Response to referee 1 (RC1)

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

The paper builds on a substantial amount of new data resulting from a large observational effort and combined with relevant numerical experiments. The addressed topic is of importance and the novel material will help identify and alleviate model deficiencies. However, the data analysis is too superficial, based on disputable methods and not fully supportive of the conclusions. Furthermore, the narrative tends to present fundamental concepts and new ideas during the course of the paper instead of clearly separating aims, methods and results. Thus, substantial work is required to deepen the analysis, strengthen the interpretation and clarify the presentation. A long list of specific comments is given below to improve the paper.

We thank the reviewer for providing constructive comments on our manuscript and we refer to the general response to all reviewers for an overlook of our improvements.
Below we answer to the reviewer's general and specific comments explaining the details of what we have changed to deepen the analysis.

General comments

1. Assessing ERA5 against observations that are assimilated in the reanalysis is questionable. It makes sense to compare the quality of forecasts and analysis to disentangle the origin of errors between initial conditions and model physics but in this case the operational analysis would be more meaningful for consistency.

We understand and share the doubts about the assimilated observations in ERA5, for this we dedicate a session of the manuscript (session 5.2) to the "Influence of sounding assimilation". We also added a reference from a recent paper on the impact of dropsondes in ECMWF IFS analysis (Stipo Sentić et al. 2022).
The use of ERA5 instead of the operational analysis is motivated by the popularity of ERA5 in the literature. We preferred ERA5 also for its hourly resolution compared to the 6-houlry operational analysis that we have available. In the manuscript we also provide information

about the mean bias for the analysis experiments done with a model cycle different from ERA5 (Figure 9).

2. Most of the results are based on mean biases only and do not discuss the statistical distribution of model and observations. This is a clear oversimplification and likely obscures a large part of the actual data content.

We understand the concern and we share the idea that a simplistic analysis might be showing only part of the picture. To strengthen and support our analysis we have edited the manuscript and added Figure 8 which contains information about the statistical distribution of the model error and its spatial variability (the latter addressed in the general comment 3).

3. The spatial variability is not mentioned and the temporal variability is assessed through the mean diurnal cycle only. This does not support the discussion of small-scale convective processes in Section 6.

This is a very valid point. As we stated in our general letter, we deepened the analysis and included a few new figures:

- In Figure 1 we show a global overview of the near surface bias with respect to ASCAT observations.
- In Figure 8 we show the spatial variability, as a statistical distribution for our study area, of the difference between forecast and ERA5.

4. The paper needs restructuring by presenting observations first, then possibly the impact of data assimilation, before discussing the quality of analysis and forecast data.

We adopted your suggestion, see above.

5. Numerous repetitions and inaccuracies make the data and methods often confusing. They must be described once and if relevant only.

We had a critical look at the entire manuscript and revised much text, as you can see from the track changes.

6. Figures tend to be misleading for negative wind components, partly due to the use of the same color bars as for wind speed. Overall, displaying wind speed and direction would be much easier to interpret.

This is true and we have evaluated what is the best solution. Because the model adopts equations for the vector components of the winds, observations usually lend themselves better for analyzing wind as a scalar quantity. There is no right or wrong way of doing this, but we should attempt to make the results as intuitively as possible. We strongly agree, the sign of the wind complicates the presentation. We decided to maintain our presentation of the vector components alongside the wind speed, but we have changed several colors maps to avoid confusion, such as:

- Wind speed is now always presented with a scale of greens (Figure 4, Figure 6, Figure 7).
- The error in wind speed now ranges from green (too weak) to purple (too strong) (Figure 10, Figure 11).
- For meridional and zonal components we have maintained the color maps from blue to red, where blue represent negative values.
- For the error in meridional and zonal components we have not changed the color map but we have flipped it for a more intuitive comparison with wind speed (Figure 10, Figure 11).

Specific comments

l.1 tropical ocean

Added

l. 7 typo: RMSE

Fixed

l. 17–18 Why?

We have improved the sentence. See lines 18-19.

l. 18–24 This comes too early as neither the ocean nor the Tropics have been mentioned yet.

Reorganized

l. 28–31 Does it reach the lower troposphere?

Yes, made it clearer in the text.

l. 32–33 Are these observations not assimilated?

We have specified this. Lines 34-35

l. 39 What is the definition of transient here?

We have clarified it. Lines 41-42

l. 55 “the largest”: ever?

Fixed

l. 68–72 Open questions rather than yes/no?

Fixed

l. 91 EUREC4A already mentioned several times

Removed

l. 93 rather defines the studied domain

Improved

l. 116 What is the “Boulevard des Tourbillons”?

Removed

l. 120 black square

Fixed

l. 121 typo: were

Fixed

l. 149 Why these 61 (arbitrary) points? Fig. 1 suggests that only 1 point is taken from ERA5.

Some text is added to clarify (lines 157-158). Figure 2 is improved. The points are indeed arbitrary and could be chosen differently, we constrained 20 of these points to be on the EUREC4A-circle (where dropsondes were launched) and placed the other pints to have a good spatial coverage of the domain.

l. 154 horizontal resolution (and already stated several times)

Fixed and reorganized

l. 158–160 either give more details or omit day 4

We have removed redundant information about day 4.

l. 163 spell out ERA5?

Fixed

l. 167 with more observations and a longer assimilation time window?

The reanalysis uses more or less the same observations and assimilation window, it is just an analysis made with a consistent version of the IFS for the past 50-70 years. In the operational analysis, the IFS changes about once a year so the analysis for 2022 will be produced with a slightly different system than the one in 2021. In the reanalysis, instead, the same model and data assimilation system is used for the entire period.

l. 168 Remove “for example” if no other reanalysis is used here

Fixed

l. 172 either give more details or omit the operational analysis

We have removed redundant information about the operational analysis and focused on ERA5.

l. 173 ERA5 available hourly

Fixed

l. 176–177 The question is confusing

Rephrased. See lines 183-184

l. 177 Why this resolution? It differs from both reanalysis and forecast

We understand the doubt, nevertheless we believe it is not reason for concern. There are two main reasons for performing experiments at 40 km resolution.

1. Because of the limited computational resources, this is the standard resolution at which research tests are conducted at ECMWF.
2. We see that the spatial resolution of the model has little impact on the results after comparing the operational forecast (9km) and ERA5 (32km) to the control forecast and analysis experiments (40 km). Figure 5 D,E,F compared to Figure 9.

l. 179 What is the need of 10-day forecasts to look at day 2?

The model experiments we use are part of a larger set of experiment. Other studies (some still unpublished) look at different ranges (e.g. Sandu et al. (2020)). We analyzed day-2 and day-4 forecast, although we opt for only showing day-2 because it is easier to disentangle sources of errors at short lead times.

l. 180–181 Are EUREC4A measurements assimilated in the operational analysis and/or reanalysis otherwise? This is a crucial point for the paper!

We agree that this is a crucial information. We have now made it clear in the data section (line 181). Whereas drop- and radio- sondes are assimilated in ERA5 and in the operational analysis, the lidar measurements are not assimilated.
As mentioned in the answer to the general comment 1, Section 5.2 is dedicated to the influence of sounding assimilation and we also added a reference from a recent paper on the impact of dropsondes in ECMWF IFS analysis (Stipo Sentić et al. 2022).

l. 185 how is shallow convective momentum transport accounted for in the model?

It is parameterized with a mass-flux entraining-detraining plume scheme. We recognize the importance of stating the type of parameterization used, so we have added some text with this information. See lines 194-197.

l. 190 did you run model experiments at different resolutions?

We only performed experiments at 40 km horizontal resolution. Please see answer to I.177 for the motivation.

l. 193–194 Why these numbers? What is the vertical resolution and range of the different datasets?

The vertical grid is arbitrary but reasonable compared to the vertical resolution of the different datasets which have now been described in Section 2.
The lower limit is chosen at 150 m to deal with the lower number of drop- and radio-sondes which provided measurements near the surface.
The upper limit is 5 km because above this level there are almost no convective clouds and our study focuses on the lower troposphere, up to the cloud top.

l. 199–214 This paragraph is confusing and largely repeats Section 2. Please clarify and merge.

We have reorganized Section 2 and Section 3 to avoid repetition and rephrased several sentences.

Fig. 3 What is “wspd”?

Fixed, defined at line 217 and in Figure 4

l. 219 mid January to mid February

Fixed

l. 220, 226 Where is the cloud base and top?

We have added indications of the mean sub-cloud layer top and inversion height in the figures and in the text. Their definition is described in the methods (lines 211-214).

l. 235 already stated above

Chapter 4 reorganized to avoid repetitions. See comments to l. 240-242

l. 238 refer to dropsondes rather than JOANNE?

Fixed. Indeed, sometimes it's clearer to refer to dropsondes rather than JOANNE.

l. 240–242 repeats l. 219–220. Again, where are the clouds?

Chapter 4 reorganized to avoid repetitions.

For an indication of the clouds please see comments to l. 220, 226.

l. 242–243 described below

The entire section is reorganized and this sentence rewritten as part of it.

l. 250–257 This belongs to the Methods

Moved to the Methods.

l. 252 instantaneous or daily averaged variable?

Instantaneous. Made it clearer in the text (please see lines 203-204).

l. 260–261 Does it mean the wind direction is wrong?

Yes, when the zonal and meridional components have biases of the opposite sign and similar in magnitude, the bias is reflected on the wind direction rather than on the wind speed. We have added lines 275-277 in a completely revised paragraph.

l. 267–268 show the error distribution?

We have added Figure 8 which shows the error distribution.

l. 272–273 It rather suggests the opposite!

We have removed this speculative sentence.

l. 276 The title does not seem appropriate as nothing is said about the synoptic situation

We have largely reorganized the text and the sections. We now improved the description of the synoptic wind state in section 4.1: Wind profile and synoptic variability.

l. 278–279 One would expect observations first, possibly combining the different sources

As mentioned for the general comment 4, we now start presenting the results from observations (Section 4: Observed Winds during EUREC4A). We have replaced the old Figure 5 with Figure 6 in the revised manuscript, which now shows observational data from radiosondes.

l. 280–281 This is not easy to infer from Fig. 5

Removed

l. 282–284 This statement is subjective. For instance, a larger bias in wind speed around 02-08 appears related to the larger wind speed at that time.

We have removed the subjective statement.

l. 288–289 This is speculative

Removed

l. 296 typo: composites

Fixed

l. 309–311 similar to l. 278–279: why not show observations?

We agree that observations should be used where possible. Differently from Figure 5, the 3-hourly resolution of the radiosondes is not enough to produce a smooth figure of the diurnal cycle. The lidar measurements, instead, do not reach 5km height.
As done in figure 8 (D,E,F), here in Figure 7 (D,E,F) we use ERA5 as our best guess, without overinterpreting it.

l. 318–324 move to section 6.1

Done

l. 330 This is unclear as both u and v show positive bias during daytime

We have corrected the sentence, the statement is valid only below 2km. Please see lines 334-336.

l. 331–333 This sounds speculative

Removed

l. 336 The phase shift was not mentioned before

We changed the text and made sure the phase shift in the forecast is mentioned. See lines 321-324.

l. 337 It is surprising this is not discussed earlier. Again, where is the cloud layer?

We now anticipate this in section 4.1 when presenting the observed winds. We also include an indication of the cloud layer as explained for the comment to l. 220, 226.

l. 344–345 This contradicts l. 334 above

We have merged and rephrased the sentences (see lines 340-343)

l. 350–351 This fundamental information (1) must be stated earlier and (2) questions the soundness of the above assessment of ERA5 winds

We agree. Please see comment to l.180-181.

l. 360–362 What about RMSE, diurnal cycle, etc. discussed in the previous section?

We have included the RMSE in Figure 9 for completeness and comparison with Figure 5. We have improved the text concerning the analysis of Figure 9. Nevertheless, we opted for not showing in the text the diurnal cycle and the statistical distribution of the error for all sensitivity experiments. We believe that our choice keeps the narrative more fluent without omitting information that would change the analysis. The reviewer is invited to see the figures below in this document.
Figure A1 shows that the diurnally of the bias is present also in the denial experiments. Removing drop- and radio- sondes has only brings a slight deterioration during daytime on the analysis, as one would expect given the temporal availability of the dropsondes.

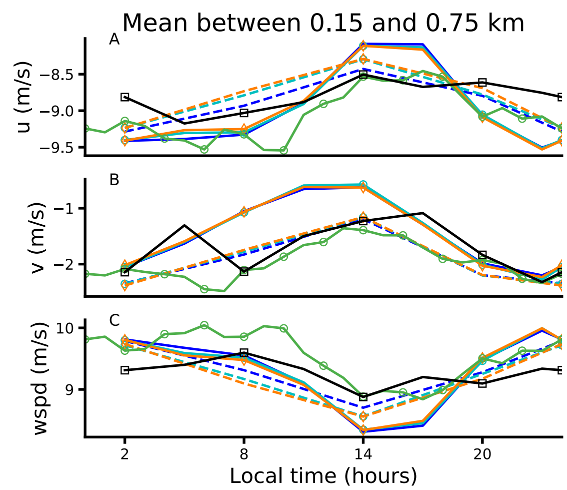

Figure A 1 Diurnal cycle in the layer between 0.15 km and 0.75 km of zonal wind (A), meridional wind (B) and wind speed (C). Black for radiosondes, green for lidar. Dashed for the analysis, solid for the forecasts. Blue: CTRL, cyan: Exp1, orange: Exp2.

Figure A2 shows that the statistical distribution of the errors in the experiment without drop- nor radio- sondes is similar to the one seen in Figure 8 for the high-resolution forecast.

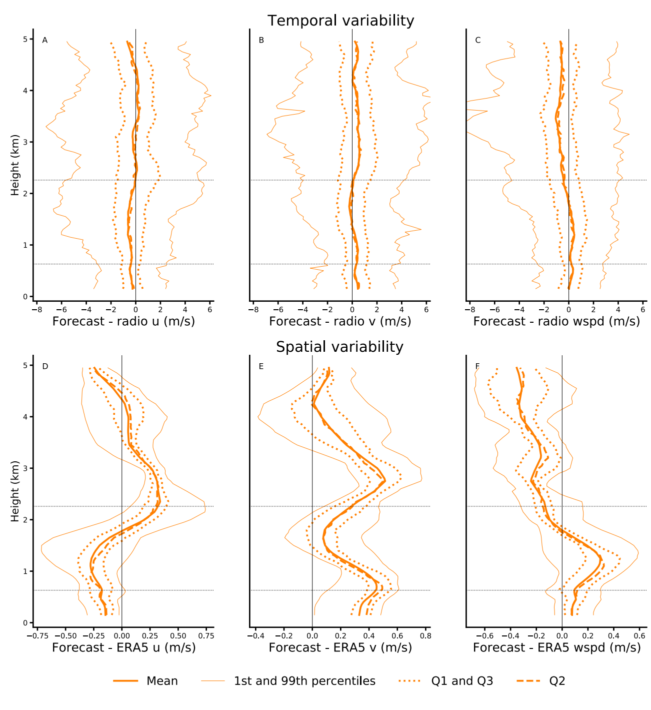

Figure A 2 Statistical temporal distribution of the error for Exp2_fc with respect to radiosondes (A,B,C), and the statistical spatial distribution of the difference between Exp2_fc and ERA5 (D, E, F).

l. 375 Modeled versus observed?

Edited

l. 376–384 More details on the methods are needed here. A paper “under preparation” is not of much use.

References to this unpublished work have been significantly reduced and the text revised to make it independent. The mentioned paper is under (minor) revision and expected to be accepted before this manuscript.

l. 386 Where is the boundary layer?

Please see answer to l. 220, 226.

l. 389 Where is the reduction in the large-scale pressure gradient to be seen?

We now, more correctly, refer to a reduction in the large-scale dynamical forcing. See lines 380-381.

l. 391 Which forcing?

The dynamical and frictional forcings. Rephrased for clarity.

l. 397 slow bias

Fixed

l. 406–408 the names are not consistent with the legend of Fig. 12 (NoCMT)

Fixed

l. 413–414 the mean bias in the zonal wind is strongest at night

We have made this clearer. Please see lines 411-412.

l. 418–426 The paragraph appears to mix shallow and deep convective momentum transport

The section has been majorly revised.

l. 429 “what we have called”: is it not the actual operational forecast?

Made it more clear using cycle names. Cycle 47r2 was operational at the time of the field campaign, while cycle 47r3 became operational on the 12th of October 2021.

l. 431–433 please be more explicit

We have added a reference (Bechtold et al., 2020) for details about the moist physics upgrade in IFS cy47r3. We have also revised the entire section.

l. 433–436 Does it matter here?

This section has been entirely rewritten to remove unsubstantiated statements.

l. 438 why the subtropics?

The entire sentence is removed.

l. 443–444 Repetition of l. 437–438

Fixed removing a sentence (see comment to l. 438).

l. 444–445 The clear improvement in u and v barely affects the wind speed. This suggests that the mean bias discussed here is only part of the picture.

To understand why an improvement in u and v does not necessary implies an improvement in the wind speed we bring the attention to our methods. The mean wind speed is not computed from the mean zonal and mean meridional components, instead it is computed independently as a mean of the wind speed at each time. This is especially relevant when u and v can have alternating sign, As an example, we show here in Figure A3 the statistical distribution of the wind components at 3 km for the forecast.

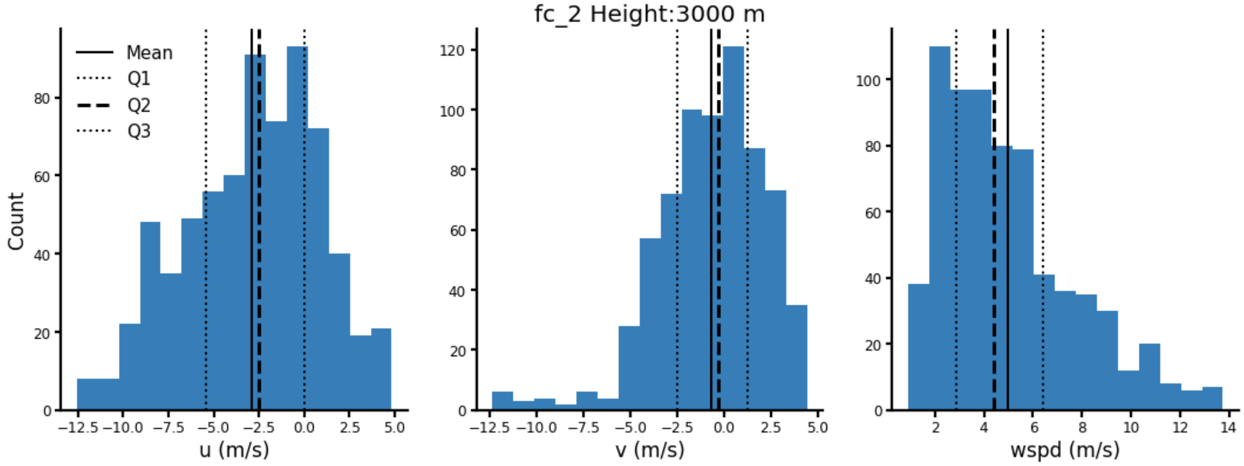

Figure A 3: Statistical distribution of the wind components at 3 km for the forecast cy 47r2.

In the table below we report a hypothetical case where an improvement in the forecasted u and v component does not lead to any improvement in the wind speed. This can occur frequently at heights where the mean meridional wind is close to 0 m/s. For example, at 3 km the v is almost half of the times positive and half of the times negative in the forecast (See Figure A3).

	u (m/s)	v (m/s)	Wind speed (m/s)
Observations	- 4	- 1	4.123
Cy47r2	- 3	+ 1	3.16
Bias in cy47t2	+ 1	+ 2	- 0.963
Cy47r3	- 3.16	0	3.16
Bias in cy47t3	+ 0.84	+ 1	- 0.963

l. 445–448 This appears speculative and contrasts with the above discussion of shallow convective transport in the lower troposphere

We agree this section has been entirely rewritten to remove unsubstantiated statements. See answer to "l. 433-336".

l. 470–471 not explicitly shown in the paper

Please see comment to l.336

l. 474–475 smaller than what?

Fixed

l. 476–477 not sufficiently supported

We agree, please see the revised text.

l. 483–484 reference?

Added: Helfer et al. (2021)

Please label all panels and refer to A, B, C, … in the text

Fixed

Displaying all figures at the end of the paper would facilitate the review process

Fixed

Response to referee 2 (RC2)

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

The authors have done a great job at exploiting the very rich EURECA4 dataset and designing a clever model-observation intercomparison strategy that is worthy of publication by itself. Some of the sensitivity experiments are relevant for the work (impact of data assimilation, role of CMT) while others, i.e. pertaining to the new moist physics, leaves the reader with the impression that the paper may be trying to cover too much ground overall.

Overall, I consider that the paper should be published in fine, provided that the authors consider taking into account the suggestions and general comments below.

We thank the reviewer for providing constructive comments on our manuscript and we refer to the general response to all reviewers for an overlook of our improvements.
Below we answer to the reviewer's general comments explaining the details of what we have changed to deepen the analysis.

General comments

1. The overall quality of the English could be improved by reducing lengthy sentences and repetitions in many parts of the MS. Please try to trim the text to improve the readability. A lot of expressions used are not plain English and should be reworded…

We fully agree on the importance of a clear and readable manuscript. We had a critical look at the entire manuscript and revised much text. Please see the general response.

2. Additional near-surface wind measurements were conducted during EUREC4A, such as ship-borne kite-based observations, or even saildrone measurements… Did you consider including these in your analysis? And if so, why did not you use them in the end?

Rather than focusing on the near surface winds or a single height, the interest of our study is on profiles for the lower 5 km of the atmosphere. Furthermore, not all datasets from EUREC4A provide sufficient spatial and temporal coverage throughout the campaign. For these reasons we believe that saildrone and kite measurements would not find the right place in our study.

3. Can the authors elaborate on the generalization of their results in the abstract and conclusion? And to what extent their results can be representative of other regions of the world?

We have edited the manuscript and the abstract to address this comment. In particular, we have added a figure showing the global surface wind bias in the forecast with respect to ASCAT observations. The conclusion have been majorly rewritten.

4. Why haven't you looked at the vertical component of wind as well? This should be feasible using the drosponde data released from the DLR Falcon 20 while it was performing the circles east of Barbados, using a strategy designed to observationally assess large scale vertical motion in the domain covered by the circle.

We agree and share the interest for the vertical component of the winds. This can indeed be retrieved from the dropsondes thanks to their launch strategy. An in-depth analysis of the vertical motion during EUREC4A is beyond the scope of this study and is being conducted in a separate study by some of the authors.

5. Some of the conclusions drawn in the paper, especially the ones pertaining to model physics are based on comparisons of momentum budget modelled and derived from observations. The methodology for the latter is described in a paper that is under preparation, and hence not yet published and not citable… The authors should describe here what is behind the observed tendencies as derived in Nuijens et al., 2021.

Conclusion based on this unpublished work have been significantly reduced and the text revised. The mentioned paper is under (minor) revision and expected to be accepted before this manuscript. Appropriate citation will thus be available by the time of publication.

6. The authors discuss the impact of the so-called 'cumulus friction effect' on the modelled wind profiles systematic errors. Can you elaborate on the role of oceanic waves in shaping the lower tropospheric wind profile over the ocean? How is this accounted for in ECMWF' IFS? Is the reported near-surface excessive easterly flow of the IFS (Belmonte Rivas and Stoffelen, 2019) due to friction induced by the sea state?

In the IFS the uncertainty due to the drag parametrization over ocean is marginal at wind speeds typical of the trades. Belmonte Rivas and Stoffelen (2019) already associate the excessive easterly flow to convective processes. The experiments without convective momentum transport (CMT) in our study show that convection acts as a connector between the upper layers and the surface bringing excessive easterlies to the surface. When CMT is turned off the surface bias is almost entirely corrected, thus we argue that the bias is not related to the surface state.

7. The rationale behind the sensitivity experiments pertaining to the 'new moist physics' in the paper is no so clear and the conclusion that you draw from them regarding the role of tropical convection on the winds in the Barbados region is over-interpreted… I would suggest to remove it or make it much more consistent than it is now.

We have critically analyses the section pertaining cycle 47r3 and it's moist physics. We believe that omitting this would be dropping important information as it concerns the most recent operational IFS cycle. Nevertheless, we understand the concern regarding its cohesion with the rest of the manuscript and it interpretation. For this reason, we have largely rewritten the section "New moist physics" and adjusted it interpretation.

8. The authors discuss the impact of large scale dynamical forcing on the wind profiles in the vicinity of Barbados… How are tropical atmospheric waves accounted for in ECMWF' IFS? Can their interaction with the mean flow impact the wind profile in the lower troposphere? Could they contribute to the systematic errors in the IFS (re)analyses and forecasts?

We find this comment very interesting, we agree that gravity waves can play a role in the momentum budget above 3 km, whereas within the boundary layer momentum diffusion is too strong for them to play a significant role.
The extent to which gravity waves impact the winds in the troposphere is currently being investigated using simulations by a colleague at MPI-M (Claudia Stephan).
Before the output of the simulations are ready, we cannot say if the IFS is inaccurate in representing the effect of gravity waves on winds above 3 km.