The representation of winds in the lower troposphere in ECMWF forecasts and reanalyses during the EUREC4A field campaign

Alessandro C.M. Savazzi[1], Louise Nuijens[1], Irina Sandu[2], Geet George[3], and Peter Bechtold[1]

[1]Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, The Netherlands
[2]European Centre for Medium Range Weather Forecasts (ECMWF), Reading, UK
[3]Max-Planck-Institut für Meteorologie, Hamburg, Germany

Correspondence: Alessandro C.M. Savazzi (A.C.M.Savazzi@tudelft.nl)

Abstract. The characterization of systematic forecast errors in lower-tropospheric winds over the ocean is a primary need for reforming models. Winds are among the drivers of convection, thus an accurate representation of winds is essential for better convective parameterizations. We focus on the temporal variability and vertical distribution of lower-tropospheric wind biases in operational medium-range weather forecasts and ERA5 reanalyses produced with the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). Thanks to several sensitivity experiments and an unprecedented wealth of measurements from the 2020 EUREC4A field campaign, we show that the wind bias varies greatly from day to day, resulting in RSME's up to 2.5 $\mathrm{m\,s^{-1}}$, with a mean wind speed bias up to -1 $\mathrm{m\,s^{-1}}$ near and above the trade-inversion in the forecasts and up to -0.5 $\mathrm{m\,s^{-1}}$ in reanalyses. The modeled zonal and meridional wind exhibit a too strong diurnal cycle, leading to a weak wind speed bias everywhere up to 5 km during daytime, turning into a too strong wind speed bias below 2 km at nighttime. The biases are fairly insensitive to the assimilation of sondes and likely related to remote convection and large scale pressure gradients. Convective momentum transport acts to distribute biases throughout the lowest 1.5 km, whereas at higher levels, other unresolved or dynamical tendencies play a role in setting the bias. Below 1 km, modelled friction due to unresolved physical processes appears too strong, but is (partially) compensated by dynamical tendencies, making this a challenging coupled problem.

1 Introduction

Accurate wind predictions are vital for renewable wind energy generation, which has experienced substantial growth in the last decade (Foley et al., 2012). An improvement in the representation of horizontal winds is also necessary for a stepwise change in the realism of climate projections (Bony et al., 2015). Motivated by this need to improve the representation of winds in weather and climate models, this study presents an analysis of wind biases in medium-range forecasts and reanalysis produced with the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). This analysis relies on a wealth of wind profiling observations, gathered during the recent EUREC4A field campaign (Stevens et al., 2021), in the oceanic trade-winds region east of Barbados, where no in-situ observations are regularly made. In this paper we thus take a fresh look at one of the most systematic and longstanding biases in forecasts of near-surface weather, i.e. the biases in lower-tropospheric winds (Hollingsworth, 1994; Brown et al., 2005, 2006; Sandu et al., 2013).

25   The characterization of systematic forecast errors in tropospheric winds over the ocean, and the understanding of their causes, is largely limited by the availability of observations of the wind profile. Apart from island radiosonde launches and near-surface measurements from buoys, there are no regular wind profiling observations over the oceans, including the tropical Atlantic Ocean (Brown et al., 2005). Only the Aeolus satellite mission provides since 2019 (Stoffelen et al., 2005; Rennie et al., 2021) a global coverage of winds, but with a footprint on the order of 100 km, a vertical resolution on the order of 500 m, and

30   systematic errors of $\sim 2\ \mathrm{m\,s^{-1}}$ (Witschas et al., 2020), its resolution and accuracy is hardly sufficient to evaluate the forecast wind biases in the lower troposphere.

       The 10-m wind speeds over the oceans were shown to be biased by up to $0.5\ \mathrm{m\,s^{-1}}$ compared to ASCAT scatterometer observations in the ECMWF reanalyses: ERA-Interim (Dee et al., 2011) and ERA5 (Hersbach et al., 2020). In particular the reanalyses show excessive mean easterlies and too weak mean meridional winds in the trade region (Belmonte Rivas and

35   Stoffelen, 2019). These biases may seem small, but they can introduce a large bias in the wind stress, which is a function of the wind speed squared. Such a wind stress bias could result in significant errors in ocean-atmosphere coupling and climate prediction (Chelton and Freilich, 2005).

       Belmonte Rivas and Stoffelen (2019) also demonstrated that errors in the mean surface 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 underesti-

40   mation of transient wind activity. 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).

       Although successive changes to the ECMWF IFS reduced the near-surface wind error over the oceans throughout the years, its typical global signature remains (Sandu et al., 2020). Sandu et al. (2020) analysed in more detail the wind profile forecast errors over the trade-winds region east of Barbados, on which we also focus in this study. They showed that the model analysis

45   (initial condition of the forecasts) is uncertain in the lowest part of the troposphere, particularly in the cloud layer, where it is most poorly constrained by observations. The IFS wind errors develop in the first 12 hours of the forecast and do not grow significantly until day five. Excessive zonal surface winds are not a widespread characteristic of day 5 forecasts, as is the case in short-range forecasts. This suggests that the cause of the bias lies in processes that act on fast time scales. Sandu et al. (2020) also explored the influence of convective momentum transport (CMT) by the abundant shallow convection in this region and

50   showed that it plays an important role in communicating wind biases that are present at cloud levels towards the surface, hinting that the biases may be established at levels above the surface layer.

       Here we exploit a unique opportunity offered through the EUREC4A field campaign (Stevens et al., 2021) to assess wind biases in forecasts and reanalyses produced with the IFS, not only at the surface but throughout the lower troposphere. Between January and February 2020 the EUREC4A field campaign provided benchmark measurements for a new generation of models

55   and scientific discoveries. EUREC4A is the largest observational field campaign of the coupled atmosphere-ocean system and the duration of the campaign and the large communal effort resulted in an unprecedentedly comprehensive record of tropospheric winds in the trades. In particular, during EUREC4A more than 1200 dropsondes, 800 radiosondes and a total of six wind lidars were deployed (Stevens et al., 2021), allowing a detailed study of the vertical structure of the winds and circulations in the boundary layer.

60   Some aspects of the systematic error in surface winds from weather models have been described in the literature, for instance the insufficient mesoscale variability in the extratropics (Gille, 2005), the lack of small-scale features relevant for sea surface temperature (SST) gradient effects (Chelton et al., 2004; Risien and Chelton, 2008) and the generally excessive zonal winds (Chaudhuri et al., 2013; Belmonte Rivas and Stoffelen, 2019; Sandu et al., 2020).

In this study we focus on the representation of the vertical profile of winds during EUREC4A in operational forecasts and
65   the ERA5 reanalyses produced with the ECMWF IFS. Our objectives are to:

a) Combine various wind profiling observations to investigate the temporal variability of the wind bias in the operational ECMWF high-resolution deterministic short and medium-range forecasts (approximately 9km at the Equator).

b) Evaluate whether the analyses and reanalyses, which blend observations and short-range forecasts through data assimilation to create the initial conditions of the forecasts, or the best representation of the past atmospheric state, show the
70   same bias as the forecasts.

c) Assess whether the assimilation of observations gathered during EUREC4A helped improve the analyses and forecasts performed with the IFS.

d) Explore the origin of the wind bias through the use of additional model sensitivity experiments.

After a description of the data (section 2) and of the methods used to derive and compare statistics of the wind profiles
75   (section 3), we address the following questions: What is the mean vertical distribution of the wind bias in forecasts and reanalyses produced with the IFS? (sections 4.1); What is the temporal variability of the wind bias? (section 4.2 and section 4.3); How much are the analyses constrained by the assimilation of radio- and dropsondes during EUREC4A? (section 5). In section 6 we then evaluate the influence of model physics, in particular the role of convection and turbulence representation. Our results are summarised and discussed in section 7.

80 **2   Data**

This study focuses on an area of about 350 km x 350 km, between 55.8W and 59.25W and between 11.4N and 14.7N. This area is located in the trade-winds region over the Atlantic Ocean, eastward of Barbados. Our study samples 29 days during the boreal winter, from the 18th of January 2020 to the 15th of February 2020, when EUREC4A took place. During the boreal winter the Inter-Tropical Convergence Zone (ITCZ) is typically located at lower latitudes, and the area east of Barbados
85   experiences undisturbed trade-winds from an east to northeast direction, with prevalence of cumulus clouds confined to the lower troposphere, moderate large-scale subsidence and an inversion around 800 hPa (Stevens et al., 2016; Brueck et al., 2015; Nuijens et al., 2014).

Several observational datasets (section 2.1) are used to evaluate the forecasts and (re)analyses produced with ECMWF IFS (section 2.2).

2.1 Observations

90

The measurements used in this study have been taken during the EUREC4A campaign (Stevens et al., 2021). Within EU-REC4A, a region of intensive measurements was defined in support of studies of cloud-circulation interactions, cloud physics, and factors influencing the mesoscale patterning of clouds. This region largely corresponds with the domain of our study (Figure 1) and it is situated near the western end of the 'Tradewind Alley', an extended corridor across the Atlantic (see Figure 1 in Stevens et al. (2021)) with its downwind terminus defined by the Barbados Cloud Observatory (BCO). The instruments used in this study are dropsondes, radiosondes and a ship-borne wind lidar system.

95
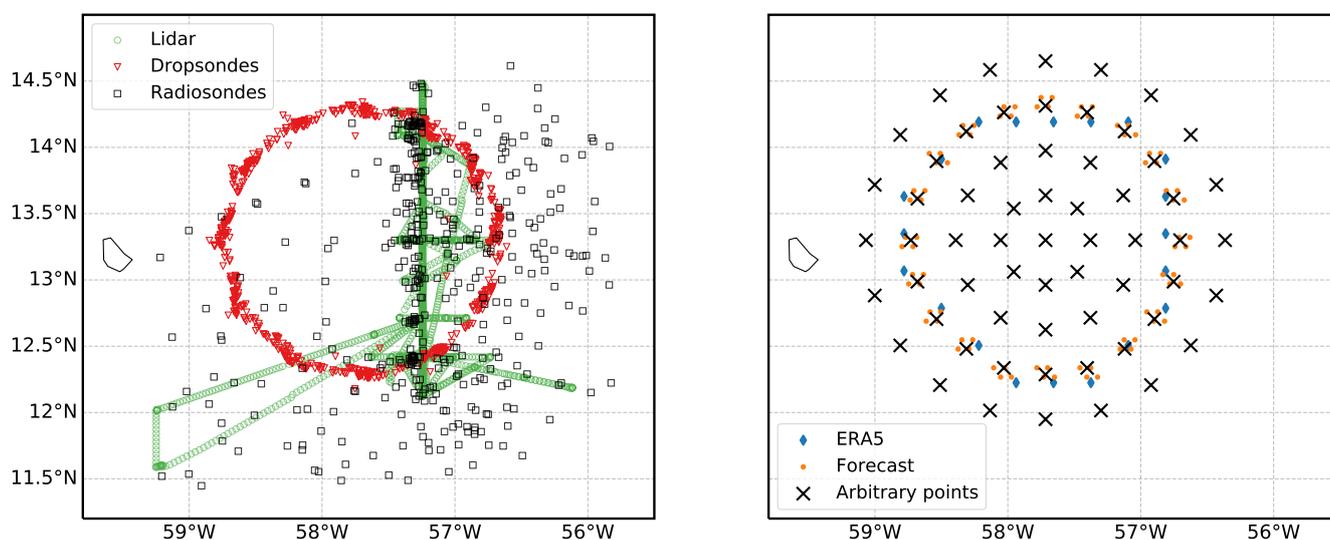
Figure 1. Overview of the spatial coverage of different datasets. Left panel illustrates the observational datasets: 3169 lidar measurements from the Meteor research vessel (green circles), 444 radiosondes from research vessels (black squares), 799 dropsondes from JOANNE (red triangles). Right panel illustrates the points in which profiles were retrieved from IFS forecasts and (re)analyses. Model grid points are shown only for the second most external ring: see text for an explanation of different modelling datasets and resolutions.

2.1.1 JOANNE

We use EUREC4A dropsonde measurements from the Joint dropsonde Observations of the Atmosphere in tropical North at-laNtic meso-scale Environments (JOANNE) dataset (George et al., 2021). The primary strategy of the EUREC4A dropsondes launches was to sample atmospheric profiles along a 222 km diameter circle centred at 13.3°N, 57.7°W. Following Stevens et al. (2021), we call this as the EUREC4A-circle. The majority of the dropsondes over the EUREC4A-circle were launched from the German high altitude and long-range research aircraft HALO, a few complementary flights being also performed by the American WP-3D Orion research aircraft. Typically, a flight over the EUREC4A-circle took one hour and 12 dropsondes were launched per circle, although the number of profiles per circle is often less than 12 due to either instrument or operator

100

errors. An overview of the circles and corresponding dropsondes is outlined in George et al. (2021). For our study, we use sounding profiles from 799 dropsondes (see red dots in Figure 1) launched from 73 EUREC4A-circles spread over 13 days between the 18th of January 2020 and the 15th of February 2020. In this study, we refer to the days with dropsonde measurements as flight-days, and we use flight-hours for the hours with dropsonde measurements within the flight-days. Figure 2 schematically represents the temporal availability of the 3 EUREC4A datasets that we use in this study. Black stripes indicate hours sampled in the corresponding dataset. It can be seen that the dropsondes sampled mostly daytime conditions.

2.1.2 Radiosondes

Radiosondes considered in this study were launched from four research vessels (RV) over the northwestern tropical Atlantic eastward of Barbados: two German research vessels, Maria S. Merian (Merian) and Meteor; a French research vessel, L'Atalante (Atalante); and a United States research vessel, Ronald H. Brown (Ron Brown). The Meteor operated between 12.5°N and 14.5°N along the 57.25°W meridian. The Ron Brown measured air-masses along the Tradewind Alley, while the Merian and Atalante vessels mainly sailed the Boulevard des Tourbillons southward of Barbados (see Figure 1 in Stevens et al. (2021)). Most radiosondes recorded information both in the ascent and descent sections, with descending radiosondes falling by parachute for all platforms except for the Ron Brown.

This study makes use of 444 radiosondes (258 in ascending mode and 186 in descending mode) within the study domain defined above, as documented in Stephan et al. (2021). Each black dot in Figure 1 (left) refers to a radiosonde either in ascending or descending mode. Only radiosondes within the domain where considered. Thus, radiosondes drifting outside of the area of interest are considered only when inside the domain, and radiosondes launched outside and drifting inside the domain are also considered only where relevant.

There are about 2 radiosondes per hour available in our selected domain during the studied period. Here, we average wind profiles from all the radiosondes available in our selected domain in each three hour period to obtain mean profiles that are more representative of the entire domain. Radiosondes provide a regular and comprehensive measurement during all days of the study as it can be seen in Figure 2.

2.1.3 WindCube Long-range Wind Lidar

A Leosphere long-range Windcube (WLS70) on board of the Meteor research vessel performed measurements at 20 different height levels, between 100 m and 2000 m. The WLS70 device has a sampling rate of approximately 6 s and measures the line-of-sight radial velocity successively at four azimuthal positions along a cone angle of 14.7°, thus every 360° scan takes around 24 s.

The radial velocities are corrected for ship motions using a simplified correction methodology using an internal GPS system of an accompanying short-range WLS7 WIndCube, which uses a combination of an xSEns MTi-G attitude and heading reference sensor (AHRS) and a Trimble SPS361 satellite compass. The simple motion correction applied to the LOS velocities takes into account the translational ship motions and the yaw information, as explained in Wolken-Möhlmann et al. (2014) and Gottschall et al. (2018). The pitch and roll information is not used, since according to previous studies (Wolken-Möhlmann

et al., 2014), the effect of these tilt motions are less relevant for relatively stable platforms. After corrections, the wind vector is retrieved and the data is averaged to 1 hourly values.

140   The left panel of Figure 1 shows for each 10 minutes, in green, the location of the RV Meteor carrying the WindCube. In the time span of 1 hour the Meteor research vessel only covers a limited area of the studied domain. For this reason, hourly mean profiles from lidar measurements are not considered to be representative of the entire domain (see section 3). Figure 2 shows that wind profiles from lidar measurements are available continuously from the 25th of January to the 15th of February.
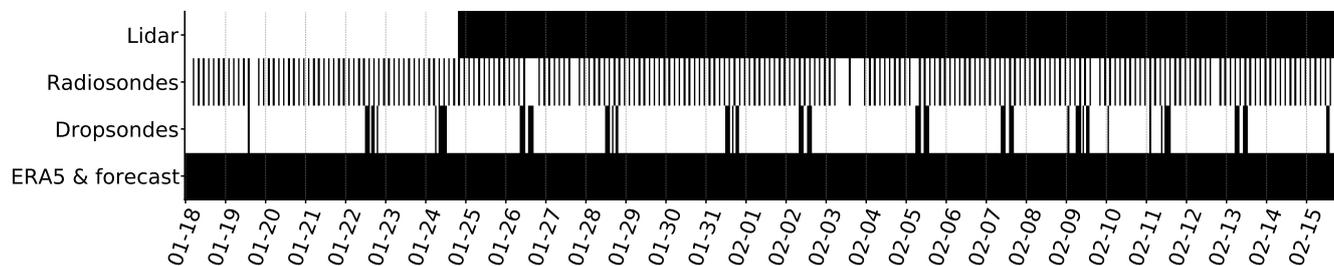
Figure 2. Overview of the temporal coverage of different observational datasets from EUREC4A which are used in this study. A black stripe indicates the availability of data for the corresponding hour.

2.2 Modelling datasets

145   We use the observations to validate several modelling datasets produced with the ECMWF IFS, covering the period from the 18th of January to the 15th of February 2020.

These datasets comprise the operational deterministic high-resolution (9km) forecasts and analysis (or initial conditions from which the forecasts start), the ERA5 reanalysis (Hersbach et al., 2020), and several experiments at coarser resolution. For each of these datasets, model output was extracted at the nearest four neighbours of 61 points placed concentrically around the

150   centre of EUREC4A-circle. Each group of four points was then used to interpolate the model values to the locations of the 61 arbitrary points using an inverse distance weighting method. The location of these arbitrary points is shown on the right panel of Figure 1 with black crosses. The second most external ring of points coincides with the EUREC4A-circle.

2.2.1 Forecast

For the operational ECMWF deterministic ten-days forecasts and analysis, which have a resolution of 9km, the extracted

155   model grid points for the EUREC4A-circle are marked in orange in Figure 1 (right). For clarity, we avoid showing the rest of the extracted model grid points. We extract hourly output for day two of the forecasts (a leadtime of 24 to 48 hours) and 3-hourly output for day four forecasts (a leadtime of 72 to 96 hours).

The wind errors in the short-range (day two) forecasts are very similar to the errors in the medium-range (day four) forecasts. This corroborates the findings of Sandu et al. (2020) which showed that over this trade-winds region the errors in wind profiles

160   develop in the first 12 hours of the forecast and do not grow significantly until day five.

Hereafter we will therefore focus on the forecasts with a lead time of two days, which we for simplicity refer to as forecast.

2.2.2 ERA5

The newest reanalysis produced by ECMWF for the Copernicus Climate Change Service, ERA5 reanalysis is widely used for model evaluation, and often it is used as a proxy for observations. Similar to the operational analysis, ERA5 is produced with

165 ECMWF IFS by optimally combining short-range forecasts and observations through data assimilation (as it is done to create the analysis, or initial condition of the forecasts). While operational analyses are not consistent in time because of regular upgrades to the forecasting system, reanalyses are produced with an unique version of the forecasting system. This leads to a consistent time series which allows one to monitor environmental changes. For example, ERA5 is produced with the IFS model cycle 41r2, at a resolution of approximately 32 km, and covers the period 1950 to present (Hersbach et al., 2020).

170 Here we exploit EUREC4A observations to evaluate the quality of the wind profiles also in the ERA5 reanalysis. The extraction points for ERA5 corresponding to the EUREC4A-circle are shown in blue in Figure 1 (right panel). In the sections below we focus on the wind profiles from ERA5, rather than from the operational analysis, because the differences in wind profiles over the EUREC4A region between the operational analysis (which is available 6 hourly) and ERA5 are marginal (not shown).

175 2.2.3 Sensitivity Experiments

A question often asked, and which also motivates looking at the performance of ERA5, is whether a reanalysis is close to reality because observations are assimilated. To answer this question, several sensitivity experiments were performed at 40 km resolution and outputs saved every 3 hours for the day two forecasts and every 6 hours for the analyses.

First, a control analysis (CTRL_an) and corresponding ten-days control forecasts (CTRL_fc) initialized from it were per-

180 formed at this resolution. Second, so-called data denial experiments were performed in which measurements made during EUREC4A are not assimilated when creating the initial conditions of the forecasts. These experiments consist of: (a) an analysis experiment in which the EUREC4A dropsondes are not assimilated, and corresponding ten-days forecasts (Exp1_an, Exp1_fc); (b) an analysis experiment where neither EUREC4A dropsondes nor radiosondes are assimilated, and corresponding ten-days forecasts (Exp2_an, Exp2_fc). Lastly, we performed an analysis experiment, and corresponding ten-days forecasts,

185 where shallow convective momentum transport is switched off (Exp3_an, Exp3_fc). These sensitivity experiments allow us to explore the origin of the IFS wind bias. For all experiments the analysis covers the period 15 January to 15 February 2020, and the ten-days forecasts were initialized daily at 00UTC from the respective analysis.

For the above mentioned sensitivity experiments we also extracted outputs at the four nearest model grid points to the 61 arbitrary points shown in Figure 1 and performed an inverse distance weighting to the location of our arbitrary points. We found

190 that the spatial resolution of the different model experiments has little impact on the results of this study, where we compare means obtained from several locations in a 350 km x 350 km domain.

3 Methods

Mean wind profiles are derived using the datasets described above. All profiles are interpolated to a grid of 50 m vertical resolution between 0.15 km and 5 km. The wind vectors are decomposed into zonal (u) and meridional (v) components and

195 analysed at different hours of the day using hourly and 3-hourly composites. While the modelling datasets directly provide vectorial wind components, the observations measure scalar quantities such as wind speed and wind direction. In this study we retrieve the corresponding meridional and zonal components for each radiosonde and dropsonde and for each of the 10 minute lidar winds.

As described in section 2, all dropsondes launched during one flight circle are averaged to a timestamp of the nearest full

200 hour. Radiosondes are averaged over the domain within three hours intervals. Lidar profiles are averaged within one hour intervals. While model outputs uniformly sample the entire domain at each time step, observations only sample one location at the time. To partially account for these differences in the data sets and the spatial variability of the trades, when we derive the forecast and (re)analysis errors with respect to the different observational datasets, we sample the model output to match the sampling we have chosen above for the respective data set. For example, when we compare to the radiosondes, we average the

205 model profiles extracted for the 61 points and over 3-hourly intervals. When we compare to the dropsondes, we average the model points extracted along the EUREC4A-circle at the hour during which the circle was flown. In the case of the wind lidar, the observed hourly mean profiles represent a small area of the domain and for this reason we use the closest point from the forecast/(re)analysis dataset to compute the model errors with respect to the lidar. When the model is simultaneously compared to multiple observational datasets (e.g. top row in Figure 4 and left panel in Figure 7), we show the model mean over the

210 selected domain, obtained from all the 61 points and with the temporal resolution available for the model output.
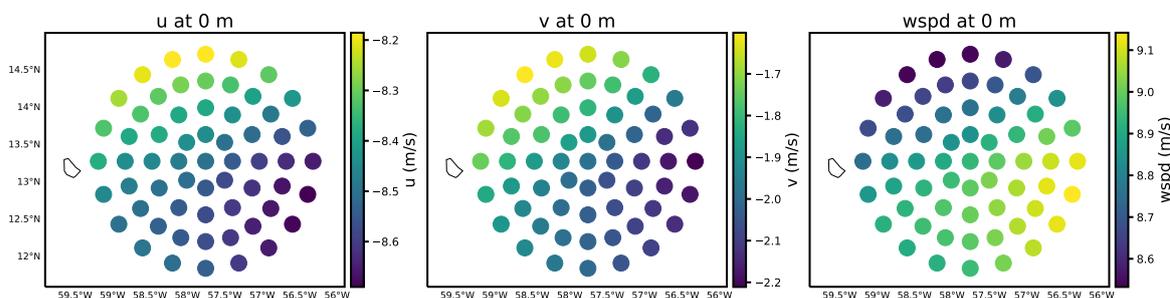
Figure 3. Spatial variability of wind components at 500 m derived from ERA5 for the whole period and at all hours.

Figure 3 helps quantifying the spatial variability of winds in the study area and motivates our choice of the spatial matching of the observations and the model output. It shows that there is a NW to SE gradient in wind, whereby the south-east region

of the domain experiences winds about 0.5 m s^{-1} stronger than the average of the domain. The lidar samples this region more frequently than the north-west area where weaker winds prevail. Thus, we expect the wind-lidar winds to generally be stronger.

215   As mentioned above, the spatial variability of winds in the domain is taken into account by calculating mean dropsonde profiles for each flight-circle. Similarly for the radiosondes, we assume that the launch locations over three hours are sufficiently dispersed to provide a good representation of the entire domain.

4   Temporal Variability of the Wind Bias

January and February 2020 are characterised by low-level north-easterly winds (Figure 4 (top row)). The (zonal) wind speed
220   typically peaks near cloud base and decreases aloft, establishing a so-called backward sheared wind profile. This structure is typical for the trades, as documented in earlier field studies (Riehl et al., 1951; Brümmer et al., 1974) and more recently using the BCO climatology alongside ERA-Interim (Brueck et al., 2015). A recent study using north-Atlantic wide Large Eddy Simulations with ICON (hindcasts performed for the pre-EUREC4A NARVAL campaign period) suggests that the local maximum in zonal wind near cloud base results from efficient turbulent diffusion in the sub-cloud layer, but little if any cumulus
225   friction at cloud base Helfer et al. (2021). There is also an apparent layer with counter-gradient momentum transport near cloud tops, which suggests that clouds tend to enhance the vertical wind gradient, rather than reduce it Larson et al. (2019); Dixit et al. (2021).

    Using JOANNE, Nuijens et al. (2021, under preparation) show that winds were relatively weak with strong backward shear during the final two weeks of January 2020, transitioning to a period with stronger winds and weaker shear during the first
230   week of February 2020, and that the campaign ended with several days with strong winds and strong backward shear.

    In the next subsections we analyse the ability of the IFS to model the mean winds and their variability during EUREC4A. First we look at the overall mean profiles during EUREC4A (section 4.1). Next, we analyse the day-to-day synoptic variability (section 4.2). Lastly, we analyse the diurnal cycle of winds on sub-daily time scales (section 4.3).

4.1   Monthly Mean

235 The top row of Figure 4 shows the mean wind profiles from different data sets for the period and domain analysed in this study (see section 2). Observations are spatially averaged as detailed in section 3 and temporally averaged over all days and all hours of the day (0 to 23 LT). The model profiles are spatially averaged over all 61 extraction points and also temporally averaged over all days and all times of the day. We have not included JOANNE in this first comparison because of the limited number of flight-days, and because JOANNE does not sample all hours of the day.

240   The mean wind speed (rightmost panel) during EUREC4A is about 9 m s^{-1} at 150 m, it slightly increases in the lower 800 meters and sharply reduces to 6 m s^{-1} in the cloud layer, between 1 km and 2 km. The zonal component is the largest contributor to the total wind speed and their profiles are similar in shape. Both ERA5 and the forecast capture well the mean zonal wind profile, particularly below 2km. The mean meridional wind component has smaller values than the zonal component, it is strongest at 150 m, about -2 m s^{-1}, and it decreases to -0.5 m s^{-1} at 2 km. Between 2 km and 3 km the meridional component

245 strengthens by more than $0.5\ \mathrm{m\,s^{-1}}$. This wind turning behaviour near the cloud top is present in the radiosonde and ERA5 data, but it is poorly captured by the forecast (solid blue line). The forecast suggests weaker meridional winds at all heights and in particular near 0.15 km and 3 km.
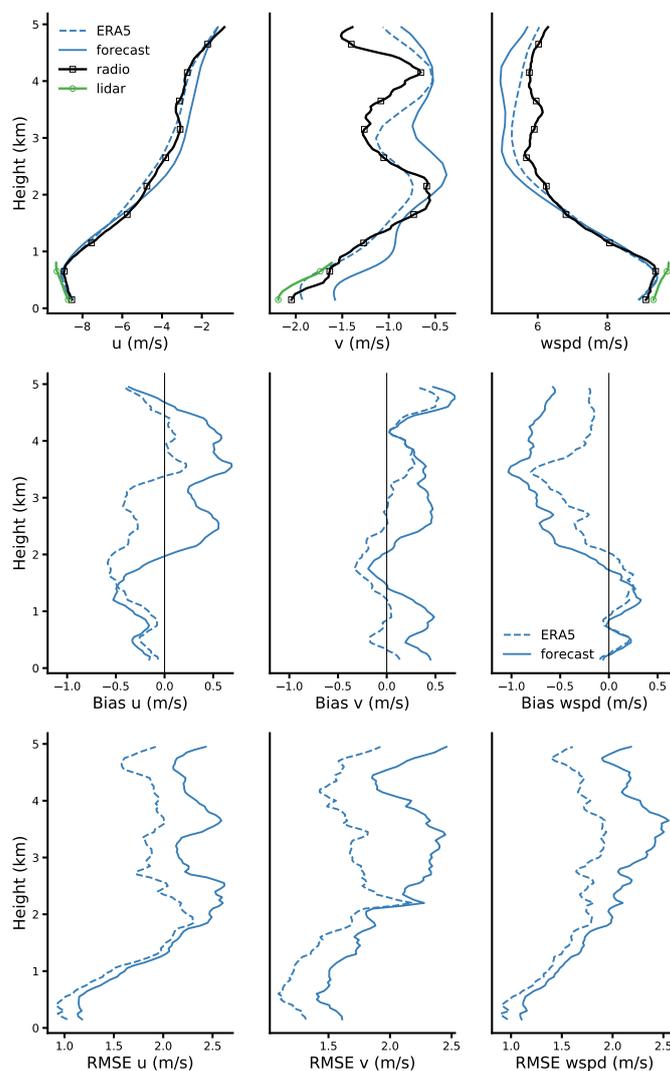
Figure 4. Mean profiles of zonal wind (left), meridional wind (middle) and wind speed (right) during EUREC4A. In the top row are monthly profiles retrieved from lidar (green circles), radiosondes (black squares), ERA5 reanalysis (solid blue), and day 2 forecast (dashed blue). The middle and bottom rows show the monthly biases, and root mean square error, of the forecast and ERA5 with respect to radiosondes.

The lidar measured stronger winds in the sub-cloud layer, as it probed the atmosphere in a region where winds were generally strong, as discussed in section 3.

250   We can further quantify the differences between the modelled and observed winds by defining the model bias and the root mean square error (RMSE) as:

$$\text{Bias} = \overline{\Theta_{\text{mod}} - \Theta_{\text{obs}}}, \tag{1}$$

$$\text{RMSE} = \sqrt{\overline{(\Theta_{\text{mod}} - \Theta_{\text{obs}})^2}}, \tag{2}$$

where the overbar represents the arithmetic mean and Θ is any modeled (mod) or observed (obs) variable. The model and different observational datasets are matched as described in section 3.

While the RMSE measures model accuracy independent of the sign of the error, the model bias takes into account the sign
255   of the errors and can be used to study the distribution of the error. The skewness of the error distribution is important for the bias: large errors that are normally distributed result in large values of RMSE, but a bias that is approximately zero. Otherwise said positive and negative errors can compensate each other and result in a nearly zero mean bias.

The biases of the forecast (solid blue) and ERA5 (dashed blue) compared to radiosondes are shown in the middle row of Figure 4. Below roughly 2 km (near the trade-inversion) the wind speed bias is relatively small (~ 0.1 m s^{-1}). This may be
260   partly because of compensating positive and negative errors for different days, but, especially for the forecast, it is largely the result of opposite zonal and meridional wind biases. In contrast, above 2 km the zonal and meridional winds are both too weak and result in a net underestimation of the wind speed.

As one would expect, ERA5 shows a better agreement with the observations than the forecast, both for the bias and the RMSE (middle and bottom row). The first two rows suggest that both ERA5 and the forecast perform considerably well in
265   mimicking the radiosondes and lidar monthly mean profiles, especially in the lower 2 km. Nevertheless, the bottom row of Figure 4 shows that the RMSE between forecast/ERA5 and radiosondes is as large as 1 m s^{-1} at 250 m and 2.5 m s^{-1} between 3 km and 4 km, for all components. This suggests that the positive and negative errors are fairly normally distributed and cancel out when calculating the model bias and the mean profiles.

In line with previous studies (Belmonte Rivas and Stoffelen, 2019), we find that on average ERA5 has slightly too strong
270   easterly winds near the surface, with a too strong easterly component and a mean bias close to zero in the northerly component. Nevertheless, here the forecast shows mean meridional winds almost 0.5 m s^{-1} too weak near the surface. A new finding here is that errors near the surface are relatively small compared to errors at higher levels. This suggests that the surface bias may be tied to processes further aloft. Before we explore this in more detail in section 5, we take a closer look at errors established at individual times, which as suggested by the relatively large RMSE should be pronounced (section 4.2). We will also show
275   that the bias near the surface and higher up critically depends on the time of the day (section 4.3).

4.2   Day-to-day Synoptic Variability

Although the trade-winds are generally steadier than midlatitude flows, they still exhibit significant variability on a daily and even hourly basis. Figure 5 shows wind components (zonal, meridional and wind speed) at hourly resolution from ERA5,

averaged over the selected domain. The first half of EUREC4A is characterised by days with wind speed up to $10\,\mathrm{m\,s^{-1}}$, while
280   the second half has stronger winds. Between Jan-27 and Jan-30 and between Feb-11 and Feb-14, winds turned south-westerly
above the cloud layer at about 3.5 km, leading to strong vertical wind shear.
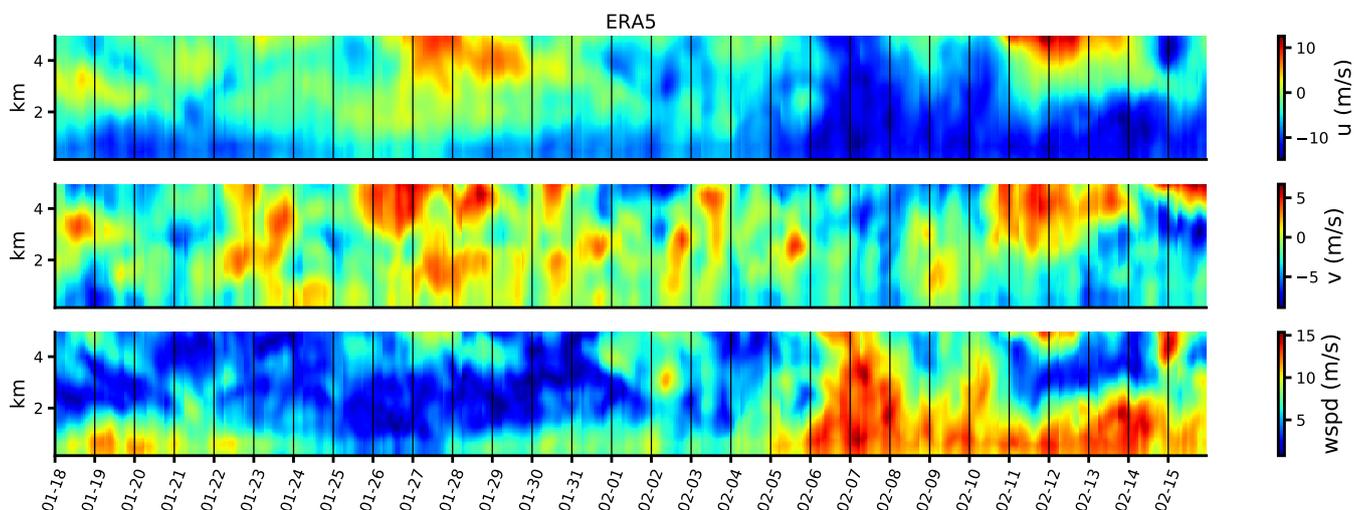
Figure 5. Time series of hourly ERA5 zonal wind (top), meridional wind (middle) and wind speed (bottom) averaged over the whole domain
during EUREC4A.

We are interested in whether the IFS error is larger on days associated with specific wind conditions. The 3-hourly bias
of ERA5 compared to the radiosondes (Figure 6) does not suggest a strong correlation between the reanalysis bias and the
wind speed or direction. The sign and magnitude of the 3-hourly biases are relatively similar in the first and second half of the
285   EUREC4A period, with positive and negative values of up to $2\,\mathrm{m\,s^{-1}}$ in both the zonal and the meridional wind components.
A similar pattern to the one in Figure 6, only with larger values, is obtained for the 3-hourly forecast bias with respect to
radiosondes (not shown).

A prolonged and significant overestimation of the zonal jet (around 2 km) can be noticed on Jan-22 and Jan-23. This could
be related to the misrepresentation of a large scale phenomenon acting on multi-diurnal time scales.
290   Figure 5 also reveals a large variability in the meridional wind component (Figure 5, middle panel) on hourly timescales. At
2 km, diurnal variations of more than $3\,\mathrm{m\,s^{-1}}$ are registered (*e.g.*, on the 2nd and 5th of February). The ERA5 3-hourly bias
in Figure6 also shows large differences within individual days. At all heights, ERA5 can overestimate and underestimate the
winds within the same day. In the following sections we investigate the diurnal cycle of winds and of the wind bias in forecasts
and reanalysis in more detail.
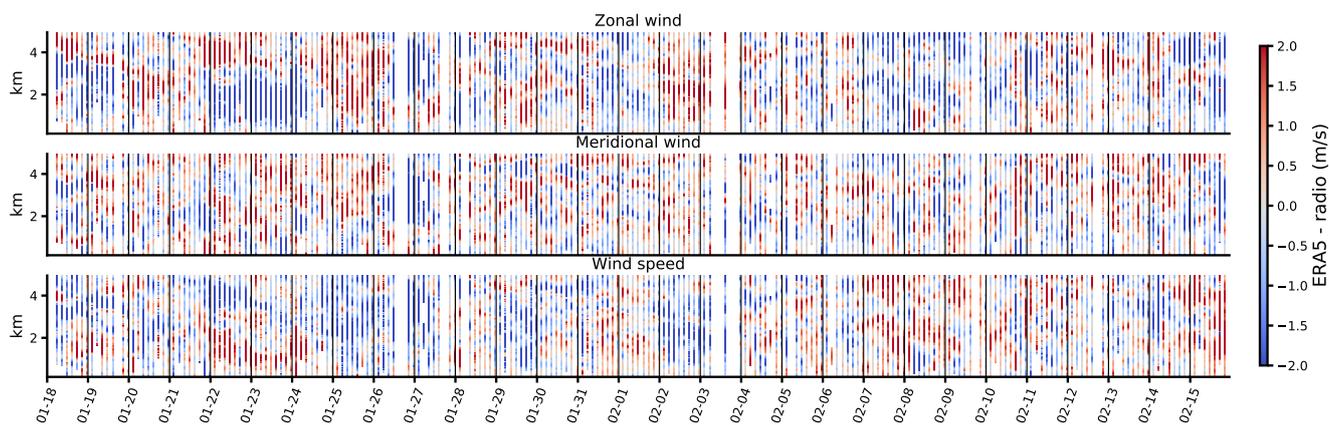
Figure 6. Difference between ERA5 and radiosonde wind profiles averaged over the whole domain and over 3-hourly intervals. From top to bottom: zonal, meridional winds and wind speed.

4.3 Diurnal Cycle

To analyse the diurnality of the winds we create composites for all the data sources as described in section 3. Hourly composite are obtained from dropsondes and lidar profiles, while 3-hourly composites from radiosondes. For ERA5 and the forecast we use hourly composites unless they are compared to radiosondes, then 3-hourly composites are used.

The mean diurnal cycle of the layer between 0.15 km and 0.75 km is depicted in Figure 7 (left), where all available datasets show a pronounced diurnal cycle with the weakest wind speeds during the day and the strongest winds at night. In the observations, the amplitude of the diurnal cycle is about 1 m s^{-1} in both the meridional and zonal component. This diurnality is significantly overestimated by the forecast with an amplitude almost twice that of the observations. At 15 LT the zonal wind bias is largest: the forecast underestimates the magnitude of the zonal wind component by 1 m s^{-1} with respect to both lidar and radiosondes measurements. This daytime bias is less pronounced, but still present in ERA5. Instead, in the late night and early morning the forecast biases are most pronounced in the meridional wind (middle left panel in Figure 7): the forecast exaggerates and anticipates the morning weakening of the meridional wind. ERA5 is notably better at capturing the amplitude of the diurnal cycle in the meridional component, confirming that the data assimilation plays an important role here, something we explore further in section 5 when we test the sensitivity to the assimilation of the radiosondes and dropsondes.

Although ERA5 itself is not bias free, we can use it to explore the vertical structure of the diurnality in winds. In the right panel of Figure 7 the wind diurnal cycle from ERA5 is represented between 150 m and 5 km. The contours show that the diurnality is present at all heights of the vertical domain, not only in the sub-cloud layer, where turbulence plays a major role.

In particular, above 2 km, the zonal and total wind speed variations (right panel of Figure 7) suggest a semi-diurnal cycle of the zonal winds, with weakest winds in the first few hours of the day and around 16 LT. A semi-diurnal cycle in zonal winds (and diurnal cycle in meridional winds) has been found over the tropical oceans in earlier studies (Dai and Deser, 1999;
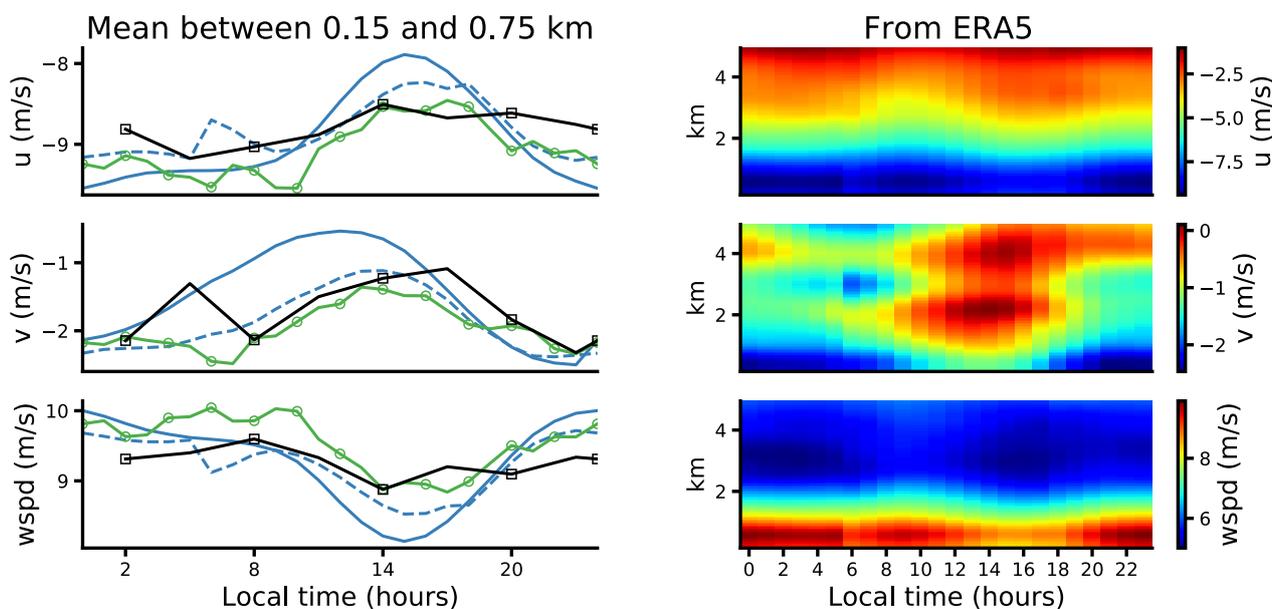
Figure 7. Diurnal cycle of zonal wind (top row), meridional wind (middle row) and wind speed (bottom row). The left column refers to the layer between 0.15 km and 0.75 km with values from radiosondes (black squares), lidar (green circles), ERA5 (dashed blue) and forecast (solid blue). The right column refers to multiple levels from surface to 5 km with values from ERA5 only.

Ueyama and Deser, 2008) and has been linked to semi-diurnal atmospheric thermal tides generated by the absorption of solar radiation by ozone in the stratosphere and water vapor in the troposphere. These tides travel downward and affect sea level pressure, whose tidal amplitudes appear mostly semi-diurnal.

The diurnal meridional wind variations are not fully understood, but Ueyama and Deser (2008) showed that over the tropical Pacific such variations agree very well with pressure-derived wind diurnality, suggesting that the pressure gradient force plays a dominant role in setting the diurnality, with a smaller role for boundary layer stability. As will be shown later in section 6.1, Barbados experiences only half of the dynamical forcing at daytime compared to nighttime in both the zonal and meridional component (Figure 11 middle compared to bottom panel). The diurnality in pressure gradients may be linked to a diurnality in the Hadley cell, where the diurnal cycle of deep convection in the ITCZ and over the South American continent, which respectively peak in the early morning and in the afternoon, may play a role.

Figure 8 quantifies the mean model bias as a function of height and time of day with respect to radiosondes (top and middle rows), while Figure 9 shows the mean bias during flight hours (top row), during daytime (between 10 LT and 16 LT) and nighttime (between 22 LT and 4 LT).

These figures reveal that a too strong easterly wind in the IFS during nighttime (as found near the surface in Figure 7) is present throughout the lower 2 km of the atmosphere. During daytime and during flight hours (which are predominantly during
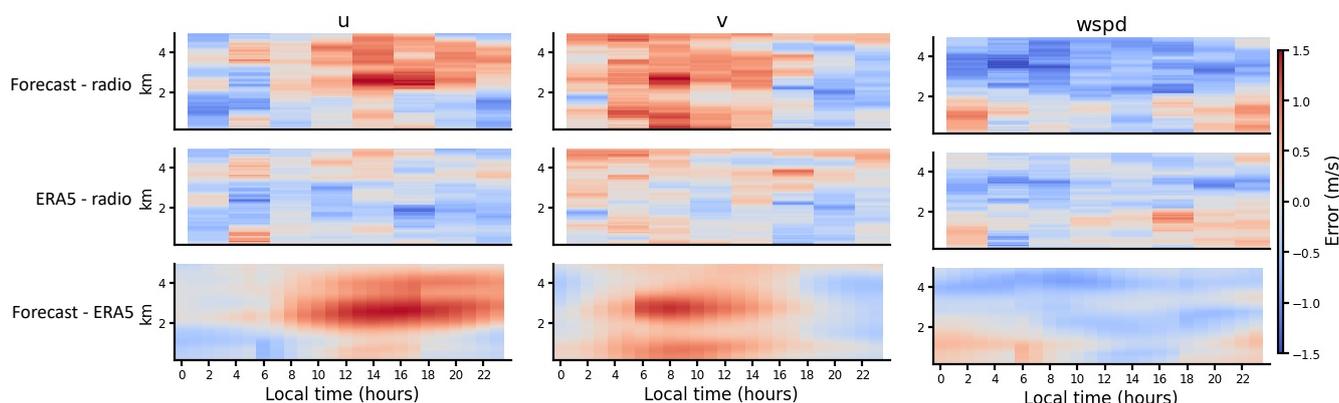
Figure 8. Diurnal cycle of the forecast bias with respect to radiosondes (top row), the ERA5 bias with respect to radiosondes (middle row), and the forecast bias with respect to ERA5 (bottom row). From left to right, columns refers to the biases in zonal wind, meridional wind and wind speed. Red regions are related to a positive bias (e.g. too weak negative zonal wind), blue regions are related to a negative bias (e.g. too weak wind speed).

330   daytime), the meridional wind component contributes most to the weak wind speed bias in the forecasts, not only below 2 km, but also above 2 km, where it is accompanied by a too weak easterly wind (Figure 8 and 9). This suggests that at least one contributor to a weak wind speed bias during daytime, seen here throughout the lower troposphere, may be a bias in the thermal wind of the Hadley cell (vertical wind shear), as influenced by deep convection away from the trades.

Another important aspect to notice is that the nighttime bias is similar for both the forecast and ERA5 while the daytime
335   bias is only visible in the forecast; this can be traced back to what is seen in Figure 7, where both the forecast and reanalysis overestimate the amplitude of the diurnal cycle but only ERA5 captures the phase of the cycle.

From Figure 9 we can also infer that the dropsondes and radiosondes agree fairly well, apart for the zonal wind in the cloud layer. Here, at about 1.5 km, the radiosondes show zonal winds $\sim$1 m s^{-1} stronger than the dropsondes. These differences may be due to differences between the descending and ascending radiosondes. Descending radiosondes tend to show stronger
340   winds above 1.5 km. Excluding the 186 descending radiosondes produces a better agreement with the dropsondes above 2 km (not shown). However, around 1 km the descending radiosondes match the dropsondes considerably better than the ascending radiosondes. We also notice that the number of operating dropsondes reduces at lower altitudes. The differences between the radio- and dropsondes and the lidar winds are, as discussed in section 3, largely due to the spatial variability of winds.

Lighter colours in the middle row of Figure 8 compared to the top row show that ERA5 performs much better than the
345   forecast at all hours of the day. Nevertheless the pattern in the right most panel (wind speed) suggests that the reanalysis only reduces the magnitude of the bias, without eliminating the fundamental causes of an overestimated diurnal wind cycle below 1 km. In the next section we explore to what extent the assimilation of radiosondes and dropsondes explain the better performance of ERA5.
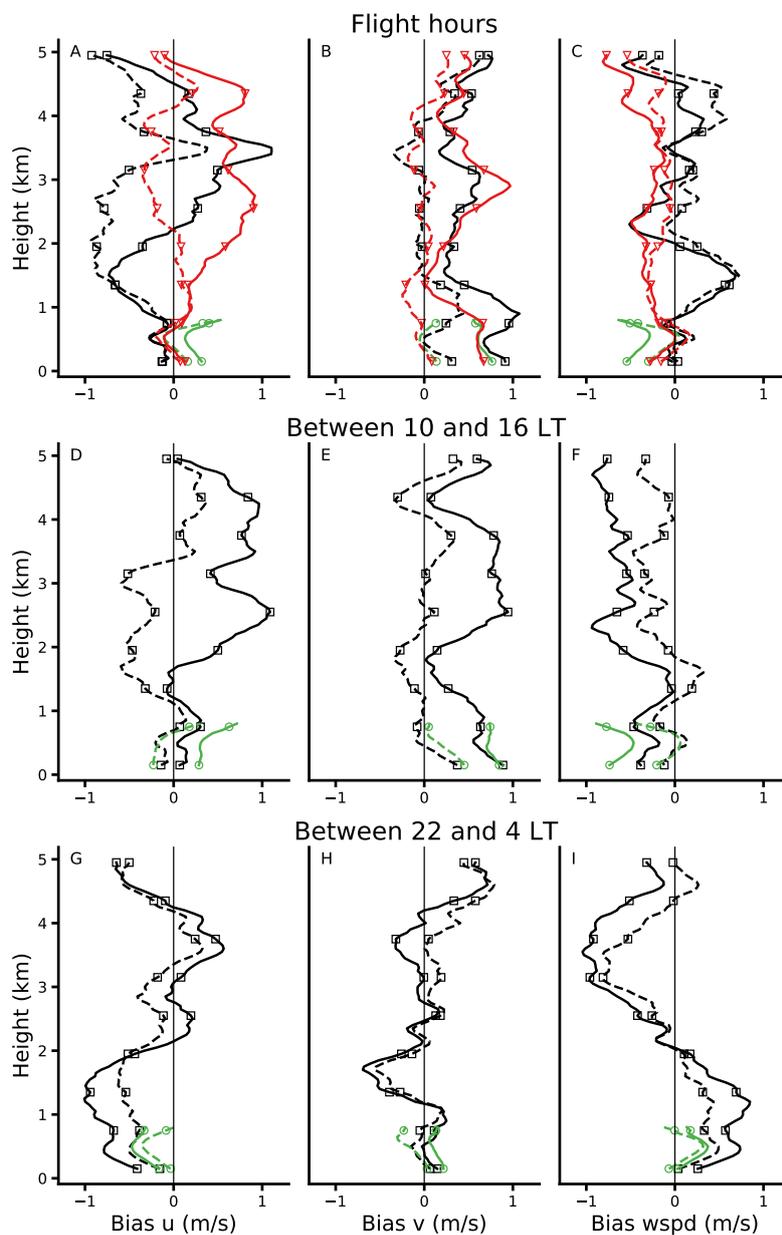
Figure 9. Mean model bias (forecast in solid and ERA5 in dashed) during flight hours (top row) and during the whole EUREC4A campaign, measured separately for daytime between 10 and 16 LT (middle row), and for nighttime between 22 and 4 LT (bottom row). The bias is calculated with respect to radiosondes (black squares), lidar measurements (green circles), and dropsondes (red triangles). From left to right the columns refer to the bias in the zonal wind (u), meridional wind (v), and wind speed.

5 Model Sensitivity to Sounds Assimilation

350 In the data assimilation process of ERA5 and the operational analysis, the radiosondes and dropsondes launched during EU-REC4A were used. In this section we investigate to what extent the assimilation of these observations has influenced the performance of the analysis and the forecast.

We performed extractions from the IFS analysis and forecast of a control experiment (CTRL_an, CTRL_fc), of an experiment without assimilating dropsondes (Exp1_an, Exp1_fc), and an experiment without assimilating dropsondes nor radioson-
355 des (Exp2_an, Exp2_fc). For each of the mentioned experiments the monthly mean bias is calculated over all hours of the day, with respect to the radiosondes, as done in the middle row of Figure 4. The results are shown in Figure 10, where the dashed lines refer to the analyses and the solid lines to the forecasts.
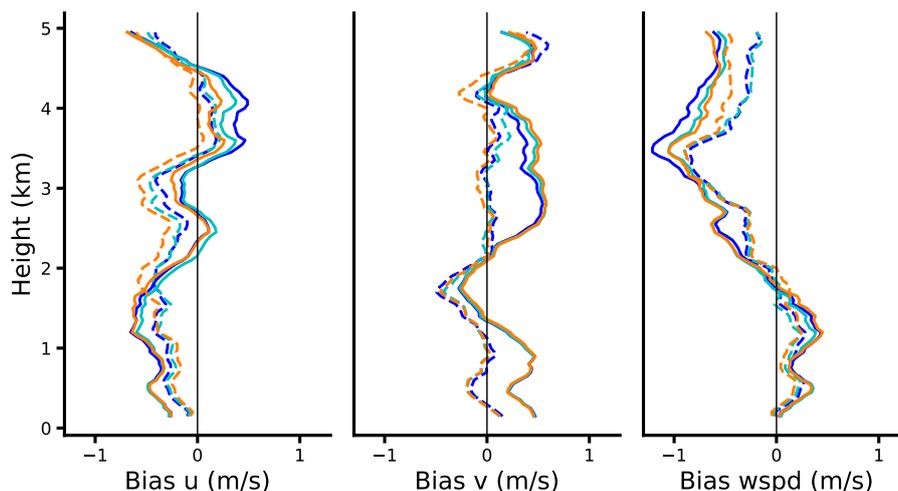
Figure 10. Monthly mean IFS bias against radiosondes as in middle row of Figure 4, forecasts are in solid and analyses in dashed. In the control experiment (blue) both dropsondes and radiosondes from EUREC4A are assimilated. In the first experiment (cyan) dropsondes are excluded from the assimilation. In the second experiment (orange) neither dropsondes nor radiosondes are assimilated.

Evidently, all analysis and forecast experiments remain considerably close to the corresponding control experiment (blue lines): the differences are everywhere small, and almost non-existent below 2 km. The sign, shape and magnitude of the profiles
360 in Figure 10 confirm the results described in previous sections (e.g. see Figure 4) and support the idea that the mean wind bias does not increase with coarser model runs (40 km spatial resolution and 3 hours temporal resolution of the model output). This also suggests that assimilating the local soundings do not alleviate the existing biases.

The fact that the analysed wind profile error 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. Typically, when one observing system is withdrawn
365 from the data assimilation system, other observing systems take over in constraining the analysis Sandu et al. (2020).

6 Influence of Model Physics on the Wind Bias

Previous sections highlighted that a wind speed bias exists throughout the lower troposphere, with a pronounced diurnality that has a different character in the zonal and meridional directions.

An important question is whether the IFS wind bias is related to a bias in large-scale dynamics (the pressure gradient force
370 that drives the geostrophic wind), a bias in boundary layer processes, including small-scale turbulent diffusion and convective momentum transport, or something else.

In the following, we will explore the momentum budget in the IFS and in the observations and perform sensitivity experiments that remove the momentum transport by shallow convection. Lastly we show how the latest revision to the representation of moist physical processes in IFS (cycle 47r3) improves the wind predictability.

375 ### 6.1 Observed Versus Modeled Momentum Budget

The mean tendencies in the momentum budget are shown in Figure 11 for flight hours only (top row), where they are compared against the observed tendencies as derived in Nuijens et al. (2021, under preparation) from the JOANNE dataset, and for all EUREC4A days, but then for the model tendencies only (middle and bottom row). In both the observations (solid lines) and in the IFS (dashed lines), the advection, pressure gradient and Coriolis force are combined into one "dynamical" large-
380 scale forcing. The budget residual, which we loosely refer to as "frictional" force, is comprised of convective and turbulent momentum transport. Horizontal and vertical advection of the mean wind are on average close to zero (not shown), so that the large-scale forcing consists mainly of the pressure gradient and the Coriolis force. Because the flight data used to derive the observed tendencies were mostly during morning and daytime hours, as described in detail in Nuijens et al. (2021, under preparation), the flight hour budget is most comparable to the tendencies between 8 and 14 LT (middle row).

385 There is remarkable agreement between the general structure and magnitude of the tendencies in the observations and the IFS in the boundary layer, providing confidence in the method used to estimate the budget from observations, as well as in the ability of the IFS to reproduce the different processes at play. There is a non-negligible positive net tendency in the zonal direction (red), in agreement with a slow-down of the easterly wind in the morning and during daytime, which appears driven by a reduction in the large-scale pressure gradient (Figure 11, middle compared to bottom row).

390 Compared to the observations, the IFS has larger dynamical and frictional tendencies in the zonal component in the sub-cloud layer up to 0.75 km, where the observations suggest a gradual weakening of the forcing with height. Because the turbulent friction and the large-scale pressure gradients are coupled through the circulation, it is hard to disentangle which of the two is at cause. However, a similar picture can be seen in the meridional component, where the model and observations agree on the dynamical forcing driving northerly winds below 500 m, while the IFS overestimates the frictional force.

395 Between 1.5 and 3 km the frictional force is near zero in the IFS, but the observations suggest a layer with negative frictional force (i.e., an acceleration of the easterly flow) that is near cumulus tops. As such there is a larger net deceleration of easterly winds in the IFS, consistent with the finding that the IFS has a weak zonal wind bias at those heights during flight hours (Figure 9, top row). In the meridional component, the IFS predicts zero frictional force above roughly 500 m, while the observations

suggest that small frictional effects are present between 1 and 2 km, and an acceleration of northerly winds is seen above 2

400   km. There might be an important role of convectively driven circulations that change winds even at levels above the sub-cloud layer, which are not present in the IFS. The tendency produced by the cumulus parameterization (in blue, Figure 11) is very small for the meridional component.

In the next section we carry out one experiment removing the momentum transport by convection, to investigate at what heights this may contribute to the bias.

405   **6.2   Removing Shallow Convective Momentum Transport**

To investigate the role of shallow convection momentum transport we compare Exp3_an and Exp3_fc, in which shallow convective momentum transport is turned off, with the same control experiment as in section 5 (CTRL_an, CTRL_fc) (see Figure 12).

CMT acts to mix winds between the surface and the cloud layer. If the wind speed increases with height, as is typically true

410   for the sub-cloud layer, this would result in an increase of wind speed near the surface and a decrease in wind speed in the cloud layer (the so-called "cumulus friction effect"). Figure 12 confirms that CMT by shallow convection acts to weaken the dominant easterly winds in the cloud layer, while strengthening winds near the surface. Without shallow CMT, in Exp3_an and Exp3_fc, a too strong easterly and total wind jet develops around 1 km during both day and night, while the mean bias in the zonal wind reduces to near zero at night and the meridional wind near the surface misses a frictional effect However, at levels

415   above 2 km, the weak wind speed bias, evident in both the zonal and meridional components particularly at daytime, remains. There is little difference between the black lines (Exp3_an and Exp3_fc) and the blue lines (CTRL_an and CTRL_fc) above 2 km in Figure 12. At these height levels, convective tendencies in the IFS are small or negligible (Figure 11).

As suggested in Sandu et al. (2020), CMT may not underlie previously established errors in near-surface winds, but may rather act to communicate erroneous winds at upper levels to the surface. Nevertheless, it appears plausible that the IFS has

420   biases in other components of the residual "frictional" force (e.g. the turbulent diffusion or the frictional force produced by the convection scheme). Using ICON-LEM hindcast runs over the North Atlantic for twelve days corresponding to the NARVAL1 (winter) and NARVAL2 (summer) flight campaigns Dixit et al. (2021) and Helfer et al. (2021) show that resolved convective circulations play an important role in driving winds southwards throughout a deep layer up to 5 km. They also reveal the presence of convectively driven horizontal circulations and associated counter-gradient momentum flux near cloud tops. The

425   IFS hardly has any convective tendencies above 1 km in the meridional component, and none above 1.5 km in the zonal component and we therefore hypothesize that model physics does play a role in the wind biases.

6.3   New moist physics

In this section we test our hypothesis on the role of model physics in the wind biases. To do so, we use a model experiment with the most recent IFS cycle (47r3) (Forbes et al., 2021) and compare it with what we have called operational forecast at the

430   time of the field campaign.

In the upgraded 47r3 cycle the main revisions concern the moist physics ensuring that the interactions between turbulence in the lowest part of the atmosphere, convective motions and the cloud physics are described as simply, efficiently, accurately, and scale-independently as possible. The moist physics upgrade also addresses errors in the parametrization of deep convection, especially for the representation of propagating mesoscale convective systems and their diurnal cycle. In particular, insufficient

435   night-time convection over land has been identified as a major shortcoming in IFS forecasts of convective activity (Becker et al., 2021; Forbes et al., 2021).

The evaluation of the IFS with the revised deep convection is ongoing, here we show how the upgrade leads to improved predictability of tropospheric winds in the subtropics. The red lines in Figure 13 indicate that the mean wind bias with respect to radiosondes is largely reduced during daytime and above 2 km. The solid blue lines refer to the operational forecast while

440   the dashed blue lines to ERA5. We present separate panels for the EUREC4A mean over all hours of the day (top row), for daytime (middle row), and for nighttime (bottom row). The upgraded model improves the wind forecast everywhere except for a slight deterioration of the zonal component below 1.5 km during daytime and above 2 km during nighttime.

Although the overall mean wind profiles are similar for the operational and the most recent forecast (see first row in Figure 13), there is a remarkable reduction of the daytime meridional wind bias (see panel E). With the upgraded model, the forecast

445   becomes closer to the observations and to ERA5 at all levels. This is in line with our hypothesis that the IFS wind bias is, at least in part, related to deep convection taking place at lower latitudes and thus, to the large scale dynamics. Furthermore, with this experiment we show that biases in near surface subtropical winds are dependent on biases in the convectively driven tropical free tropospheric flow.

7   Discussion and Conclusions

450   In this study we exploited multiple measurements from the EUREC4A field campaign to assess the lower tropospheric wind bias in the operational forecasts and ERA5 reanalyses performed with the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecast (ECMWF). We focused on a 350 km x 350 km domain in the trade-winds region eastward of Barbados and investigated wind profiles extending up to 5 km height during a month-long period during boreal winter. To the authors' knowledge, this is the first time that observational vertical profiles of wind fields are available over

455   ocean for such an extended period of time and from various instruments.

Our analysis shows that the structure and variability of the trade-winds are reproduced quite well in the IFS, although there are biases both at the surface and throughout the troposphere, with the largest values of the bias near and above the mean trade-inversion (~ 2.3km). In a monthly average the forecast underestimates the meridional wind component by about 0.5 m s^{-1} in the layers below 1 km and between 2.5 and 4 km. The zonal wind component is also about 0.5 m s^{-1} too weak between 2.5

460   and 4 km, while it is slightly overestimated below 1 km, in line with the known near-surface excessive easterly flow of the IFS (Belmonte Rivas and Stoffelen, 2019). The RMSE of the forecasts is larger: it increases with height from 1 m s^{-1} near the surface to 2.5 m s^{-1} near 3.5 km in all wind components (see Figure 4). The RMSE is independent of the sign of the error

and thus also measures positive and negative random errors that can otherwise compensate in computing the mean bias. As expected, the wind bias is smaller in ERA5 with the RMSE peaking at about 2 m s^{-1}.

465 Our study further shows that the wind bias is not constant throughout the day, but it exhibits a diurnal cycle, as previously documented in Vial et al. (2019). In the lower 0.75 km winds are weakest during the day, at 14 LT, with the wind speed being ∼9 m s^{-1} on average. The strongest wind speed is registered around midnight with a mean value of ∼10 m s^{-1}. The modelled diurnal cycle of the winds is overestimated in the lower 0.75 km, with too weak winds during the day and too strong winds during the night, particularly in the forecasts (see Figure 7). While ERA5 captures the phase of this cycle, the daytime
470 weakening of the meridional component is too early in the forecasts. This results in a too weak wind turning away from the geostrophic wind direction.

An analysis of the impact of the assimilation of the EUREC4A soundings shows that the IFS (re)analysis and forecasts are not very sensitive to the assimilation of local wind information in these undisturbed trade-winds conditions, and are apparently well constrained through large-scale dynamics and the other observing systems. In ERA5 the lower tropospheric wind bias is
475 smaller, but still present.

The underestimation of the easterly wind at all levels during daytime in the IFS suggests that the equatorial deep convection influencing the thermal winds of the Hadley cell is one of the contributors to the bias. Evidence of the important role played by large scale processes in determining wind diurnality in the trades is derived from a comparison of the momentum budget and from a model experiment with improved moist physics. Although it is challenging to point to one process being responsible for
480 the biases, because small-scale momentum mixing and large-scale dynamical forcing are closely coupled, we do find biases in the momentum tendencies that are consistent with biases in an overall "frictional" force. It appears that convective tendencies in the model are negligible above 1.5 km in the zonal direction, and above 500 m in the meridional direction, while observations suggest that a net acceleration of winds exists. LES hindcasts of a previous campaign east of Barbados suggests that horizontal circulations associated with convection play an important role in driving winds near the surface and near cloud tops. At the
485 moment, such large domain LES hindcasts are also developed for EUREC4A and will soon be used to disentangle which processes are responsible for biases in winds on individual days.

Data availability. The observational data used in this publication was gathered in the EUREC4A field campaign. JOANNE dataset is publicly available (https://doi.org/10.25326/221), the radiosondes dataset is publicly available (https://doi.org/10.25326/62), the WindCube lidar dataset is publicly available (https://eurec4a.aeris-data.fr/).
490 The data for the sensitivity experiments performed with the IFS used in this study are available at the following DOIs. CTRL_an: https://doi.org/10.21957/4vgx-3f28 , CTRL_fc: https://doi.org/10.21957/240p-1k07 , Exp1_an: https://doi.org/10.21957/zfxz-3h02 , Exp1_fc: https://doi.org/10.21957/nv0f-pr71 , Exp2_an: https://doi.org/10.21957/7zx9-6084 , Exp2_fc: https://doi.org/10.21957/mgrt-pp74 , Exp3_an: https://doi.org/10.21957/2t2w-wy02 , Exp3_fc: https://doi.org/10.21957/af7h-bf97 .

ERA5 was produced by ECMWF as part of implementing the Copernicus Climate Change Service on behalf of the European Union and
495 is made publicly available through the Copernicus Data Store (https://cds.climate.copernicus.eu/cdsapp#!/search?type=dataset).

Author contributions. AS, LN, and IS designed the study. AS performed the analysis and wrote the paper. LN and IS revised the paper, the figures and wrote some text. GG pre-processed the dropsondes measurements and produced the JAONNE dataset, GG also helped writing section 2.1.1. PB performed the model experiments with the most recent IFS cycle (47r3) and helped writing section 6.3. GG and PB also provided critical review of the paper.

500 *Competing interests.* This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Starting grant agreement no. 714918).

EUREC4A is funded with support of the European Research Council (ERC), the Max Planck Society (MPG), the German Research Foundation (DFG), the German Meteorological Weather Service (DWD) and the German Aerospace Center (DLR).

The authors declare that they have no conflict of interest.

505 *Disclaimer.* The data from the IFS sensivity experiments and from the ECMWF operational deterministic ten-days forecast are published under a Creative Commons Attribution 4.0 International (CC BY 4.0). Copyright: "© 2021 European Centre for Medium-Range Weather Forecasts (ECMWF)". ECMWF does not accept any liability whatsoever for any error or omission in the data, their availability, or for any loss or damage arising from their use. Source: www.ecmwf.int.

Acknowledgements. The authors would like to thank Sebastien Massart and Mohamed Dahoui for helping to set up the data assimilation
510 experiments in which the EUREC4A sondes were not assimilated and Julia Windmiller, Jakob Mann, Hugo Rubio Hurtado, Soeren Lund, Geiske de Groot and Kevin Helfer for their help with the deployment of the WindCubes on the RV Meteor and processing of the WindCube measurements.

References

Becker, T., Bechtold, P., and Sandu, I.: Characteristics of Convective Precipitation over Tropical Africa in Storm-Resolving Global Simula-
515   tions, Quarterly Journal of the Royal Meteorological Society, 147, 4388–4407, https://doi.org/10.1002/qj.4185, 2021.

Belmonte Rivas, M. and Stoffelen, A.: Characterizing ERA-Interim and ERA5 Surface Wind Biases Using ASCAT, Ocean Science, 15,
        831–852, https://doi.org/10.5194/os-15-831-2019, 2019.

Bony, S., Stevens, B., Frierson, D. M. W., Jakob, C., Kageyama, M., Pincus, R., Shepherd, T. G., Sherwood, S. C., Siebesma,
        A. P., Sobel, A. H., Watanabe, M., and Webb, M. J.: Clouds, Circulation and Climate Sensitivity, Nature Geoscience, 8, 261–268,
520   https://doi.org/10.1038/ngeo2398, 2015.

Brown, A. R., Beljaars, A. C. M., Hersbach, H., Hollingsworth, A., Miller, M., and Vasiljevic, D.: Wind Turning across the Marine At-
        mospheric Boundary Layer, Quarterly Journal of the Royal Meteorological Society, 131, 1233–1250, https://doi.org/10.1256/qj.04.163,
        2005.

Brown, A. R., Beljaars, A. C. M., and Hersbach, H.: Errors in Parametrizations of Convective Boundary-Layer Turbulent Momentum Mixing,
525   Quarterly Journal of the Royal Meteorological Society, 132, 1859–1876, https://doi.org/10.1256/qj.05.182, 2006.

Brueck, M., Nuijens, L., and Stevens, B.: On the Seasonal and Synoptic Time-Scale Variability of the North Atlantic Trade Wind Region and
        Its Low-Level Clouds, Journal of the Atmospheric Sciences, 72, 1428–1446, https://doi.org/10.1175/JAS-D-14-0054.1, 2015.

Brümmer, B., Augstein, E., and Riehl, H.: On the Low-Level Wind Structure in the Atlantic Trade, Quarterly Journal of the Royal Meteoro-
        logical Society, 100, 109–121, https://doi.org/10.1002/qj.49710042310, 1974.

530   Chaudhuri, A. H., Ponte, R. M., Forget, G., and Heimbach, P.: A Comparison of Atmospheric Reanalysis Surface Products over the Ocean
        and Implications for Uncertainties in Air–Sea Boundary Forcing, Journal of Climate, 26, 153–170, https://doi.org/10.1175/JCLI-D-12-
        00090.1, 2013.

Chelton, D. B. and Freilich, M. H.: Scatterometer-Based Assessment of 10-m Wind Analyses from the Operational ECMWF and NCEP
        Numerical Weather Prediction Models, Monthly Weather Review, 133, 409–429, https://doi.org/10.1175/MWR-2861.1, 2005.

535   Chelton, D. B., Schlax, M. G., Freilich, M. H., and Milliff, R. F.: Satellite Measurements Reveal Persistent Small-Scale Features in Ocean
        Winds, Science, 303, 978–983, https://doi.org/10.1126/science.1091901, 2004.

Dai, A. and Deser, C.: Diurnal and Semidiurnal Variations in Global Surface Wind and Divergence Fields, Journal of Geophysical Research:
        Atmospheres, 104, 31 109–31 125, https://doi.org/10.1029/1999JD900927, 1999.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer,
540   P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haim-
        berger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz,
        B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim Reanalysis:
        Configuration and Performance of the Data Assimilation System, Quarterly Journal of the Royal Meteorological Society, 137, 553–597,
        https://doi.org/10.1002/qj.828, 2011.

545   Dixit, V., Nuijens, L., and Helfer, K. C.: Counter-Gradient Momentum Transport Through Subtropical Shallow Convection in ICON-LEM
        Simulations, Journal of Advances in Modeling Earth Systems, 13, e2020MS002 352, https://doi.org/10.1029/2020MS002352, 2021.

Foley, A. M., Leahy, P. G., Marvuglia, A., and McKeogh, E. J.: Current Methods and Advances in Forecasting of Wind Power Generation,
        Renewable Energy, 37, 1–8, https://doi.org/10.1016/j.renene.2011.05.033, 2012.

Forbes, R., Laloyaux, P., and Rodwell, M.: IFS Upgrade Improves Moist Physics and Use of Satellite Observations, ECMWF Newsletter, pp.
550   17–24, 2021.

George, G., Stevens, B., Bony, S., Pincus, R., Fairall, C., Schulz, H., Kölling, T., Kalen, Q. T., Klingebiel, M., Konow, H., Lundry, A., Prange,
        M., and Radtke, J.: JOANNE : Joint Dropsonde Observations of the Atmosphere in Tropical North atlaNtic Meso-Scale Environments,
        Earth System Science Data Discussions, pp. 1–33, https://doi.org/10.5194/essd-2021-162, 2021.

Gille, S. T.: Statistical Characterization of Zonal and Meridional Ocean Wind Stress, Journal of Atmospheric and Oceanic Technology, 22,
555   1353–1372, https://doi.org/10.1175/JTECH1789.1, 2005.

Gottschall, J., Catalano, E., Dörenkämper, M., and Witha, B.: The NEWA Ferry Lidar Experiment: Measuring Mesoscale Winds in the
        Southern Baltic Sea, Remote Sensing, 10, 1620, https://doi.org/10.3390/rs10101620, 2018.

Helfer, K. C., Nuijens, L., and Dixit, V. V.: The Role of Shallow Convection in the Momentum Budget of the Trades from Large-Eddy-
        Simulation Hindcasts, Quarterly Journal of the Royal Meteorological Society, n/a, https://doi.org/10.1002/qj.4035, 2021.

560   Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Sim-
        mons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. D., Dahlgren,
        P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J.,
        Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Vil-
        laume, S., and Thépaut, J.-N.: The ERA5 Global Reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049,
565   https://doi.org/10.1002/qj.3803, 2020.

Hollingsworth, A.: Validation and Diagnosis of Atmospheric Models, Dynamics of Atmospheres and Oceans, 20, 227–246,
        https://doi.org/10.1016/0377-0265(94)90019-1, 1994.

Houchi, K., Stoffelen, A., Marseille, G. J., and De Kloe, J.: Comparison of Wind and Wind Shear Climatologies Derived from High-
        Resolution Radiosondes and the ECMWF Model, Journal of Geophysical Research, 115, D22 123, https://doi.org/10.1029/2009JD013196,
570   2010.

Larson, V. E., Domke, S., and Griffin, B. M.: Momentum Transport in Shallow Cumulus Clouds and Its Parameterization by Higher-Order
        Closure, Journal of Advances in Modeling Earth Systems, 11, 3419–3442, https://doi.org/10.1029/2019MS001743, 2019.

Nuijens, L., Serikov, I., Hirsch, L., Lonitz, K., and Stevens, B.: The Distribution and Variability of Low-Level Cloud in the North Atlantic
        Trades, Quarterly Journal of the Royal Meteorological Society, 140, 2364–2374, https://doi.org/10.1002/qj.2307, 2014.

575   Rennie, M. P., Isaksen, L., Weiler, F., de Kloe, J., Kanitz, T., and Reitebuch, O.: The Impact of Aeolus Wind Retrievals on ECMWF Global
        Weather Forecasts, Quarterly Journal of the Royal Meteorological Society, n/a, https://doi.org/10.1002/qj.4142, 2021.

Riehl, H., Yeh, T. C., Malkus, J. S., and la Seur, N. E.: The North-East Trade of the Pacific Ocean, Quarterly Journal of the Royal Meteoro-
        logical Society, 77, 598–626, https://doi.org/10.1002/qj.49707733405, 1951.

Risien, C. M. and Chelton, D. B.: A Global Climatology of Surface Wind and Wind Stress Fields from Eight Years of QuikSCAT Scatterom-
580   eter Data, Journal of Physical Oceanography, 38, 2379–2413, https://doi.org/10.1175/2008JPO3881.1, 2008.

Sandu, I., Beljaars, A., Bechtold, P., Mauritsen, T., and Balsamo, G.: 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,
        https://doi.org/10.1002/jame.20013, 2013.

Sandu, I., Bechtold, P., Nuijens, L., Beljaars, A., and Brown, A.: On the Causes of Systematic Forecast Biases in Near-Surface Wind Direction
585   over the Oceans, ECMWF Technical Memorandum, 866, https://doi.org/10.21957/wggbl43u, 2020.

Stephan, C. C., Schnitt, S., Schulz, H., Bellenger, H., de Szoeke, S. P., Acquistapace, C., Baier, K., Dauhut, T., Laxenaire, R., Morfa-Avalos, Y., Person, R., Quiñones Meléndez, E., Bagheri, G., Böck, T., Daley, A., Güttler, J., Helfer, K. C., Los, S. A., Neuberger, A., Röttenbacher, J., Raeke, A., Ringel, M., Ritschel, M., Sadoulet, P., Schirmacher, I., Stolla, M. K., Wright, E., Charpentier, B., Doerenbecher, A., Wilson, R., Jansen, F., Kinne, S., Reverdin, G., Speich, S., Bony, S., and Stevens, B.: Ship- and Island-Based Atmospheric Soundings from the 2020 EUREC⁴A Field Campaign, Earth System Science Data, 13, 491–514, https://doi.org/10.5194/essd-13-491-2021, 2021.

Stevens, B., Farrell, D., Hirsch, L., Jansen, F., Nuijens, L., Serikov, I., Brügmann, B., Forde, M., Linne, H., Lonitz, K., and Prospero, J. M.: The Barbados Cloud Observatory: Anchoring Investigations of Clouds and Circulation on the Edge of the ITCZ, Bulletin of the American Meteorological Society, 97, 787–801, https://doi.org/10.1175/BAMS-D-14-00247.1, 2016.

Stevens, B., Bony, S., Farrell, D., Ament, F., Blyth, A., Fairall, C., Karstensen, J., Quinn, P. K., Speich, S., Acquistapace, C., Aemisegger, F., Albright, A. L., Bellenger, H., Bodenschatz, E., Caesar, K.-A., Chewitt-Lucas, R., de Boer, G., Delanoë, J., Denby, L., Ewald, F., Fildier, B., Forde, M., George, G., Gross, S., Hagen, M., Hausold, A., Heywood, K. J., Hirsch, L., Jacob, M., Jansen, F., Kinne, S., Klocke, D., Kölling, T., Konow, H., Lothon, M., Mohr, W., Naumann, A. K., Nuijens, L., Olivier, L., Pincus, R., Pöhlker, M., Reverdin, G., Roberts, G., Schnitt, S., Schulz, H., Siebesma, A. P., Stephan, C. C., Sullivan, P., Touzé-Peiffer, L., Vial, J., Vogel, R., Zuidema, P., Alexander, N., Alves, L., Arixi, S., Asmath, H., Bagheri, G., Baier, K., Bailey, A., Baranowski, D., Baron, A., Barrau, S., Barrett, P. A., Batier, F., Behrendt, A., Bendinger, A., Beucher, F., Bigorre, S., Blades, E., Blossey, P., Bock, O., Böing, S., Bosser, P., Bourras, D., Bouruet-Aubertot, P., Bower, K., Branellec, P., Branger, H., Brennek, M., Brewer, A., Brilouet, P.-E., Brügmann, B., Buehler, S. A., Burke, E., Burton, R., Calmer, R., Canonici, J.-C., Carton, X., Cato Jr., G., Charles, J. A., Chazette, P., Chen, Y., Chilinski, M. T., Choularton, T., Chuang, P., Clarke, S., Coe, H., Cornet, C., Coutris, P., Couvreux, F., Crewell, S., Cronin, T., Cui, Z., Cuypers, Y., Daley, A., Damerell, G. M., Dauhut, T., Deneke, H., Desbios, J.-P., Dörner, S., Donner, S., Douet, V., Drushka, K., Dütsch, M., Ehrlich, A., Emanuel, K., Emmanouilidis, A., Etienne, J.-C., Etienne-Leblanc, S., Faure, G., Feingold, G., Ferrero, L., Fix, A., Flamant, C., Flatau, P. J., Foltz, G. R., Forster, L., Furtuna, I., Gadian, A., Galewsky, J., Gallagher, M., Gallimore, P., Gaston, C., Gentemann, C., Geyskens, N., Giez, A., Gollop, J., Gouirand, I., Gourbeyre, C., de Graaf, D., de Groot, G. E., Grosz, R., Güttler, J., Gutleben, M., Hall, K., Harris, G., Helfer, K. C., Henze, D., Herbert, C., Holanda, B., Ibanez-Landeta, A., Intrieri, J., Iyer, S., Julien, F., Kalesse, H., Kazil, J., Kellman, A., Kidane, A. T., Kirchner, U., Klingebiel, M., Körner, M., Kremper, L. A., Kretzschmar, J., Krüger, O., Kumala, W., Kurz, A., L'Hégaret, P., Labaste, M., Lachlan-Cope, T., Laing, A., Landschützer, P., Lang, T., Lange, D., Lange, I., Laplace, C., Lavik, G., Laxenaire, R., Le Bihan, C., Leandro, M., Lefevre, N., Lena, M., Lenschow, D., Li, Q., Lloyd, G., Los, S., Losi, N., Lovell, O., Luneau, C., Makuch, P., Malinowski, S., Manta, G., Marinou, E., Marsden, N., Masson, S., Maury, N., Mayer, B., Mayers-Als, M., Mazel, C., McGeary, W., McWilliams, J. C., Mech, M., Mehlmann, M., Meroni, A. N., Mieslinger, T., Minikin, A., Minnett, P., Möller, G., Morfa Avalos, Y., Muller, C., Musat, I., Napoli, A., Neuberger, A., Noisel, C., Noone, D., Nordsiek, F., Nowak, J. L., Oswald, L., Parker, D. J., Peck, C., Person, R., Philippi, M., Plueddemann, A., Pöhlker, C., Pörtge, V., Pöschl, U., Pologne, L., Posyniak, M., Prange, M., Quiñones Meléndez, E., Radtke, J., Ramage, K., Reimann, J., Renault, L., Reus, K., Reyes, A., Ribbe, J., Ringel, M., Ritschel, M., Rocha, C. B., Rochetin, N., Röttenbacher, J., Rollo, C., Royer, H., Sadoulet, P., Saffin, L., Sandiford, S., Sandu, I., Schäfer, M., Schemann, V., Schirmacher, I., Schlenczek, O., Schmidt, J., Schröder, M., Schwarzenboeck, A., Sealy, A., Senff, C. J., Serikov, I., Shohan, S., Siddle, E., Smirnov, A., Späth, F., Spooner, B., Stolla, M. K., Szkółka, W., de Szoeke, S. P., Tarot, S., Tetoni, E., Thompson, E., Thomson, J., Tomassini, L., Totems, J., Ubele, A. A., Villiger, L., von Arx, J., Wagner, T., Walther, A., Webber, B., Wendisch, M., Whitehall, S., Wiltshire, A., Wing, A. A., Wirth, M., Wiskandt, J., Wolf, K., Worbes, L., Wright, E., Wulfmeyer, V., Young, S., Zhang, C., Zhang, D., Ziemen, F., Zinner, T., and Zöger, M.: EUREC⁴a, Earth System Science Data, 13, 4067–4119, https://doi.org/10.5194/essd-13-4067-2021, 2021.

Stoffelen, A., Pailleux, J., Källén, E., Vaughan, J. M., Isaksen, L., Flamant, P., Wergen, W., Andersson, E., Schyberg, H., Culoma, A., Meynart, R., Endemann, M., and Ingmann, P.: The Atmospheric Dynamics Mission for Global Wind Field Measurement, Bulletin of the American Meteorological Society, 86, 73–88, https://doi.org/10.1175/BAMS-86-1-73, 2005.

Ueyama, R. and Deser, C.: A Climatology of Diurnal and Semidiurnal Surface Wind Variations over the Tropical Pacific Ocean Based on the Tropical Atmosphere Ocean Moored Buoy Array, Journal of Climate, 21, 593–607, https://doi.org/10.1175/JCLI1666.1, 2008.

Vial, J., Vogel, R., Bony, S., Stevens, B., Winker, D. M., Cai, X., Hohenegger, C., Naumann, A. K., and Brogniez, H.: A New Look at the Daily Cycle of Trade Wind Cumuli, Journal of Advances in Modeling Earth Systems, 11, 3148–3166, https://doi.org/10.1029/2019MS001746, 2019.

Witschas, B., Lemmerz, C., Geiß, A., Lux, O., Marksteiner, U., Rahm, S., Reitebuch, O., and Weiler, F.: First Validation of Aeolus Wind Observations by Airborne Doppler Wind Lidar Measurements, Atmospheric Measurement Techniques, 13, 2381–2396, https://doi.org/10.5194/amt-13-2381-2020, 2020.

Wolken-Möhlmann, G., Gottschall, J., and Lange, B.: First Verification Test and Wake Measurement Results Using a SHIP-LIDAR System, Energy Procedia, 53, 146–155, https://doi.org/10.1016/j.egypro.2014.07.223, 2014.
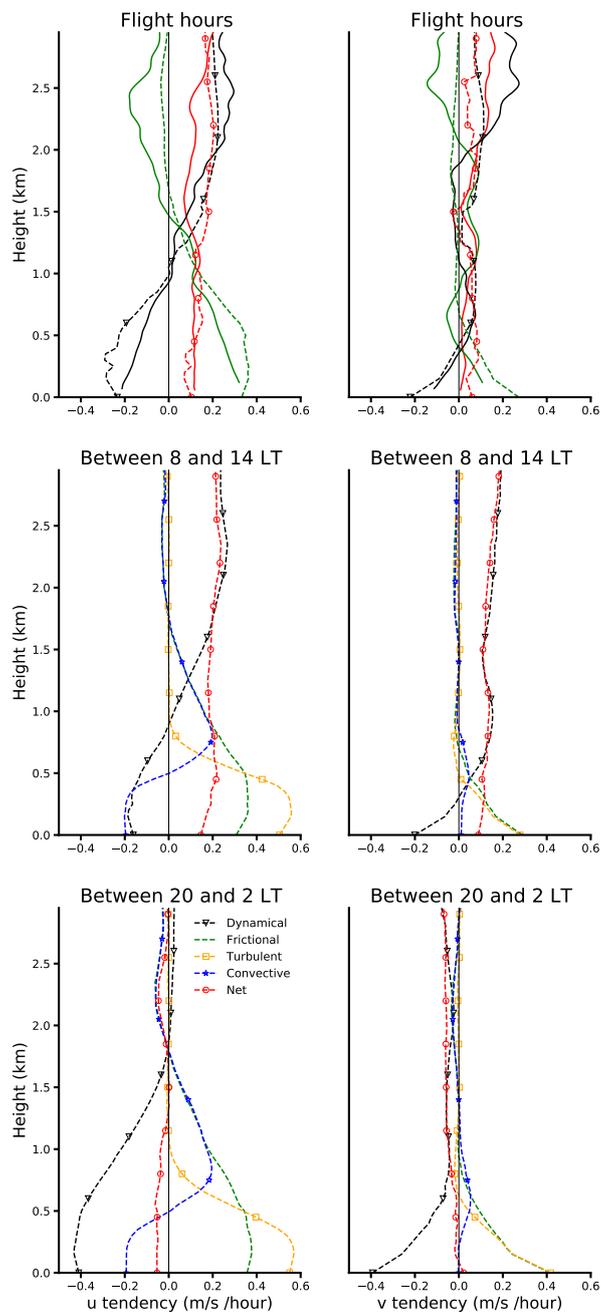
Figure 11. Components of the momentum budget retrieved from the dropsondes (solid) and the forecast (dashed). The net tendency (red circles) balances the dynamical force (black triangles) and the frictional force (green). The latter is split into turbulent and convective for the forecast. The top row refers to days and hours sampled by the dropsondes (flight hours). The middle row refers to the hours between 8 and 14 LT during all EUREC4A days. The bottom row refers to the hours 20 to 2 LT during all EUREC4A days.
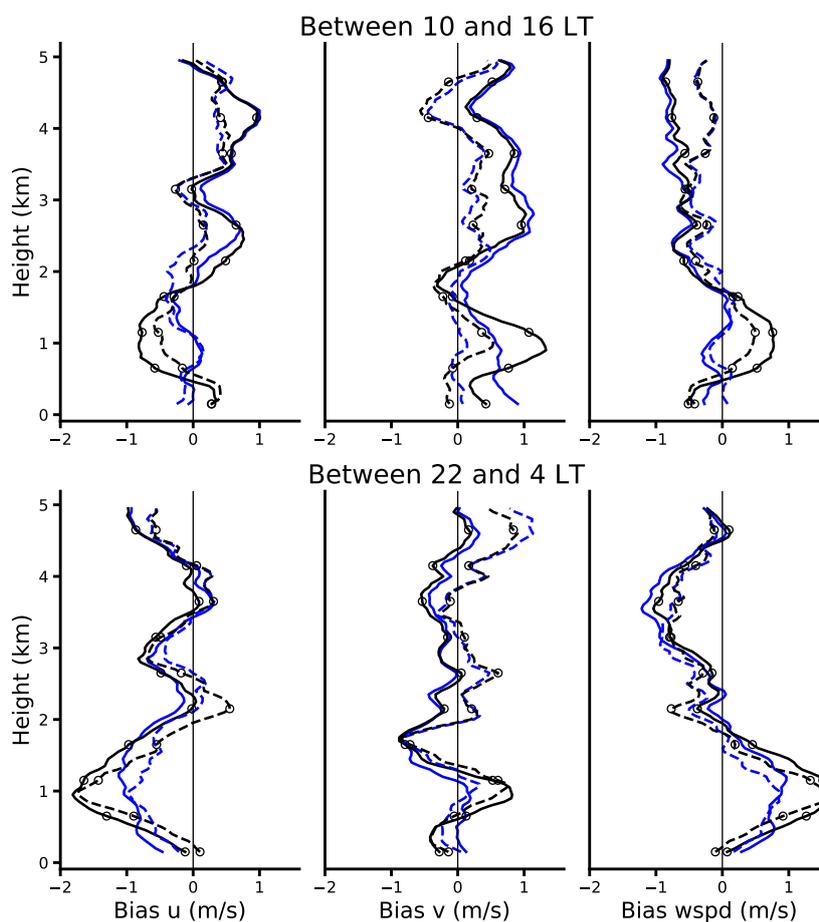
Figure 12. Mean IFS bias (forecast in solid and analysis in dashed) during EUREC4A with respect to radiosondes. Blue refers to the control experiment while black circles to the experiment with convective momentum transport turned off (NoCMT). Top and bottom row respectively for bias between 10 and 16 LT and between 22 and 4 LT.
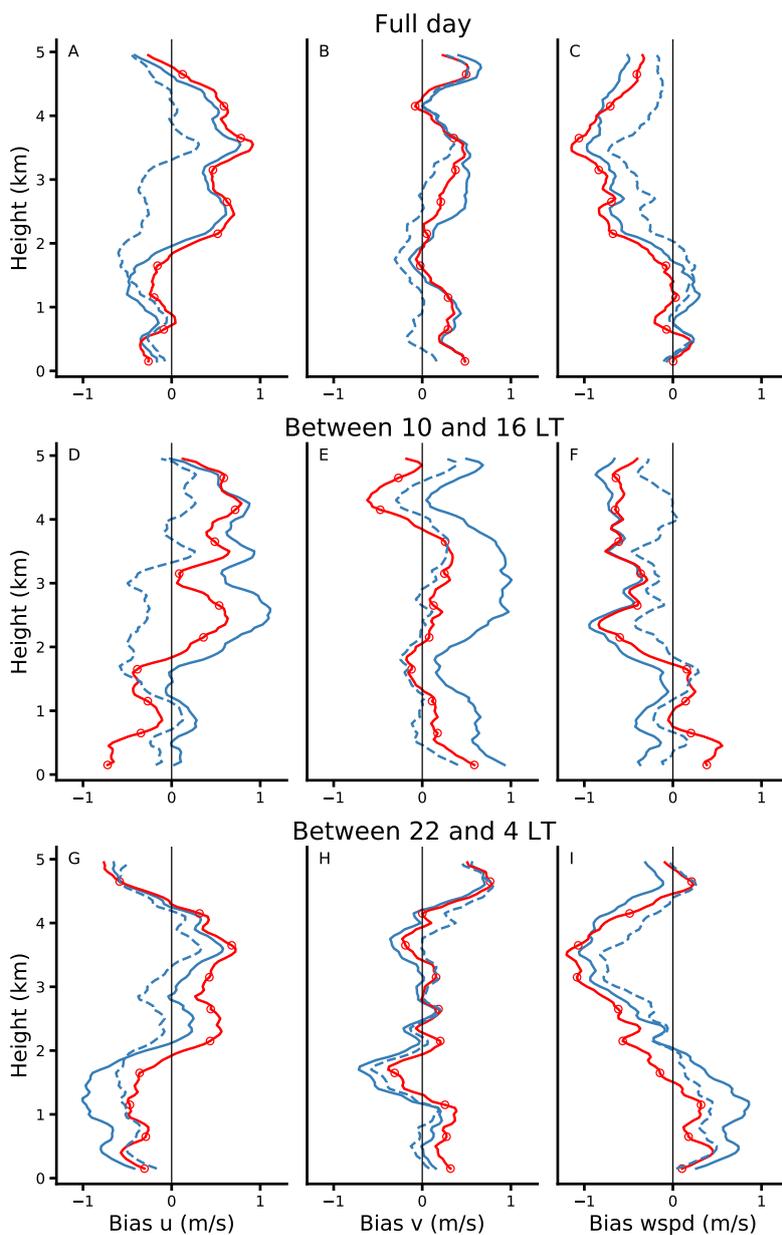
Figure 13. Mean model bias for ERA5 (dashed blue), the operational forecast (solid blue), and a forecast with the new model cycle 47r3 (red circles). The bias is shown separately for all hours of the day (top), for daytime between 10 and 16 LT (middle row), and for nighttime between 22 and 4 LT (bottom row). The bias is calculated with respect to radiosondes. From left to right the columns refer to the bias in the zonal wind (u), meridional wind (v), and wind speed.