at various lead times, verifying at 00 UTC, are fairly similar and they are similarly distant from the respective verifying analyses (dashed black lines). However, the bottom panels of Figure 11 show that there is a pronounced diurnal cycle in wind speed and in its components both in the forecasts and in the analysis, particularly in the lowest 1.5 km. The spread among the 0, 6, 12 and 18 UTC wind profiles in the lowest 1.5km is almost as big in the analysis as it is in the forecasts. The magnitude and even sign of the forecast error against the analysis in this layer depends on the time of the day, but given that the analysis seems to be only weakly constrained by observations it is hard say whether it provides a ground truth here. EUREC4CA observations are being used at the moment to evaluate the quality of wind profiles both in the forecasts and in the analysis, and will be reported on in a follow-on study. The reasons for this strong diurnal cycle in the boundary layer winds are not clear and will be also investigate in future work.

6 Conclusions

Surface wind direction biases are among the most resilient biases in the ECMWF IFS forecasts and have been known since the early 1990's (*Hollingsworth* 1994). Short-range forecast verification against ASCAT observations over ocean and synop observations over land shows a veering of the forecasted surface wind direction with respect to observations in the NH (and a backing in the SH), leading to a smaller wind turning angle between the forecasted surface wind and the forecasted geostrophic wind than that seen in observations. Despite continuous improvements to the forecasting system throughout the last decades, these short-range forecast biases did not entirely dissapear although they were continuously reduced. Although the main focus at ECMWF is on the quality of medium-range forecasts, eliminating biases in short-range forecasts is important both for avoiding non-linear error growth at longer time ranges, and because a bias free model (in the short-range) is key for making the best use of observations in the data assimilation process. Moreover, further reducing the surface wind biases is not only important for improving the quality of wind forecasts at local scale, but also for narrowing down the uncertainty

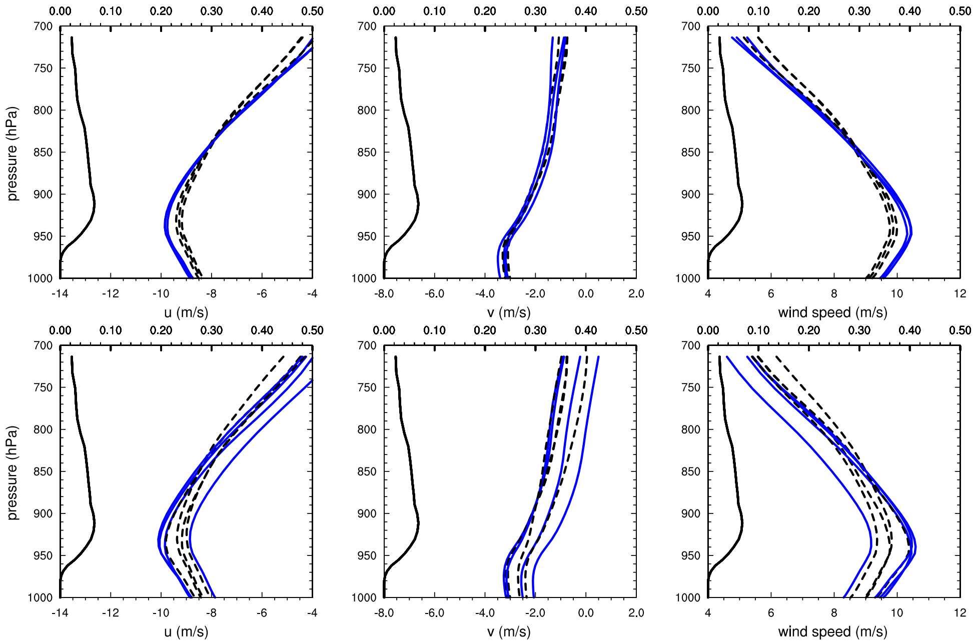

Figure 11: *Mean profiles of zonal (left), meridional wind (middle) and wind speed (right) for the forecasts and verifying analysis of the CTL experiment for December 2017, at different times. Top: forecast leadtimes 24, 48 and 144 hours (full blue lines) and verifying analysis at 00 UTC (black dashed lines). Bottom: forecast lead times 24, 30, 36 and 42 and verifying analysis at 00, 06, 12, 18 UTC (dashed black lines); the lowest wind speed here in both the forecast and the analysis is at 18 UTC. As in Fig. 8 the profiles represent averages over the 31 daily forecasts (and the corresponding 31 verifying analyses) over a region east of Barbados (10-15N, 49-59W). The cloud fraction from the 12 hours forecasts of the CTL experiment is also shown (black line on the left of each plot with values shown by the top x-axis).*

in the representation of surface drag which is strongly related to the representation of (boundary layer) winds.

In this study, we examined the evolution of the short-range surface wind direction biases in operational ECMWF deterministic forecasts over the years and related it to some of the changes to the IFS known to have improved the representation of near-surface winds. Through a conditional analysis of the short-range (first-guess) forecasts errors in surface winds against ASCAT observations we demonstrated that the remaining errors over the oceans are mostly related to the representation of unstable boundary layers, although the representation of turbulent diffusion in stable conditions also plays a role in particular in the summer hemisphere mid-latitude storm tracks.

By examining the wind profiles in a region upstream of Barbados, typical of undisturbed trade-winds, we showed that the zonal wind component gradually accelerates in the lowest part of the shallow cumulus layer in the first 12 hours of the forecasts (with respect to the verifying analysis). This is due to an inbalance between the acceleration due to dynamics and the deceleration due to the shallow convection and turbulence schemes. When the transport by shallow convection is active, the increasingly faster winds near cloud base are transported towards the surface which results in the too fast easterlies, and the too strong veering of the surface winds indicated by the comparison against the ASCAT observations. This hypothesis is further demonstrated through an experiment in which the momentum transport by shallow convection is switched off. This change leads to an even larger error at cloud base (with respect to the analysis), but the faster winds from that region are not transported to the surface and so the near-surface wind errors in both the zonal wind component and in wind direction are smaller. Turning off shallow convection momentum transport is not a desirable change of course, as it leads to a significant deterioration of the forecast performance throughout the ten-days. However this experiment highlights that the errors in near-surface and cloud base winds are most sensitive to the momentum transport in the lower part of the boundary layer and in particular to the balance between the different terms near cloud base. These errors could be related to the representation of momentum transport in either clear and cloudy updrafts, or to issues in the partition between the two. These processes are uncertain and research to use LES and observations to evaluate and improve the parametrizations used to represent them in the IFS is ongoing, in the framework of the CloudBrake project at TU Delft (*Saggiorato et al.* 2020).

Our analysis showed that upstream of Barbados, a fairly strong diurnal cycle in the boundary layer wind profiles is present in both the forecast and in the analysis, and the magnitude and even sign of the forecast errors (with respect to the analysis) depends on the time of the day. For example, the forecasted winds are faster than the analysis winds below 850 hPa at 00 and 12 UTC. At these times, the observations act to slow down the zonal winds in the lower troposphere, although it is difficult to say which observations play the biggest role based on OSEs. The fact that both the forecast and the analysis winds are quite sensitive to the representation of certain processes (i.e. the momentum transport by shallow convection) in the cloud base layer, suggests that the analysis is not very well constrained by observations (at least prior to the assimilation of Aeolus winds). In the lack of direct wind profile observations it is difficult to know how much we can trust the forecast errors estimated by comparison with the verifying analysis in the lowest 1.5 km. An attempt was made to compare the forecasts and analysis profiles shown in Figure 8 with wind profiles from radiosounds launched at the Barbados airport, but the island effect on near-surface winds makes the comparison difficult. Work is ongoing in collaboration with TU Delft and Max Planck Institute for Meteorology to examine the wind errors in the IFS in the same region by comparison with hundreds of dropsondes launched during the recent EUREC4A field experiment.

In the Barbados region, the errors in the wind profiles develop in the first 12 hours of the forecasts and do not grow further (until day 5), although their magnitude and even sign (with respect to the analysis) depends on the time of the day. In other regions, the day 5 forecast errors in surface winds are not necessarily similar with the short range errors. In particular, too strong surface easterlies and westerlies are not a widespread characteristic of day 5 forecasts, as is the case in short-range forecasts. This is likely because while short-range forecast errors are driven by errors in momentum transport in unstable boundary layers, or possibly to errors in surface drag, which act on fast time scales, errors in the medium-range are related to large-scale circulation errors which gradually develop during the forecast. Another hyphothesis for the fast change (or adjustment) of the boundary layer winds in the first 12 hours of the forecasts, is that this is a response to an erronated large-scale pressure gradient in the initial conditions. This hypothesis will also be further explored with EUREC4A observations which provide for the first time a ground truth for collocated wind, temperature, and humidiy profile observations in a wide region (a circle of about 100km radius) in the undisturbed trade-winds region.

Acknowledgements

The authors would like to thank Peter Bauer, Katie Lean and Elias Holm for discussions on the use and impact of observations and balance operator in the data assimilation in the IFS, as well as Mark Rodwell and Jean Bidlot for discussions regarding wind biases and surface drag over the oceans.

References

Bauer, P., A. Thorpe, and G. Brunet, The quiet revolution of numerical weather prediction, *Nature*, *525*(7567), 47, 2015.

Bechtold, P., M.Kohler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. Rodwell, F. Vitart, and G. Balsamo, Advances in simulating atmospheric variability with the ecmwf model: From synoptic to decadal time-scales, *Quart. J. Roy. Meteor. Soc.*, *134*(634), 1337–1351, doi:10.1002/qj.289, 2008.

Bechtold, P., N. Semane, P. Lopez, J. P. Chaboureau, A. Beljaars, and N. Bormann, Representing equilibrium and nonequilibrium convection in large-scale models, *J. Atmos. Sci.*, *71*(2), 734–753, doi: 10.1175/JAS-D-13-0163.1, 2014.

Beljaars, A., and P. Viterbo, *The role of the boundary layer in a numerical weather prediction model in: Clear and cloudy boundary layers*, 287-304 pp., Royal Netherlands Academy of Arts and Sciences, North Holland Publishers, Amsterdam, 1998.

Beljaars, A., A. Brown, and N. Wood, A new parameterization of turbulent orographic form drag, *Quart. J. Roy. Meteor. Soc.*, *130*, 1327–1347, 2004.

Belmonte Rivas, M., and A. Stoffelen, Characterizing era-interim and era5 surface wind biases using ascat, *Ocean Science*, *15*(3), 831–852, doi:10.5194/os-15-831-2019, 2019.

Bony, S., B. Stevens, and P. C. e. a. F. Ament, S. Bigorre, Eurec4a:a field campaign to elucidate the couplings between clouds, convection and circulation, *Surveys in Geophysics*, *38*, 1529–1568, doi: 10.1007/s10712-017-9428-0, 2017.

Bormann, N., H. Lawrence, and J. Farnan, Global observing system experiments in the ecmwf assimilation system, (839), doi:10.21957/sr184iyz, 2019.

Brown, A., and A. Grant, Non-local momentum mixing in the convective boundary layer, *Boundary-Layer Meteorology*, *84*, 1–22, 1997.

Brown, A., A. Beljaars, H. Hersbach, A. Hollingsworth, M. Miller, and D. Vasiljevic, Wind turning across the marine atmospheric boundary layer, *Quart. J. Roy. Meteor. Soc.*, *131*, 1233–1250, 2005.

Brown, A., A. Beljaars, and H. Hersbach, Errors in parametrizations of convective boundary layer turbulent momentum mixing, *Quart. J. Roy. Meteor. Soc.*, *132*, 1859–1876, 2006.

Chen, G., I. M. Held, and W. A. Robinson, Sensitivity of the latitude of the surface westerlies to surface friction, *J. Atmos. Sci.*, *64*, 2899–2915, 2007.

Haiden, T., M. Janousek, F. Vitart, L. Ferranti, and F. Prates, Evaluation of ecmwf forecasts, including the 2019 upgrade, (853), doi:10.21957/mlvapkke, 2019.

Hersbach, H., B. Bell, P. Berrisford, A. Horanyi, and J. M. S. et al., Global reanalysis: goodbye era-interim, hello era5, *ECMWF Newsletter*, *159*, 2019.

Hollingsworth, A., Validation and diagnosis of atmospheric models, *Dyn, Atmos. Oceans*, *20*, 227–246, 1994.

Holtslag, A. A. M., G. Svensson, P. Baas, S. Basu, B. Beare, A. C. M. Beljaars, F. C. Bosveld, J. Cuxart, J. Lindvall, G. J. Steeneveld, M. Tjernstrm, and B. J. H. Van De Wiel, Stable atmospheric boundary layers and diurnal cycles: Challenges for weather and climate models, *Bulletin of the American Meteorological Society*, *94*(11), 1691–1706, doi:10.1175/BAMS-D-11-00187.1, 2013.

Houchi, K., A. Stoffelen, G. J. Marseille, and J. De Kloe, Comparison of wind and wind shear climatologies derived from high-resolution radiosondes and the ecmwf model, *Journal of Geophysical Research: Atmospheres*, *115*(D22), doi:10.1029/2009JD013196, 2010.

IFS documentation, CY46r1, *Tech. rep.*, ECMWF, available at: https://www.ecmwf.int/en/elibrary/19308-part-iv-physical-processes, 2019.

Klein, S. A., and J. R. Norris, On the relationships among low-cloud structure, sea surface temperature, and atmospheric circulation in the summertime northeast pacific, *J. Climate*, *8*, 1140–1155, 1995.

Koehler, M., M. Ahlgrimm, and A. Beljaars, Unified treatment of dry convective and stratocumulus-topped boundary layers in the ecmwf model, *Quart. J. Roy. Meteor. Soc.*, *137*, 43–57, 2011.

Lawrence, H., N. Bormann, I. Sandu, J. Day, J. Farnan, and P. Bauer, Use and impact of arctic observations in the ecmwf numerical weather prediction system, *Quart. J. Roy. Meteor. Soc.*, *145*(725), 3432–3454, doi:10.1002/qj.3628, 2019.

Lean, P., M. Bonavita, E. Holm, N. Bormann, and T. McNally, Continous data assimilation for the ifs, *ECMWF Newsletter*, *158*, 2019.

Lott, F., and M. Miller, A new subgrid orographic drag parameterization: Its formulation and testing, *Quart. J. Roy. Meteor. Soc.*, *123*, 101–127, 1997.

Rabier, F., H. Jrvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, The ecmwf operational implementation of four-dimensional variational assimilation. i: Experimental results with simplified physics, *Quart. J. Roy. Meteor. Soc.*, *126*(564), 1143–1170, doi:10.1002/qj.49712656415, 2000.

Saggiorato, B., L. Nuijens, A. Siebesma, S. de Roode, I. Sandu, and L. Papritz, The influence of convective momentum transport and vertical wind shear on the evolution of a cold air outbreak, *Journal of Advances in Modeling Earth Systems*, *n/a*(n/a), e2019MS001,991, doi:10.1029/2019MS001991, 2020.

Sandu, I., A. Beljaars, P. Bechtold, T. Mauritsen, and G. Balsamo, Why is it so difficult to represent stably stratified conditions in numerical weather prediction (nwp) models?, *Journal of Advances in Modeling Earth Systems*, *5*, 117–133, 2013.

Sandu, I., A. Beljaars, and G. Balsamo, Improving the representation of stable boundary layers, *ECMWF Newsletter*, *138*, 2013/2014.

Sandu, I., A. van Niekerk, T. G. Shepherd, S. B. Vosper, A. Zadra, J. Bacmeister, A. Beljaars, A. R. Brown, A. Dörnbrack, N. McFarlane, et al., Impacts of orography on large-scale atmospheric circulation, *npj Climate and Atmospheric Science*, *2*(1), 10, 2019.

Schlemmer, L., P. Bechtold, I. Sandu, and M. Ahlgrimm, Uncertainties related to the representation of momentum transport in shallow convection, *Journal of Advances in Modeling Earth Systems*, *9*(2), 1269–1291, doi:10.1002/2017MS000915, 2017.

Simpson, I. R., J. Bacmeister, I. Sandu, and M. J. Rodwell, Why do modeled and observed surface wind stress climatologies differ in the trade wind regions?, *Journal of Climate*, *31*(2), 491–513, doi:10.1175/JCLI-D-17-0255.1, 2018.

Stevens, B., F. Ament, S. Bony, S. Crewell, F. Ewald, S. Gross, A. Hansen, L. Hirsch, M. Jacob, T. Klling, H. Konow, B. Mayer, M. Wendisch, M. Wirth, K. Wolf, S. Bakan, M. Bauer-Pfundstein, M. Brueck, J. Delano, A. Ehrlich, D. Farrell, M. Forde, F. Gdde, H. Grob, M. Hagen, E. Jkel, F. Jansen, C. Klepp, M. Klingebiel, M. Mech, G. Peters, M. Rapp, A. A. Wing, and T. Zinner, A high-altitude long-range aircraft configured as a cloud observatory: The narval expeditions, *Bulletin of the American Meteorological Society*, *100*(6), 1061–1077, doi:10.1175/BAMS-D-18-0198.1, 2019.

Svensson, G., and A. Holtslag, Analysis of model results for the turning of the wind and the related momentum fluxes and depth of the stable boundary layer, *Boundary-Layer Meteorology*, *132*, 261–277, 2009.

Troen, I., and L. Mahrt, A simple model of the atmospheric boundary layer; sensitivity to surface evaporation, *Boundary-Layer Meteorology*, *37*, 129–148, doi:https://doi.org/10.1007/BF00122760, 1986.

van Niekerk, A., I. Sandu, and all, Constraining orographic drag effects (coorde): a model comparison of resolved and parametrized orographic drag,, *Journal of Advances in Modeling Earth Systems*, p. submitted, 2020.

Viterbo, P., A. Beljaars, J.-F. Mahfouf, and J. Teixeira, The representation of soil moisture freezing and its impact on the stable boundary layer, *Quart. J. Roy. Meteor. Soc.*, *125*, 2401–2426, 1999.

Wood, R., and C. S. Bretherton, On therelationship between stratiform low cloud cover and lower tropospheric stability, *Journal of Climate*, *19*, 6425–6432, 2006.

Zadra, A., WGNE Drag Project:An inter-model comparison of surface stresses, *Tech. rep.*, available at http://collaboration.cmc.ec.gc.ca/science/rpn/drag_project/documents/wgne_drag_project_report01.pdf, 2015.

Zhu, P., On the mass-flux representation of vertical transport in moist convection, *J. Atmos. Sci.*, *72*(12), 4445–4468, doi:10.1175/JAS-D-14-0332.1, 2015.