Response to Referee Comment (RC2) on

Investigation of dynamical scenarios leading to particularly high impact of Aeolus on NWP forecasts

https://doi.org/10.5194/egusphere-2022-1150

We are grateful for the referee's careful reading and the detailed and insightful discussion on our manuscript. The responses to the individual comments and the corresponding changes in the manuscript are presented in the following.

General Comment:

This is a very nice study which aims to link episodes of enhanced forecast impact from Aeolus to particular dynamical features or changes in circulation pattern. This complements standard overall forecast impact evaluations of Aeolus which report on average impact. As for any feature-based forecast impact evaluations conducted over relatively short periods (3 months in this case), I have some reservations about the robustness of the findings given the relatively low sample size for the cases evaluated. However, the authors present their results in the context of clear positive overall impact and present them primarily as illustrative examples of potential mechanisms, rather than firmly established links. The authors are also sufficiently frank about the sampling limitations of their study and state that longer experimentation would be needed to firm up conclusions. I think the paper can therefore be published after addressing the comments below in a minor revision

Specific Comments:

Comment #1:

Title: As some of the causal relationships remain speculative, I feel a more "neutral" title may be more appropriate. Something along the lines of "Investigation of links between dynamical scenarios and particularly high impact of Aeolus on NWP forecasts"

Response to Comment#1:

We see the point that the title should convey a more "neutral" message and find your suggestion very good. Therefore, we have changed the title accordingly:

"Investigation of links between dynamical scenarios and particularly high impact of Aeolus on NWP forecasts"

Comment #2:

Section 2.1: I had expected to read a sentence or two about the biases that have been observed in Aeolus data in this section. They (or at least their corrections) are mentioned later (around line 125), but I feel some mention here would be worthwhile. They are relevant in the context of the bias impact presented later.

Response to Comment #2:

Thank you very much for your comment. Indeed, some information about the Aeolus bias would be useful in Section 2.1.. Therefore, we have added the following sentences at the end of the description of the Aeolus HLOS observations:

The Aeolus data processing of the OSE period includes a NRT bias correction method. During the first part of the mission, validation studies showed large systematic differences, which vary seasonally, spatially and with orbital phase - particularly pronounced for the Rayleigh wind observations (e.g., Martin et al., 2021a; Rennie et al., 2021). These detected bias dependencies were found to be related to long-wave and solar radiation fluctuations and the radiative response to which the spectrometers of the DWL are very sensitive. The operationally implemented bias correction is based on a multiple linear regression method of ECMWF O-B statistics and the thermistors of the telescope M1-mirror, eliminating most part of the bias (Weiler et al., 2021).

Weiler, F., Rennie, M., Kanitz, T., Isaksen, L., Checa, E., de Kloe, J., Okunde, N., and Reitebuch, O.: Correction of wind bias for the lidar on board Aeolus using telescope temperatures, Atmospheric Measurement Techniques, 14, 7167–7185, https://doi.org/10.5194/amt-14-7167-2021, 2021.

To avoid duplications, we changed the part around line 125 in the end of Section 2.2.1 accordingly:

The OSE for this study covers the Northern Hemisphere summer, July 2020 - October 2020. Although the operationally implemented telescope primary mirror M1 bias correction is very effective, the DWD system still shows a small residual bias that depends on altitude for the Rayleigh wind observations. Therefore, a model-based bias correction scheme has been applied…

<u>*Comment #3:*</u>

L87-88 ("This combination … than pure 3D-Var could (?).": I don't quite understand what is meant with this sentence, in particular what is meant by a "stable analysis" and "incorporating more information from the observations". I would think the main aspect is that the flow-dependent background error covariance will allow a better weighting of observations and the background.

<u>Response to Comment #3:</u>

We understand that this sentence needs further explanation. We like your suggestion and decided to focus more precisely on the aspect of the advantages of the flow-dependent background error covariance of the LETKF. Therefore, we have replaced the sentence as follows:

The core module of the global data assimilation system is a Hybrid Variational Ensemble Kalman Filter (VarEnKF) combining the flow-dependent background error covariance matrix from a Local Ensemble Transform Kalman Filter (LETKF) with the static covariance matrix from the 3-dimensional Variational Data Assimilation system (3D-Var) (Rhodin et al., 2017). This combination allows the inclusion of the time-varying background error structures and thus a better weighting of observations and background.

<u>*Comment #4:*</u>

Related to the point above: Here or elsewhere the authors might want to address to what extent their results are affected using a 3D-Var-based assimilation system, rather than, for instance, a 4D-Var system or equivalent. I would expect that a 4D-Var system would be more able to derive wind information from satellite radiances (through the tracing effect). I would expect this to affect to what extent the system without Aeolus data is able to correct some of the error features described in section 4.

Response to Comment #4:

We understand the point that better adjusting winds with mass information in 4D-Var than in 3D-Var may play a role. Nevertheless, we think that directly comparable OSEs and comparable sensitivity studies from different centers would be required to draw robust conclusions on this aspect as currently available OSEs and DA systems differ in many respects besides the underlying DA method. Thus, we prefer not to add speculative comments on this and hope that dedicated future studies shed some light on this interesting aspect.

Comment #5:

Section 2.2: The reader is referred to Martin et al (2022) for some important details about the experiment configuration. Unfortunately, the Martin et al (2022) reference is incomplete in the reference list, and I was unable to look up these details. I suggest including some more details in the present paper (in particular, a brief outline of what other observations are included in the CTRL and what values are assigned for observation errors for Aeolus).

Response to Comment #5:

Unfortunately the study Martin et al. 2022 is still under review. We've now added the link to the submitted manuscript with more detailed information about the experiment configuration in the reference. Furthermore, we've added information on the percentage of operationally assimilated observations at DWD:

As described in Martin et al. (2022), the control run (CTRL) was performed without Aeolus but with all other operationally used observation types assimilated. The observations assimilated operationally are mainly radiances that account for ~64% of the total observations. Winds from scatterometers, satellite imagery, and GNSS signals together constitute about 18%. Conventional observations from aircraft reports, radiosondes, surface stations, buoys, pilot and wind profiler represent ~7% of the total number of observations. The proportion of assimilated wind profiles from the spaceborne lidar of the Aeolus mission is about 2% (~20.500 HLOS wind observations per assimilation cycle).

Martin, A., Weissmann, M., and Cress, A.: Impact of assimilating Aeolus observations in the global model ICON: A global statistical overview, https://ucloud.univie.ac.at/index.php/s/DIP0XxSgKA6eXek, 2022. ------------------------------
----

Comment #6:

L126: A reference regarding the M1-mirror bias correction method should be added.

Response to Comment #6:

We kindly refer to the response to Comment #2. The reference **Weiler et al. 2021** has been added regarding the M1-mirror bias correction method.

Comment #7:

Section 2.2.2: A reference for ERA5 should be added. It would also be useful to discuss the quality of ERA5, particularly for the tropical stratospheric wind field, since it is used as verification reference. Given the lack of direct wind observations in this area, might ERA5 data be prone to biases in these regions?

Response to Comment #7:

We have added a few sentences about the quality of the ERA5 data in section 2.2.2 to explain in more detail why we think ERA5 is a good verification dataset, but also to provide more information about the uncertainties in the stratosphere. Furthermore, two references have been added.

The ERA5 output is produced using the 4D-Var data assimilation of the ECMWF IFS at a horizontal resolution of 31 km and with 137 vertical model levels up to the height of 80 km (from 1000 hPa to 1 hPa, with 40 levels below 5 km) (Hersbach et al., 2020). As ERA5 reanalyses are based on a different model with different resolution compared to the OSE, they provide a relatively independent data source. The higher vertical resolution of ERA5 allows finer details of atmospheric phenomena to be resolved, such as a more realistic representation of atmospheric waves and their interaction with the mean flow, which is especially crucial for the study of QBO in Sec. 4.1. Furthermore, ERA5 assimilates a partly different set of observations than the global data assimilation system in the ICON model (e.g. more satellite radiances) and does not use the Aeolus observations. It is well known that NWP models in the stratosphere are typically subject to large uncertainties. ERA5 was found to have a cold bias in the lower stratosphere and a warm bias near the stratopause (Hersbach et al., 2020). However, the increased number of assimilated GPSRO bending angles in ERA5 since 2006 has significantly reduced this model bias, increasing confidence in using the stratospheric reanalyses for verification (Laloyaux et al., 2020).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., 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., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Laloyaux, P., Bonavita, M., Dahoui, M., Farnan, J., Healy, S., Hólm, E., and Lang, S. T. K.: Towards an unbiased stratospheric analysis, Quarterly Journal of the Royal Meteorological Society, 146, 2392–2409, https://doi.org/10.1002/qj.3798, 2020.

In Section 4.1, we have modified the second paragraph to point out that the results may be affected by model biases in the region:

It should be taken into account that both the ERA5 reanalysis used for verification and the global model ICON exhibit large uncertainties in the tropical stratosphere, probably contributing to the pronounced impact of Aeolus observations. The QBO is mainly driven by a combination of upward-propagating low-frequency equatorial waves and inertia-gravity waves from the troposphere that dissipate and deposit momentum to the upper level zonal-mean zonal winds (Shepherd et al.,2018). However, a realistic representation of the wave, mean-flow interaction behind the QBO is typically limited by insufficient vertical model resolution, uncertainties in parameterized processes such as tropical convection, and the sparseness of direct wind measurements in the tropics. Given the lack of direct wind observations in the area, the models there could be prone to biases.

Shepherd, T., Polichtchouk, I., Hogan, R., and Simmons, A.: Report on Stratosphere Task Force, ECMWF Technical Memoranda, https://doi.org/10.21957/0vkp0t1xx, 2018

--

Comment #8:

L156 ("… 120 h are reduced by 2 up to almost 5% on average…"): The "2" seems grammatically out of place. At the same time, 2% appears a more typical number to quote here, given the values in the table.

Response to Comment #8:

Thank you for spotting this error. We've added the missing "**%**" after the number:

Within the tropical band, the forecast errors of forecast lead times from 24 to 120 h are reduced by 2% up to almost 5% on average (Table 1).

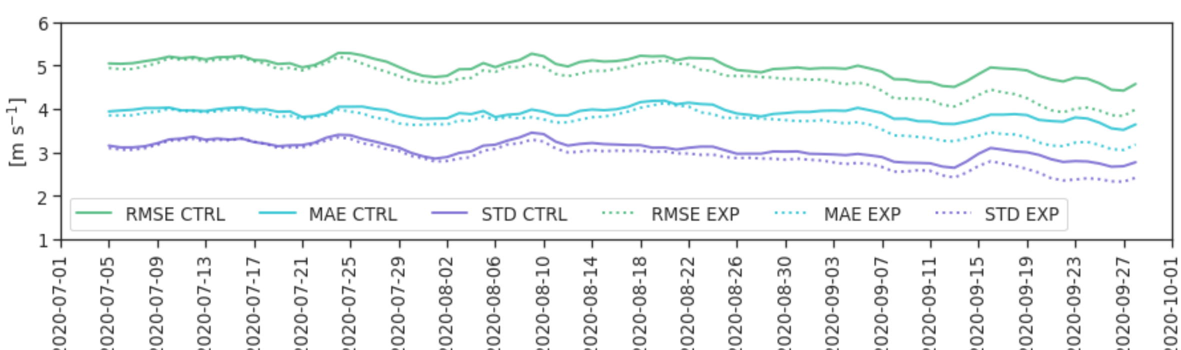
--

Comment #9:

Fig. 2 and accompanying discussion: It appears that the improvement in RMSE for the 2nd half of the period shown is primarily due to a reduction in the mean error. This could be spelled out more clearly in the text. I would be curious to know whether there is also a reduction in the standard deviation of the forecast error between the two experiments. Also, since the Aeolus data have been bias-corrected towards the background (L127), I wonder how the change in the mean wind comes about. Do Aeolus departures indicate this bias in the background? Are wind biases in ERA5 sufficiently small for the tropical stratosphere to be certain that this mean change is indeed an improvement?

Response to Comment #9:

The figure below shows the 72-hour RMSE and its error components (mean absolute error MAE and standard deviation STD), with the dotted lines representing the error metrics of the Aeolus assimilation run. It can be seen that the standard deviation and mean error are reduced almost equally in the second half of the period due to Aeolus. We appreciate your comment on this issue, but we think the interpretation of the two error components is non-trivial and depends on the definition. The equatorial stratosphere is dominated by a large-scale slow phenomenon, the QBO phase change. On smaller timescales, such as those considered here, we believe that the reduction in MAE is due to the observations of Aeolus rather than bias. Bias correction should play a minor role here, since the underlying lookup table is based on a preceding period and larger areas. Furthermore, the analysis bias is corrected, which differs from the forecast bias.



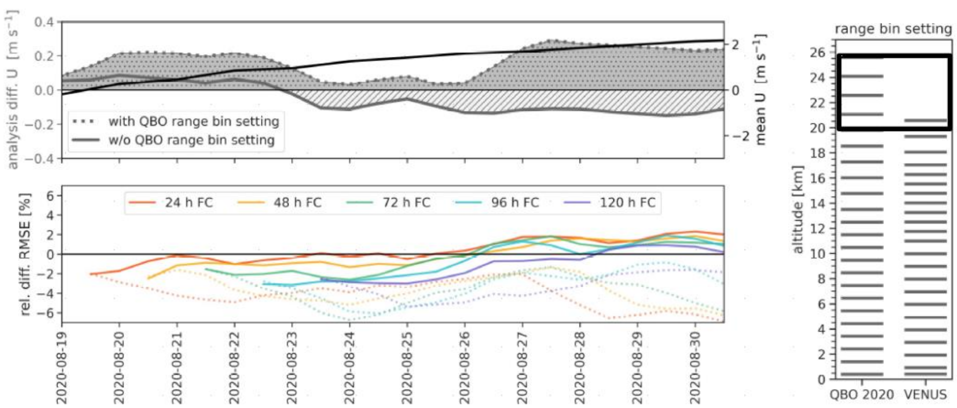
--

Comment #10:

Fig. 3: It would be helpful for the reader to mark the typical location of the 30-50 hPa layer in the plot of the range-bin-settings. I suspect it will be largely in the range of 20-26 km, and this information should help the associated discussion. Do the authors think that the primary mechanism for the improvement shown for 30-50 hPa for the "with QBO range bin settings" version is that this setting covers the layer in question, rather than, say, vertical resolution considerations?

Response to Comment #10:

We appreciate your suggestion and added the 30-50 hPa layer in the plot of the range-bin-settings. We don't think that the primary mechanism for the improvement is necessarily the Aeolus measurements directly in the 30-50 hPa. Instead, this may also be related to the higher resolution in the upper troposphere, that allows for finer details of atmospheric phenomena and better representation of vertically propagating waves in the upper troposphere.



Comment #11:

L224 ("... to get a more reliable statement"): I suggest rephrasing to "... to be able to draw firm conclusions".

Response to Comment #11:

Thank you very much for the good suggestion. We changed the wording acordingly:

As the stratosphere is characterized by large model uncertainties, a longer OSE would be useful to draw robust conclusions.

Comment #12:

L243 ("The largest error reduction coincides with the strongest negative slope in the SST anomaly."): This is true, but similarly large error reductions are also found during the period 15-19 September when the SST anomaly is flat. I am not fully convinced that there is a connection, based on this single case.

Response to Comment #12:

We modified the last sentence of the first paragraph in Section 4.2. to remind the reader that the Figure shows the 48 h RMSE reduction at the forecast and not at initialization time. The largest error reduction is not at the same time point as the largest slope, but about 48 h later. With this information, the relationship between the SST anomaly and the effects of Aeolus should become clearer for the September 15-19 period.

The largest error reduction in the 48 h forecast occurs about 48 h after the strongest negative increase in the SST anomaly, which corresponds to the time of initialization.

Furthermore, we changed the phrase “as function of time and pressure [hPa]” in the caption of Figure 4 accordingly:

as function of forecast time and pressure [hPa]

Comment #13:

L339/340 (“Selz et al (2022) found … the error growth.”): I struggle to follow this sentence.

Response to Comment #13:

Unfortunately, there was an error in the structure of the sentence. Thank you for bringing it to our attention. We have reworded it as follows:

Selz et al. (2022) found that latent heat release in convective systems and the divergent component of the atmospheric flow dominate the error growth with respect to physical processes.