



# ESA Contract Report

Aeolus DISC contract No. 4000126336/18/I-BG

Contract Report to the European Space Agency

## The NWP impact of Aeolus Level-2B winds at ECMWF

WP-5500, Aeolus DISC (Data Innovation Science  
Cluster) Phase E2

Authors: Michael Rennie and Lars Isaksen

January 2024

Series: ECMWF - ESA Contract Report

A full list of ECMWF Publications can be found on our web site under:

<http://www.ecmwf.int/publications>

Contact: [library@ecmwf.int](mailto:library@ecmwf.int)

© Copyright 2024

European Centre for Medium Range Weather Forecasts  
Shinfield Park, Reading, Berkshire RG2 9AX, England

Literary and scientific copyrights belong to ECMWF and are reserved in all countries. The content of this document is available for use under a Creative Commons Attribution 4.0 International Public License. See the terms at <https://creativecommons.org/licenses/by/4.0/>.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error, omission and for loss or damage arising from its use.

## Abstract

This report documents the research done at ECMWF over the period 2019-2023 on the data assimilation, monitoring and NWP (Numerical Weather Prediction) impact of the wind retrievals from ESA's Earth Explorer mission Aeolus. This work was done as part of ESA's Aeolus DISC (Data Innovation Science Cluster) project.

Aeolus, the world's first Doppler wind lidar in space, produced wind profile retrievals with sufficient quality to improve global Numerical Weather Prediction forecasts by a useful magnitude. This has been demonstrated with an assessment of data quality via comparisons of the Aeolus Level-2B (L2B) HLOS (horizontal line-of-sight) wind data with the ECMWF model equivalents and by NWP impact assessment with the OSE (Observing System Experiments) and the FSOI (Forecast Sensitivity Observation Impact) methods. Aeolus was operationally assimilated at ECMWF from 9 January 2020 until 30 April 2023 (due to end of mission).

The estimated precision (1 standard deviation) of the HLOS winds varied considerably during the mission and with geolocation, season, processing software version and range-bin settings, particularly for the Rayleigh-clear HLOS winds varying from 4-7.5 m/s in the troposphere. The precision was more stable for the Mie-cloudy varying from 2.5-3.6 m/s in the troposphere. After L2B processor corrections the systematic errors were typically within  $\pm 1$  m/s (daily averages). The HLOS winds have sufficient information content to improve the ECMWF forecasts despite the random and systematic errors being larger than pre-launch expectations.

Positive impact was demonstrated by OSEs for most periods of the mission, however the magnitude of impact was greater during periods of relatively strong signal levels, which varied a lot during the mission: due to laser energy variations and signal loss within the instrument unique to the FM-B laser.

The largest positive impact of the HLOS winds occurs in the tropical upper troposphere; particularly just after the FM-B laser was switched on in July-December 2019. This period showed statistically significant and good magnitude positive impact on vector wind, temperature, geopotential and humidity forecasts in the tropical and polar troposphere and lower stratosphere by 0.5-2% in root mean square error up to 3-4 days. The impact in polar tropospheric regions tends to fade after 4 days, whilst the positive impact extends to 9-10 days in tropics and Southern Hemisphere extratropics at 100 and 50 hPa (15-20 km). The largest tropical impact at 100 hPa occurs in the east Pacific – a region with the largest wind background errors. The next best impact was shown using the FM-A laser (2018/2019 and 2022/2023), with similar spatial patterns of positive impact to that shown with FM-B. The impact also improved as the ground processing algorithms improved during the mission e.g. going from 1<sup>st</sup> to 2<sup>nd</sup> reprocessing campaign for FM-B.

When signals were relatively strong and the noise was reasonable, the Rayleigh-clear winds provided a greater proportion of the positive impact than the Mie-cloudy, presumably due their much greater spatial coverage than Mie-cloudy winds. Mie-cloudy winds tend to provide more of the polar region impact. Mie-cloudy impact was improved using a more realistic assigned observation error modelling (including representativeness error).

Given good signal levels, Aeolus impact is of comparable magnitude to several other important (operational) satellite observing systems at ECMWF which is a positive result for a demonstration mission; considering Aeolus accounted for fewer than 0.5% of the observations assimilated i.e. there is a strong impact per observation.

The relative FSOI for Aeolus was 5% in early FM-B reprocessed dataset in 2019 but reduced to 2% in mid-2022 due to the then very poor signal levels, with the Mie-cloudy impact exceeding the Rayleigh-clear in 2022. This relative FSOI increased to 3-4% in late 2022/early 2023 with the switch back to the FM-A laser. At its peak for the early FM-B period, Aeolus's relative FSOI is similar in magnitude to a MetOp IASI instrument, ranking amongst the highest FSOI per satellite instrument and it had a similar impact to the radiosonde network.

Consistent spatial patterns in how Aeolus modifies the analysis fields were found: the non-systematic changes tend to be largest in convective areas at ~200-70 hPa; mostly correcting the larger random background forecast errors associated with convection (as seen in the EDA spread). Systematic changes are consistently largest in the tropical upper troposphere and lower stratosphere, and there are also some consistent patterns towards the poles (perhaps fighting other biases in the analysis/model); comparisons to radiosondes suggest Aeolus is not degrading biases.

The magnitude of Aeolus' positive impact was found to be strongly dependent on the implementation of a bias correction of the HLOS winds derived using the ECMWF model as a reference. The biases have a complicated behaviour (with geolocation and time), meaning it took a great effort by the DISC team to understand and correct them. In Autumn 2019 an explanation for a significant contributor to the bias was found, via careful assessments of the ECMWF O-B statistics with reference to the satellite's housekeeping information. The Rayleigh biases are predominantly dependent on gradients in temperature across the instrument's primary mirror. The temperature gradients influence the receiver response via angular changes of the light onto the spectrometers, due to varying focus from the parabolic mirror. The variations in temperature gradients are driven by varying top of atmosphere radiation and the mirror's on-board thermal control mechanism. A bias correction using the NRT primary mirror thermistor reading as predictors was implemented in the operational ground processing chain since 20 April 2020 (and in reprocessed data), providing an accurate bias correction.

The NWP impact of the demonstration mission Aeolus is promising for obtaining a strong and consistent impact from the proposed operational EUMETSAT follow-on mission (EPS-Aeolus) which aims for significantly improved precision HLOS winds compared to Aeolus and operational robustness.

## Table of Contents

Abstract .....	1
1 Introduction.....	7
2 Methods .....	9
2.1 Aeolus Level-2B wind data .....	9
2.1.1 Commissioning Phase NRT FM-A dataset.....	14
2.1.2 NRT 2019 FM-A dataset .....	16
2.1.3 NRT 2019 FM-B dataset .....	17
2.1.4 NRT mid-2020 FM-B dataset.....	21
2.1.5 First reprocessed FM-B dataset: July to December 2019 .....	21
2.1.6 Second reprocessed FM-B dataset: July 2019 to October 2020 .....	21
2.1.7 Late 2021/2022 NRT FM-B dataset .....	22
2.1.8 Third reprocessed FM-A dataset: September 2018 to June 2019 .....	22
2.1.9 Late 2022/early 2023 NRT FM-A dataset .....	22
2.2 ECMWF’s HLOS wind observation operator.....	23
2.3 Observation-minus-background departure statistics .....	27
2.4 Observing System Experiments.....	32
2.4.1 Commissioning Phase NRT FM-A dataset.....	33
2.4.1.1 Experimental set-up .....	33
2.4.1.2 Quality control decisions and corrections.....	34
2.4.2 NRT 2019 FM-A dataset .....	36
2.4.2.1 Experimental set-up .....	36
2.4.2.2 Quality control decisions and corrections.....	36
2.4.3 NRT 2019 FM-B dataset .....	36
2.4.3.1 Experimental set-up.....	36
2.4.3.2 Quality control decisions and corrections.....	37
2.4.4 NRT mid-2020 FM-B dataset.....	38
2.4.4.1 Experimental set-up .....	38
2.4.4.2 Quality control decisions and corrections.....	38
2.4.5 First reprocessed FM-B dataset: July to December 2019 .....	39
2.4.5.1 Experimental set-up .....	39
2.4.5.2 Quality control decisions and corrections.....	39
2.4.6 Second reprocessed FM-B dataset: July 2019 to October 2020 .....	39
2.4.6.1 Experimental set-up .....	39
2.4.6.2 Quality control decisions and corrections.....	40
2.4.7 Late 2021/2022 NRT FM-B dataset .....	40
2.4.7.1 Experimental set-up .....	40

2.4.6.2	Quality control decisions and corrections.....	40
2.4.7	Third reprocessed FM-A dataset: September 2018 to June 2019 .....	40
2.4.7.1	Experimental set-up .....	40
2.4.7.2	Quality control decisions and corrections.....	41
2.4.8	Late 2022/early 2023 NRT FM-A dataset .....	41
2.4.8.1	Experimental set-up .....	41
2.4.8.2	Quality control decisions and corrections.....	41
2.4.9	Summary table of OSEs for different periods and L2B datasets .....	42
2.5	Forecast Sensitivity to Observation Impact (FSOI) .....	44
3	Assessment of L2B data quality by NWP-based monitoring .....	46
3.1	Monitoring of Aeolus for the Commissioning Phase using the NRT dataset .....	46
3.1.1	Rayleigh-clear HLOS wind O-B time variations and anomalies .....	46
3.1.2	Mie-cloudy HLOS wind O-B time variations and anomalies.....	51
3.1.3	Comparing the quality of NRT data from mid-September 2018 to early January 2019.....	55
3.1.3.1	L2B Rayleigh-clear .....	55
3.1.3.2	L2B Mie-cloudy.....	58
3.2	An explanation for the systematic error variations with orbital phase — M1 mirror temperature-dependence .....	62
3.3	NWP monitoring of NRT Aeolus data.....	72
3.3.1	Trends in L2B Rayleigh-clear statistics for NRT data.....	72
3.3.2	Trends in L2B Mie-cloudy statistics for NRT data.....	78
3.4	First reprocessed dataset (early FM-B) monitoring .....	81
3.5	Second reprocessed dataset (FM-B) monitoring.....	83
3.6	Third reprocessed dataset (FM-A) monitoring .....	87
3.7	Whole mission monitoring: combining the third and second reprocessing with the NRT dataset monitoring.....	90
3.7.1	Summary of Aeolus L2B wind random error statistics.....	91
4	Results of Aeolus NWP impact assessment at ECMWF.....	95
4.1	Commissioning Phase NRT FM-A OSEs.....	95
4.2	2019 NRT FM-A OSEs .....	102
4.2.1	Experiments without bias correction .....	102
4.2.2	Experiments with a bias correction scheme.....	106
4.3	2019 NRT FM-B dataset OSEs.....	111
4.3.1	Bias correction scheme.....	111
4.3.2	Analysis changes from assimilating Aeolus HLOS winds .....	113
4.3.3	Short-range forecast impact.....	118

4.3.3.1	Rayleigh-clear and Mie-cloudy .....	118
4.3.3.2	Mie-cloudy only .....	122
4.3.3.3	Rayleigh-cloudy only .....	123
4.3.4	Medium-range forecast impact .....	124
4.3.4.1	Rayleigh-clear and Mie-cloudy .....	124
4.3.4.2	Impact of Aeolus in the stratosphere .....	130
4.3.4.3	Mie-cloudy only .....	132
4.3.4.4	Improving Mie-cloudy impact via the inclusion of representativeness error .....	133
4.3.4.5	Rayleigh-cloudy only .....	135
4.4	Mid-2020 NRT FM-B dataset OSEs .....	137
4.4.1	Bias correction scheme .....	137
4.4.2	Analysis changes due to assimilating Aeolus HLOS winds .....	137
4.4.3	Short-range forecast impact .....	142
4.4.3.1	A comparison of bias correction methods: geolocation-based look-up table versus the M1 temperature-dependent .....	142
4.4.3.2	Rayleigh-clear and Mie-cloudy with M1 temperature-dependent-bias correction combined with improved Mie assigned observation error .....	143
4.4.4	Medium-range forecast impact .....	147
	Rayleigh-clear and Mie-cloudy with M1 T-dependent-bias correction combined with improved Mie assigned observation error .....	147
4.5	First reprocessed dataset (FM-B) OSEs .....	151
4.5.1	Atmospheric temperature-dependent-bias correction for Rayleigh-clear winds ..	151
4.5.2	Analysis changes due to assimilating Aeolus HLOS winds .....	152
4.5.3	Short-range forecast impact .....	155
4.5.4	Medium-range forecast impact .....	159
4.5.5	Tropical cyclone impact .....	164
4.6	Second reprocessed dataset (FM-B) OSEs .....	166
4.6.1	Analysis changes due to assimilating Aeolus HLOS winds .....	166
4.6.2	Short-range forecast impact .....	169
4.6.3	Medium-range forecast impact .....	172
4.7	Late 2021/2022 NRT FM-B dataset OSE .....	177
4.7.1	Short-range forecast impact .....	177
4.7.2	Medium-range forecast impact .....	181
4.8	Third reprocessed dataset (FM-A) OSEs .....	183
4.8.1	Short-range forecast impact .....	184
4.8.2	Medium-range forecast impact .....	186

4.8.3	Tropical cyclone impact .....	190
4.9	Late 2022/early 2023 NRT FM-A dataset OSEs .....	191
4.9.1	Short-range forecast impact .....	192
4.9.2	Medium-range forecast impact .....	196
4.10	A consolidation of OSE impact from Aeolus .....	199
4.11	Aeolus impact assessment by FSOI .....	201
4.11.1	Operational ECMWF data assimilation FSOI results .....	201
4.11.1.1	Relative FSOI comparison of Aeolus L2B HLOS winds to other observation types .....	201
4.11.1.2	Time-series of Aeolus L2B wind FSOI results .....	204
4.11.1.3	Geographical distribution of Aeolus L2B HLOS wind FSOI .....	210
4.11.2	FSOI results for the first reprocessed early FM-B dataset .....	216
4.11.3	FSOI results for the second reprocessed FM-B dataset .....	220
5	Discussion .....	225
6	Conclusions .....	232
7	Acknowledgements .....	234
8	References .....	235
9	Appendix .....	240
9.1	Acronyms .....	240
9.2	Estimating metrics of Aeolus observation errors from O-B departure statistics .....	244
9.3	Simple scalar data assimilation analysis for assessing Aeolus Rayleigh wind impact .....	244
9.4	Rayleigh HLOS wind error propagation including solar background noise .....	246
9.5	The dependence of Rayleigh HLOS wind standard error on incidence angle .....	248
9.6	The sensitivity of Aeolus L2B O-B statistics to a model averaging versus point observation operator .....	253
9.7	Does atmospheric refraction matter for the Aeolus wind geolocations? .....	261
9.8	Modelling of relative FSOI versus assigned observation error .....	265
9.9	Variational Quality Control for Aeolus in the IFS .....	267
9.10	Some notes on the horizontal scale of L2B winds and N/P settings .....	272
9.11	History of changes in AUX_PAR_2B in NRT operational ground segment processing .....	274
9.12	Geometry of Aeolus for the derivation of the Aeolus (H)LOS wind observation operator .....	278

## 1 Introduction

This ESA contract report is a conversion of a Technical Note (TN, AED-TN-ECMWF-NWP-025--20230809\_v7.0) has been prepared as part of the Aeolus DISC (Data Innovation Science Cluster) Phase E2 contract, to allow the content to be referenced. It consists of the research done for the Aeolus DISC project on the Numerical Weather Prediction impact of Aeolus wind data at ECMWF over the period 2019-2023.

We focus on ECMWF's assessment of the quality of the Aeolus Level-2B HLOS (horizontal line-of-sight) wind data and its impact in ECMWF's global NWP system. The assessment of the data quality is performed using ECMWF's global model short-range wind forecasts as a reference and NWP impact is assessed via the verification of the effect on analysis and forecast skill from the data assimilation of Aeolus winds i.e. Observing System Experiments. The TN provided guidance to other NWP centres (in the first years of the mission) and the scientific community regarding the data quality and suggested usage of the Aeolus wind observations. It has been updated at roughly six-month intervals during the ESA funded Aeolus DISC Phase E2 contract period; evolving with our developing understanding of this Earth Explorer mission. The definitions of commonly used acronyms can be found in the Appendix.

In 2018 wind profiles were listed as the "highest priority critical atmospheric variable not adequately measured by the then current or planned observing systems" in the World Meteorological Organisation (WMO) Rolling Review of Requirements (WMO, 2018). Many studies have provided evidence for the potential benefits of more wind profile observations for NWP, such as: Stoffelen et al. (2006), Marseille et al. (2008), Baker et al. (2014), Horányi (2015a), Tan et al. (2007), Weissman and Cardinali (2007) and Illingworth et al. (2018).

The European Space Agency's Aeolus satellite was launched on 22 August 2018 with a payload of the first Doppler Wind Lidar (DWL) in space and the first European lidar in space. Aeolus partly fills the wind data gap in the Global Observing System (GOS) and hence it is of great interest for demonstrating the value of satellite wind profiles for NWP. However, being a single satellite instrument with limited spatial and temporal coverage and with some data quality issues (as reported in this TN) it only partially addresses the gap.

Aeolus is an Earth Explorer (EE) mission in ESA's Living Planet Programme (Ingmann and Straume, 2016). The mission's main objective is to provide profiles of high-quality wind component retrievals from the surface up to the lower stratosphere, using a DWL instrument (Atmospheric Laser Doppler Instrument (ALADIN)) in a polar sun-synchronous, 320 km altitude, dawn-dusk orbit with a 18:00<sup>1</sup> local solar time ascending node equator crossing (and hence 06:00 local solar time for the descending node equator crossing). The mission's most important product consists of slant path (off-nadir) profiles of the horizontal line-of-sight (HLOS) wind component, pointing perpendicularly to the satellite-earth surface velocity vector. The mission lifetime was designed to be at least three years, which was exceeded. A detailed description of the mission and the ALADIN instrument can be found in ESA's ADM-Aeolus Science Report (ESA, 2008).

The ALADIN instrument is a powerful and highly frequency-stabilised pulsed ultraviolet (UV, 354.8 nm wavelength) DWL. The instrument laser emits pulses of around 40-150 mJ at rate of 50.5 Hz through the 1.5 m diameter Cassegrain telescope into the atmosphere. The light is scattered by air molecules (Rayleigh scattering), particles and hydrometeors (Mie scattering), and a very small fraction of the scattered light is backscattered to the instrument where it is received via the same Cassegrain telescope. The atmospheric signal is sampled in two channels; one designed for detecting the backscatter

---

<sup>1</sup> The LTAN (Local Time at Ascending Node) has been permitted to drift within  $\pm 10$  minutes of 18:00.

from the clear-air molecules (a Double Fabry-Pérot spectrometer, called the Rayleigh Channel) and one for the particulate backscatter (a Fizeau spectrometer, called the Mie channel). The backscatter signal is Doppler shifted (frequency shifted) due to the motion of the atmospheric scatterers along the instrument's line-of-sight (and due to satellite-earth relative motion, which can be corrected), which is determined by comparing the backscatter frequency to the instrument emit frequency. Wind is the is a natural movement of air or other gases relative to a planet's surface and hence Aeolus measures Doppler frequency shifts due to the wind. The Rayleigh channel in clear-air (only molecules) is measuring the movement of the gas-phase and hence matches the definition of wind. The Mie channel is measuring the movement of particulates (e.g. cloud ice/water) which are assumed to be advected with the wind, at the same speed as the gas-phase.

The Level-2B (L2B) processing team (ECMWF and KNMI) have collaborated closely with the Level-1 processing team (DLR, DoRIT, Météo-France), ESA and Industry as part of the Aeolus DISC project to characterize, calibrate and validate the ALADIN instrument and the Aeolus data products with the aim of maximising the data quality.

Short-range NWP forecasts have been extensively used in the past to detect observation quality issues e.g., Hollingsworth et al. (1986), Stoffelen (1999) and Lu et al. (2011). This same principle has been applied for Aeolus, via comparisons of the winds to the short-range forecast. ECMWF are in the unique position of being closely involved in the data monitoring/assimilation testing and the L2B processor "wind retrieval" development. Being data user and provider has led to a fast feedback loop from the detection of anomalies via ECMWF monitoring, to improvements in the L1 and L2 processor via the DISC team. This has allowed a reasonably quick transition to decent quality data and operational assimilation (since January 2020) for a demonstration mission, given the evidence of a positive NWP impact. The Aeolus HLOS winds data quality has been assessed over four and a half years, covering the period from September 2018 until April 2023.

In section 0 we describe the methods applied for data monitoring and assimilation: information such as which L2B datasets were used for different periods (due to the quickly evolving ground processing chain), how comparisons of Aeolus L2B HLOS winds to the ECMWF model are made and the settings and quality control decisions for Observing System Experiments (OSE) used for assessing Aeolus' NWP impact. In section 3, the quality of Aeolus L2B HLOS wind observations is assessed using the ECMWF global model as a reference over the mission. In section 4, the results of the many NWP impact assessment results are presented. This is followed by the discussion, conclusions and the appendix.

## 2 Methods

### 2.1 Aeolus Level-2B wind data

Aeolus wind products are the output of a multiple stage ground processing chain. The different processing levels are described in ESA (2008) and, with more up to date material provided in the Algorithm Theoretical Baseline Documents (ATBDs) for the various product levels (Level-1B and Level-2B).

The Aeolus L2B HLOS wind product contains retrieved winds suitable for use in NWP and scientific research, due to several important additional corrections which are applied relative to the Level-1B (L1B) HLOS wind products<sup>2</sup>. The Aeolus ground segment processing chain has been developed since the early 2000's by various groups: ESA, Airbus DS, DLR, MDA, DoRIT, KNMI, Météo-France and ECMWF. Tan et al. (2008) explains the original concept of the L2B processor; however, this reference is now partially obsolete due to several significant algorithm updates since then. The L2B processor ATBD is the most up-to-date description (but it also needs updating).

Four types of L2B wind observation are available: Rayleigh-clear, Rayleigh-cloudy, Mie-clear and Mie-cloudy. For NWP, the focus has been on the use of the Rayleigh-clear and Mie-cloudy winds, which by design are of better quality than the other two types. If the L2B processor classification of measurement-bins into clear and cloudy conditions is performed perfectly, then no Mie-clear winds should be retrieved. In practice there is usually a small sample of Mie-clear winds due to the sometimes-significant noise present at the measurement-bin scale L1B scattering ratio or L1B Mie SNR values (1 range-bin thickness in depth, 1 measurement in width: particularly noisy with the original N=30, P=19 settings). The Rayleigh-cloudy winds are derived from Rayleigh channel signal for which there is a large component of Mie scattering (cross-talk) on top of the Rayleigh-Brillouin spectrum. Naturally the Rayleigh-cloudy and Mie-cloudy winds sample same areas in the atmosphere, particularly when Rayleigh and Mie range-bin settings are collocated (as has been mostly true during the mission). It was expected that Rayleigh-cloudy winds would provide little additional benefit to NWP due to their collocation with significantly less noisy and less biased Mie-cloudy winds. However, they are derived from independent signals and should be considered for NWP. A limited assessment of Rayleigh-cloudy wind impact is reported in this TN; partly because bias issues have only been improved in NRT data in the last few months of the mission. OSEs with future reprocessing will investigate this more.

An example plot of Aeolus L2B HLOS wind observations, both L2B Rayleigh-clear and Mie-cloudy, during the early Commissioning Phase<sup>3</sup> (CP) from an orbit on 15 September 2018 is shown in Figure 1. One-BRC (~88 km) horizontal averaging was applied in the L2B processor for both the Mie and Rayleigh winds, due to the L2B processor settings at the time. There were ~5.5 times more Rayleigh-clear winds than Mie-cloudy winds due to the RBS and spatial coverage of attenuated backscatter when operating from space. To help interpret the wind features in the HLOS wind "lidar curtain" in Figure 1, the strongest winds (jet streams) are annotated. ALADIN, being an active space-borne optical instrument, is totally attenuated by optically thick ice clouds, water clouds and aerosol, hence there are areas (shown as black due to an absence of winds) with no wind observations available, particularly closer to the surface. Away from the poles, during ascending orbits, Aeolus measures the predominantly westerly extratropical jet streams as positive HLOS winds, and during descending orbits the westerly extratropical jet streams are measured as negative HLOS winds.

<sup>2</sup> The website <https://confluence.ecmwf.int/display/AEOL> provides the recent version of L2B processor software and its documentation for download (including ATBDs).

<sup>3</sup> The Commissioning Phase covered from launch until end of January 2019.

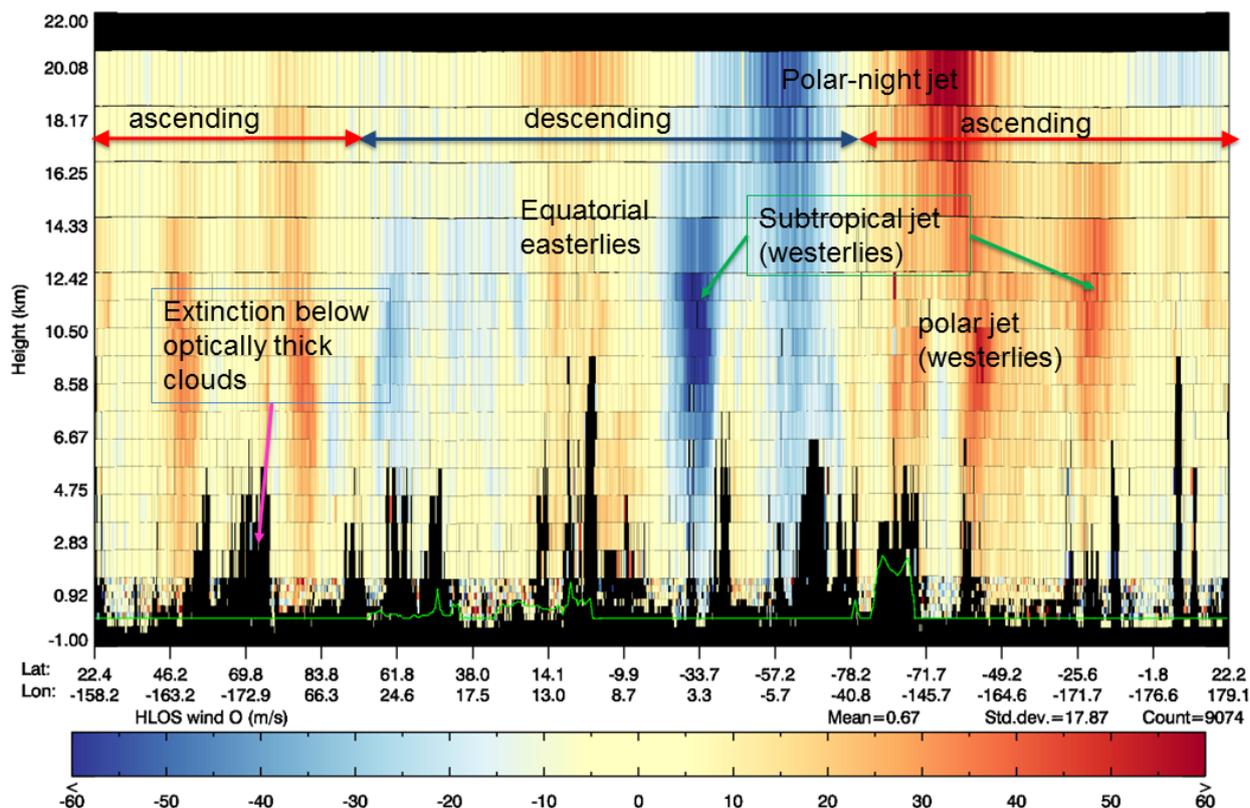


Figure 1. L2B Rayleigh-clear and Mie-cloudy HLOS wind observations for one orbit from 15 September 2018. The vertical axis is the geometric height relative to EGM96 (Earth Gravity Model from 1996) geoid and the horizontal axis is time along the orbit (but labelled with latitude/longitude geolocations along the orbit for ease of understanding). The colour scale indicates the HLOS wind speed (units: m/s). Winds blowing away from the satellite produce positive HLOS winds and winds blowing towards the satellite produce negative HLOS winds. To visualise lower wind speeds more clearly, the scale saturates at  $\pm 60$  m/s, despite the speed in some jet streams reaching larger values (e.g. 100 m/s). Black areas indicate an absence of observations. Each HLOS wind observation is plotted as a coloured rectangle with boundaries indicating the spatial extent of the measurement-bins used in the retrieval, hence the vertical sampling (range-bin thicknesses) is evident (the horizontal extent of an observation is not easily seen for this whole orbit plot). The Digital Elevation Model i.e. Earth’s surface is shown in green (Antarctica is visible as higher ground at the lowest latitudes).

ALADIN’s line-of-sight is pointed roughly perpendicular to the satellite’s velocity and is approximately 10 degrees off the zonal direction for a large fraction of an orbit<sup>4</sup>. The sensitivity (partial derivative) of HLOS wind with respect to the zonal ( $u$ ) and meridional ( $v$ ) wind component, derived using the reported L2B wind azimuthal angles, is shown in Figure 2. There are two branches to the  $u$ -component sensitivity for ascending and descending orbits, whereas there is only one branch for the  $v$ -component sensitivity — meaning that the  $v$  wind component affects the HLOS wind similarly whether ascending or descending. The  $u$ -component wind sensitivity reaches its maximum absolute value in the latitude range  $-50$  to  $+70^\circ$ . The  $v$ -wind component sensitivity peaks as Aeolus approaches the poles

<sup>4</sup> The  $\sim 10$  degrees from zonal direction assertion applies well for the latitude range  $-25$  to  $+55$  degrees latitude. Beyond this range it becomes increasingly meridional towards the poles. The asymmetry with latitude is due to the chosen orbit inclination combined with Aeolus pointing to the right with respect to the orbital velocity.

whereas the u-wind component sensitive tends to zero. The asymmetry with the line-of-sight pointing to the right of the orbital velocity means there is an asymmetry in the HLOS wind u/v component sensitivity between the Southern and Northern Hemispheres, hence on average it is more (less) sensitive to the u (v) component in the Northern Hemisphere than the Southern Hemisphere.

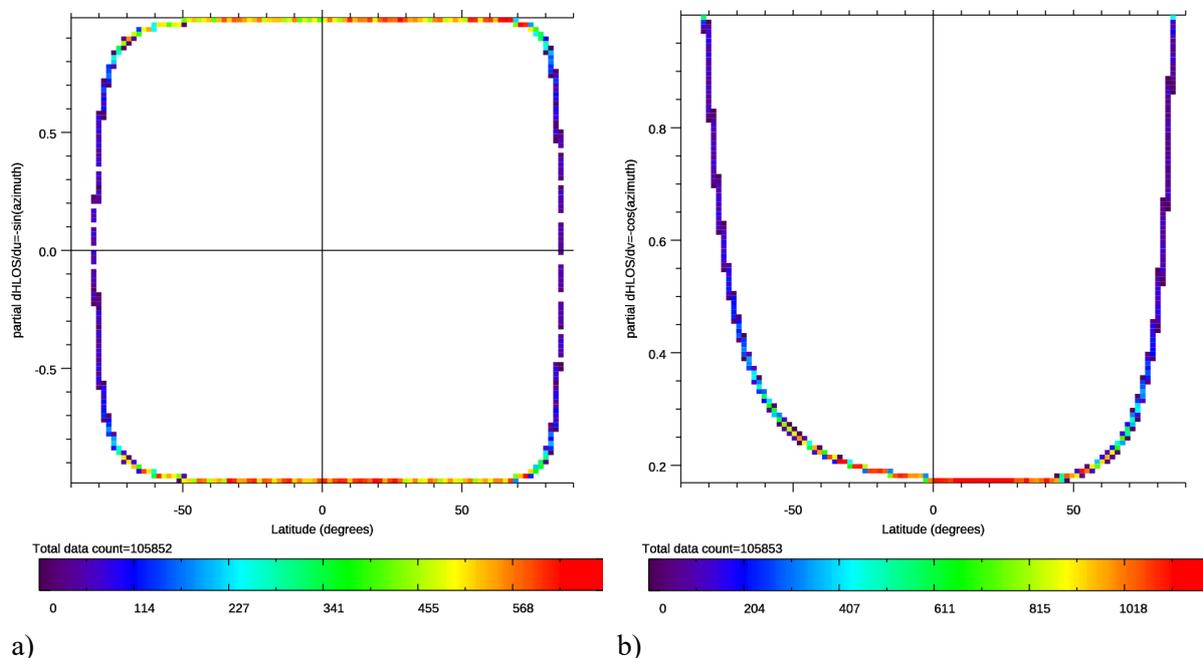


Figure 2. The sensitivity of HLOS wind to a) zonal wind component i.e.  $\frac{\partial HLOS}{\partial v}$  and b) meridional wind component i.e.  $\frac{\partial HLOS}{\partial u}$  as derived from measured Aeolus azimuthal angle data as a function of latitude (and using equation (1) for the HLOS wind). Colours indicate the data count in this 2D histogram.

Figure 3 helps to visualise the geometry of the orbit and hence why the meridional component sensitivity peaks at the poles. The greater spatial and time sampling over the North Pole is evident due to the near polar orbit.

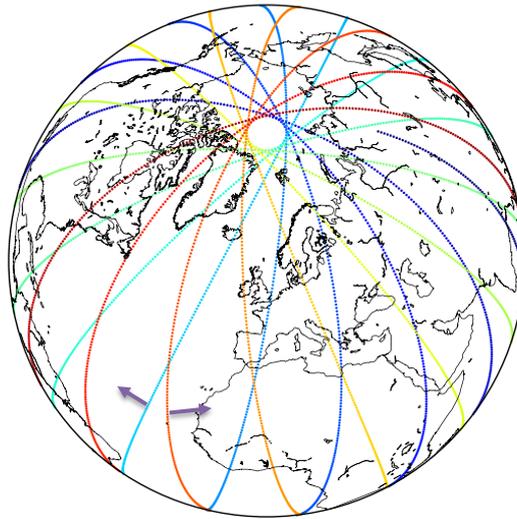


Figure 3. An example of Aeolus' typical near surface geolocations in one day. The line-of-sight pointing direction of ALADIN is indicated by the purple arrows for an ascending (eastward pointing) and a descending (westward pointing) orbit phase.

The L2B processor settings (via the AUX files) determine many characteristics of the L2B wind observations. For example, one important setting is the maximum horizontal averaging length-scale. During the CP NRT processing, this was set to be one Basic Repeat Cycle (BRC) for both the Rayleigh and the Mie observations, meaning the HLOS winds were produced from up to thirty smaller horizontal-scale (~2.9 km) measurements (covering 12 seconds), resulting in observations of up to ~88 km horizontal extent. See appendix 9.10 for details on the horizontal extent of L2B winds, which varied according to the instrument's N/P settings (see Appendix 9.10).

The L2B wind vertical resolution is limited by the range-bin settings (RBS), selected by mission experts and commanded on-board the satellite. The thickness of the 24-available range-bins can vary with altitude from 250 m to 2 km, in 250 m steps. Typically, the RBS become thicker with height to compensate for the decreasing signal levels with altitude (at least for Rayleigh clear-air attenuated backscatter profiles). The Mie and Rayleigh channels can be commanded to have different range-bin settings; however, they have tended to be rather similar during the mission for the benefit of Rayleigh classification into clear/cloudy by use of the Mie channel – thus requiring collocated range-bins.

The range-bin settings remained fixed during most of the CP, focusing on a very high sampling of range-bins near the ground (0.25 km thick range-bins) to ensure decent ground returns. Ground returns can be used to obtain a zero-wind reference, which was considered before launch to be an important source of wind calibration information, therefore the choice of high sampling near the ground during the CP. The RBS have evolved significantly during the mission, in line with improved knowledge of where in the vertical the largest NWP impact is obtained and due to requests by scientists for RBS beneficial to specific CAL/VAL campaigns or to observer natural phenomena (such as Quasi-Biennial Oscillation (QBO) or volcanic eruption plume monitoring).

The NWP impact of L2B winds has been assessed over many time periods and with many processor baselines as the mission progressed. The data quality has evolved significantly during the mission. NRT processing algorithms, instrument performance and data availability has varied a lot, leading to some natural time boundaries and processor version choices for the assessment. A complete mission reprocessing with the most recent (and therefore best) processing algorithms is required to reduce such boundaries; this is targeted for the Aeolus DISC Phase F (after end of mission).

There have been many changes in the operational processing chain during the mission that affect the quality of the L2B winds, some highlights of most relevance to L2B winds are shown on the timeline of Figure 4.

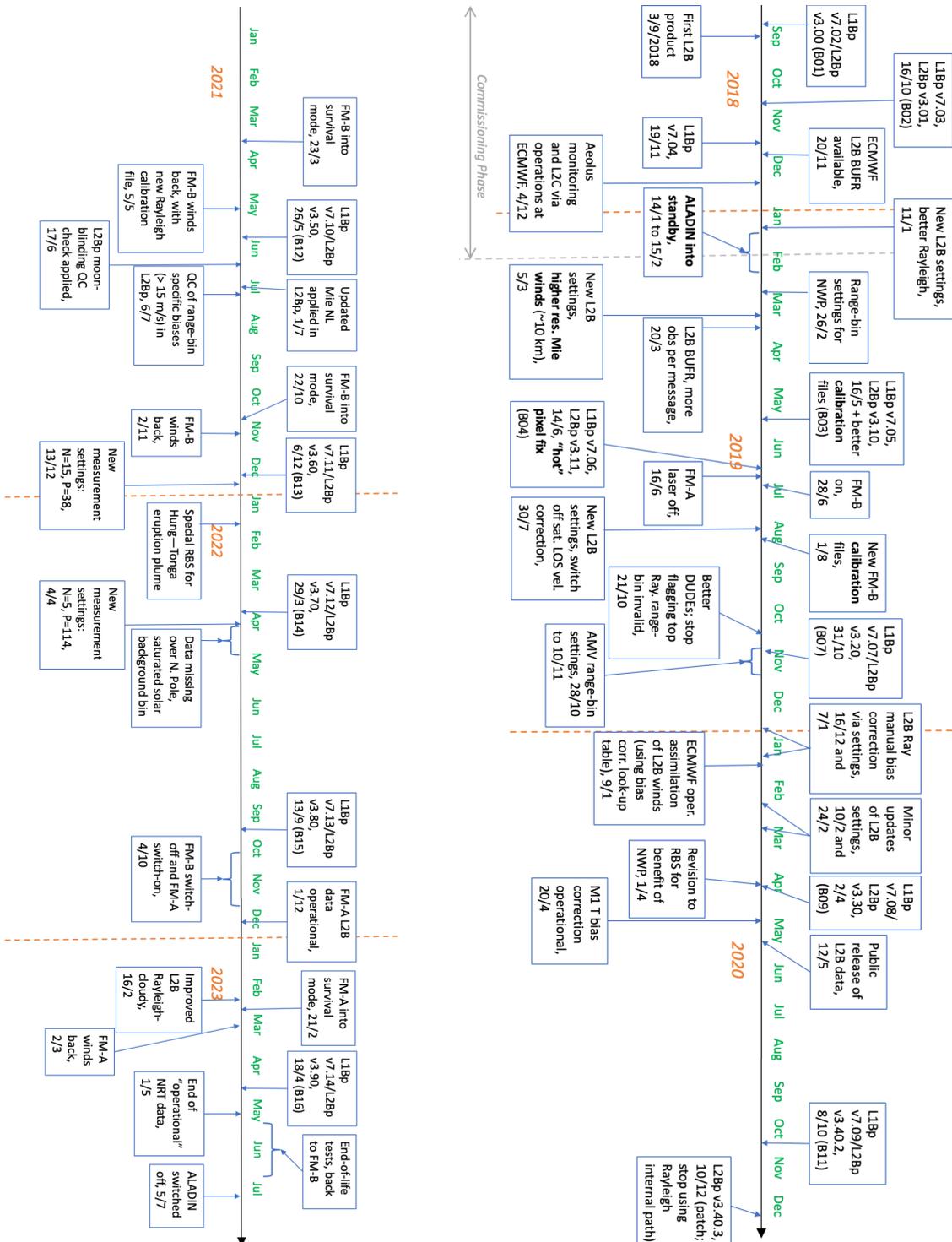


Figure 4. A timeline of the main changes in the Aeolus Level-2B dataset for NRT operational processing during the mission.

This TN retains some obsolete NWP impact assessments using NRT and non-official reprocessed datasets, to help provide some historical context (shows how the impact improved with better processing baselines, quality control (QC) and data assimilation settings). The first data assessed for Aeolus was during the mission's CP (which used the FM<sup>5</sup>-A laser). For an OSE from April to June 2019 FM-A data, the L2B dataset was produced via ECMWF's own offline reprocessing mechanism. This had specially defined QC settings in the AUX\_PAR\_2B to try to minimise the effect of unique problems for the dataset at the time e.g. biased winds associated with hot pixels<sup>6</sup> (see Weiler et al. (2021)) affecting specific range-bins.

Several NRT datasets with the FM-B laser have been tested. The first used PDGS's operational NRT processing chain, from 2 August 2019 onwards when the quality was sufficient (FM-B wind data started on 28 June 2019) because earlier NRT data applied inappropriate calibration files. A mid-2020 FM-B period covered from when the M1 telescope temperature bias-corrected winds became available i.e. April 2020 until September 2020 (see Section 3.2). This used nearly identical data to the PDGS's operationally produced NRT dataset but was produced by the back-up server at ECMWF's L2/Met PF (Level-2 Meteorological Processing Facility; operated as part of the DAMI contract). This was to allow a slightly longer OSE period with M1 temperature biased corrected winds, which became operationally available on 20 April 2020, whereas they became available on 4 April 2020 on the back-up server for testing of the bias correction scheme.

A NRT FM-B L2B dataset OSE was run covering 13 December 2021 to 28 September 2022. This was useful for assessing if positive impact was still possible, given the continuing decline in SNR for FM-B which strongly affects the Rayleigh-clear wind random error. Finally, since the re-switch-on of FM-A in October/November 2022, a NRT FM-A dataset OSE has run covering 1 December 2022 until 30 April 2023. The data quality improved significantly with the switch back to FM-A and was the best NRT dataset of the mission in terms of random errors (see section 3.3). 30 April 2023 was the last day of "operational" L2B data, to be distributed to NWP centres. It was followed by an end-of-life (EOL) testing phase and then the deorbiting of Aeolus with it finally hitting the Earth at ~19 UTC on 28 July 2023.

This TN includes NWP impact assessment using ESA's official reprocessed datasets. The first reprocessing covered from 29 June to 31 December 2019 FM-B data and became available in late 2020; this was the first official (ESA) reprocessed dataset covering that period; and was demonstrated to be of better quality than the NRT dataset; see Abdalla et al. (2020). A second reprocessing, covering FM-B from June 2019 to October 2020 was produced; again demonstrating benefits in data quality; see Abdalla et al. (2021). The third reprocessing was the first official reprocessing of the early part of the mission when the FM-A laser was in operation. It was released by ESA in November 2022, covering the sensing period of September 2018 to mid-June 2019. The DISC has confirmed this to be a significant improvement in the data quality with respect to the NRT dataset; see Abdalla et al. (2023).

### 2.1.1 Commissioning Phase NRT FM-A dataset

This dataset was a combination of reprocessed (as of 2018 i.e. not official reprocessing activities, covering 3 September to 19 November 2018) and operational near real-time (NRT) produced data, both

<sup>5</sup> Flight Model. There are two available lasers on ALADIN: FM-A (nominal) and FM-B (redundant).

<sup>6</sup> "Hot" pixels are increased dark current rates for specific ALADIN ACCD detector pixels, which can cause large biases in HLOS winds if not corrected for. Information on when they occurred for specific range-bins can be found here (CAL/VAL access only): <https://www.aeolus.esa.int/confluence/display/CALVAL/Calval+Black+List>

produced by ESA's Payload Data Ground Segment (PDGS). The data was produced with the L1B processor version 7.04 and L2B processor version 3.01, which was baseline 2B02. The L2B EE format products were converted to L2B BUFR products by the L2B processing team. In NRT operations, ECMWF's L2/Met PF, which is part of the Ground Segment, provided the L2B EE format and corresponding BUFR data.

The vertical range-bin settings<sup>7</sup> applied during most of the CP are shown for the example data of Figure 1. The RBS varied from 250 m for range-bins near the surface up to ~1 km altitude, then 1 km thickness up to ~13 km altitude, followed by to 2 km thickness to ~21 km altitude (the lower stratosphere). This RBS was applied from the start of the mission until 25 February 2019; a phase in which instrumental calibration had priority over NWP impact. The near-surface 250 m range-bins are very noisy for the Rayleigh-clear due to the short-time accumulation and attenuation of signal by clouds above, but the Mie -cloudy winds are mostly low noise, due to the relatively large backscatter from the (well sampled) boundary layer water cloud tops.

To retrieve HLOS wind observations, a calibration of the Rayleigh and Mie spectrometer instrument responses as a function of backscattered light frequency is required (ESA, 2008). The CP NRT dataset was generated using a fixed set of instrument response calibration files: to aid the characterisation and quantification of the different contributors to the Aeolus instrument drift; to verify the satellite Attitude and Orbital Control Systems (AOCS) and the calculations of the satellite attitude. Weekly calibration instrument modes were performed during the CP (and throughout most of the mission so far). Another source of wind calibration information is the ground return winds, which can be used to correct atmospheric winds for issues such as e.g. unknown satellite miss-pointing; thermal and satellite to target range dependency of the spectrometer response leading to short (within an orbit); and long (seasonal effects) term bias variations. The optimal calibration strategy for Aeolus was still being worked on during the mission e.g. see section 3.2.

A critical calibration for the L2B processor Rayleigh winds is provided by the Rayleigh-Brillouin Correction look-up table (AUX\_RBC\_L2 file) which is produced by the Calibration Suite software. For the CP NRT dataset the AUX\_RBC\_L2 was derived from an Instrument Spectral Response (ISR) (ESA, 2008) measured on 19 September 2018; an Instrument Response Calibration (IRC) was not directly used (via the so-called update Corrected Spectral Registration (CSR) procedure, which is an optional part of the Calibration Suite processing). This was because testing using the IRC led to significantly biased winds (e.g. -7 m/s) when applying it to the update CSR software (the algorithms have improved since then). Reprocessing of FM-A data (3<sup>rd</sup> reprocessing) has been shown by the DISC to produce significantly better data quality in terms of biases and random errors.

To account for an observed frequency offset between the internal reference and atmospheric path response curves as seen in the Rayleigh Response Calibration (RRC) after the instrument was launched, the atmospheric response curves were shifted by approximately 160 MHz in the AUX\_RBC\_L2 file produced by the Aeolus Calibration Suite. The resulting updated AUX\_RBC\_L2 file was used in the L2B wind processing, but this led to an almost constant global bias of several m/s with respect to ECMWF model equivalent HLOS winds during the September 2018 period. Hence, it was decided, as a preliminary measure, to adjust the frequency shift such that the L2B Rayleigh-clear global mean departures with respect to ECMWF's short-range forecast became close to zero in that period, which was achieved by changing the shift to 155.4 MHz. This tuning, using global ECMWF statistics was

---

<sup>7</sup> Vertical range-bin sampling intended to be more suitable for NWP exploitation were applied since 26 February 2019. This meant reducing the range-bin resolution near the ground (for ground returns) as was applied in the CP.

deemed necessary at the time to get an early dataset suitable for early NWP impact studies<sup>8</sup>. It should be noted that evaluations demonstrate that the ECMWF short-range forecast has small global averaged wind bias when compared to high quality observations such as radiosondes (e.g. less than 0.3 m/s for zonal wind component). In February and March 2019, several improvements in the Calibration Suite were made, and reliance on ECMWF as a reference for calibration was temporarily avoided. Improvements in the AUX\_RBC\_L2 file generation were sought. For example, Météo-France developed a method to obtain the appropriate frequency shift from the difference between internal reference and ground return Rayleigh response. However, as will be shown in later sections a reliance on the ECMWF model as a global bias reference remains the best option; however, ground return zero-winds can do a reasonable job.

The applied L2B processor Mie wind calibration information (the Mie Response Calibration (MRC)) was the same as that used by the L1B product Mie wind observations. It was chosen by the Level-1 team expert (Uwe Marksteiner, DLR) based on assessments of the quality of the then weekly MRCs (based on ground returns over polar regions in nadir-pointing mode). An MRC valid on 15 October 2018 was chosen and applied in the processing/re-processing of the whole CP dataset. Nominally the L2B processor obtains the MRC information via the L1B wind mode product, but an option to directly read the MRC information (from AUX\_MRC\_1B file) was implemented. Testing showed that it is possible to tune the MRC parameters to minimise the Mie wind bias with respect to the ECMWF model, however because the Mie bias was relatively small for the CP, it was not done to avoid a possible risk of aliasing NWP model biases into the satellite observation dataset. Improvements in the Mie calibration have been found in the subsequent years using NWP as a reference; see Marseille et al. (2022), in determining the non-linearity of the MRC.

### 2.1.2 NRT 2019 FM-A dataset

This dataset covered the period from 2 April to 14 June 2019 and was chosen for assessing NWP impact given the possibility to run OSEs with ECMWF's modified L2B data. It was produced via L1Bp v7.04 and L2Bp v3.01 patch 1. The L2B processing was produced at ECMWF from a non-operational set-up (L2/Met PF back-up server) with bespoke AUX\_PAR\_2B settings to manually avoid biased HLOS wind on specific range-bins (due to so called "hot" pixels). This was necessary because the solution for hot pixels (regular dark-current measurements) was not commanded in the operational processing chain until 14 June 2019 (see Figure 4)<sup>9</sup>. The AUX\_RBC\_L2 and AUX\_MRC\_1B files applied resulted in small global average biases for this period. The end of the experimental period was determined by the switch from laser FM-A to FM-B (because of a continuing decrease of laser pulse energy with FM-A).

Some of the features of the L2B wind quality for this period are listed below:

- Following a gradual decline, the FM-A laser energy reached a rather low value of ~40-45 mJ per pulse.
- The number of uncorrected hot pixels increased with time causing an increasing number of range-bins with large, mostly systematic wind errors. However, ECMWF's offline processed dataset flagged the winds invalid for the affected range-bins, via the AUX\_PAR\_2B. This

---

<sup>8</sup> Unfortunately, we have not been able to calibrate Aeolus L2B winds without the aid of NWP model equivalents for a global bias correction; see section 3.2.

<sup>9</sup> This was achieved by the application of so-called DUDE (Down Under Dark Experiment) calibration data four times per day (initially) in combination with L1B processor v7.06. This led to a dramatic improvement in the quality of the L2B winds.

decreased wind counts by ~20-25% winds compared to the nominal situation. The range-bins rejected for the Mie channel had IDs (with 1 as the top altitude range-bin); 24, 16, 13, 5 and 2, and for the Rayleigh channel: 20, 15, 11, 5 and 1.

- Updated instrument response calibrations i.e. AUX\_RBC\_L2 and AUX\_MRC\_1B were applied, which improved biases.
- Commanded range-bin settings aimed at improving NWP impact were applied i.e. thinner range-bins in the upper troposphere and lower stratosphere and fewer 250 m range-bins near the surface compared to the early CP.
- Higher horizontal resolution and hence increased data count of Mie-cloudy winds, via tuning of the L2Bp grouping algorithm settings (in AUX\_PAR\_2B). The Mie-cloudy winds were produced with a measurement grouping maximum size of 10 km (advanced setting) during this period, compared with up to 1 BRC (~88 km) in the CP. The change in Mie horizontal averaging setting occurred on 5 March 2019 (see Figure 4).

### 2.1.3 NRT 2019 FM-B dataset

This dataset covered the period from 2 August until 31 December 2019. It consisted of the operationally produced L2B products from the NRT PDGS processing i.e. operationally produced L2B BUFR data using L2BP v3.11 and L1Bp v7.06 (B06) followed by L1Bp v7.07 and L2Bp v3.20 from 31 October 2019 (B07). The start date was determined by when reasonable FM-B calibration files entered operations following the FM-B switch-on, resulting in smaller global-daily bias. The calibration files did not change over this period. Reasonably small global average biases were found during August 2019; however, as shown in section 3.2, the bias subsequently increased with time. There was also a decrease in precision for Rayleigh winds during the test period due to a decline in atmospheric path signal levels – a trend which went on to affect the entire period of FM-B; see section 3.3.

The range-bin settings were changed on several occasions during this period<sup>10</sup>. The initial range-bin settings were more suited for NWP than those applied in the CP but did not vary as a function of latitude until 21 October 2019; see Figure 5 for an example of the RBS applied in August 2019. By January 2020, the RBSs varied considerably with latitude to try to maximise NWP impact, see Figure 6. A period of modified RBS to maximise collocations with Atmospheric Motion Vector (AMV) winds for CAL/VAL purposes was implemented between 28 October 2019 and 10 November 2019. It increased the resolution of the range-bins significantly in parts of the troposphere, at the price of a reduced top range-bin altitude. RBS were adjusted for another test period, aimed at improving collocations with the Strateole-2 campaign data (high altitude super-pressure balloons) with the top range-bin set to be higher than 20 km in the tropics from 10 November 2019 onwards (which is evident in the tropical latitudes of Figure 6).

L2B data properties to note for this test period:

- Before 21 October 2019 all Rayleigh range-bin 1 (top altitude range-bin) HLOS winds were flagged as invalid, due to a limitation in the dark current calibrations (DUDEs) for this range-bin.

---

<sup>10</sup> The CAL/VAL website provided details: <https://www.aeolus.esa.int/confluence/display/CALVAL/Vs+002>

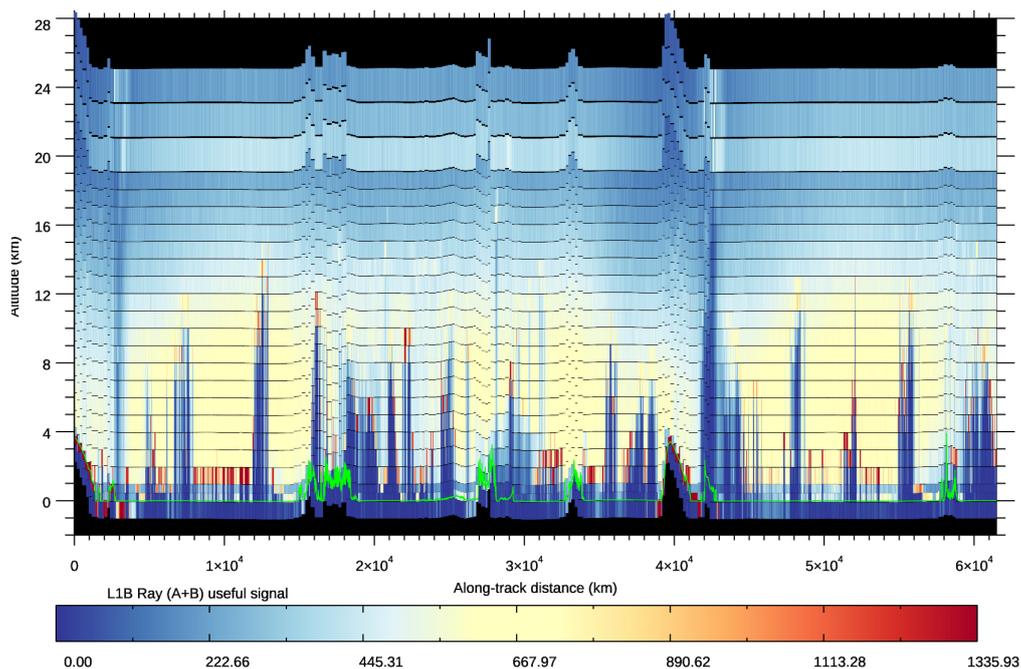
- The AOCS LOS velocity correction<sup>11</sup> was switched off (via AUX\_PAR\_2B), as it was shown to degrade, rather than improve wind biases.
- There was an inadvertent reduction in the Mie wind observation counts with a processor baseline update on 31 October 2019. Also, the Mie horizontal maximum accumulation length (advanced setting) was increased to 12 km and the Rayleigh was changed to exactly 1 BRC (~83 km with N=30, P=19); previously it was set to a specific distance rather than an exact BRC, causing occasional small observation lengths.
- On 16 December 2019 the AUX\_PAR\_2B was updated to apply a manual bias correction of the Rayleigh winds to compensate for a global average bias drift that was rather large in FM-B data; section 3.3. The manual bias correction was set to +4 m/s.
- A L1B processing bug caused the L2B Rayleigh estimated instrument wind error to be overestimated in the summer polar regions by ~20-30%, particularly at high altitudes. It was fixed in the operational L1B baseline introduced on 2 April 2020. It was due to the solar background noise term being too large in the L1B Rayleigh SNR (signal to noise ratio) calculation<sup>12</sup>.

---

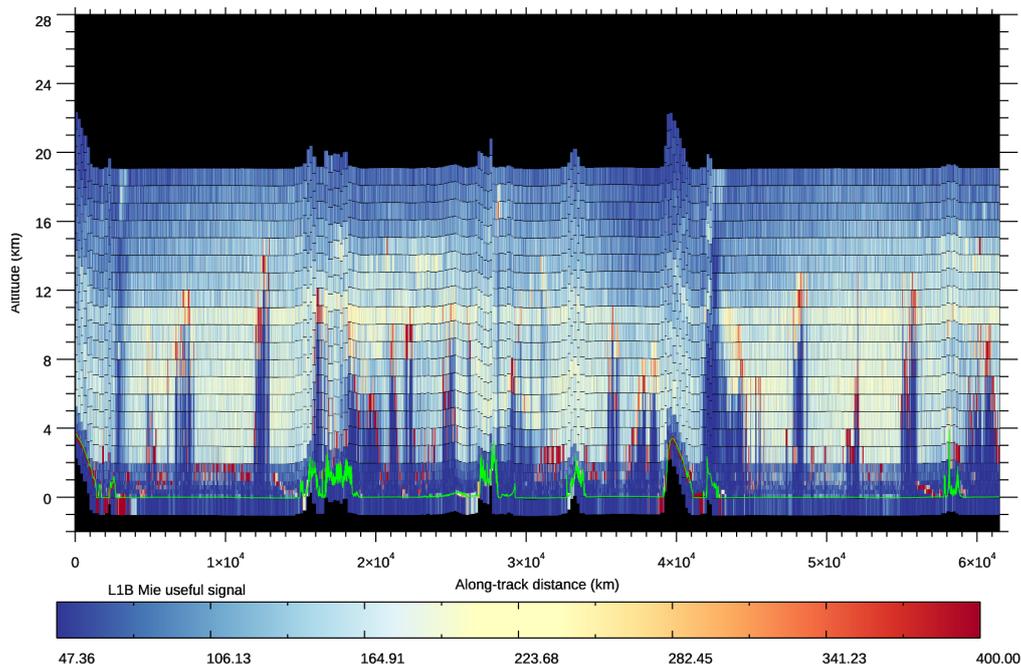
<sup>11</sup> The AOCS LOS velocity correction has never been applied in the L2B processing since because no clear benefit could be found.

<sup>12</sup> This is explained in the DISC confluence item:

[https://csde.esa.int/confluence/display/AEOLUSDISC/dL1B\\_013](https://csde.esa.int/confluence/display/AEOLUSDISC/dL1B_013)

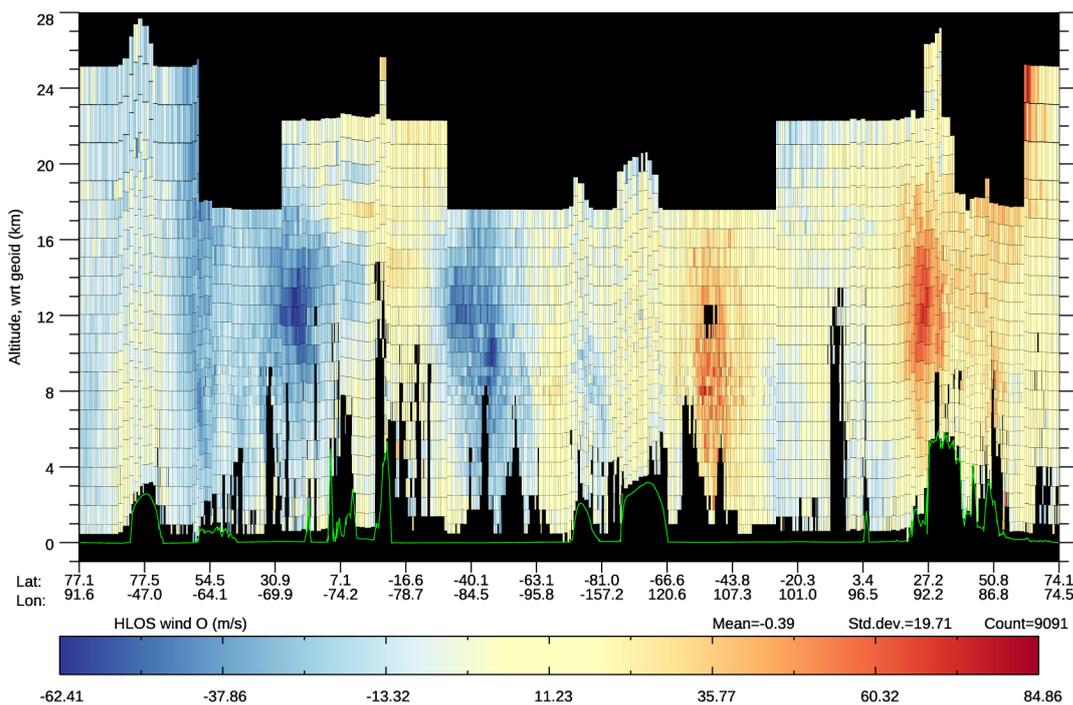


a)

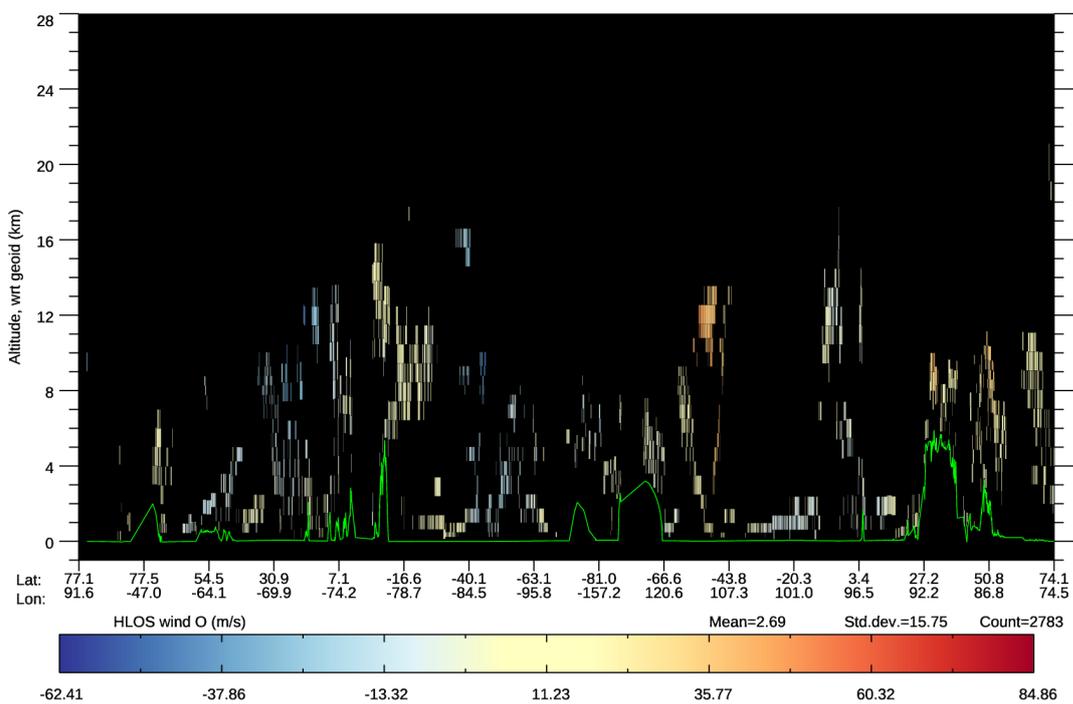


b)

Figure 5. Examples of the range-bin settings applied for a) the Rayleigh channel and b) the Mie channel in August 2019 for the early FM-B period, as shown via plots of the L1B useful signal levels at measurement level for one L1B file.



a)



b)

Figure 6. Example of the range-bin settings applied for a) L2B Rayleigh-clear and b) L2B Mie-cloudy on 22 January 2020 for the FM-B period, as shown via plots of the L2B HLOS winds for one L2B file.

#### 2.1.4 NRT mid-2020 FM-B dataset

This period from 4 April until 20 September 2020 was chosen because an important L2B processor bias correction was applied: the M1 mirror-temperature-dependent HLOS wind bias correction (see section 3.2 and Weiler et al. (2021)). The processing chain applied was: L1Bp v7.08 (operational data from PDGS) and L2Bp v3.30 produced on the ECMWF L2/Met PF back-up server (which was an almost exact copy of the operational server). The back-up server L2B data was used because the M1 mirror temperature-dependent-bias correction scheme was available 16 days earlier than in operations (a pre-operational testing period); thus providing a larger dataset. The L2B data is equivalent to the operationally produced NRT L2B data since 20 April 2020 (when the M1 T-bias-correction went operational); baselines 09 and 10. L1Bp 7.08 and L2Bp v3.30 from operations (baseline 08) for the period 4-19 April 2020 i.e. without the M1 T-bias-correction, was also used to verify the impact of the M1 T-bias-correction compared to the existing ECMWF bias correction method (look-up table as function of argument of latitude and longitude).

Miscellaneous information to note on the L2B data for this test period:

- Instrument testing periods:
  - An Instrument Telescope Refocus (ITR) took place on 25-29 May 2020. This led to a period of very large biases in the HLOS winds.
  - An M1 mirror temperature thermal control test on 6-10 July 2020. This led to a period of very large biases in the HLOS winds.

#### 2.1.5 First reprocessed FM-B dataset: July to December 2019

The DISC and PDGS produced a first officially reprocessed L2B dataset using L1Bp v7.08 and L2Bp v3.30 (baseline 2B10) covering the early FM-B period from 29 June to 31 December 2019. This was released in L2B BUFR format in late 2020, and hence applied in OSEs. The early FM-B NRT ground processing issues (listed in section 2.1.3) were resolved. Appropriate L2B Rayleigh (AUX\_RBC\_L2) and Mie calibration (AUX\_MRC\_1B) files were applied for this dataset, generated from early FM-B period calibration information. This dataset applied the M1-temperature-based bias correction (see section 3.2), updated daily to remove any global offset bias trends and the M1 T related sub-orbital bias variations for both Rayleigh-clear and Mie-cloudy winds. The DISC assessments of the O-B statistics show that the reprocessed Rayleigh-clear winds are improved relative the NRT dataset for both systematic and random error statistics (Abdalla et al. (2020)). The Mie-cloudy winds have similar random error statistics, but improved bias (however wind-speed dependent biases (imperfect MRC non-linearity) were not yet resolved for this reprocessing). This first reprocessing covers an important period for assessing the NWP impact of Aeolus HLOS winds because the FM-B laser had its largest atmospheric path signal of the mission, and hence the largest NWP impact for Aeolus (as demonstrated in Section 4.5).

#### 2.1.6 Second reprocessed FM-B dataset: July 2019 to October 2020

The DISC and PDGS produced the second reprocessed L2B dataset in summer 2021 using baseline 11: L1A/L1B v7.09.1 and L2Bp v3.40.2 i.e. the processor deliveries from summer 2020. It covered the FM-B period from 28 June 2019 to 10 October 2020. This was made available to ECMWF in L2B BUFR format in August 2021, and hence applied in the OSEs. The L2B data maximum horizontal averaging length-scale is set to exactly one Basic Repeat Cycle (BRC) for Rayleigh observations (classic grouping), meaning the HLOS winds are produced from up to 30 smaller horizontal-scale (~2.8 km)

measurements (covering 12 seconds), resulting in observations of up to ~83 km horizontal extent (given the N=30, P=19 setting applied at the time; see appendix 9.10). For Mie observations the maximum horizontal averaging length-scale was set to 12 km (allowing accumulations of up to 4 measurements).

Appropriate L2B Rayleigh (AUX\_RBC\_L2) and Mie calibration (AUX\_MRC\_1B) files were applied for this dataset, both generated from early FM-B period calibration information. This dataset, like the first reprocessed dataset, applied the M1-temperature-based bias correction (see section 0), with daily updates to remove any global offset bias trends and the M1 temperature related sub-orbital bias variations for both Rayleigh-clear and Mie-cloudy winds. The Mie-cloudy wind-speed dependent bias (Mie non-linearity) was not yet detected and corrected.

### 2.1.7 Late 2021/2022 NRT FM-B dataset

This dataset covered the period from 13 December 2021 until 28 September 2022. It used L2B products derived from: baseline 13 (L1Bp v7.11/L2Bp v3.60); baseline 14 (L1Bp v7.12/L2Bp v3.70) since 29 March 2022 and baseline 15 (L1Bp v7.13/L2Bp v3.80) since 13 September 2022. These were produced in NRT by the operational Ground Segment. The Mie wind-speed dependent bias improved on 1 July 2021 with the application of the atmospheric path Mie-nonlinearity correction derived via NWP.

### 2.1.8 Third reprocessed FM-A dataset: September 2018 to June 2019

This dataset covered the period from 4 September 2018 until 16 June 2019. It was the first official reprocessing of the early FM-A period, which was the Commissioning Phase until end of February 2019. It used L2B products derived from: baseline 14 (L1A/L1B v7.12 and L2B v3.70) which was the early 2022 delivery. Note that there is a data gap due to the switch-off of the laser from 14 January 2019 until 15 February 2019 (see Figure 4). The Mie-cloudy wind-speed dependent bias was resolved in this reprocessing campaign via the application of the atmospheric path Mie-nonlinearity correction. Also, worth noting is that the flagging of winds as invalid during periods of moon-blinding was switched on in the L2B processor to avoid sporadic biases associated with de-pointing. This QC results in gaps in the tropical coverage every two weeks.

### 2.1.9 Late 2022/early 2023 NRT FM-A dataset

This data was produced after the switch-off of the FM-B laser (4 October 2022) and the second switch-on of the FM-A laser during October and November 2022. This dataset covered 1 December 2022 until 30 April 2023. The FM-A data being provided to NWP users without a L2B blocklist<sup>13</sup> defines the start date. The end date is when the mission came to an end in terms of the provision of NRT “operational” L2B BUFR data for NWP use. There was a data gap of around 9 days after 20 February due to the instrument going into Standby Mode (a blocklist was removed on 2 March 2023). L2B data was produced at baseline 15 (L1Bp 7.13/L2Bp v3.80) and at baseline 16 (L1Bp v7.14/L2Bp v3.90) since 20 April 2023. This period contains the best random errors of the mission for NRT data (see Section 3.3). However, the increasingly large number of hot-pixels causes sporadic biases on a selection range-bins (despite 8 DUDes per day). Also a new issue on a specific set of pixels (issues with quantum efficiency, O. Lux (DLR) reports) has led to a large bias on range-bin 16 for Rayleigh-clear winds since 22

<sup>13</sup> The L2B blocklist refers to a method of flagging (via AUX\_PAR\_2B) all HLOS wind results invalid during a time-period in the L2B EE product (and consequently L2B BUFR), so that non-nominal data is not operationally assimilated by NWP centres.

December 2022. The L2Bp QC flags the range-bin invalid due the range-bin specific mean(O-B) check. Therefore, the dataset has a missing range-bin at around 6-10 km altitude (depending on the RBS) since 22 December 2022. However, when B16 came into operations the dark-current correction was switched off (via L1Bp) for the specific pixel of concern, which led to the data for range-bin 16 being used since the bias was sufficiently small (but not zero).

## 2.2 ECMWF's HLOS wind observation operator

The Aeolus HLOS (horizontal line-of-sight) wind observation operator (forward model) applied in the ECMWF Integrated Forecasting System (IFS) is:

$$v_{HLOS} = -u \sin \phi' - v \cos \phi' \quad (1)$$

A derivation is provided in Appendix 9.12. The HLOS wind is a linear function of the NWP model's zonal wind component ( $u$ ) and meridional wind component ( $v$ ) interpolated to the observation geolocation point.  $\phi'$  is the spherical coordinate system azimuthal angle<sup>14</sup>, describing the line-of-sight pointing of the laser projected onto the horizontal plane. The observation geolocation information, as provided via the L2B BUFR data, is that applicable at the observation's horizontal centre-of-gravity for time, latitude and longitude and the vertical centre-of-gravity for the altitude (geometric height with respect to EGM96 geoid) and azimuthal angle.

The  $u$  and  $v$  wind components of the ECMWF model at the observation geolocation are obtained by the default IFS data assimilation interpolation method. The time-interpolation is done by assigning observations to 30-minute time slots during the 4D-Var window, hence a set of meteorological fields (on 137 vertical levels) is available from the non-linear trajectory every 30 minutes at the centre of the timeslot. Therefore, the worst case the time difference between observation and model equivalent is 15 minutes. Note that the high-resolution (~9 km horizontal grid) model trajectory runs with a time-step of 7.5 minutes. The horizontal spatial interpolation of these model fields to the observation latitude and longitude uses bi-dimensional twelve-point interpolation for the full resolution outer loop trajectories and bi-linear four-point interpolation for lower resolution 4D-Var minimisation trajectories.

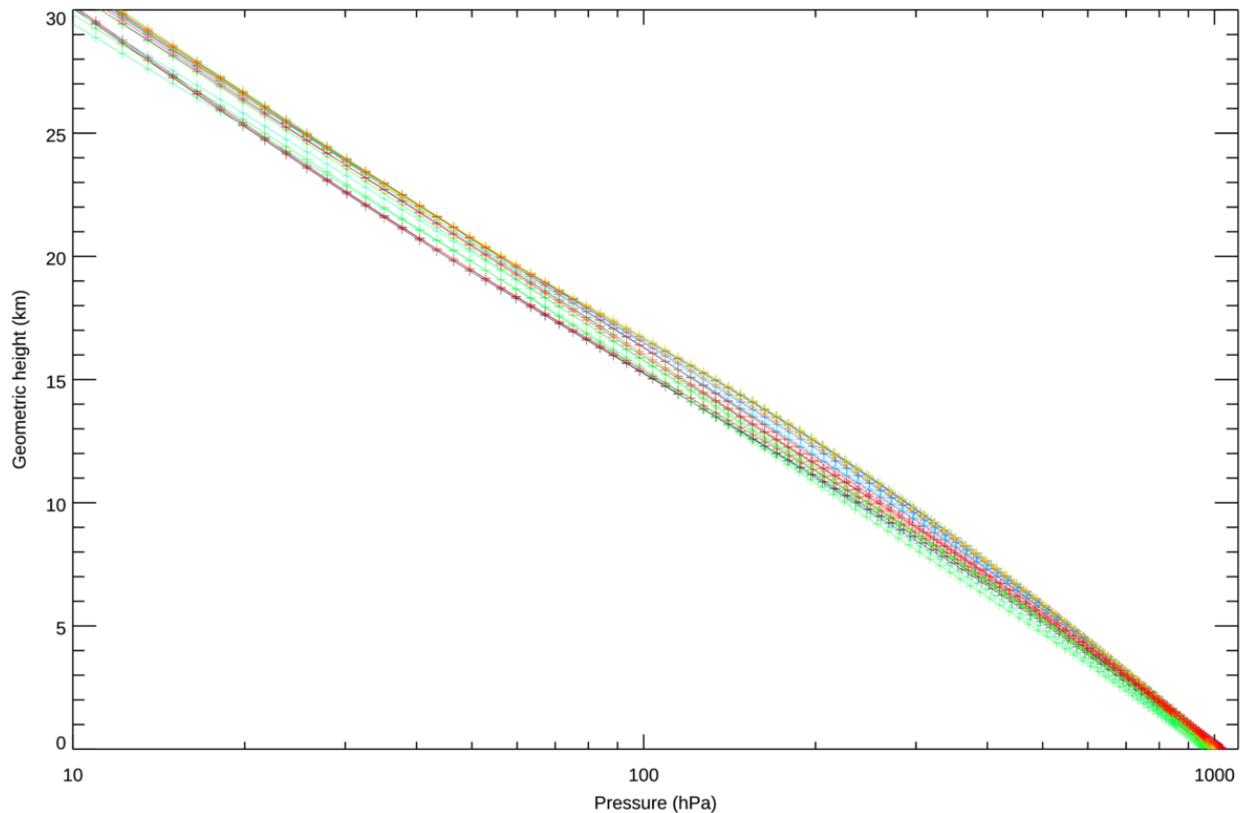
The resultant model vertical profile of  $u$  and  $v$  components as a function of pressure at the observation's time and horizontal position is vertically interpolated to the L2B vertical centre-of-gravity geometric height. This is done by converting the observation geometric height to a pressure value. The observation's geometric height is first converted to geopotential using a normal gravity formula (Somigliana's equation) and then the geopotential is converted to a pressure for the observation (using standard ECMWF data assimilation conversions, see ECMWF (2018)). Next, the model profile of  $u$  and  $v$  components as a function of pressure are linearly interpolated as a function of the logarithm of pressure to the observation pressure; hence the HLOS wind observation operator (equation 1) is applied to the interpolated model  $u$  and  $v$  components, treating the observation as a point-like wind.

The reliance on a priori background forecast information could be reduced if geometric height is used as the vertical co-ordinate rather than using the model to convert it to a pressure value. That is, if we forward modelled the geometric heights of model levels and then interpolated the model winds to the observation as a function of geometric height. This should be implemented as an option in the IFS, so the impact can be tested. One reason for using pressure as the vertical coordinate is that other wind

<sup>14</sup> The azimuthal angle is measured clockwise from North and applies to the horizontal projection of the target to satellite pointing vector. The convention for Aeolus is that (H)LOS wind is positive (negative) if wind blowing away (towards) the instrument.

observations in the IFS, e.g. radiosondes, aircraft, Atmospheric Motion Vectors also use pressure, making comparisons of O-B statistics from different measurement techniques easier to compare.

To aid the reader in converting from geometric height to pressure, Figure 7 shows the relationship between pressure and the calculated geometric height for the ECMWF model for a selection of meteorological conditions across the globe.



*Figure 7. The variation of geometric height (altitude) with pressure from ECMWF's IFS global model (which has 137 vertical levels) for a sample of meteorological conditions across the globe, along a simulated Aeolus orbit (different coloured lines). This can be referred to throughout the document to help convert pressure to altitudes.*

The geometric height vertical spacing of the ECMWF model's 137 levels is shown in Figure 8. In the IFS 137-level configuration, the vertical level spacing varies with altitude, being: < 200 m in the lower troposphere (<3 km); 300 m in the upper troposphere; 400 m at 50 hPa (~20 km), 800 m at 10 hPa (~30 km); and 3 km near the model top (~80 km). This vertical grid is more finely spaced than the typical Aeolus RBS, however the effective vertical resolution of the model is probably several vertical levels.

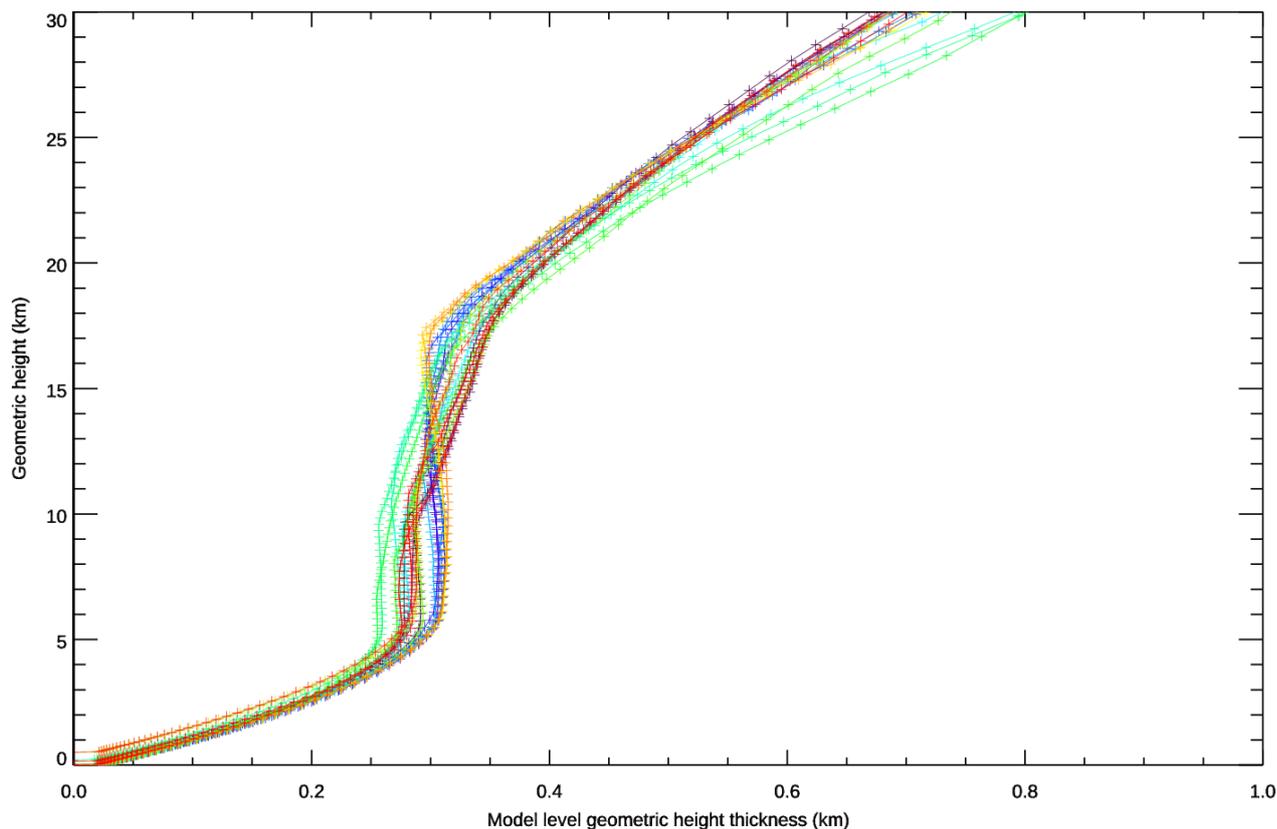


Figure 8. The vertical spacing between ECMWF's IFS 137 levels in geometric height (km) versus geometric height (km) for a sample of meteorological conditions around the globe (different colours).

The vertical wind component,  $w$ , is ignored in the nominal IFS HLOS wind operator because it is assumed to be negligible in the L2B retrieval to produce HLOS winds. However, it could be included in the IFS HLOS wind operator using the following equation:

$$v_{HLOS} = -u \sin \phi' - v \cos \phi' - w \tan \theta_e \quad (2)$$

Where  $\theta_e$  is the elevation angle (see Appendix 9.12), which is available via the L2B BUFR files.

The ECMWF model is thought to have a rather inaccurate representation of the generally small term  $w$  (compared to  $u$  and  $v$ ), certainly on small horizontal scales associated with convection, and it is therefore perhaps of little benefit to include it (although this has not been investigated in detail<sup>15</sup>). The vertical wind component is not used by any IFS observation operators at the time of writing. However, the potential impact of vertical motion on the Aeolus HLOS winds should be considered for future assessment. In terms of the Aeolus observation,  $w$  may contribute significantly to the LOS wind (and hence an error in the L2B HLOS wind) in specific conditions: strong convergence/divergence such as convective areas or in strong gravity waves. However, most of the vertical motion associated with strong convective, large-scale frontal ascent or even terminal (sedimentation) velocity of hydrometeors occurs within or below optically thick clouds, which Aeolus cannot measure, due to the optical signal being strongly attenuated. Ignoring the vertical wind component, with an elevation angle of 53 degrees (typical) leads to a bias in HLOS wind of approximately  $-1.33w$ .

The IFS HLOS wind observation operator effectively assumes that Aeolus winds are point-like (Dirac delta function) wind component measurements. This is probably a reasonable approximation in the horizontal dimension because the effective resolution of ECMWF's global model is on the order of

<sup>15</sup> Some preliminary investigation was done in Appendix 9.4 of (Rennie, 2016).

4-8 times the grid spacing (see Abdalla (2013)). The current ECMWF high resolution grid spacing at the time of writing  $\sim 9$  km and even less in the Ensemble of Data Assimilations (EDA), Isaksen et al. (2011). The 4D-Var inner loop (analysis increments) resolution is significantly coarser than the outer loop. The 4D-Var outer loop effective resolution is in reasonable agreement with the along-track horizontal resolution of Rayleigh-clear winds ( $\sim 83$ - $88$  km); note that across-track the laser beam has a diameter of only 10 m (somewhat complicating the picture). Note that the Mie-cloudy winds at typically 10-20 km along-track horizontal resolution (varied in NRT data during mission) and will typically have higher horizontal resolution than the ECMWF model.

The point-wind assumption is thought to be a poorer assumption in the vertical dimension, given that even the model's winds can be seen to vary significantly over the thickest (1-2 km) Aeolus range-bin setting, despite vertical wind shear being underestimated in the ECMWF model; see Houchi et al. (2010). In the tropical upper troposphere and along tropospheric frontal zones the ECMWF model's HLOS wind vertical shear is often more than 20 m/s per km ( $0.02 \text{ s}^{-1}$ ) and regularly  $\sim 0.03 \text{ s}^{-1}$  in the tropical upper troposphere, for example see Figure 9.

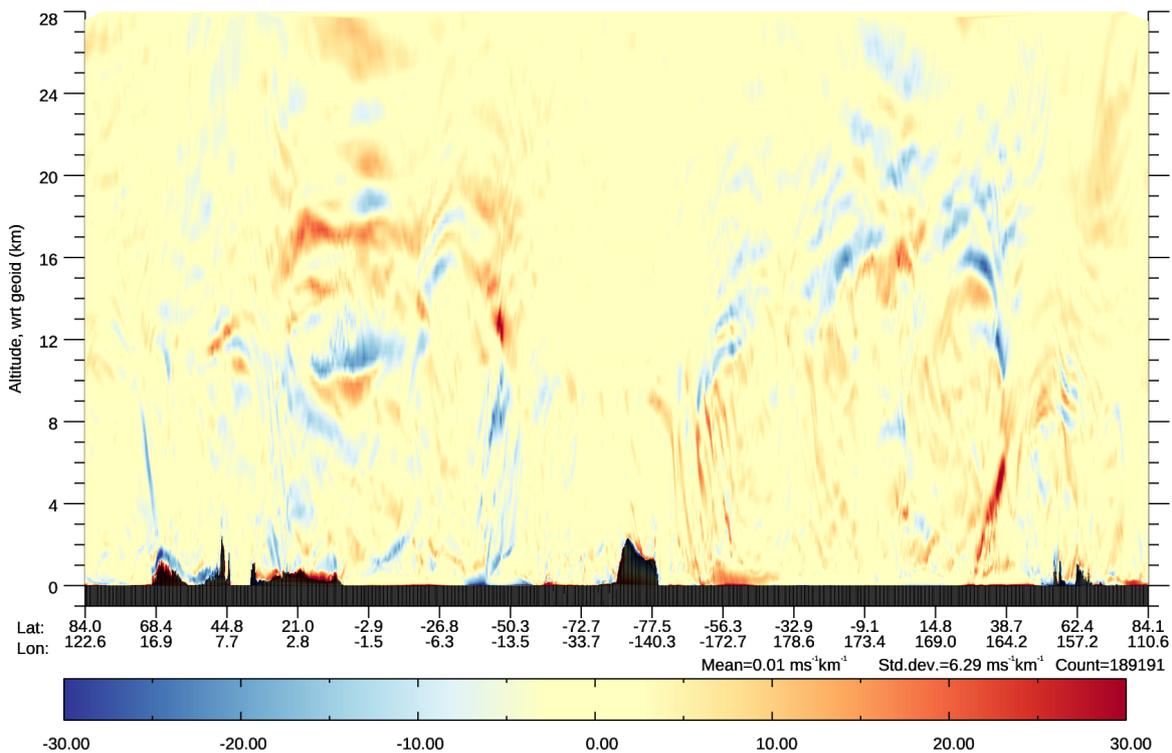


Figure 9. An example of the vertical shear of ECMWF model simulated HLOS wind ( $\frac{\partial HLOS}{\partial z}$ ) along an Aeolus orbit. The HLOS wind vertical shear has units of  $\text{ms}^{-1}\text{km}^{-1}$ .

An observation operator that accounts for the averaging kernel derived by the lidar equation should be more realistic. Such an operator for the Rayleigh channel in clear air conditions would account for the molecular attenuated backscatter variation throughout the vertical range-bin and horizontal extent of the observation. A solution for the Mie-cloudy winds (or Rayleigh-cloudy) relies on the ECMWF model's prediction of where particulate backscatters lie within the bin, which may have significant errors for clouds (and the information is absent for aerosol in NWP). Commanding thinner range-bins alleviates this issue in the vertical dimension, but at the cost of reduced overall vertical coverage given

the 24 range-bin limitation. An initial investigation of an observation operator accounting for the attenuated backscatter averaging kernel is covered in appendix 9.6; so far this only showed relatively small improvements in the differences of forward modelled wind to the observation.

Theoretically, the optimal extraction of Aeolus information would involve forward modelling the L1B Rayleigh channel A and B useful signals and the Mie useful signal levels (per pixel) via a full lidar equation-based observation operator (see appendix 9.6), i.e. using both the Doppler frequency shift and signal amplitude information of Aeolus. But this is significantly more complicated. Forward modelling the ALADIN instrument accurately enough to avoid biases with respect to the L1B useful signals is probably difficult (as demonstrated by the DISC's end-to-end simulator (E2S) and its imperfect representation of real Aeolus data despite the complexity of the simulation). Also, such a forward model would require good representations of aerosol and cloud optical properties from the IFS (complicated by Aeolus' co-polar backscatter), so perhaps this an area for research. Machine learning could perhaps be used for such a forward model if an unbiased forward model using the conventional approach proves to be difficult.

### 2.3 Observation-minus-background departure statistics

As discussed in the introduction, the comparison of the L2B HLOS winds to a state-of-the-art data assimilation system and forecast model is a very powerful method for determining the observation data quality. We use the ECMWF short-range forecast model equivalent HLOS winds to compute observation-minus-background (O-B) departure statistics.

The ECMWF background forecasts are of course not the truth, but they have a high level of accuracy (low bias) when averaged over large spatio-temporal scales. They also have a good precision, relative to Aeolus HLOS winds, thanks to the assimilation of millions of observations per day from the Global Observing System; being short-range forecasts and being based on accurate atmospheric modelling developed over many decades. Estimates of the magnitude of the forecast wind errors (both random and systematic) via O-B statistics for high quality measurements: such as radiosonde, aircraft and radar wind profile winds demonstrate this, see Figure 10. Global O-B biases are less than 0.3 m/s in u-wind and less than 0.15 m/s for v-wind (not shown). The standard deviations between these reference observations and the background are ~2.2-2.8 m/s; most of which can be ascribed to small-scale variability present in the observations that the model cannot represent i.e. representativeness error.

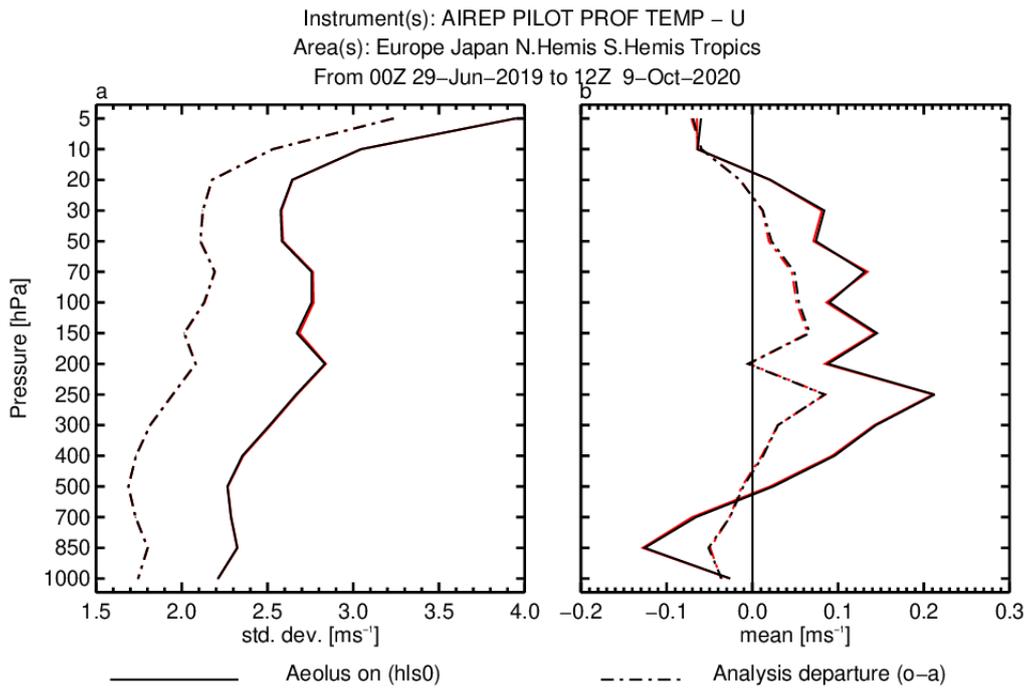


Figure 10. Zonal wind component, global O-B (solid lines) and O-A departure statistics (dot-dash lines) for conventional wind data (aircraft, radiosonde and radar wind profilers) for the period 29 June 2019 to 9 October 2020. The left plot is the standard deviation of departures, and the right plot is the mean of departures.

The spatial variations of background random errors and their correlations are reasonably well known via ECMWF’s EDA method for provides flow dependent inputs to the background error covariance matrix (B matrix): an example map of the average EDA spread for zonal wind at 200 hPa is shown in Figure 11. The background errors are largest near the tropics and in areas of convergence (convection). However, the magnitude of the EDA spread ( $1-\sigma$ ) tends to be underestimated compared to what we believe is the magnitude of the real background errors due to the low resolution of the dataset and a tendency for the ensemble to be under-dispersive.

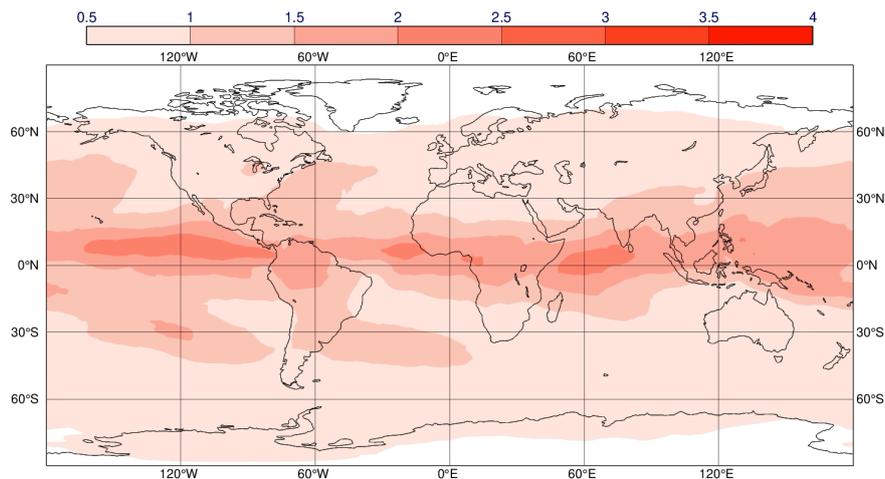


Figure 11. Average of the operational EDA spread (proxy for short-range forecast random error  $1-\sigma$ ) for u-wind component for the period 29 June 2019 to 30 December 2019 for model level 74 (which is roughly 200 hPa). Units: m/s.

Increasingly, knowledge on the model error component (Q matrix) in data assimilation is being exploited, via weak constraint 4D-Var (applied in the stratosphere at ECMWF at the time of writing). However, in areas with limited wind profile observation coverage (still large areas of tropics, SH and oceans), then the forecast wind error statistics are more uncertain.

An indication of the magnitude and spatial variation of the errors for u-wind component in the ECMWF analysis is provided from statistics of differences with respect to another provider of global NWP forecasts i.e. the Met Office analysis; see Figure 12. This indicates that systematic differences (and by assumption errors) are larger in the tropics and polar regions than elsewhere (reaching 2 m/s at smaller spatial scales). Random differences (multiply by  $\frac{1}{\sqrt{2}} \approx 0.7$  on assumption errors are of each NWP centre are similar and uncorrelated) are also larger in areas of convection (reaching 3-4 m/s), with some positive correlation with the spatial patterns of EDA spread (Figure 11), but larger in magnitude.

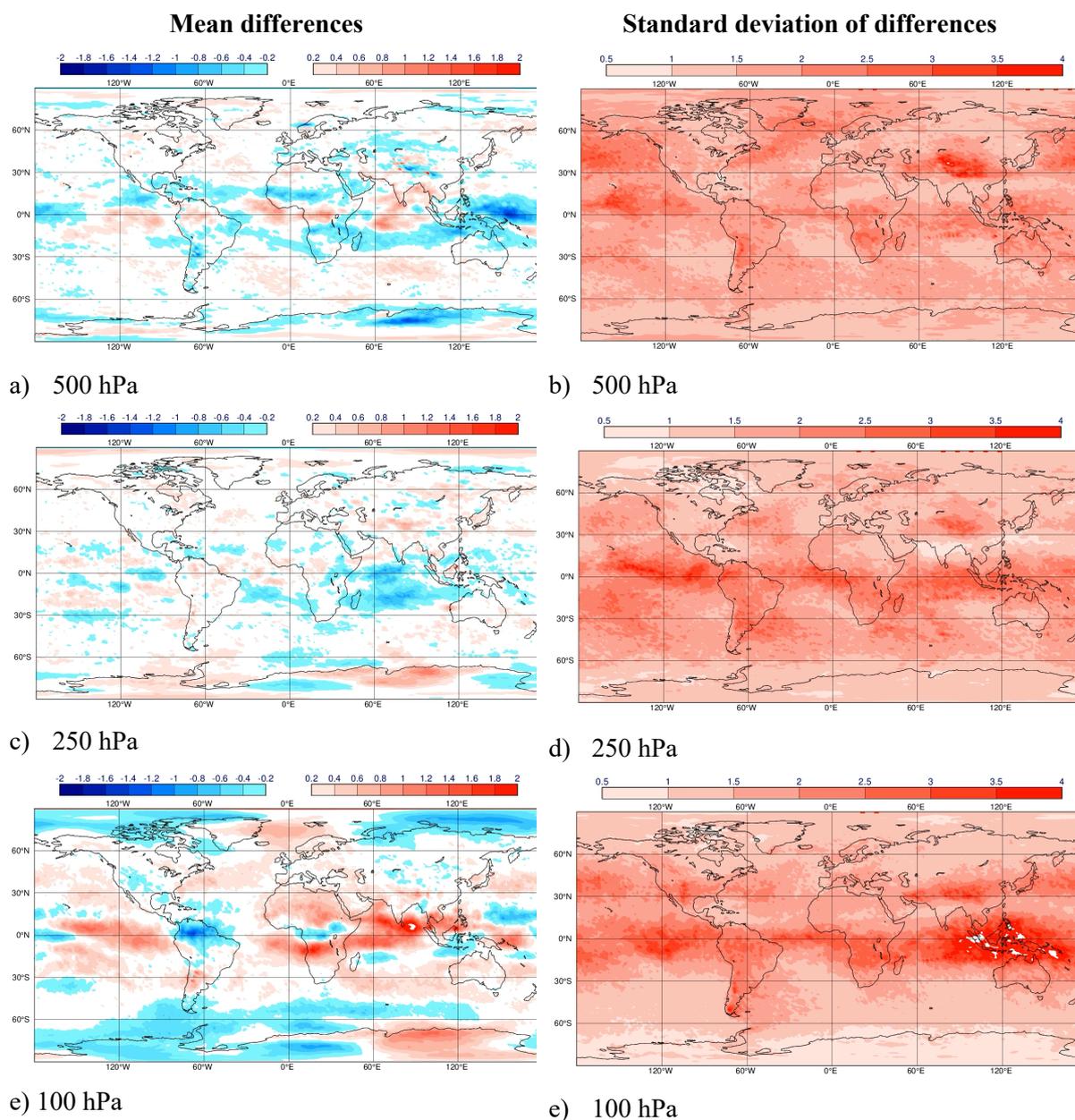


Figure 12. An impression of analysis  $u$ -component errors in NWP models via statistics of the differences between ECMWF and the Met Office for the period 1 January to 20 February 2023. Left column: mean(ECMWF – Met Office), right column: standard deviation(ECMWF – Met Office). Pressure level: 500 hPa for a) and b); 250 hPa for c) and d); 100 hPa for e) and f). Units: m/s.

The precision and accuracy of the L2B HLOS winds (the O in the O-B departures) depends on many factors, such as: atmospheric signal levels (i.e. laser energy, radiometric efficiency of the instrument and atmospheric attenuated backscatter); the accuracy of applied calibrations; Level-1 and -2 processor algorithms and quality control.

Two methods have been used to calculate O-B departure statistics: the first is bespoke to Aeolus using the auxiliary meteorological file (AUX\_MET\_12 file, shortened to AUX\_MET in the rest of the TN). The AUX\_MET contains vertical profiles of ECMWF model fields at predicted ground-track geolocations (as produced in NRT operations). The AUX\_MET is produced at high horizontal resolution T<sub>CO1279</sub> (see Malardel et al. (2016)), which corresponds to a grid spacing of around 9 km, sampled every 3 seconds (~22 km) along a predicted orbit, at 137 vertical levels.

The AUX\_MET is an input to the L2B processor. It is required to perform the Rayleigh-Brillouin Correction (Dabas, 2008). The L2B processor uses the nearest neighbour temperature and pressure forecasts from the AUX\_MET to select an appropriate atmospheric path Rayleigh-response versus frequency calibration curve from the AUX\_RBC\_L2 look-up table. The AUX\_MET also provides some other variables (not used in the wind retrieval) such as the  $u$  and  $v$  wind components as a function of geometric altitude, which can be forward modelled to HLOS wind (as in equation 1); therefore, providing a convenient and collocated wind reference to calculate O-B departure statistics. Since the AUX\_MET data is already interpolated to the predicted ground-track geolocations (using the same method as the Aeolus observation operator in the IFS), the nearest-neighbour profile (within a defined time-window) is selected for the O-B departure statistics. The NRT AUX\_MET data production (done during the mission via the L2/Met PF) contains forecasts from 0 to 30 hours which are updated every 12 hours. The timeliness of the forecast was about 3 hours – generated from the 6-hour (short cut-off) assimilation cycle analyses.

In this TN, the AUX\_MET derived O-B statistics are calculated and plotted via bespoke Aeolus monitoring tools that have been developed over several years pre- and post-launch, originally developed for assessing Aeolus E2S simulated EE data via the chain-of-processors (Rennie, 2016). These tools are best suited for small datasets (up to ~1 day), due to slow performance when handling the large data files (using large ASCII dumps of L1B/L2B data via a tool in the L2B processing software package).

L2B product data quality flags are used for some Quality Control (QC) decisions. The most important of which is the overall data quality flag for each wind result: if invalid, then the HLOS wind is rejected from the statistics. To remove gross errors which strongly affect the non-robust metrics (e.g. mean and standard deviation) some further quality control (QC) has been found to be useful predictor, especially for the Mie-cloudy winds (prior to B16). Particularly the L2Bp estimated HLOS wind error (standard error), which is derived via error propagation from L1B useful signal levels to L2B HLOS wind using shot-noise (Poisson statistics) assumptions (see L1B and L2Bp ATBDs). QC using the size of the background departure (O-B), is avoided for this method of verification to maintain some independence from NWP (apart from special range-bin specific QC for the detection of biases associated with hot-pixels). QC thresholds for L2Bp estimated error are chosen subjectively, based on finding a balance between the fraction of observations that pass QC and the intrinsic quality of the dataset. Thresholds for estimated observation errors of around 8-12 m/s were found to be appropriate for the Rayleigh-clear HLOS winds and 4-5 m/s for the Mie-cloudy. However, this type of fixed threshold QC has its limitations due to varying useful signal levels associated with instrumental drifts during the

mission and hence varying fractions of data passing the QC and varying processing baselines affecting the L2Bp estimated error magnitude. The QC is particularly important for the Mie winds which had a significant fraction of gross errors<sup>16</sup> (even in winds flagged overall valid by the L2B processor) which the QC helps to remove. The Rayleigh observations have relatively few gross errors but do suffer from rather large random errors when strong attenuation by clouds above leads to low signal levels in the lower troposphere. This is a relatively small fraction of the data and is not very representative of the overall dataset, and hence it is useful to remove it from the statistics.

The second method of deriving departure statistics is the traditional NWP method, in which the L2B BUFR data is ingested into the ECMWF data assimilation analysis and the O-B and O-A (observation-minus-analysis) departure statistics are calculated. The data assimilation analysis can be a Research Department (RD) experiment or the operational (deterministic or high-resolution) data assimilation system. Aeolus was “switched on” for monitoring purposes in early December 2018, but blacklisted (not assimilated in the NRT operational analysis, meaning the data had no influence on the analysis state, but could still be compared to the operational background and analysis model state). Since 9 January 2020 the L2B Rayleigh-clear and Mie-cloudy winds were operationally assimilated at ECMWF, meaning the O-B, and especially O-A, statistics lost some independence with respect to the Aeolus data itself, but O-B is still valid for monitoring the quality of the data. ERA-5 did not assimilate nor monitor Aeolus, due to it using an older IFS cycle incompatible with Aeolus.

ECMWF’s generic observation monitoring software (obstat) can be used to calculate the O-B statistics from the archived data files (ODB) from operational and RD experiments. Also, the ODB SQL commands can be used, for example, to produce time-series of O-B statistics to be plotted with one’s tool of choice (Python has been used). ODB SQL can handle very large datasets, e.g. years of data, for generating the statistics required for whole- mission time-series plots.

In RD experiments, the non-linear forecast model is typically run at a lower horizontal resolution e.g. T<sub>CO</sub>399 (~29 km) than operations T<sub>CO</sub>1279 (~9 km) to save computational and archiving costs. Operational Aeolus monitoring statistics were available during the mission’s lifetime on ECMWF’s [website](#). Human-interpreted monthly L2B monitoring reports (using ECMWF’s O-B statistics) are available via ESA’s [website](#).

The QC applied in this type of monitoring is flexible: e.g. it can check of the size of the absolute O-B departure; the overall L2B validity flag or the L2B estimated error magnitude. Typical QC decisions found to be appropriate for the time-series plots of section 3.3, are to reject data in which the absolute value of O-B is greater than 25 m/s for the Rayleigh-clear and 10 m/s for the Mie-cloudy (both of which are reasonably relaxed levels of QC, on the order of  $\sim 3\sigma$ ) in addition to the overall L2B validity flag. This choice avoids dependence on the L2Bp estimated error magnitude, which has limitations as already discussed.

In data assimilation at ECMWF, the model background (or first-guess) check rejects observations when the O-B departure is greater than so many standard deviations of the predicted error (from the assigned observation error and the EDA spread derived background error estimate) and is aimed at removing outliers (see Järvinen and Andersson, 1999). This factor was nominally set to five for Aeolus L2B winds (a 5- $\sigma$  check) i.e. reject observation if:

$$O - B = y - H(x_b) > 5\sqrt{\sigma_O^2 + \sigma_B^2} \quad (3)$$

<sup>16</sup> A new method of QC using the Mie-core fit “residual error” significantly reduces the fraction of gross errors and hence reliance on L2Bp estimated error. This option is possible with B16.

The reason for the factor of 5 is unclear but it seems to have been inherited from the settings applied for conventional wind data i.e. radiosondes. There has been some benefit found with a stricter first-guess check with the OSEs for the third reprocessing.

Aeolus also applies variational quality control (VarQC) in data assimilation, using the assumption of a Gaussian plus flat gross-error distribution for the departures in the 4D-Var cost function. The Aeolus settings for VarQC result in a down-weighting of observations in the data assimilation if normalised departures are between  $4\text{-}5\text{-}\sigma$  (half-weight at  $4.71\text{-}\sigma$ ). The VarQC method and settings for Aeolus are discussed in appendix 9.9.

Further validation of the Aeolus data quality through comparisons with collocated observations from ground-based and airborne (remote sensing) instrumentation has been performed by the Aeolus Calibration and Validation teams (CAL/VAL teams) and the research community since the data became publicly available in May 2020. The CAL/VAL teams have made comparisons of Aeolus observations directly to other wind observations available in the Global Observing System (GOS), such as radiosondes, wind profilers, AMVs and to other NWP models from Meteorological Centres world-wide. In this way, the HLOS wind observation quality has been well characterized through many sources, and it also allows the influence of NWP model errors on the retrieval to be assessed. Other CAL/VAL teams have performed NWP impact assessments of Aeolus data using their own NWP models, which have been presented at many workshops/conferences over the years, and in peer-reviewed journal papers. The CAL/VAL team's verification methods have shown consistency with the assessment of L2B data quality done via the ECMWF model O-B departure statistics, providing confidence in our own verification methods. The results of the CAL/VAL validations are not discussed further here.

## 2.4 Observing System Experiments

The standard technique in NWP for evaluating the impact of a new observing system is to perform an Observing System Experiment (OSE). In such an experiment the data assimilation system is run twice, once including the observing system under investigation, and once excluding the data. The two sets of analyses and forecasts initialised from those analyses are then compared to investigate the impact on forecast skill of the observing system being studied (Andersson et al. (1991)). OSEs are the most reliable method to assess the forecast skill due to a change to the data assimilation system. Therefore, OSEs have been performed to assess the NWP impact of Aeolus L2B HLOS winds, which is a large component of this TN.

The Aeolus OSEs assimilated the full set of satellite and in situ observing systems available to ECMWF operations at the time, because we want to know the impact of Aeolus in a realistic NRT operational NWP system. When evaluating new observing systems, experiments may be done using a reduced observing system (i.e. taking out observations to leave the bare minimum) to see the influence more clearly; this has not been done for Aeolus. OSEs have been run for many periods over the whole mission, either with NRT datasets (or locally produced offline reprocessing) and using the official reprocessed datasets. Note that a peer-reviewed publication (Rennie et al. (2021)) using the OSE results from NRT early FM-B and mid-2020 FM-B data is available, but this TN is more up to date. Also, an ECMWF newsletter feature article was written which assessed the impact with the second reprocessed dataset (Rennie et al. (2022)).

OSE verification involves comparing the forecast skill of the changed observing system to that without the change. Forecasts are compared to a reference analysis which is considered the best estimate of the atmospheric state at a given time. Differences between forecasts and the reference analysis at the validity time of the forecast are assumed to represent the forecast error. When a significant change to

the analysis is made from adding a new observation type, the choice of reference analysis can be critical to how the impact is perceived. For example, with the “own analysis” method the analyses of the experiment are used to calculate the experiment’s forecast errors and the analyses of the control are used to calculate the control’s forecast errors. Adding new observations can increase the variability (activity) in the forecasts which may verify as a degradation compared to the control which did not use the observations, due to the control having more consistency between forecasts and analyses. This apparent negative impact becomes negligible after around two days in the extratropics, by which time forecast errors have grown sufficiently with time to be much greater than analysis errors. However, “own analysis” artefacts may persist for longer in the tropics due to slower error growth rates compared to the extratropics.

The ECMWF operational analysis can be used as the reference analysis (for both experiment and control). The operational analysis is usually run at higher resolution (and hence typically more accurate) than the than research mode OSEs. Also, the operational analysis uses the operationally applied observing system and EDA derived background errors of the day appropriate to that observing system, which should improve accuracy. For Aeolus, it was noticed that verification against operational analysis verification is sensitive to whether it was assimilated or not in the operational analysis. It has been shown that the short-range forecast impact for Aeolus “improved” when verified against operational analysis, for dates after Aeolus went operational (on 9 January 2020) compared to beforehand. Therefore analysis-based verification should be ignored at day 1 forecasts or earlier.

Another very important element in OSEs is verification against other observation types, in observation space i.e. O-B. O-A statistics. This involves assessing if the assimilation of Aeolus pulls the short-range forecasts (or analyses) closer to, for example: radiosondes, aircraft, passive microwave/IR sounders or GNSS radio occultation observations (amongst others). This method is critical to assessing whether the analysis-based verification at short-range forecasts is meaningful.

For Aeolus OSEs during the mission, after QC there were typically 205,000 HLOS winds assimilated per day (125,000 Rayleigh-clear and 80,000 Mie-cloudy, variable depending on signal levels and aerosol load). This is about a quarter the sample size of Atmospheric Motion Vector (AMV) winds assimilated per day (840,000 wind components in CY48R1 operations). Aeolus typically provided much less than 1% of the total number of observations assimilated at ECMWF.

#### 2.4.1 Commissioning Phase NRT FM-A dataset

##### 2.4.1.1 Experimental set-up

The first CP OSE applied the IFS code base of CY45R1<sup>17</sup>, with a 4D-Var outer loop horizontal resolution of T<sub>co511</sub> (~23 km grid spacing), inner loops of T<sub>L</sub> 255/319/399 with 137 vertical levels up to around 80 km altitude (these vertical levels are applied in all OSEs). The nominal ECMWF operational set of satellite and conventional observations was assimilated. The period for the OSE was from 12 September until 16 October 2018, which was chosen because the systematic errors in the L2B dataset were relatively stable with time and the instrument health was nominal (see section 3.1 for the O-B time series monitoring results). However, the OSE was later extended to the whole of the CP to assess the impact of an increasing observation bias. This period covers part of the Northern Hemisphere autumn and Southern Hemisphere spring.

---

<sup>17</sup> ECMWF CY45R1 was operational from 5 June 2018 until 10 June 2019. A list of official changes for this cycle is available here: <https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model>

The experiments are the same as the control, except that they additionally assimilated the Aeolus L2B HLOS wind observations. Two experiments were performed; one assimilating both the Rayleigh-clear and the Mie-cloudy observations and one assimilating only the Mie-cloudy observations; to determine their relative impact.

The assigned Aeolus data assimilation observation error were derived from the L2Bp estimated error (one value per wind result), which have been consistently shown to be meaningful by their mean values being highly correlated with the O-B standard deviation (as demonstrated in sections 3.1.1 and 3.1.2). However, they were found to be generally underestimates when used as data assimilation assigned observation errors, because they only account for instrument-based noise. Also, there are known to be some missing instrument related noise terms in the error estimate formulation e.g. effect of hot-pixels, frequency jitter, readout noise, detection chain offset correction noise<sup>18</sup>. If only an instrument noise estimate, then they do not account for model representativeness error, which varies with the NWP model and atmospheric conditions and the forward modelling assumptions, and any forward model errors. This probably account for some of the underestimation. It was therefore decided to increase the size of the L2Bp estimated errors by using a scaling factor, which was chosen to be 1.5 for the Mie-cloudy and 1.1 for the Rayleigh-clear based at the time on given experience from pre-launch simulation studies.<sup>19</sup> This initial modelling of observation error was basic and some refinements are investigated with the early FM-B data OSEs; see section 4.3.4.4.

#### 2.4.1.2 Quality control decisions and corrections

Being the first few months of NRT produced data of an Earth Explorer mission, various non-nominal corrections and QC decisions were necessary to try to maximise the usefulness of the available data. The applied QC and corrections were done in the pre-processing and screening phase before the data assimilation process, after the L2B BUFR data has been read into the IFS.

The applied QC decisions and data corrections that were appropriate for this specific dataset:

- Assimilate only Rayleigh-clear and Mie-cloudy winds (i.e., Rayleigh-cloudy and Mie-clear are rejected due to their generally poorer quality, as confirmed by O-B statistics).
- Observations with assigned observation errors (i.e. after scaling) greater than a threshold were rejected to reduce the number of gross-errors (which have some correlation with instrument noise predictions, especially for the Mie-cloudy winds) without rejecting too many usable observations:
  - Rayleigh-clear:
    - If pressure < 90 hPa, reject if assigned observation error > 11 m/s
    - If 90 < pressure < 200 hPa, reject if assigned observation error > 7.7 m/s
    - If pressure > 200 hPa, reject if assigned observation error > 6.6 m/s

<sup>18</sup> Readout noise and detection chain offset noise have been included in the L1B Rayleigh SNR since B16.

<sup>19</sup> More recent FM-B testing suggests 1.4 to be more optimal for the Rayleigh winds and 1.75 for the Mie-cloudy winds (however this was based on new smaller scale horizontal averaging of the Mie winds), and this has been further modified for the Mie to include representativeness error. It should be noted that investigations have shown that the L2B Rayleigh-clear estimated errors are too large in the polar summer by 20-30% due a bug in the L1B processors' estimate of the Rayleigh SNR (an erroneous factor in front of the solar background noise term). Therefore, in our OSEs we will have been giving too little weight to the Rayleigh winds in polar summer. This was resolved with L1Bp v7.08 and L2Bp v3.30.

- Mie-cloudy: reject if assigned observation error > 4.5 m/s
- The HLOS wind observation geometric heights required a correction for this period: by adding 250 m. This was due to a known error in the LOS (Line-of-sight) pointing knowledge during the CP (a Star Tracker calibration issue<sup>20</sup>).
- The Mie-cloudy winds were bias corrected by -1.35 m/s (a global constant offset) for improved global agreement with the ECMWF model winds. This correction was only valid from 12 September to 16 October 2018. The bias was likely caused by imperfect calibration e.g. noise during the calibration procedure or errors in the processing algorithms. Other NWP centres in the CAL/VAL have since confirmed that this correction was necessary (ruling out an ECMWF model bias).
- Specific pressure ranges for the Rayleigh channel were blacklisted to try to avoid range-bin dependent biased HLOS winds (which was caused by hot pixels). This method worked reasonably well but was imperfect, due to the difficulty of fixed altitude range-bins varying in pressure space along the orbit. The blacklisting performed was:
  - Rayleigh-clear:
    - Avoid range-bin 11 by rejecting data between 400-500 hPa for the whole period
    - Avoid range-bin 5 by rejecting data between 150-200 hPa after 4 November 2018
    - Avoid range-bin 15 by rejecting data between 700-800 hPa after 24 November 2018
  - Mie-cloudy:
    - Avoid range-bin 13 by rejecting data between 600-750 hPa after 21 October 2018
- Winds within 20 hPa (~160 m) of the ECMWF model's orography were discarded because occasionally ground returns were wrongly classified as wind observations with the version of the L2Bp used (fixed since then).
- Rayleigh winds with range-bin thicknesses of 250 m were rejected due to excessive noise. There seemed to be no benefit from this data.
- Rayleigh winds with horizontal accumulation lengths less than 60 km and Mie less than 5 km (1 measurement at the time) were rejected since they tended to have more gross-errors in O-B statistics.
- Specific periods, when the satellite AOCS (the Star Trackers) were commissioned and switched to the redundant side for calibration, had biased winds. These periods were blacklisted in the IFS:
  - Data from 03:00 UTC on 25 September 2018 until 14:51 UTC on 26 September 2018
  - Data on 9 November 2018 between 09:25 UTC and 15:25 UTC
- A method was employed to avoid duplicate observations present due to overlaps between orbital dumps. The first occurrence of the observation is selected for use (this is also applied in other test periods). Such duplicate checks are applied in all OSEs in this TN.

---

<sup>20</sup> This error was resolved on 26 February 2019 in operational data and was corrected in the third reprocessed data (released in late 2022).

Spatial thinning of the observations was not applied (as is done for many other satellite observations) due to insufficient evidence to justify its necessity (observation error scaling could mitigate any spatially correlated observation errors, which thinning typically tries to mitigate). That is, no evidence of strong spatial (horizontal or vertical) error correlations has been observed so far. As mentioned in section 2.2, a  $5\text{-}\sigma$  background forecast departure QC check is applied during the screening phase of the data assimilation process (some tuning of this was considered in later OSEs e.g. for the third reprocessing which covered this period, see Section 2.4.7.).

Most of the applied QC decisions and data corrections for this early CP NRT dataset were not required for the officially reprocessed FM-A dataset (third reprocessing), see Section 2.4.7.

## 2.4.2 NRT 2019 FM-A dataset

### 2.4.2.1 Experimental set-up

The applied IFS code base was a tagged version of CY46R1 from 26 March 2019<sup>21</sup>. The model outer loop trajectory was set to  $T_{CO399}$  (~29 km) and the inner loops to  $T_L 95/159/255/255/255$ . The nominal operational set of satellite and conventional observations was assimilated. The 2.5-month period of the OSE from 2 April to 14 June 2019 was long enough to improve the statistical significance compared to the short 1-month CP period. This period is meteorological spring to early summer in the Northern Hemisphere and autumn to early winter the Southern Hemisphere. The assigned observation errors used a scaling factor of 1.4 for the Rayleigh-clear winds (derived via diagnostic methods) and a range of scaling factors were tested for the Mie-cloudy winds.

### 2.4.2.2 Quality control decisions and corrections

As already discussed in section 2.1.2, improved AUX\_PAR\_2B QC settings allowed the manual flagging of non-nominal hot pixel affected range-bins (as detected via O-B monitoring). This resulted in roughly a quarter of range-bins being rejected.

Whilst the OSEs for this period were running, much was learned about the behaviour of the Rayleigh-clear wind biases, which is discussed in section 3.2, that caused large changes in bias in the NH polar regions over a matter of days. This led to the development of a bias correction scheme using the ECMWF model as a reference, which is described in more detail in section 2.4.3.2 for the early FM-B period. In testing this period, it was found to be more reliable to verify forecasts against the ECMWF operational analysis ( $T_{CO1279}$ ) — since it is more accurate than the lower resolution OSEs ( $T_{CO399}$ ). Aeolus was still not operationally assimilated during this period.

The QC decisions were the same as for the early FM-A period, other than the correction to observation geometric height was no longer required.

## 2.4.3 NRT 2019 FM-B dataset

### 2.4.3.1 Experimental set-up

The applied IFS code base was a tagged version of CY46R1 from 3 December 2019 and the operational blacklist file dated 10 December 2019 was applied. The nominal ECMWF operational satellite and conventional observations were assimilated. A selection of experiments was performed: Rayleigh-clear

---

<sup>21</sup> Cycle 46R1 went operational at ECMWF on 11 June 2019.

plus Mie-cloudy; Mie-cloudy only; modified Mie-cloudy assigned observation error and Rayleigh-cloudy only.

The model outer loop and trajectory was set to  $T_{CO399}$  (~29 km) and the inner loops to  $T_L$  95/159/255/255/255. The period tested is from 2 August until 31 December 2019, spanning late summer to early winter in the Northern Hemisphere and late winter to early summer in the Southern Hemisphere.

The assigned observation error model was initially as follows: L2Bp estimated observation error scaling of 1.4 for Rayleigh-clear and 1.75 for Mie-cloudy. The scaling values were based on “Desroziers diagnostics” (Desroziers et al. (2005)) and some tuning via OSE impact assessment. An improvement in the Mie-cloudy assigned observation error model was developed to account for representativeness error, which is explained in section 4.3.4.4.

### 2.4.3.2 Quality control decisions and corrections

The nominal settings employed in the FM-B dataset OSEs were:

- Bias correction using the ECMWF’s HLOS wind as a reference as function of orbit phase angle (argument of latitude) and longitude (see section 4.3.1). Rayleigh-clear and Mie-cloudy winds are independently bias corrected and updated every few days based on the previous week’s O-B statistics. The Mie-cloudy bias correction did not require the longitude variation and was found to be much more stable with time than the Rayleigh-clear. Further details on the ECMWF bias correction look-up table method can be found here on the CAL/VAL website (CAL/VAL users only):

<https://www.aeolus.esa.int/confluence/display/CALVAL/L2B+Data+Quality+Handbook#L2B+DataQualityHandbook-ECMWFgeneratedbiascorrectionarrays>

- No specific time periods were blacklisted. However, in retrospective monitoring it was noticed that significant biases occurred on 3 September 2019 due to a temporary switch from Star Tracker A to Star Tracker B. See the periods that could be considered for blacklisting here (CAL/VAL users only):

<https://www.aeolus.esa.int/confluence/display/CALVAL/Calval+Black+List>

- As in the late FM-A period, forecast verification was found to be more informative using the ECMWF operational analysis ( $T_{CO1279}$ ) as the reference rather than own analysis, by more clearly revealing the Rayleigh-clear tropical impact at shorter time ranges.
- QC decisions applied:
  - Only assimilate Rayleigh-clear and Mie-cloudy winds with a “valid” overall confidence flag
  - Do not assimilate data within 20 hPa (~160 m) of the surface (to try to avoid any undetected ground returns). *This is probably no longer required due to improvements in the L2B processing since then to better avoid ground returns.*
  - Avoid Rayleigh winds at pressure > 850 hPa. It is unclear if this was necessary, but there were some indications of degradation when using boundary layer Rayleigh-clear winds, which tend to be very noisy (and perhaps biased).
  - Do not use any Rayleigh-clear winds with:
    - Estimated errors (before scaling) > 12 m/s if the pressure ≤ 200 hPa (to try to allow data in the wintertime polar vortex)
    - Estimated errors (before scaling) > 8.5 m/s if the pressure > 200 hPa

- Horizontal accumulation lengths < 60 km (to reduce outliers)
- Vertical accumulation lengths < 300 m (to reduce outliers)
- Do not use any Mie-cloudy winds with:
  - Estimated errors (before scaling) > 5 m/s. This was found to be a suitable threshold to remove many gross errors.
- For the Rayleigh-cloudy wind experiments, do not use if:
  - Estimated errors (before scaling) > 5 m/s. This was found to reasonable QC threshold for removing gross errors.

#### 2.4.4 NRT mid-2020 FM-B dataset

##### 2.4.4.1 Experimental set-up

These experiments were done to assess the impact with the operationally applied Aeolus assimilation settings and investigate if they could be improved upon. Aeolus was operationally assimilated at ECMWF from 9 January 2020. Several experiments were performed: Rayleigh-clear plus Mie-cloudy with nominal assigned observation errors; modifying the assigned observation errors and tuning the quality control. For longer experiments covering 4 April until 20 September 2020, the applied IFS code was CY47R1.1 from May 2020 (with some modifications to the Aeolus BUFR to ODB code to allow modifications to the assigned observation errors) and the operational blacklist file dated 2 July 2020 was applied (this had to be updated during the experiment).

A short OSE was performed from 4-19 April 2020 with the operational CY46R1 IFS code to compare the impact of L2B winds with the M1 temperature bias correction method applied (see section 3.2) to the then operational ECMWF look-up table bias correction method.

In all experiments the model outer loop and trajectory was set to  $T_{Co399}$  (~29 km) and the inner loops to  $T_L$  95/159/255/255/255. The nominal ECMWF operational set of satellite and conventional observations was assimilated. The test period spans late spring to late summer in the Northern Hemisphere and late autumn to late winter in the Southern Hemisphere.

##### 2.4.4.2 Quality control decisions and corrections

- The experiments in which the L2B data applied the M1 temperature dependent bias correction (see section 3.2) did not require any further bias correction using the ECMWF model. Therefore the method using the argument of latitude versus longitude look-up table could be switched off. A short reference experiment that used the ECMWF bias correction look-up table method for a comparison to the M1 temperature dependent method is described in section 2.4.3.2.
- Two periods in which the data had large wind biases due to Special Operations Requests (instrument testing) (see section 2.1.4) were backlisted. This resulted in a loss of 10 days out of the 169 days (4 April to 20 September 2020) i.e. 6% data outage.
- QC decisions:
  - The nominal QC decisions were applied, as done for the early FM-B period OSEs: see section 2.4.3.2. Any further modifications will be explained with the results.

### 1.1.1. First reprocessed FM-B dataset: July to December 2019

#### 2.4.4.3 Experimental set-up

OSEs using the first official ESA (PDGS/DISC produced) reprocessed L2B data for the early FM-B laser period were performed. Aeolus had its largest atmospheric signal levels during this period, and it is therefore an interesting period for assessing the potential of DWL, especially with the improved quality reprocessed data, in comparison to the NRT dataset.

The IFS code was a tagged version of CY47R2 from 17 December 2020 and the operational blacklist file dated 2 November 2020 was used. The nominal ECMWF operational set of satellite and conventional observations was assimilated. Several experiments were performed using Rayleigh-clear plus Mie-cloudy, but with a range of observation error scaling assigned for the Mie-cloudy winds to find the optimum. Also, a new atmospheric-temperature-dependent bias correction for the Rayleigh-clear winds was tried (due to systematic errors with altitude) and testing of modifications to VarQC to use the Huber norm distribution instead of the usual Gaussian plus flat distribution (see appendix 9.9) was performed.

The model outer loop and trajectory was set to  $T_{CO399}$  (~29 km) and the inner loops to  $T_L$  95/159/255/255/255. The experiments cover from 29 June until 31 December 2019, spanning summer to early winter in the Northern Hemisphere and winter to early summer in the Southern Hemisphere. The start date is just over a month earlier than the original NRT early FM-B period (section 2.4.3).

The assigned observation errors baseline was as follows: L2Bp estimated observation error scaling of 1.4 for Rayleigh-clear and 1.25 for Mie-cloudy. However, an additional representativeness error term was tested for the Mie-cloudy winds which was found to improve the impact.

#### 2.4.4.4 Quality control decisions and corrections

The same nominal QC decisions as in the NRT early FM-B OSE were applied, however as mentioned already, some experiments used modified QC settings, as will be explained in the results section.

Additional ECMWF IFS-based blacklisted Aeolus data periods were applied to avoid bad quality data:

- **ITR calibration and TxA sensitivity test;** between 18:17 UTC on 1 July 2019 and 12:15 UTC on 2 July 2019
- **Bias related to reconfiguration to Star Tracker B;** between 07:50 and 14:40 UTC on 3 September 2019

This was the first reprocessed L2B dataset to apply the M1-temperature-dependent bias correction (see section 3.2) and therefore no further bias correction of a similar nature, i.e. varying with geolocation/time, was deemed necessary.

### 2.4.5 Second reprocessed FM-B dataset: July 2019 to October 2020

#### 2.4.5.1 Experimental set-up

OSEs using the second official reprocessed L2B data for FM-B laser up to October 2020 were performed. The first few months of this period was when Aeolus had its largest atmospheric signal levels and is therefore of interest for reassessing the impact with improved quality reprocessed data.

The applied IFS code was a branch from CY47R2 created on 17 December 2020 (same as in first reprocessed dataset experiments) and the operational blacklist file dated 23 June 2021 was applied. The nominal ECMWF operational set of satellite and conventional observations was assimilated. A control

without Aeolus assimilation and an experiment with Aeolus Rayleigh-clear plus Mie-cloudy assimilation on was run.

The model outer loop and trajectory was set to  $T_{Co639}$  (~18 km) and the inner loops to  $T_L$  191/191/255/255/255. The period tested is from 29 June 2019 until 9 October 2020 spanning all seasons, thus making this the longest OSE for Aeolus L2B winds. N.B. the OSE was continued after the end of the second reprocessing using the NRT L2B data until 26 September 2021. This was mainly for the benefit of the ESA funded impact studies of Aeolus on extreme weather and forecast busts (work of Giovanna de Chiara at ECMWF). The extended OSE was found to be very useful for Aeolus DISC work, particularly for assessing how the impact decreased with time as the FM-B signal levels decreased.

The assigned observation error model (scaling and representativeness errors) was the same as those applied with ECMWF operations in CY47R3 (since 13 October 2021).

#### 2.4.5.2 Quality control decisions and corrections

The QC settings were the same as applied in the first reprocessed dataset OSE. However, ECMWF IFS-based blacklisting was not required due to a new feature of L2B processor: blacklisting via the `AUX_PAR_2B` file, which can set the overall confidence flag to invalid during periods of bad data quality, so that the assimilation system rejects such data.

#### 2.4.6 Late 2021/2022 NRT FM-B dataset

##### 2.4.6.1 Experimental set-up

An OSE using the operationally produced L2B dataset from 13 December 2021 (after the  $N=15$ ,  $P=38$  on-board laser pulse accumulation settings were introduced) until 28 September 2022 was performed to assess if Aeolus still had some positive impact despite the L2B winds becoming increasingly noisy. There was also a further change to  $N=5$ ,  $P=114$  on 4 April 2022.

The applied IFS code base was CY47R3.3 and the operational blacklist file dated 11 January 2022 was applied (but updated as necessary for this OSE operating close to NRT). The nominal ECMWF operational set of satellite and conventional observations was assimilated. A control without Aeolus assimilation and an experiment with Aeolus Rayleigh-clear plus Mie-cloudy assimilation on was run.

The model outer loop and trajectory was set to  $T_{Co639}$  (~18 km) and the inner loops to  $T_L$  191/191/255/255/255.

The assigned observation error scaling and representativeness errors were as applied in ECMWF operations in CY47R3.3.

##### 2.4.6.2 Quality control decisions and corrections

The operational data assimilation applied QC criteria were applied.

#### 2.4.7 Third reprocessed FM-A dataset: September 2018 to June 2019

##### 2.4.7.1 Experimental set-up

OSEs using the third official reprocessed L2B data for the FM-A laser from September 2018 to June 2019 were performed. This was the first assessment of NWP impact for the early FM-A period (part of which was the Commissioning Phase) with a decent quality processing baseline (including the M1-temperature bias correction).

The applied IFS code is CY48R1.1 from 17/04/23 and was run on the ATOS HPC with the operational blacklist file dated 29 March 2023. The nominal ECMWF operational set of satellite and conventional observations was assimilated. A control run without Aeolus assimilation and experiments with: Aeolus Rayleigh-clear plus Mie-cloudy assimilation on (some additional ones with refined QC); Mie-cloudy only; were performed.

The model outer loop and trajectory was set to  $T_{Co639}$  (~18 km) and the inner loops to  $T_L$  191/191/255/255/255. The period tested is from 4 September 2018 (when the data quality settled down) until 16 June 2019 spanning a reasonable fraction of one year. However, there is a significant data gap from 14 January 2019 (06 UTC has BUFR of size zero) until 15 February 2019 (18 UTC has non-zero BUFR) due to the unplanned switch-off (survival mode) of the FM-A laser. Therefore the verification avoided the data gap.

The assigned observation error scaling and representativeness errors were as applied in ECMWF operations (which was CY47R3 at the time).

#### 2.4.7.2 Quality control decisions and corrections

The operational data assimilation applied QC criteria were applied.

#### 2.4.8 Late 2022/early 2023 NRT FM-A dataset

##### 2.4.8.1 Experimental set-up

OSEs using the NRT L2B dataset from 1 December 2022 until 30 April 2023 (last day of “operational” L2B data) were performed to assess the impact with the improved signal levels and good Rayleigh-clear noise evident for FM-A. A data gap (FM-A into survival mode) occurred between 20 February and 2 March 2023, hence that period was not used for verification purposes. Also, there was a small data gap due to the L2B blacklist being applied when B16 was implemented around 17-19 April 2023 (this small period was used for verification).

The applied IFS code base was CY47R3.4. The blacklist was dated 29 March 2023. The nominal ECMWF operational set of satellite and conventional observations was assimilated. A control without Aeolus assimilation and an experiment with Aeolus Rayleigh-clear plus Mie-cloudy assimilation on was run. An extra experiment using in addition the Rayleigh-cloudy winds was run from 3 March, after the Rayleigh-cloudy data quality improved due to a new correction for the scattering ratio using NWP as a bias reference (see Marseille et al. 2023).

The model outer loop and trajectory was set to  $T_{Co639}$  (~18 km) and the inner loops to  $T_L$  191/191/255/255/255.

An OSE with the assigned observation error scaling and representativeness errors were as applied in ECMWF operations was done. Additional experiments were performed with modified scaling of the observation errors to assess if more impact could be obtained – in particular because of reduced readout noise with the laser setting of  $N=5$ . One experiment was run with the Rayleigh-clear L2Bp estimated error scaling factor being 1.25 (rather than the nominal 1.4) and one with the Rayleigh-clear scaling at 1.25 and the Mie-cloudy scaling modified to 1.13 with a 1.8 m/s (rather than 2 m/s) representativeness error. The new error scaling parameters were guided by the results of Desroziers method.

##### 2.4.8.2 Quality control decisions and corrections

The operational data assimilation applied QC criteria were applied.

### 2.4.9 Summary table of OSEs for different periods and L2B datasets

Table 1 lists some pertinent details regarding the OSEs performed for various time periods and L2B datasets.

Table 1. Summary of the OSE properties.

OSE period name	Start	End	L1Bp version	L2B data source	L2Bp version	PDGS baseline	Data assimilation system	Bias correction method
<b>Commissioning Phase FM-A</b>	12 September 2018	16 October 2018	7.04	PDGS reprocessed	3.01	2B02	CY45R1; T <sub>co</sub> 511 (~23 km grid)  Control ID: h3kc  Experiment ID: h3ko	Constant off set bias correction for Microcloudy.
<b>2019 NRT FM-A</b>	2 April 2019	14 June 2019	7.04	Offline ECMWF produced	3.01 patch 1	N/A	CY46R1 (26 March 2020 branch); T <sub>co</sub> 399 (~29 km grid)  Control ID: h5zf  Experiment ID: h8j1	E(O-B) look-up table method, vs argument of latitude only (see section 3.4 for details)
<b>2019 NRT FM-B</b>	2 August 2019	31 December 2019	7.06, then 7.07 since 31 October 2019	PDGS NRT dataset	3.11, then 3.2 since 31 October 2019	2B06 then 2B07 since 31 October 2019	CY46R1.2 4D-Var outer loop resolution: T <sub>co</sub> 399 (~29 km grid)  Control ID: hbnq  Experiment ID: hdai	E(O-B) look-up table method (see section 3.4 for details)
<b>First reprocessing (FM-B)</b>	29 June 2019	31 December 2019	7.08	Reprocessed dataset from ESA (DISC, PDGS)	3.30	2B10	CY47R2 (17 Dec 2020 branch); 4D-Var outer loop resolution: T <sub>co</sub> 399 (~29 km grid)  Control ID: hil4  Experiment IDs: hil5,	L2B processor with M1 temperature based bias correction (see section 3.2)

							<p>hiuf (increased Mie assigned error), hldz (atm. T bias correction), hmmr (hldz + Huber norm VarQC), hmol (hldz + use Rayleigh-clear &gt; 850 hPa)</p>	
<b>Mid-2020 NRT FM-B</b>	4 April 2020	20 September 2020	7.08	Offline L2B processed dataset at ECMWF until 19 April, then PDGS NRT dataset since 20 April	3.30	2B10	<p>CY47R1.1 4D-Var outer loop resolution: T<sub>co</sub>399 (~29 km grid)</p> <p>Control ID: hel1</p> <p>Experiment IDs: hel4, hf0w (increase weight of Rayleigh), hf7f (relax blist QC),</p>	L2B processor with M1 temperature based bias correction
<b>Second reprocessing (FM-B)</b>	29 June 2019	9 October 2020 (for reprocessing) then followed by NRT dataset until 26 September 2021	7.09.1	Reprocessed dataset from ESA (DISC, PDGS), then followed by PDGS NRT dataset	3.40.2	2B11 (reprocessing), 2B11 and 2B12 for NRT data	<p>CY47R2 (17 Dec 2020 branch); 4D-Var outer loop resolution: T<sub>co</sub>639 (~18 km grid)</p> <p>Control ID: hlpv</p> <p>Experiment ID: hls0</p>	L2B processor with M1 temperature based bias correction
<b>Late 2021/2022 NRT FM-B</b>	13 December 2021	28 September 2022	7.11 then 7.12 since 29 March 2022	PDGS NRT dataset	3.60 then 3.70 since 29 March 2022	2B13/2B14	<p>CY47R3.3; 4D-Var outer loop resolution:</p>	L2B processor with M1 temperature

							Tco639 (~18 km grid) Control ID: hopy Experiment ID: hoyc	based bias correction
<b>Third reprocessing (FM-A)</b>	4 September 2018	16 June 2019	7.12	Reprocessed dataset from ESA	3.70	2B14	CY48R1.1; 4D-Var outer loop resolution: Tco639 (~18 km grid) Control ID: hwcw Experiment ID: hwbj	L2B processor with M1 temperature based bias correction
<b>Late 2022/early 2023 NRT FM-A</b>	1 December 2022	30 April 2023	7.13/7.14	PDGS NRT dataset	3.80/3.90	2B15/2B16	CY47R3.4; 4D-Var outer loop resolution: Tco639 (~18 km grid) Control ID: i02b Experiment ID: i029 Special Rayleigh-cloudy experiment used CY47R3b; Control ID: i0rd Experiment ID: i0rb	L2B processor with M1 temperature based bias correction

## 2.5 Forecast Sensitivity to Observation Impact (FSOI)

Forecast Sensitivity to Observation Impact (FSOI) (Langland and Baker (2004), Cardinali (2009) and Janiskova and Cardinali (2016)) measures the impact of an observation in terms of how its assimilation affects the short-range forecast error growth. FSOI measures the impact of observations in the context of all the other observations assimilated i.e. it does not require the “denial” of the observation type, as is necessary for OSEs. At ECMWF, the difference of global dry energy norm error (using the analysis as reference) at 36 and 24 hour forecast ranges is projected back onto the analysis using the adjoint forecast model and adjoint of the assimilation system. It relies on the accuracy of the model’s adjoint (with simplified dry and moist physical processes) which limits it to short-range forecast assessment; OSEs are required for medium-range impact assessment. By using the analysis for short-range impact assessment, FSOI is prone to “own-analysis” verification issues (as explained for OSEs).

The FSOI is summed over time and space for different observation subsets to compute their contribution to reducing the forecast error. The “relative FSOI” impact provides the relative impact of one observation group compared to the impact of all observations and is often used to rank different observation types. The absolute FSOI with the dry energy norm has units of  $10^{-5}$  J/kg. FSOI requires fewer forecasts than OSEs to reach a robust result, since it is a short-range forecast verification method using a global verification metric.

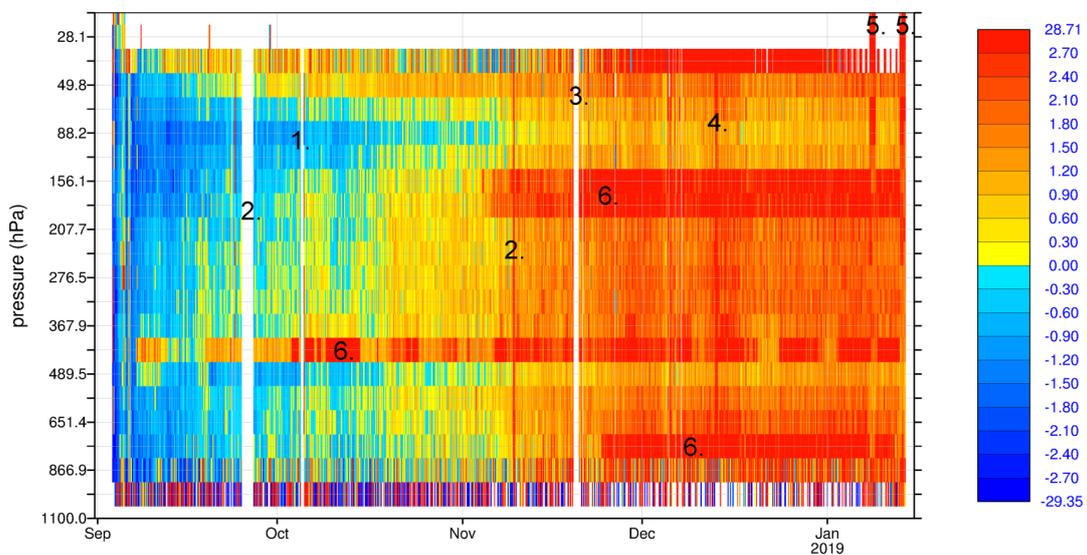
After Aeolus started to be operationally assimilated at ECMWF on 9 January 2020, it also was used in the FSOI suite (fsobs) running in parallel to operations (but behind NRT). Time-series of FSOI are useful for monitoring how Aeolus impact varied during the mission. An fsobs suite can also be run from a Research Department OSE experiment. This was done with a first reprocessed early FM-B experiment (with Rayleigh-clear and Mie-cloudy assimilated (experiment ID: hil5)) and with the second reprocessed FM-B experiment (with Rayleigh-clear and Mie-cloudy assimilated (experiment ID: hls0); to allow assessment of Aeolus in a period with its greatest signal levels.

### 3 Assessment of L2B data quality by NWP-based monitoring

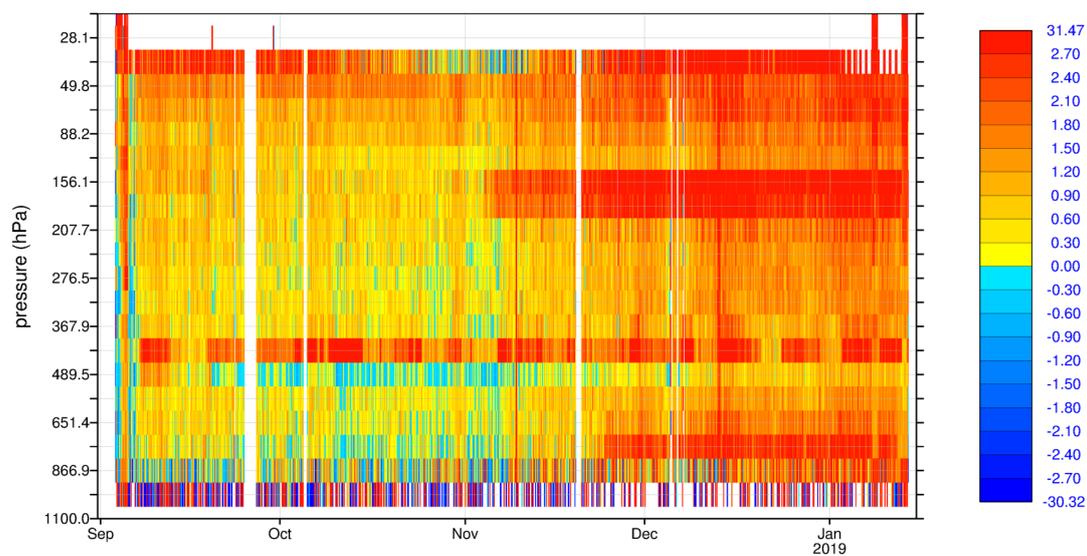
#### 3.1 Monitoring of Aeolus for the Commissioning Phase using the NRT dataset

##### 3.1.1 Rayleigh-clear HLOS wind O-B time variations and anomalies

A time-series Aeolus O-B departure statistics from NRT data were calculated via an offline assimilation experiment (before ECMWF operational monitoring, unavailable during the CP) in which the L2B data was blacklisted but Aeolus departures were still calculated (see section 2.3). Figure 13 shows the L2B Rayleigh-clear HLOS wind global statistics of mean O-B binned by time (in 3 hourly time slots) and pressure (pressure bins chosen so that the pressure axis is linearly proportional to altitude, see Figure 7 for a reference). The mean O-B was found to depend on whether the satellite was in the ascending or descending phase of its polar orbit, hence the statistics are split accordingly in Figure 13 a) and b).



a)



b)

Figure 13. Global mean(O-B) as a function of time (every 3 hours) and pressure (a selection of pressure ranges from surface to near 24 km altitude) for the L2B Rayleigh-clear HLOS winds. The colour-scale has units of m/s. a) Ascending orbit phases and b) descending orbit phases. Numbered features are referred to in the text.

Figure 13 a) labels some anomalies in the data quality and availability; the associated causes are listed below:

1. Satellite on-board software anomaly period
2. Star Tracker problem periods
3. Data gap due to a transition from reprocessed to operationally produced BUFR data (not an instrument issue)
4. Flight model A laser cold-plate temperature test period
5. Testing of modified vertical range-bin settings
6. Range-bins with increased bias due *uncorrected* enhanced dark current in memory zone pixels (“hot” pixels).

In both ascending and descending orbit phases, there was a positive trend in the mean O-B with time. No evidence was found to suggest that the ECMWF model caused the bias trend, therefore it was due to L2B HLOS wind systematic errors. Differences in bias between ascending and descending phases were noted e.g. in September 2018 the ascending phase is negatively biased whereas the descending phase is positively biased. This was found to be partially due to a HLOS wind speed dependence to the bias resulting from imperfect L2Bp Rayleigh calibration (see section 3.1.3.1). The ascending orbit phases measure on average positive HLOS winds and descending phases measure on average negative HLOS winds, due to the prevalence of westerly zonal winds in the extratropics. Therefore, a HLOS wind speed dependent bias (slow bias) manifests itself also as differences in bias between ascending (negative bias) and descending phases (positive bias). Later testing using more appropriate calibration information (Rayleigh Response Calibration in mid-September 2018) resulted in a smaller wind speed dependent bias: a linear fit of wind speed dependent bias (L2B versus background) was improved by ~2%, corresponding to 1 m/s reduction in bias for a 50 m/s HLOS wind. A contributor to FM-A’s orbital phase dependent biases of the Rayleigh was discovered for FM-B data in late 2019, as is discussed in section 3.2.

Wind biases which are present on specific vertical range-bins were seen in the O-B statistics. This led the DISC Level-1 experts’ discovery of an unexpected instrument problem, which is now referred to as “hot pixels”. Increased dark current background and noise levels were found on specific pixels of the instrument’s Accumulation Charge-Coupled Device (ACCD). They are thought to be triggered by the space environment radiation exposure (Weiler et al. (2021)). The small changes in dark current background levels, when left uncorrected, are enough to cause significant range-bin dependent biases in HLOS wind, particularly for the Rayleigh channel. Hot pixel induced wind bias tends to fluctuate with time as the level of dark current varies, apparently randomly. The pattern mostly seen (by DLR’s monitoring of this) is that the dark current levels rise to a very high level when the event is first triggered, for it then to fluctuate and stabilise after a while at a low elevation level. Some pixels however take a long time to stabilise or keep on fluctuating. The positive bias associated with range-bin 11 (~400 hPa with CP RBS) was caused by a hot pixel with particularly strong fluctuations, as can be seen in Figure 13. In June 2019, a new Level 1B processor baseline went operational, that used a new type of dark current calibration, done regularly (so called DUDE procedures) to correct for the effect. There were initially four DUDEs per day (increased to eight per day in September 2021, which was necessary due to an increasing number of hot pixels), allowing for a reasonably accurate correction of the varying dark

current. However, temporary biases on specific range-bins still occur due to hot pixel jumps between DUDEs. Also, winds cannot be measured during DUDEs.

Figure 14 shows the global L2B Rayleigh-clear standard deviation of O-B as a function of time and pressure, highlighting how the L2B HLOS wind retrieval random error changed with time, because the model short-range forecast errors are rather stable in the global average. We show only the ascending orbit phase statistics, due to the descending result being very similar. The random errors were very large near the surface because of the thin 250 m (short-time) range-bins not accumulating much signal (also due to attenuation of signal by clouds above) with the CP RBS. The smallest standard deviations occur for range-bins near the 100 hPa level (~13-16 km), because of the use of 2 km thick range-bins around this pressure level (which have more time to accumulate more signal), as compared to the 1 km thick-range-bins at lower altitudes (see Figure 1). Doubling the range-bin thickness should reduce instrument noise by ~30% (see Appendix 9.3) due to roughly doubling the signal, but at the cost of halving the vertical resolution. It seems that the reduced counting noise from more signal outweighs the loss of resolution for the variance of O-B to reduce. The mid-tropospheric 1 km range-bins typically had O-B standard deviation of 4-5 m/s.

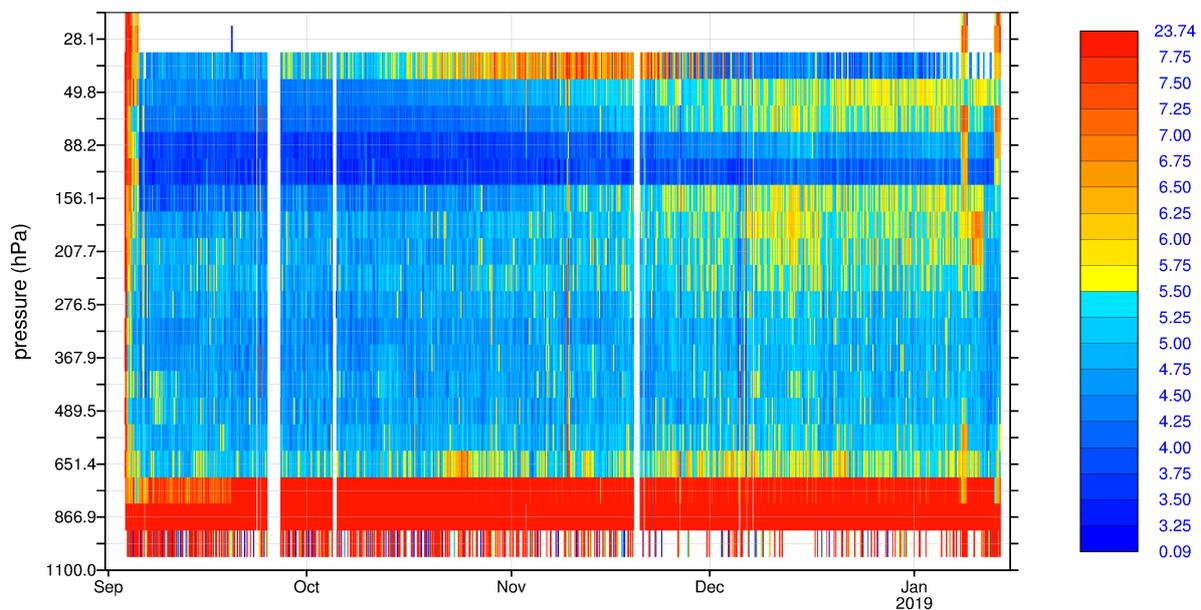


Figure 14. Global ascending orbit standard deviation of O-B as a function of time (every 3 hours) and pressure for the L2B Rayleigh-clear HLOS winds. The scale has units of m/s.

There was a general trend for the standard deviation to increase with time, particularly at the upper levels, e.g. at 200 hPa from late November 2018 onwards. This was caused mainly by a combination of increased solar background noise affecting the Southern Hemisphere observations (as the austral summer approached) and a steady decrease in the FM-A laser UV output energy with time which translates directly to a steady decrease in atmospheric path useful signal (as documented by ESA and DLR). It is not understood what caused the higher noise at the highest altitudes in November 2018.

The random errors are steadier for the mid-tropospheric levels than for higher altitudes. The generally increasing noise with altitude occurs because of the climatological profile of attenuated backscatter in clear skies; some examples of attenuated backscatter profiles (both clear and cloudy scenes) with altitude (forward modelled from AUX\_MET data) are shown in Figure 15. The signal drops off quickly above 10 km altitude (note the logarithmic x-axis).

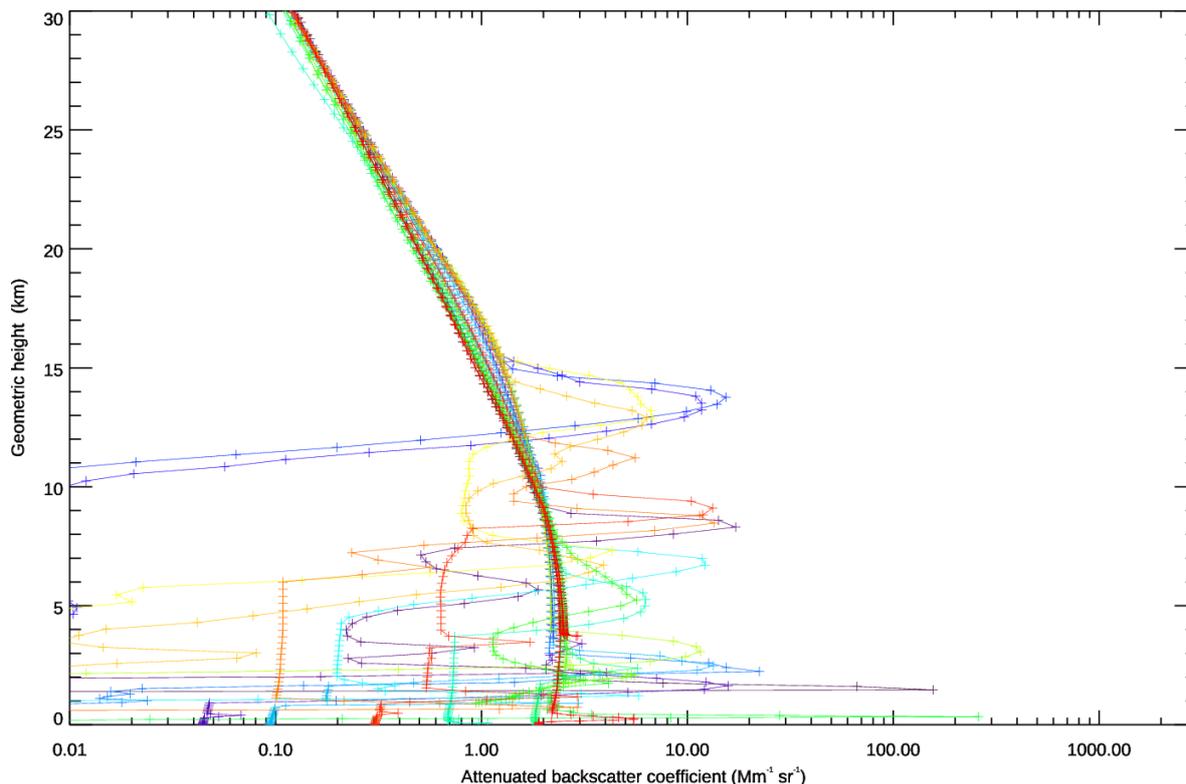


Figure 15. Examples of simulated profiles (at various locations around the globe) of attenuated backscatter coefficient as simulated from a combination of molecular and cloud particulate optical properties as derived from the ECMWF model (via temperature, pressure and cloud ice/liquid water content; AUX\_MET data). The smoother curves with altitude result from clear air (molecular scattering only); the departures from this are due to cloud backscatter and extinction.

Investigations at the time showed that the upper range-bin random error increase was strongly related to the increasingly poor L2B processor classification of measurement-level data into clear and cloudy conditions, because of increasingly noisy L1B measurement-bin scattering ratio associated with the decreasing laser energy (laser energy was around 20% lower in January 2019 than in September 2018). An improvement in the standard deviation is evident in the last few days of Figure 14 (11 January 2019), which was due to a change in the L2B processor settings which improved the clear-cloudy classification (via AUX\_PAR\_2B settings tuning; see details in the appendix 9.11).

Figure 16 shows the ascending phase Rayleigh-clear O-B statistics versus time for a mid-tropospheric pressure bin (319-368 hPa, ~8 km) which was not affected by hot pixels. The statistics are split into a) Northern Hemisphere (NH) extratropics and b) Southern Hemisphere (SH) extratropics. This pressure bin is roughly where the Rayleigh winds are most precise (of those with 1 km thick range-bins). A quantitative assessment of the bias and random error variations with time can be deduced from Figure 16. The bias increased with time, as already mentioned, as shown also in Figure 16. The bias settles in early December 2018, but to different levels in the NH (2-3 m/s) and the SH (1 m/s). After an initial decrease with time, the standard deviation of O-B settles in the NH after October 2018 to around 4.2 m/s. The initial decrease is due to the reduction in solar background noise as the boreal winter approached. The standard deviation of O-B in the SH increased a little with time; probably due to a combination of decreasing UV laser energy and increasing solar background noise as the austral summer approached.

Generally, the L2Bp estimated error (green lines of Figure 16) has similar time trends to the changes in the O-B standard deviation, implying it as a useful error statistic estimate. The Rayleigh HLOS wind errors are significantly larger than the estimated typical 1.5-2 m/s ECMWF background forecast errors, hence the standard deviation of O-B is dominated by the Rayleigh observation errors. Note that for the SH, the L2Bp estimated observation error appears to be too large in the increased solar background conditions in late November onwards i.e. it exceeds the standard deviation of O-B. An explanation for this discrepancy was discovered in mid-2019. It was due to an error in the L1B processor Rayleigh SNR formulation, which was fixed in L1Bp v7.08 and entered operations with baseline 1B08 on 2 April 2020.

The early mission bias and random error fluctuations appeared to settle down by 12 September 2018; hence this date was chosen as the start date for the CP NRT data OSE. In the NH, the increasing bias became large enough to perhaps degrade the NWP impact around 16 October 2018, hence the chosen end date for the period (section 2.4.1.1). In the SH the bias also increased with time, but at a smaller rate. The CP OSE chosen dates are also appropriate for the Mie winds, as demonstrated in the next section.

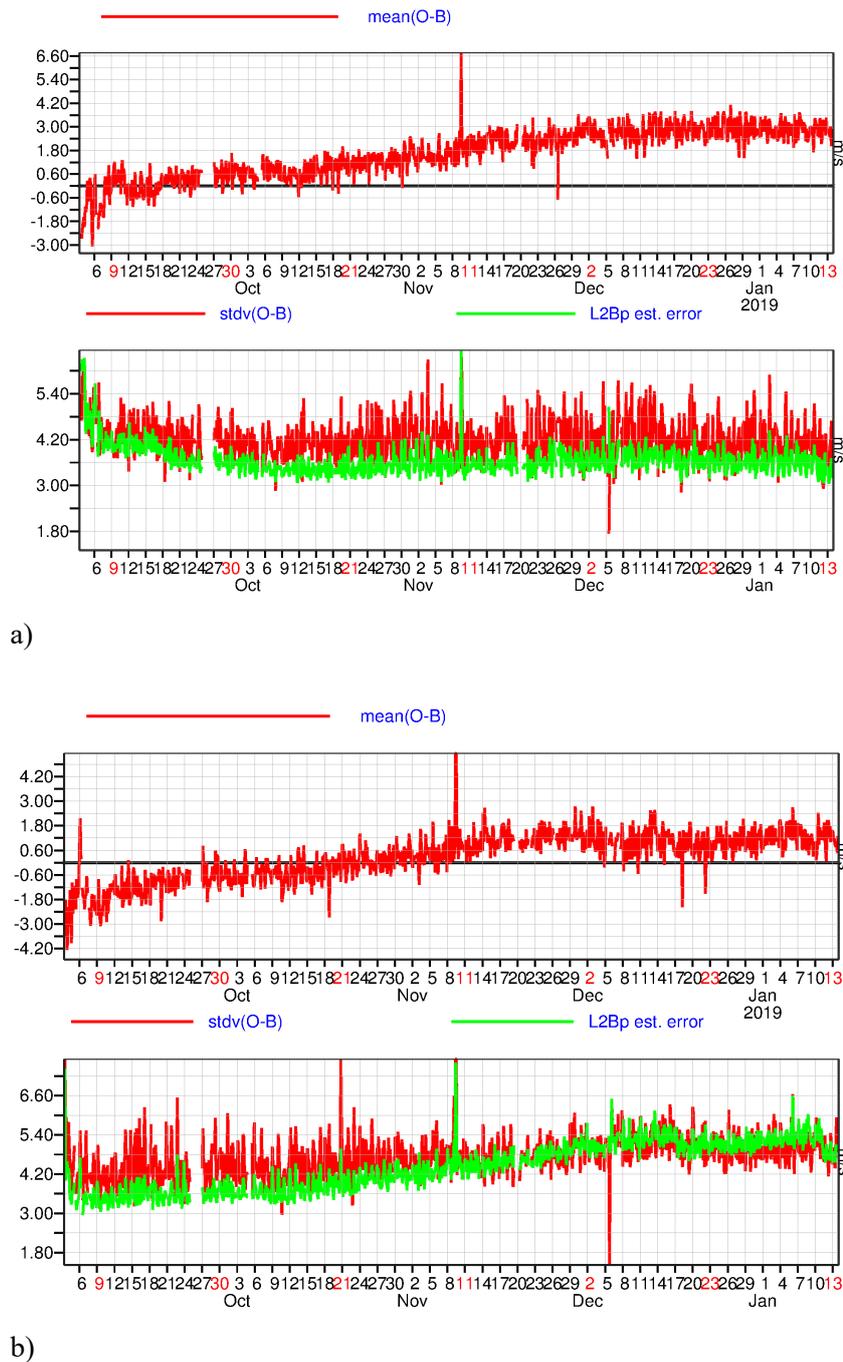
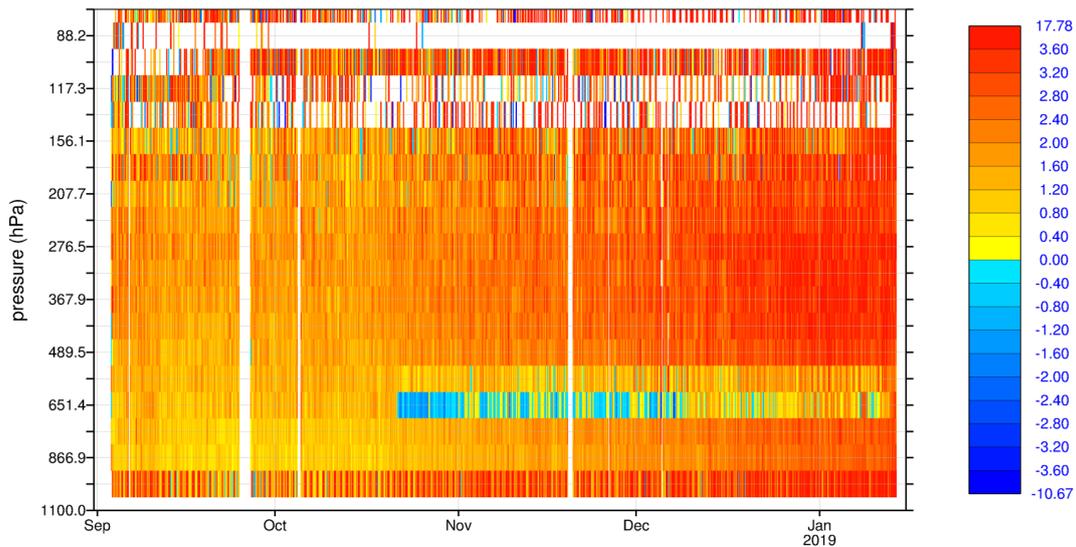


Figure 16. Time series (every 3 hours) of L2B Rayleigh-clear HLOS wind mean(O-B) (upper plot of each figure) and stdv(O-B) (lower plot of each figure) for the pressure range 319-368 hPa for ascending orbit phases only. a) Northern Hemisphere extratropics (20-90 degrees latitude) b) Southern Hemisphere extratropics (-20 to -90 degrees latitude). The green line is the L2B processor estimated error derived from signal levels assuming shot-noise and solar background noise.

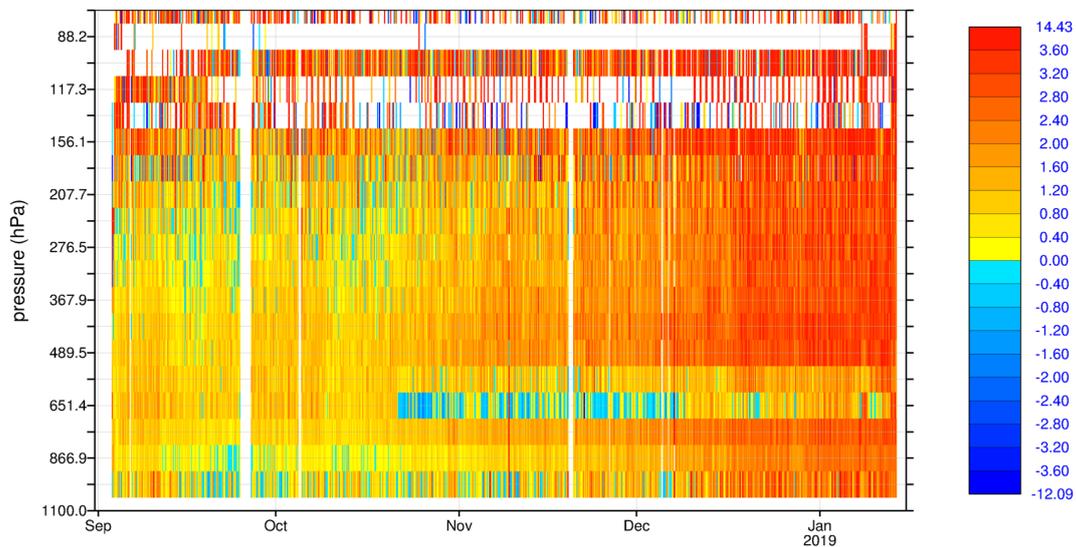
### 3.1.2 Mie-cloudy HLOS wind O-B time variations and anomalies

Figure 17 shows the L2B Mie-cloudy HLOS wind mean O-B statistics, split into ascending (a) and descending (b) orbit phases. The Mie winds are affected by the same instrumental and operational anomalies as labelled in Figure 13 a). For both ascending and descending orbits there is a general positive trend in the bias with time (as also seen for the Rayleigh) but starting from an already small

positive bias in September 2018 (there was no tuning to the ECMWF model for the Mie calibration). The descending phases are slightly less positively biased than the ascending, with the difference in September suggesting the Mie winds have a fast wind speed dependent bias with respect to the model, in contrast to the Rayleigh which had a slow bias for this period (this is confirmed in section 3.1.3). Figure 17 shows the appearance of negatively biased winds for observations at ~650 hPa from 21 October 2018 onwards, due to a hot pixel affecting Mie range-bin 13.



a)



b)

Figure 17. Global mean(O-B) as a function of time (every 4 hours) and pressure for the L2B Mie-cloudy HLOS winds (CAL/VAL dataset). The scale has units of m/s. a) Ascending orbit phases and b) descending orbit phases.

Figure 18 shows the global L2B Mie-cloudy standard deviation of O-B as a function of time and pressure. Only ascending orbit phases are shown because the descending was very similar. As expected, the Mie-cloudy winds have a greater precision than the Rayleigh-clear, by comparison to Figure 14 (note the different colour scales). The random error was smallest for the range-bins in the lower troposphere, which are 250 m thick with strong backscatter from the tops of optically thick boundary layer water clouds. The very lowest range-bins were contaminated by ground return signals, hence the larger noise; this issue was resolved with L2B processor version 3.10 (baseline 2B03). There was an obvious increase in standard deviation associated with range-bin 13's hot pixel which shows up at ~600-650 hPa. Apart from the hot pixel influence there was no significant trend in the standard deviation with time. This suggests the Mie-cloudy errors are not strongly limited by atmospheric path useful signal levels unlike for Rayleigh-clear (molecular backscatter) winds. This is due to the strong backscatter from most ice/water cloud types, see for example the size of cloud induced attenuated backscatter compared to the molecular in Figure 15; it can be a factor 10-100 greater.

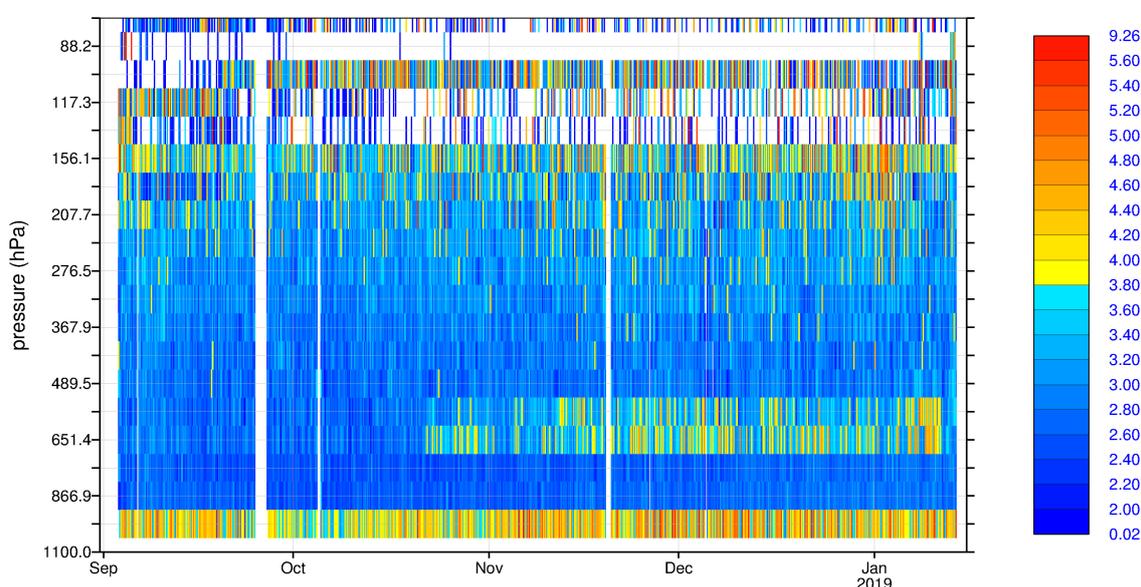
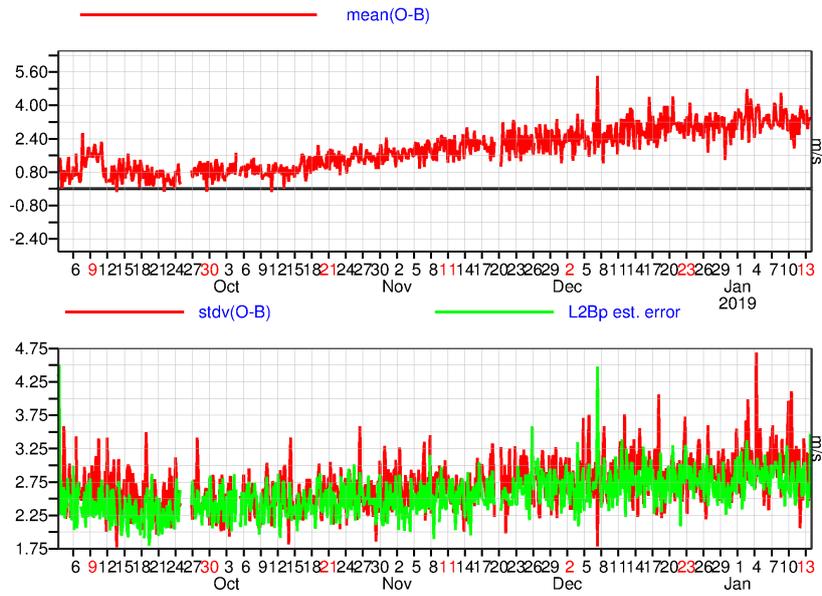
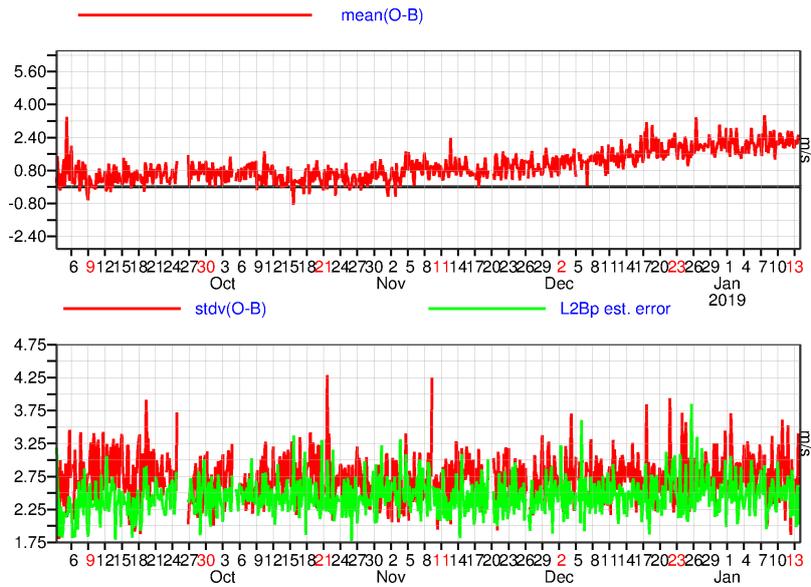


Figure 18. Global standard deviation(O-B) as a function of time (every 4 hours) and pressure for the L2B Mie-cloudy HLOS winds. The scale has units of m/s. Data for ascending orbit phases only.

Figure 19 shows statistics for a pressure bin in the lower troposphere (752-867 hPa, approximately 2 km altitude) for which there are large samples resulting from strong backscatter of planetary boundary layer (PBL) water cloud tops (and some aerosol); the figure is split into a) SH extratropics and b) NH extratropics (data for descending orbit phases only). This pressure range was not affected by hot pixels. It is evident that biases started to increase from early October 2018 in the NH, but much later in the SH; around late November. The bias settled in January to ~3 m/s in the NH and ~2 m/s in the SH. This behaviour has some similarities with the Rayleigh development, but the details differ. Some of the bias drift with time may be related to an imperfect AOCs LOS velocity correction (in the L1B processor at the time) which was applied during this period, which changes with the seasons (the same correction was applied to the Rayleigh winds). It was later turned off in operations with FM-B data due to being demonstrated to make the biases worse. The random error in the NH shows a slight increase with time, which is not understood but may perhaps be due to seasonal changes in cloud conditions. In the SH the standard deviation of O-B is steady despite decreasing laser energy. The magnitude and trends in L2Bp Mie estimated errors (green lines) strongly correlates with the standard deviation in O-B (red lines); suggesting it is a useful error estimate. Similar behaviour is found for other pressure ranges.



a)



b)

Figure 19. Time series (every 4 hours) of L2B Mie-cloudy HLOS wind mean(O-B) (top plot of each figure) and stdv(O-B) (bottom plot of each figure) for the pressure range 752-867 hPa for descending orbit phases only. a) Northern Hemisphere extratropics (20-90 degrees latitude) b) Southern Hemisphere extratropics (-20 to -90 degrees latitude). The green line is the L2Bp estimated error derived from useful signal information.

### 3.1.3 Comparing the quality of NRT data from mid-September 2018 to early January 2019

#### 3.1.3.1 L2B Rayleigh-clear

ALADIN's FM-A laser reported UV energy per pulse (PD74) dropped from approximately 61 mJ to 52 mJ from mid-September (15 September 2018) to early January 2019 (early January). Plots of O-B statistics as a function of altitude for the two periods are shown in Figure 20 (using the AUX\_MET O-B method; see section 2.3). There is a sample of around 80,000 observations in both cases (around 20 hours of data). The "robust standard deviation" in Figure 20 is the median absolute deviation (MAD) scaled by 1.4826 which is equivalent to the standard deviation for a normal distribution (Ruppert, 2010). The scaled MAD is less sensitive to outliers than the standard deviation, which is useful given that an NWP based first-guess check is not applied for these statistics.

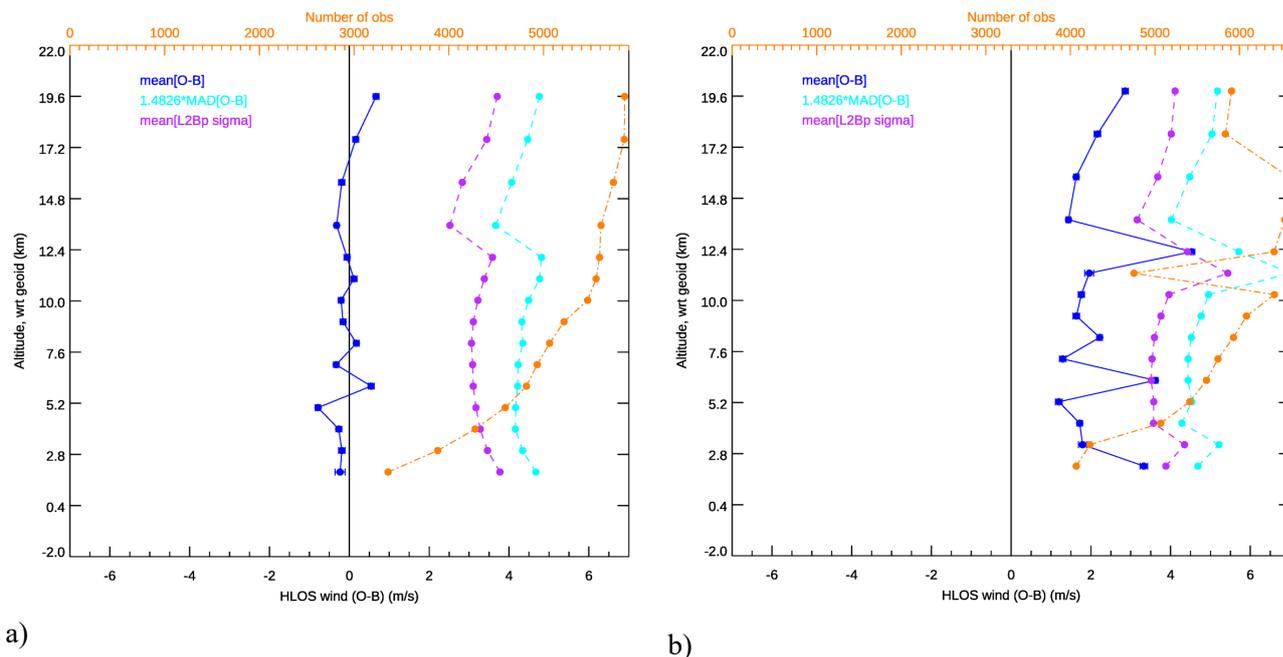


Figure 20. Global L2B Rayleigh-clear HLOS wind O-B statistics as a function of altitude. The dark blue line is the mean(O-B), the cyan line is the robust standard deviation of O-B, the purple line is the mean of the L2Bp estimated error and the orange line is the number of observations (top x axis). a) For mid-September 2018 b) for early-January 2019; both using the NRT dataset.

It is evident from Figure 20 that the accuracy and precision of the Rayleigh winds degraded with time during the CP. As already shown in Figure 13 and Figure 16, the Rayleigh winds were relatively unbiased in mid-September 2018 and became positively biased by early January 2019 (as already shown in previous plots). The bias in early January 2019 has spikes for several altitudes due the hot pixel affecting specific range-bins. The global average O-B bias in early January 2019 is about 2 m/s (discarding hot pixel affected levels). Given that global biases of ECMWF u-wind relative to radiosondes are rather small, at typically less than 0.3 m/s (as determined by radiosonde O-B departure statistics at ECMWF), then Aeolus Rayleigh observations must account for the bias change. The profile average robust standard deviation of O-B is 4.4 m/s in mid-September and 4.9 m/s in early January. Also, the counts are relatively reduced at specific range-bins in early January due to 8 m/s L2Bp estimated error QC which was applied, so rejecting more data than in September (overall 4% more rejections). Note that the L2Bp estimated random errors increase in the summer Poles due to larger solar background noise estimates, and hence the QC rejections are not a strictly fair comparison.

As mentioned earlier, some of the degradation in the Rayleigh-clear wind quality in early January 2019 was exacerbated by the L2Bp classification procedure (into clear and cloudy measurement-bins) which performed worse than in mid-September 2018, due to increased noise in the L1B processor measurement-scale scattering ratio. The quality loss was later mitigated to a reasonable extent by retuning the L2B processor scattering ratio classification thresholds. Increasing the scattering ratio threshold from 1.25 to 1.6 allowed a greater number of genuinely clear measurement-bins into the calculation of Rayleigh-clear winds, hence reducing the noise. Also, some improvement was obtained via QC of large spikes on measurement-bin useful signal levels, which are thought to be due to cosmic radiation affecting the instrument (given the tendency for this to occur in the South Atlantic Anomaly region). This led to the overall O-B robust standard deviation improvement of 0.3-0.4 m/s (as is evident in Figure 14 after the 11 January 2019).

Figure 21 shows the L2B Rayleigh-clear wind quality over the dynamic range of HLOS wind for the two periods. The Pearson's correlation coefficient for mid-September 2018 in a) is 0.96, but this reduced slightly in early January 2019 to 0.95 via b). That is, the Rayleigh winds performed well over the dynamic range in both periods. To estimate the wind-speed-dependent errors, we plot the mean(O-B) as a function of B in c) and d) of Figure 21. The use of Desroziers diagnostics on conventional u-wind observation departures (radiosondes, aircraft) gives a global average value for the background forecast random error of  $\sigma_B = 1.6$  m/s (not shown). Errors in the independent variable, in this case the background HLOS wind (B), of a regression scheme leads to biases in the estimated fit coefficients; for a simple linear regression, an underestimate of the fit coefficient known as attenuation bias or regression dilution occurs (Frost, 2000). Because for the Rayleigh winds  $\sigma_B$  is significantly smaller than  $\sigma_O$  it is reasonable to use B as the independent variable. Simulations of the attenuation bias induced by the assumed  $\sigma_B$  produce a slow bias of -1.5% (not shown). The linear fit coefficient in c) shows a slope error of -4% in mid-September 2018. This is significantly more negative than the -1.5% one would be expected without any real wind speed dependent biases and therefore it is assumed to be a real Rayleigh slow bias in mid-September 2018 of approximately -2.5%. This slow bias partially accounts for the ascending/descending orbit phase bias differences in September 2018 already noted from Figure 13. The linear fit coefficient in early January 2019 from d) is around -1%, so the slope error is closer to the expected -1.5% early January than in mid-September 2018. The reason for this is that the true Rayleigh response functions changed with time in a way which fortuitously improved the slope error with time when continuing to use the same imperfect Rayleigh calibration file (AUX\_RBC\_L2).

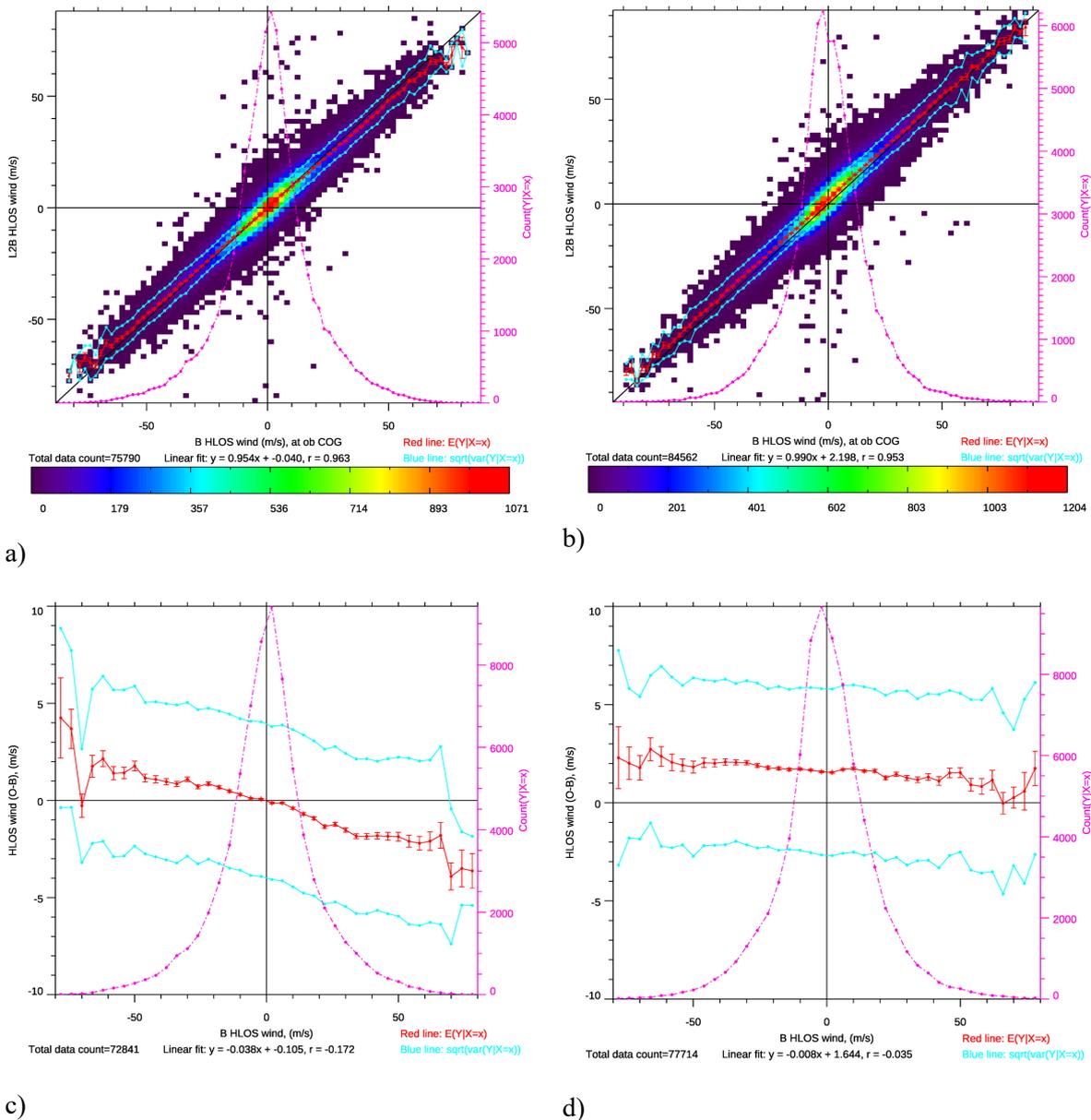


Figure 21. Global L2B Rayleigh-clear O-B statistics over the HLOS wind dynamic range. Dependence of L2B HLOS wind on background HLOS wind: a) in mid-September 2018 and b) in early-January 2019, as shown by 2D histograms. Dependence of mean(O-B) on background HLOS wind in c) mid-September 2018 and d) early-January 2019; the red-line is the mean(O-B) binned as a function of B (with the error bar showing the standard error of the mean), the cyan lines are the  $\pm 1$ -standard deviation of O-B. The pink lines are the data count.

Later in the CP it was demonstrated that the HLOS wind bias varies along the orbit (particularly for the Rayleigh-clear winds); the variation with orbital phase angle (argument of latitude) is shown in Figure 22 for the L2B Rayleigh-clear winds. Zero degrees argument of latitude corresponds to the ascending node equator crossing point. The descending phase of the orbit is between 90 and 270 degrees, elsewhere it is ascending. It can be seen in mid-September 2018 a), that the bias was more positive in the descending phase and more negative in the ascending phase; which concurs with the assessment of Figure 13 and Figure 21. There is also an orbital phase dependent bias in early January 2019 as shown in Figure 22 b) which peaks at the North Pole and is at its minimum at the South Pole. This was shown

(via L2B processor testing and personal communication with DLR) to be partially caused by an imperfect correction of the AOCS LOS velocity; but for the Rayleigh this is not the dominant source, the explanation was uncovered after the CP; see section 3.2. If the AOCS LOS velocity correction is switched off in the L2B processor the bias shows less variation with orbit phase in early-January 2019 (not shown). The reported satellite velocity correction maximum amplitude was small at  $\sim 0.16$  m/s HLOS in September 2018 but was up to 1 m/s HLOS in early January with a similar sinusoidal shape as a function of argument of latitude to the mean(O-B) bias. A similar variation of the bias on argument of latitude for the Mie winds in early-January is shown in Figure 25 b); since all winds apply the same AOCS LOS velocity correction. The cause of the imperfect satellite LOS velocity correction was found to be a bug in a coordinate transformation in the on-board software, which was corrected with L1Bp 7.09 (released in 2020).

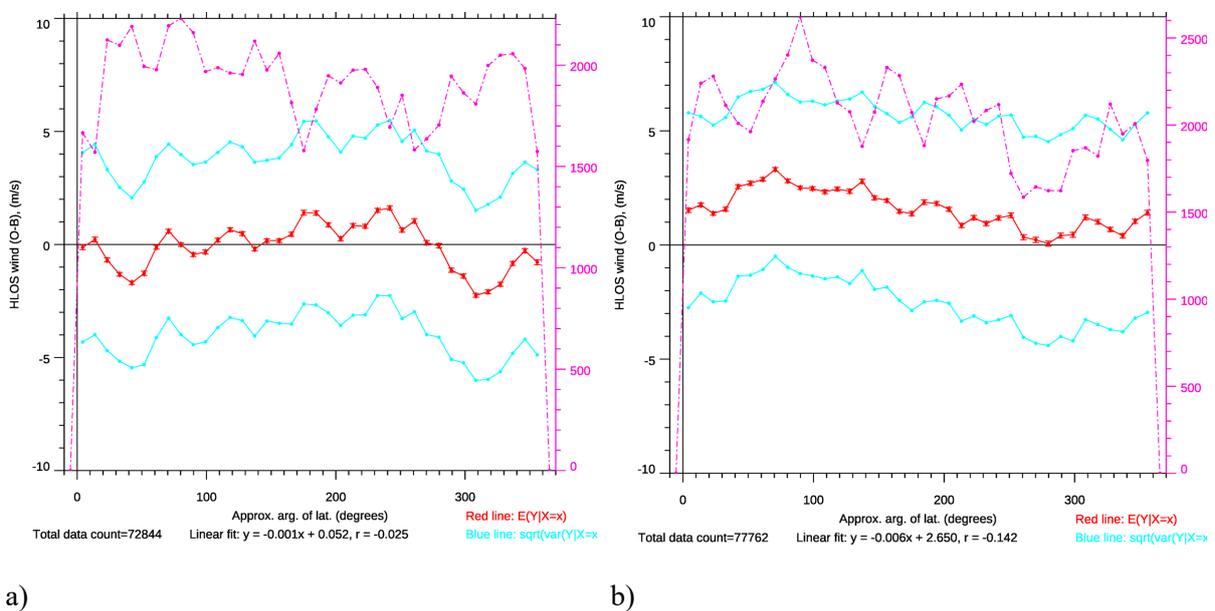


Figure 22. Dependence of the L2B Rayleigh-clear HLOS wind mean(O-B) on the orbital phase angle (argument of latitude) from zero degrees at the ascending node equator crossing point. The red-lines are the mean(O-B) binned as a function of argument of latitude (with the error bar showing the standard error of the mean), the cyan lines are the  $\pm 1$ -standard deviation of O-B and the pink line is the count. a) For mid-September 2018 b) for early-January 2019.

### 3.1.3.2 L2B Mie-cloudy

Global L2B Mie-cloudy O-B statistics for mid-September 2018 and early January 2019 are shown in Figure 23. The average bias in mid-September 2018 was 1.2 m/s, as is also evident in Figure 17. This bias grew significantly<sup>22</sup> and by early January 2019 had reached 4 m/s (after excluding the negatively biased data due to the hot pixel at roughly 3 km (range-bin 13)). The Mie-cloudy random error level as assessed by the L2Bp estimated error statistics is almost identical in mid-September 2018 and early January 2019 (at 1.65 m/s), however the profile average robust standard deviation increased from 3.25

<sup>22</sup> It will be interesting to understand the source of this bias drift for the early FM-A period once the first officially reprocessed datasets for this period become available. This was a significant bias drift which could be due to e.g. an internal path Mie response drift due to angular changes or frequency changes with time, meaning that the imperfectly known Mie non-linearity calibration was sampled differently with time.

m/s to 3.49 m/s, which may be associated with the hot pixel (see the peak in robust standard deviation at  $\sim 3$  km). To reiterate the Mie random error for the two periods is similar and does not show an obvious effect of the laser energy (and resultant atmospheric path signal) decrease, as also shown in Figure 18. As already discussed, it is assumed that the backscatter signal from clouds is sufficiently strong for the emitted signal level decrease to not be the limiting factor for Mie random errors.

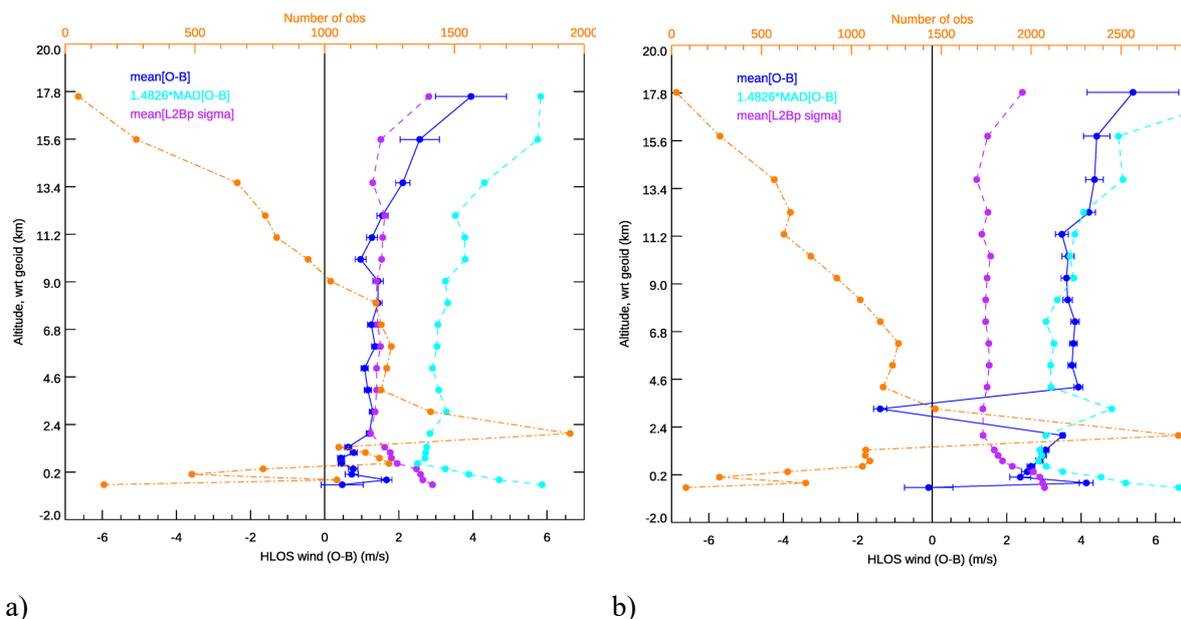


Figure 23. Global L2B Mie-cloudy HLOS wind O-B statistics as a function of altitude. The dark blue line is the mean(O-B), the cyan line is the robust standard deviation of O-B, the purple line is the mean of the L2Bp estimated error and the orange line is the number of observations (read off top axis). a) For mid-September 2018 b) for early-January 2019.

The Mie-cloudy HLOS wind quality over the HLOS wind dynamic range is shown in Figure 24. High Pearson's correlation coefficients are found: 0.97 in mid-September a) and 0.96 in early January b). The HLOS wind dependence of the bias shows a fast bias (the linear fit shows a wind speed dependent bias which has larger magnitude winds than the ECMWF model) in mid-September with a linear fit of +5%, see c). Note that the independent variable is chosen to be  $(O+B)/2$  rather than B alone, because the Mie-cloudy HLOS wind random errors are estimated to be of similar magnitude to the model background errors, therefore, to decrease the effect of errors in the independent variable for the regression it is beneficial to average. The fast bias also occurs in early-January but is smaller at +3%, see d). It is unclear if the fast bias of the Mie observations relative to the model is a problem with the Mie calibration or in the model winds in the cloudy areas that the Mie channel samples<sup>23</sup>. It should be noted that Mie-cloudy winds sample different atmospheric conditions to Rayleigh-clear winds, which should lead to differences in the O-B statistics e.g. model wind biases could differ in cloudy conditions, or in boundary layer cumulus clouds.

<sup>23</sup> In hindsight, it is more likely to be imperfections in the Mie calibration — certainly for this early mission dataset. However, the ECMWF model has been shown to be too slow in jet streams in the Northern Pacific (personal communication: comparisons to dropsondes by David Lavers).

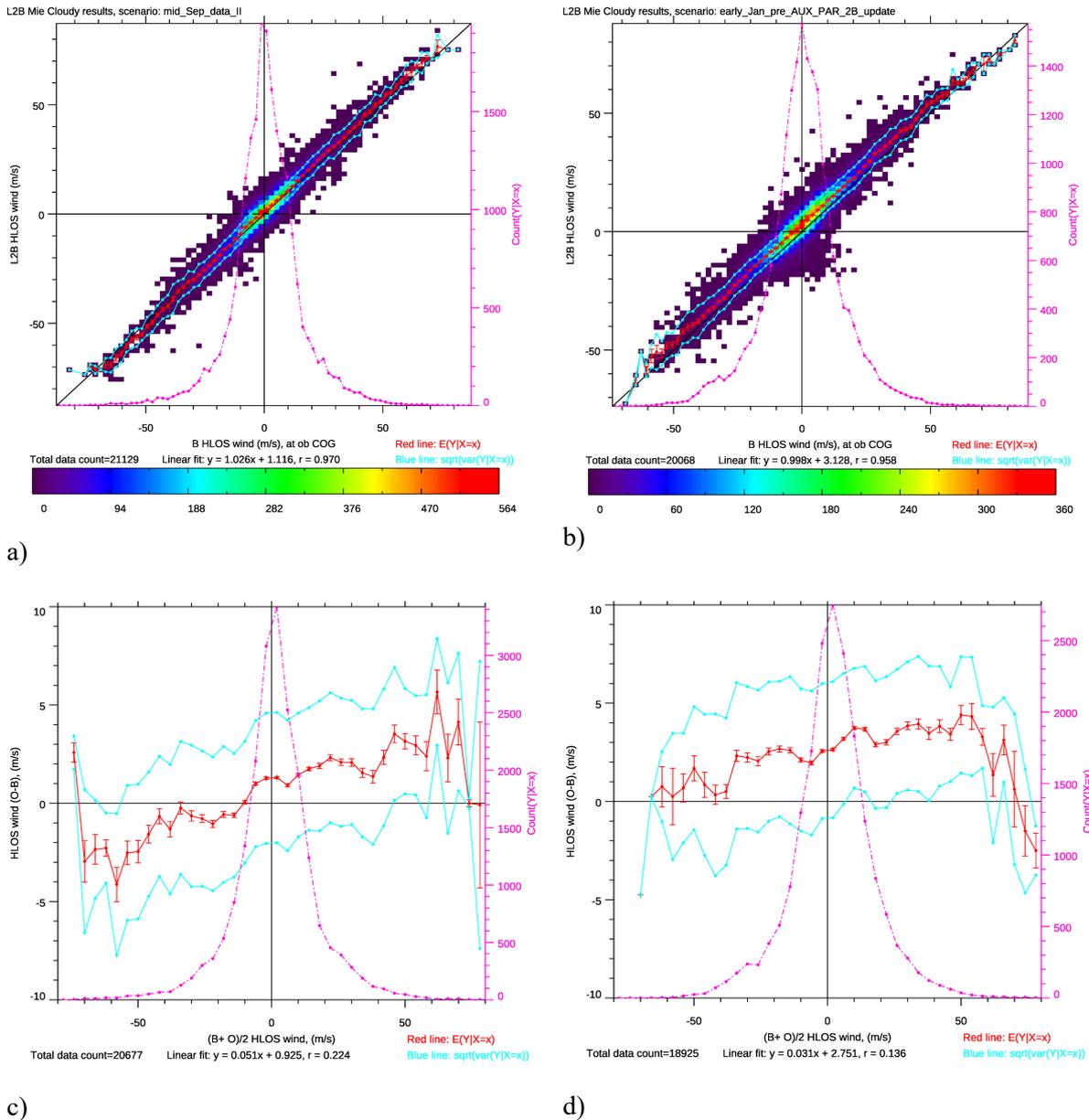
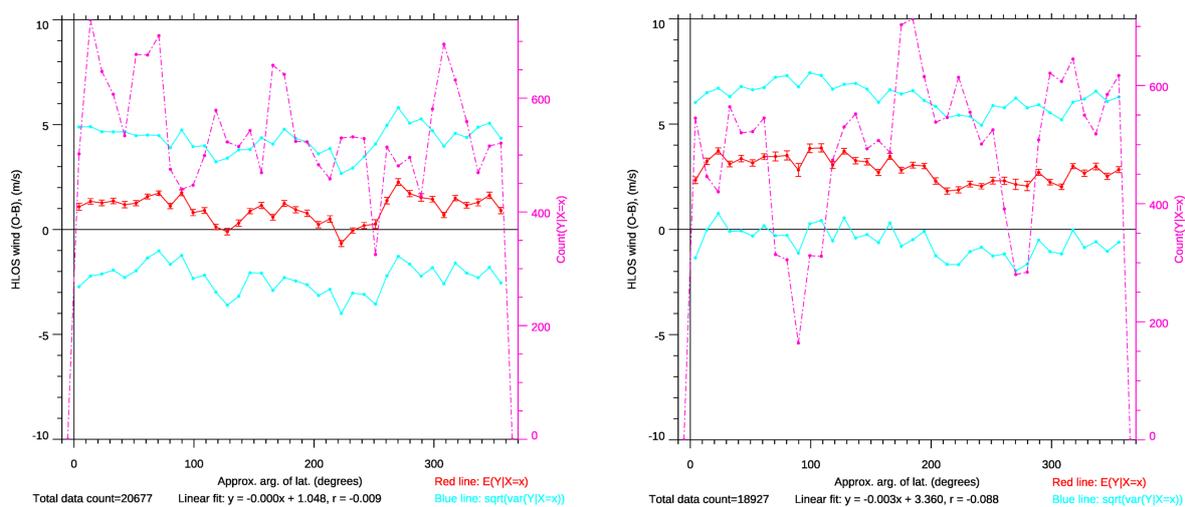


Figure 24. Global L2B Mie-cloudy O-B statistics over the HLOS wind dynamic range. Dependence of L2B HLOS wind on background HLOS wind a) in mid-September 2018 and b) in early-January 2019, as shown by 2D histograms. Dependence of mean(O-B) on background HLOS wind in c) mid-September 2018 and d) early-January 2019; the red-line is the mean(O-B) binned as a function of B (with the error bar showing the standard error of the mean), the cyan lines are the  $\pm$  standard deviation of O-B.

The Mie-cloudy bias as a function of argument of latitude is shown in Figure 25. The dependence in mid-September 2018, Figure 25 a), is thought to be mostly a result of the fast bias (reported above) given the variations in the average HLOS wind along the orbit. The behaviour in early January 2019, Figure 25 b), shares a similar pattern to the applied satellite LOS velocity correction and it thought to be mostly due to this imperfect correction (as discussed earlier for the Rayleigh).



a)

b)

Figure 25. Dependence of the L2B Mie-cloudy HLOS wind mean(O-B) on the orbital phase angle (called argument of latitude) from zero degrees at the ascending node equator crossing point. The red-lines are the mean(O-B) binned as a function of argument of latitude (with the error bar showing the standard error of the mean), the cyan lines are the  $\pm 1$ -standard deviation of O-B and the pink line is the count. a) For mid-September 2018 b) for early-January 2019.

### 3.2 An explanation for the systematic error variations with orbital phase — M1 mirror temperature-dependence

Throughout the mission, Aeolus O-B departures have shown the presence of large magnitude biases (several m/s HLOS) that differ between ascending and descending orbit phases (as is evident in Figure 26 and Figure 27), particularly for the Rayleigh-clear winds.

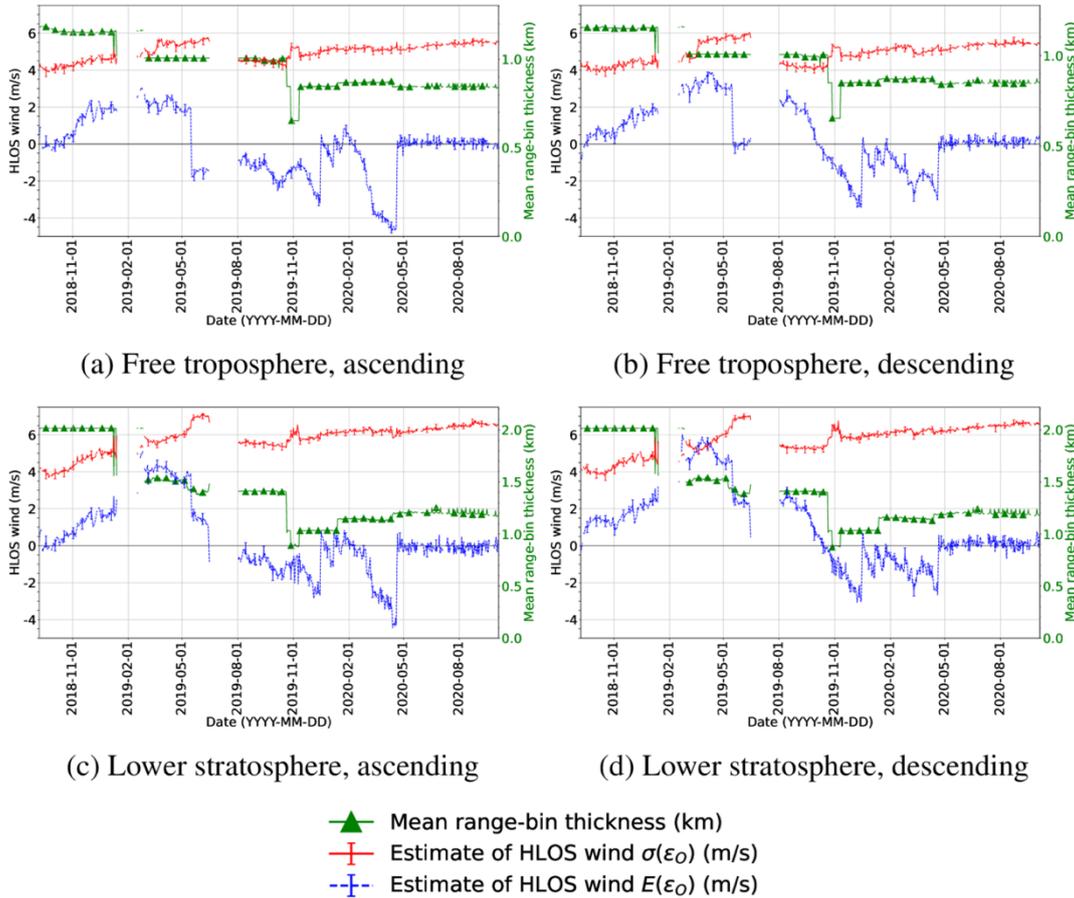


Figure 26. Time-series from 7 September 2018 until 5 October 2020 of L2B Rayleigh-clear HLOS wind estimated bias ( $E(\epsilon_0)$ ) and estimated random error ( $\sigma(\epsilon_0)$ ). Statistics for the free troposphere (100-800 hPa) are shown in the top row for a) ascending orbit phase and b) descending orbit phase. Statistics for the lower stratosphere (above 100 hPa) are shown in the bottom row for c) ascending orbit phase and d) descending orbit phase. Each datum corresponds to a global sample from one day. Also shown is the mean range-bin thickness of the wind retrievals. The error bars represent the error estimate uncertainty.

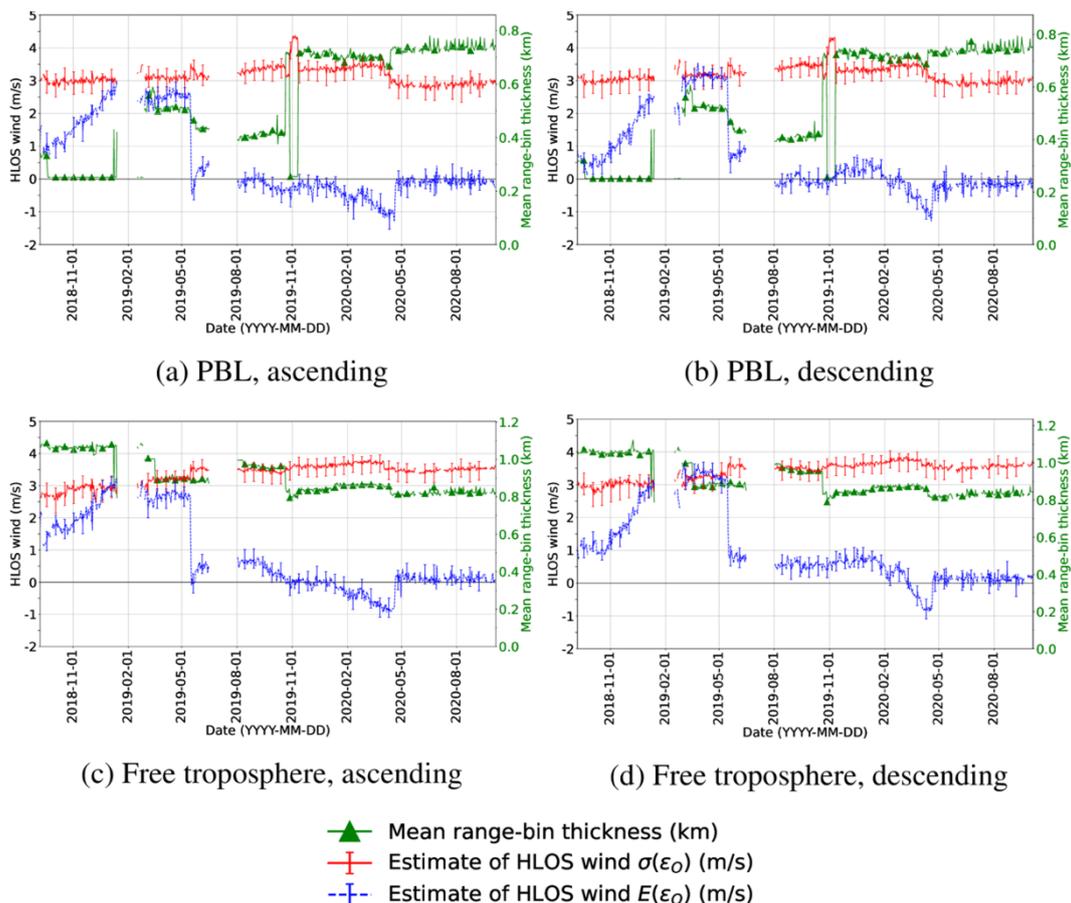


Figure 27. As in previous figure, but for L2B Mie-cloudy HLOS winds. Statistics for the planetary boundary layer (PBL, 800-1100 hPa) are shown in the top row for a) ascending orbit phase and b) descending orbit phase. Statistics for the free troposphere (100-800 hPa) are shown in the bottom row for c) ascending orbit phase and d) descending orbit phase.

The biases were confirmed to also exist with respect to radiosonde vector winds converted to Aeolus-like HLOS wind component and to another NWP assimilation system (Martin et al. (2021)) and therefore are certainly caused by Aeolus.

The Rayleigh-clear bias was found to vary more generally with geolocation as is shown in Figure 28. Here  $E(O - B)$  is binned by the orbit phase angle (argument of latitude) and by longitude from week-long samples (the bias remains roughly the same in geographical pattern over a week-long timescale). Argument of latitude is defined as  $0^\circ$  at the ascending node equator crossing and  $360^\circ$  degrees when back ascending at the equator.

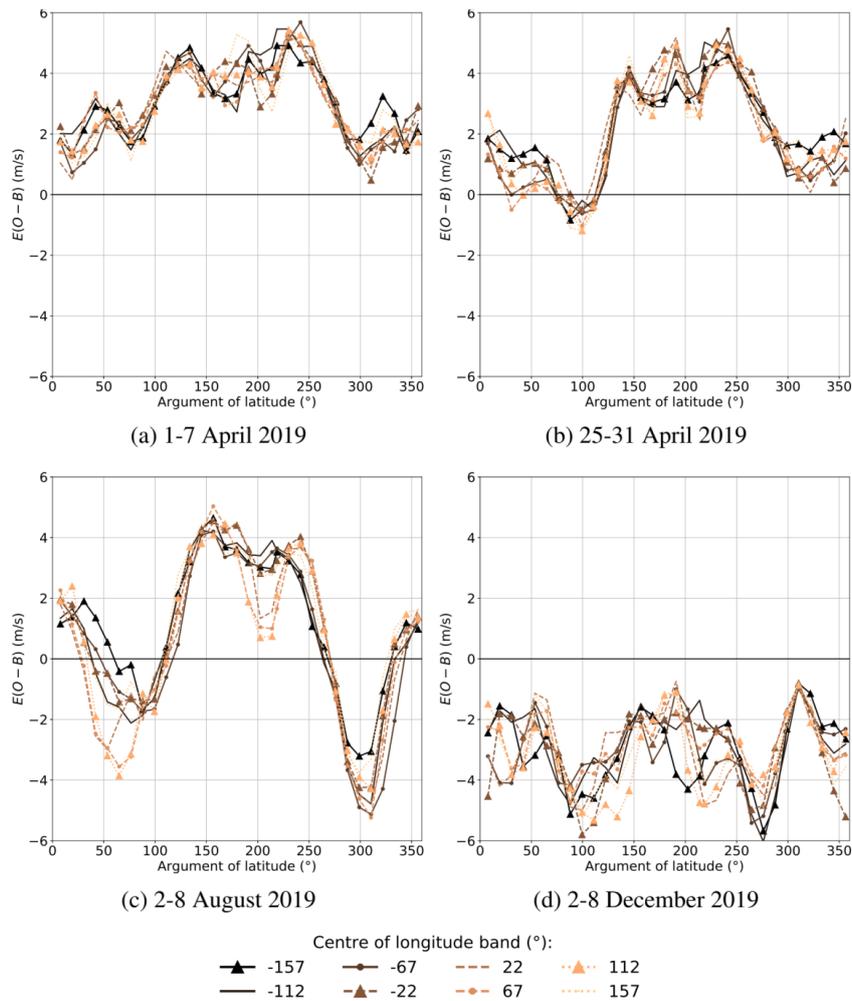


Figure 28. Rayleigh-clear bias ( $E(\mathbf{O} - \mathbf{B})$ ) binned by argument of latitude (orbit phase angle) and longitude for a) 1-7 April 2019 (FM-A) b) 25-31 April 2019 (FM-A) c) 2-8 August 2019 (FM-B) and d) 2-8 December 2019 (FM-B). There are 32 argument of latitude bins ( $11.25^\circ$  wide) and eight longitude bins ( $45^\circ$  wide). The argument of latitude and longitude bins are referenced by the value at the centre of the bin. For visualisation purposes eight longitude bins are shown, however in practical bias correction for the early FM-B OSE 10 bins was found to work best.

The Rayleigh-clear bias can change significantly over a monthly timescale e.g. it decreased by 3 m/s during April 2019 near the North Pole (argument of latitude of  $90^\circ$  in Figure 28 a) and b). In August 2019 it varied considerably with argument of latitude and longitude (Figure 28 c)); the range of bias variation is greatest in NH summer, i.e. a range of 10 m/s HLOS in August compared to 5 m/s in April and December 2019.

There is also an orbit phase dependence to the Mie-cloudy bias (Figure 29) but the bias range of  $\sim 2.5$  m/s is much smaller than for the Rayleigh-clear. Also, the longitudinal dependence was very small for the Mie-cloudy.

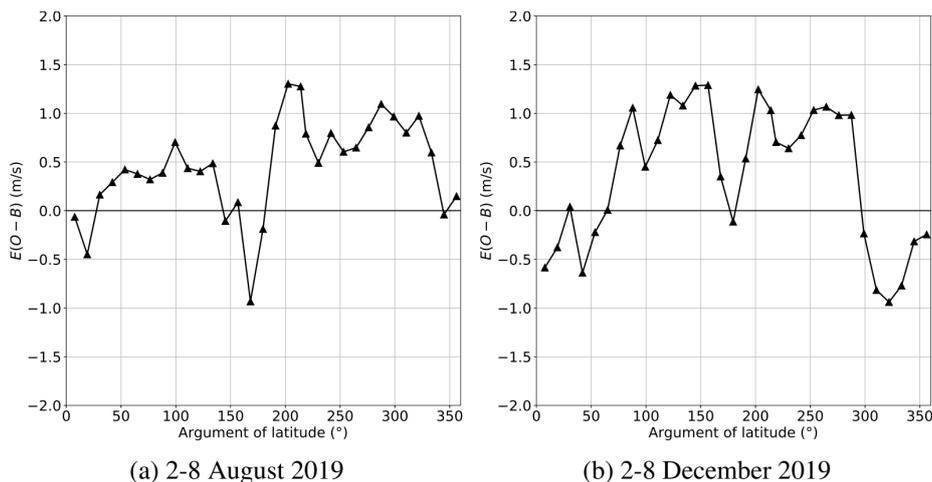
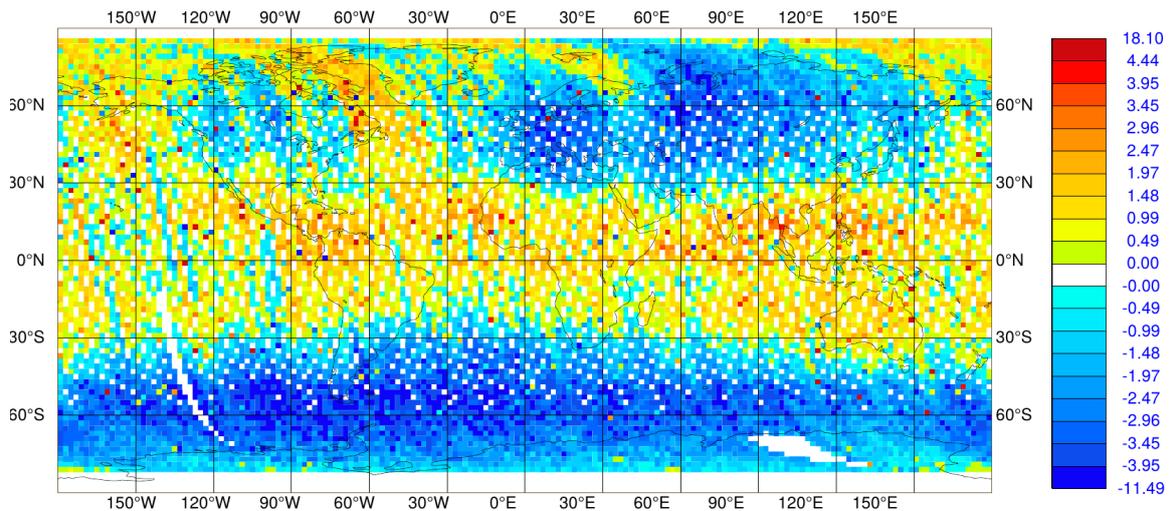
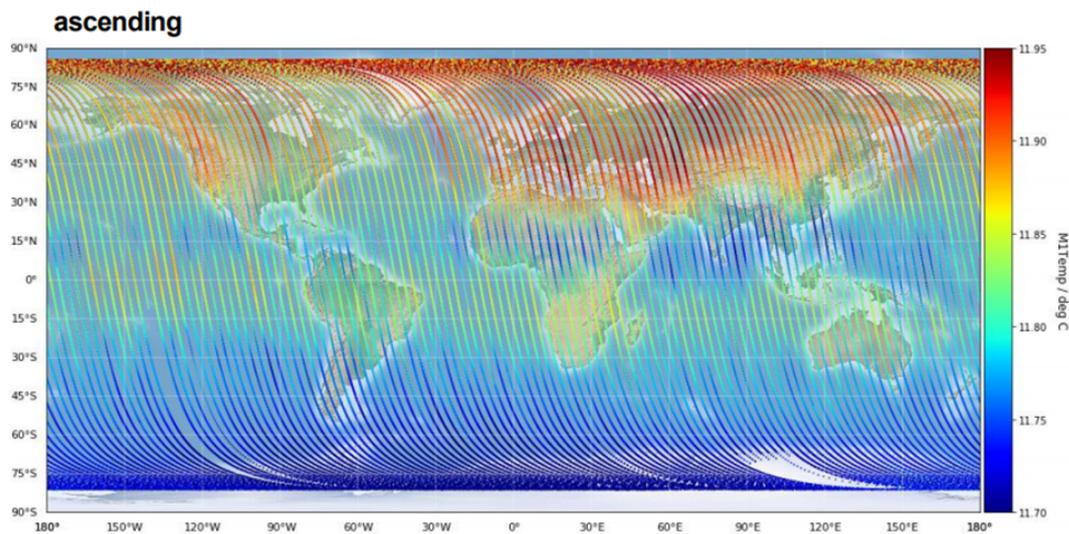


Figure 29. Mie-cloudy bias ( $E(O - B)$ ) binned by argument of latitude (orbit phase angle) for a) 2-8 August 2019 and b) 2-8 December 2019. There are 32 argument of latitude bins ( $11.25^\circ$  wide).

Inspired by the research of Bell et al. (2008) which resolved SSMIS bias issues, it was decided to compare Aeolus biases, on a sub-orbital timescale, with the available satellite housekeeping data, to try to determine the root cause of the bias. It became clear that there was a link between the ALADIN primary (M1) mirror (Korhonen et al. 2008) temperatures and the Rayleigh-clear bias when comparing maps of bias compared to mean M1 temperature in the Northern Hemispheric summer over several days e.g. Figure 30. However, the relationship to mean M1 temperature had complicated non-linear characteristics.



a)



b)

Figure 30. a) L2B Rayleigh-clear mean(O-B) from 6 August 2019 to 7 September 2019 between 0-400 hPa b) MI mirror average temperature (degrees Celsius) for a similar period (courtesy of Fabian Weiler, DLR).

The M1 temperature readings are provided by thermistors which are distributed (unevenly) across the mirror; many of which provide readings every four seconds. An example of a selection of the temperature readings during eight hours on 8 August 2019 is shown in Figure 31. There is an orbital periodicity (~1.5 hours) to the temperature readings, but never exactly repeated.

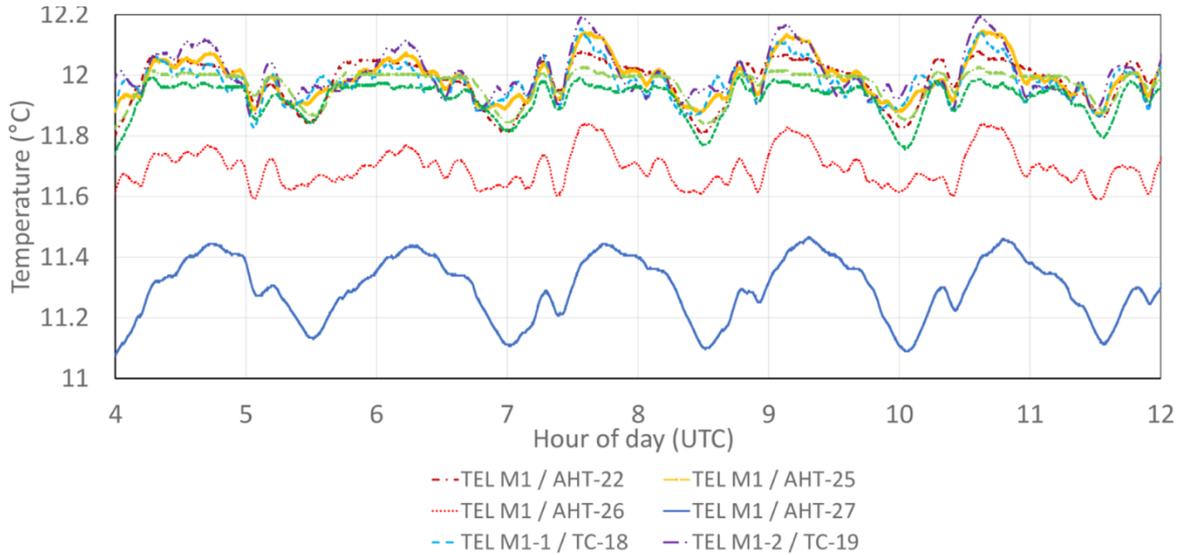


Figure 31. A selection of M1 mirror temperature readings from 4 to 12 UTC on 8 August 2019. The labels for the different thermistors are given, which identify them within ALADIN.

After careful investigation, a strong linear correlation was found between the gradients of temperature across the mirror and the L2B Rayleigh-clear bias  $E(O - B)$ .  $E(O - B)$  is derived from a 2-minute (887 km) sample, taken overall altitudes. Such a sample is typically made up of 200 winds (assuming 20 reporting vertical range-bins and 10 profiles). The difference of the mean of the mirror’s outer edge temperatures minus the mean of the near centre temperatures showed a particularly high Pearson correlation coefficient of 0.93 on 8 August 2019, see Figure 32. The Rayleigh-clear HLOS wind bias is very sensitive to the magnitude of the mirror temperature gradient i.e. 47 m/s (HLOS) per Kelvin on 8 August 2019.

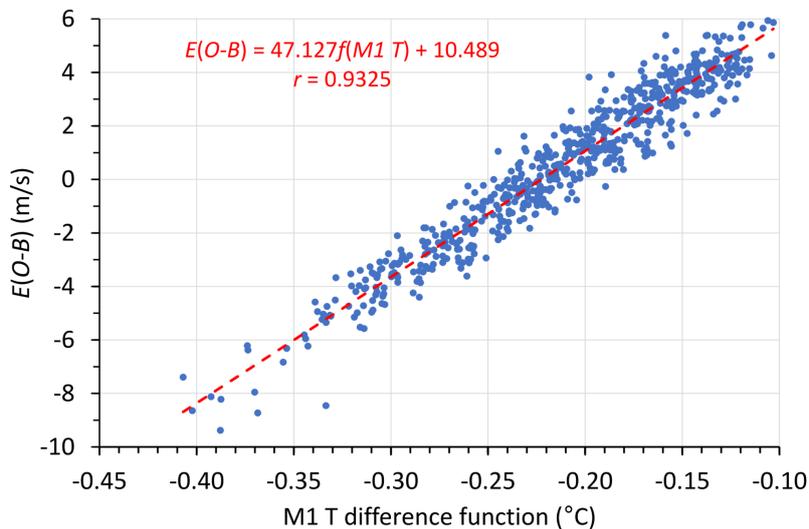


Figure 32. Aeolus L2B Rayleigh-clear HLOS wind bias  $E(O - B)$  as a function of a M1 mirror temperature function (in this example the mean outer temperatures minus mean centre temperatures) using all the data available on 8 August 2019. Each data point is derived from a 2-minute sample mean, over all altitudes. The dashed line is the linear fit and the coefficients (gradient and intercept) are written in the chart, as is the Pearson correlation coefficient  $r$ .

This apparently linear dependence allowed a bias correction method to be developed using the near real-time M1 temperature readings as predictors, and the linear regression coefficients calculated from an earlier time period (using M1 temperatures and  $E(O - B)$  for the bias). An example of how well the M1 temperature dependent bias correction works is shown in Figure 33 for 9 August 2019 (red line); here the M1 temperature bias correction linear regression coefficients were derived from independent data the day before (8 August 2019, as shown in Figure 32).

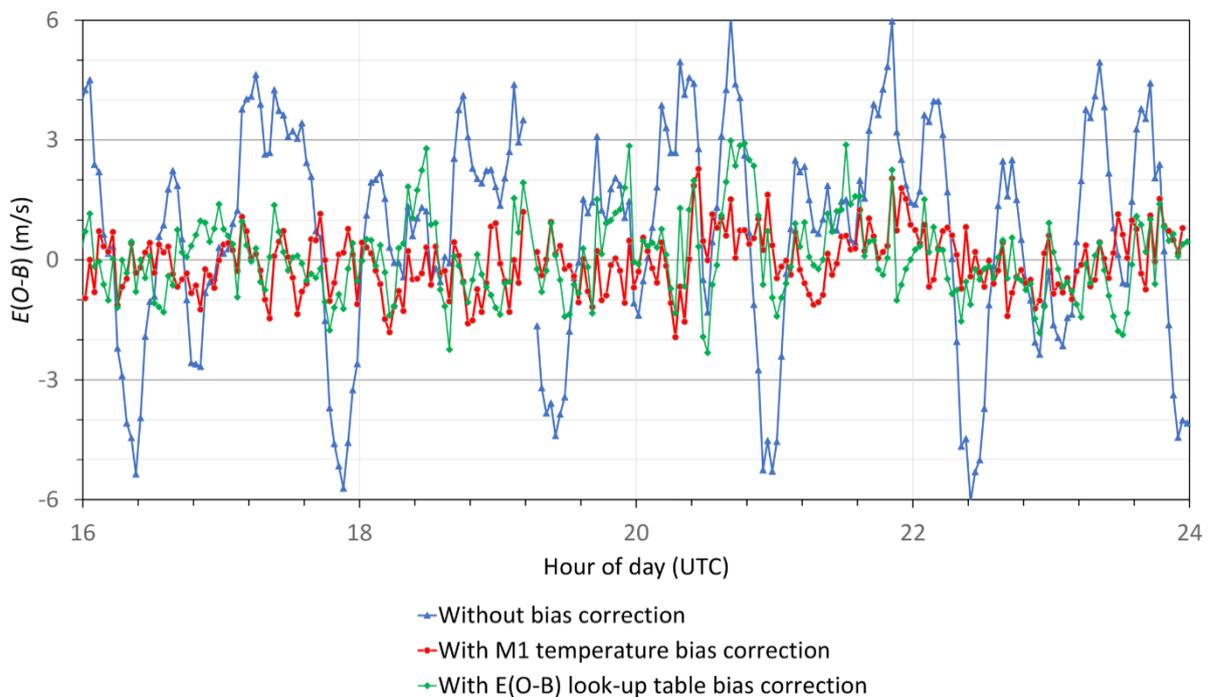


Figure 33. Time-series of Aeolus L2B Rayleigh-clear HLOS wind bias  $E(O - B)$  every 2-minutes (using all vertical levels) between 16 and 24 UTC on 9 August 2019. The blue line with triangular markers shows the bias without any corrections, the red line with circular markers shows bias after the M1 temperature bias correction is applied and the green line with diamond markers shows bias after the previous week  $E(O - B)$  look-up table bias correction method is applied. The M1 temperature bias correction linear regression was derived from the M1 temperature gradient function and O-B statistics of the previous day i.e. 8 August 2019, as shown in Figure 32.

The 24-hour sample standard deviation of the 2-minute average biases ( $E(O - B)$ ) is reduced from 2.6 m/s to 0.77 m/s with the M1 temperature bias correction. The look-up table bias correction method (described at the end of this section using the previous week sample of  $E(O - B)$ ), binned by argument of latitude and longitude is also plotted in the same figure; this method does reasonably well with a standard deviation of  $E(O - B)$  of 1.05 m/s, hence showing it was reasonable to apply in the early FM-B OSE, before the M1 temperature bias correction was available.

Investigations by the Aeolus DISC and ESA confirmed that the M1 mirror temperatures vary along the orbit in response to varying top of atmosphere radiation (short- and long-wave) and the M1 mirror's onboard thermal control mechanism in response to this. Hence the periodicity along the orbit. The likely explanation for the bias is that the temperature gradient variations (of only up to 0.3 K) cause small changes in the angle of incidence of the received atmospheric path laser light onto the spectrometers. The spectrometer response is sensitive to angular changes (particularly the Rayleigh spectrometer) which is misinterpreted as a Doppler frequency shift due to wind. Extensive evaluation

confirmed that the vast majority of the Rayleigh-clear HLOS wind bias can be explained by the M1 temperature variations, rather than other sources of bias (given that the bias remaining after the correction are typically less than 1 m/s). The Mie-cloudy winds also have an M1 temperature dependent bias, but of the opposite sign to the Rayleigh and approximately an order of magnitude (10) smaller. Despite the smaller effect it is however still beneficial to correct the Mie-cloudy winds.

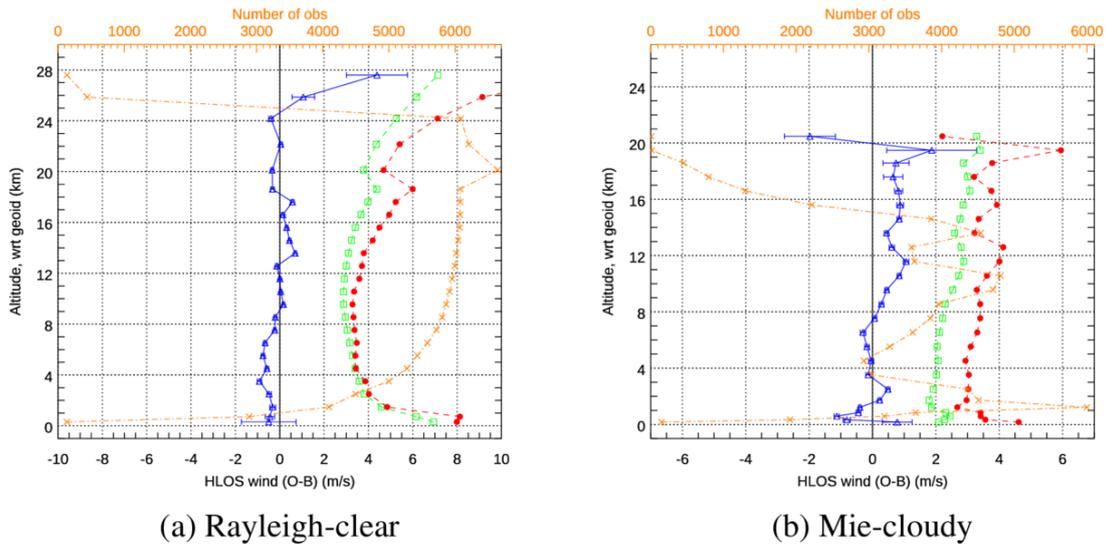
Slowly varying thermal variations were already expected as a source of bias pre-launch along with e.g. line-of-sight pointing knowledge errors, but it was assumed that the biases would be a perfectly harmonic (sinusoidal) function (to be corrected by a Harmonic Bias Estimator) with orbit phase angle (argument of latitude) and be of small magnitude. As shown above the bias is not well described by a simple harmonic function, instead it is very much scene dependent, and the biases are large for the Rayleigh-clear.

Testing by scientists from the Aeolus DISC and ESA (specifically Fabian Weiler (DLR) and Thomas Kanitz (ESA)) showed a more robust M1 temperature bias correction performance over long periods (e.g. months) is achieved with a flexible weighting of all the thermistor readings, rather than the fixed (outer minus centre) gradient function. A multiple linear regression algorithm using a flexible set of thermistors was developed for operational use. The performance of the multiple regression method and investigations using ground returns as the bias reference is described in Weiler et al. (2021).

It is not ideal that NWP model wind information from the past is used to bias correct the present Aeolus winds, however the physical explanation for the M1 temperature-dependent-bias makes it more acceptable. In effect it is another instrument calibration, but since we have only imperfect instrumental means of determining the coefficients, then NWP has been necessary for the operational correction. Any localised ECMWF model wind bias e.g. in the tropics, will act as a noise source in the global fit procedure. Any altitude bias variation is not an issue, because all the O-B values throughout the profile are used; so effectively correcting the altitude mean bias. The method can also work with Aeolus ground return winds rather than NWP model as the bias reference, as demonstrated in Weiler et al. (2021) and even better results have been demonstrated in the meantime. However, the performance has not yet reached the stability required for operational use, partly due to the limited and variable coverage, over high albedo (typically ice and snow); partly due to lower-than-expected signal levels.

The M1 temperature bias correction has been applied in operational L2B processing since 20 April 2020 with the multiple linear regression method, using all 15 house-keeping temperatures and trained with O-B departures; hence the large improvement in global average bias as evident in Figure 36 and Figure 39 from that date. Twice daily updates of the regression coefficients are performed to allow for the correction of a time varying global offset bias which as mentioned earlier is due to a drift of the internal Rayleigh response or internal Mie response and to allow a quick recovery of the bias correction following any instrument testing periods, where the L2B data is suboptimal.

An example of the quality of L2B winds as a function of altitude with the M1 temperature bias correction applied for data measured on the 2 August 2019 is shown in Figure 34. The figure shows error statistics as a function of altitude, in particular  $E(O - B)$ , estimated  $\sigma(\epsilon_o)$ , the mean of L2B processor estimated instrument noise  $\sigma(\epsilon_{o,instr})$  and the data count. Note that there are still residual biases with altitude, but they are generally small (within  $\pm 1$  m/s).



E(O-B)  
E(L2Bp  $\sigma(\epsilon_{O, instr})$ )  
Estimated  $\sigma(\epsilon_O)$

Figure 34. Error statistics of L2B HLOS winds with the M1 temperature-bias-correction applied for 14 orbits of data on 2 August 2019 as a function of altitude. a) Rayleigh-clear and b) Mie-cloudy HLOS winds. Blue triangles show  $E(O - B)$  i.e. bias; red circles show  $\sigma(\epsilon_O)$  i.e. estimated wind retrieval error; green squares show  $E(\sigma(\epsilon_{O, instr}))$  i.e. the average L2B processor estimated instrument error and orange crosses show the data count. QC of winds with L2B processor estimated error greater than 12 m/s for the Rayleigh-clear and 5 m/s for the Mie-cloudy was applied.

The bias correction applied to the L2B winds for the NRT dataset late FM-A and early FM-B period OSEs (sections 4.2 and 4.3) did not use the M1 temperature-bias-correction method, because it was not available at the time. Instead, bias correction was done via look-up table of  $E(O - B)$  binned by argument of latitude and longitude. For the Rayleigh-clear winds, 32 argument of latitude bins ( $11.25^\circ$  each) and 10 longitude bins ( $36^\circ$  each) was found to work well. For the Mie-cloudy winds we found that the longitude dimension to the look-up table was not required. The bias correction look-up tables were updated once per week, using the past week's O-B statistics. This was the method applied when Aeolus was initially operationally assimilated at ECMWF (from January to April 2020).

A comparison of the linear regression coefficients derived for the mean(outer) minus mean(inner) temperature gradient method using the 3<sup>rd</sup> (early FM-A) and 2<sup>nd</sup> reprocessing (FM-B) datasets is shown in Figure 35. This demonstrates that the wind bias was more sensitive to the M1 mirror temperature gradients for FM-B than for FM-A both for the Mie and the Rayleigh channel. The FM-B Rayleigh sensitivity increased with time, ending up being about twice as sensitive as FM-A.

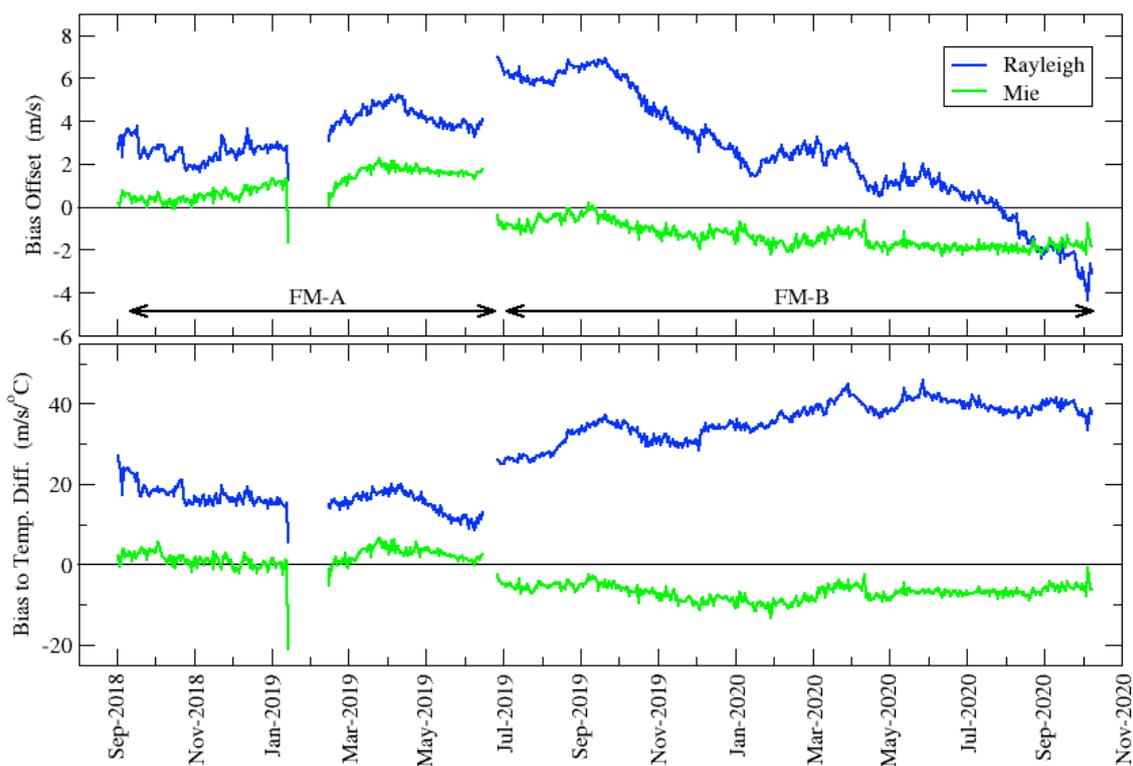


Figure 35. Time-series of linear regression fit coefficients (from AUX\_TEL files) using the mean(outer) minus mean(inner) thermistor temperatures readout as the predictor. Top panel is the constant (offset bias of LOS wind) of the linear fit and bottom is the gradient (LOS wind bias sensitivity to temperature gradients of M1 mirror) of the linear fit. This figure was produced by Saleh Abdalla (ECMWF).

### 3.3 NWP monitoring of NRT Aeolus data

NRT produced Aeolus L2B winds were introduced into the ECMWF operational data assimilation for monitoring purposes in early December 2018. The data was blacklisted in operations, meaning it had no influence on the ECMWF operational analysis. The ECMWF archived operational ODB O-B departures are an important tool method investigating the quality of the NRT data versus time, as illustrated in this section. From early September to 2 December 2018 an offline (research) data assimilation experiment (again with Aeolus blacklisted) was run at operational resolution ( $T_{CO}1279$ ) to fill the early mission gap in the “operational” NRT monitoring.

#### 3.3.1 Trends in L2B Rayleigh-clear statistics for NRT data

Monitoring of L2B Rayleigh-clear HLOS wind O-B statistics for the whole mission’s NRT data, is shown in Figure 36. This covers until the last data of the mission on 5 July 2023 (last data was at 14:30 UTC, orbit 28189). The statistics are computed from daily, global, all vertical levels (profile average) samples.

The applied QC rejects data if  $|O - B| > 25 \text{ m/s}$  and if wind results are flagged invalid<sup>24</sup>. This O-B threshold is quite relaxed (relative to the typical profile average random errors of  $1\sigma \sim 6 \text{ m/s}$ ). Since the Aeolus wind noise characteristics have degraded with time, particularly for FM-B data from July 2019 to early October 2022, then a more relaxed QC threshold (allowing more data into the statistics; grey line in plot) resulted in a fairer comparison of the data quality over the whole mission. The magnitude of the standard deviation (a non-robust statistic) is very sensitive to the QC threshold value. For example, using a 25 m/s threshold compared to a 15 m/s threshold increases the standard deviation of O-B by  $\sim 2 \text{ m/s}$ . The reason that standard deviation rather than, for example, the scaled median absolute deviation was chosen, is because the ODB SQL tool used to generate the statistics does not have a median function available. The L2Bp estimated standard error is not used in QC here but is included in the plots as an additional metric for the noise (mean value per day plotted in orange). Note that due to processor baselines updates the L2Bp estimated error underwent some step-changes.

---

<sup>24</sup> The L2B processor provides with each wind result an overall confidence flag. Setting this to invalid for all winds has been the method (since  $\sim$ October 2020) to avoid NWP centres operationally assimilating the data during Special Operations Requests (SOR) when there is a risk of the NRT data being of poor quality (so called “blocklisting”). Hence there can be gaps in the monitoring time-series after this period due to numerous SOR.

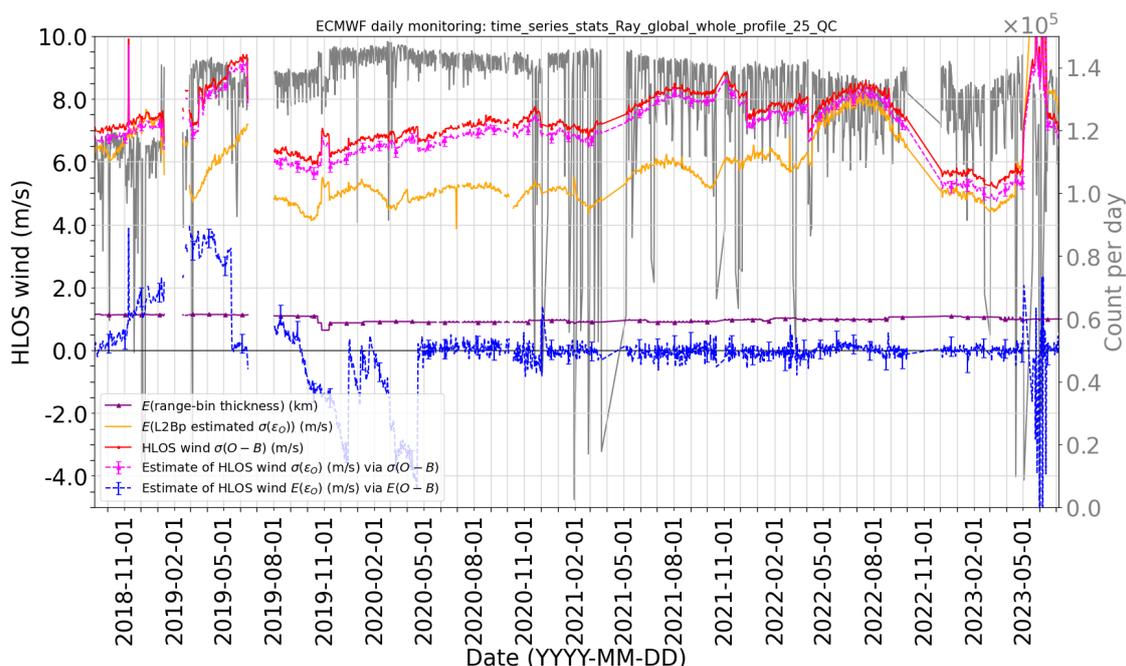
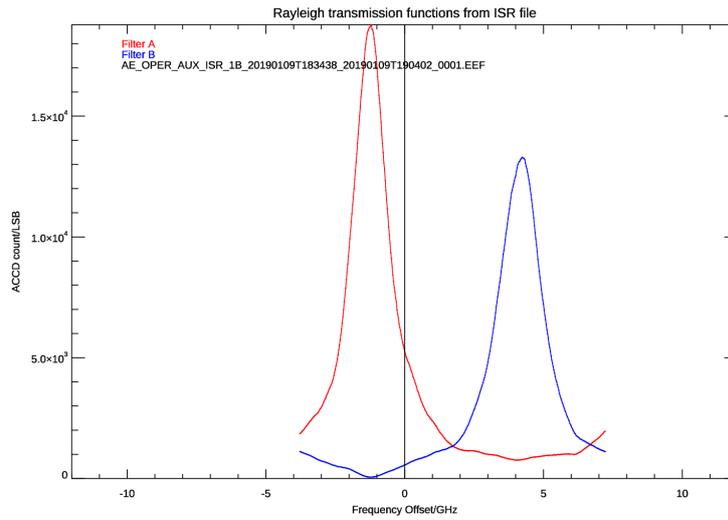


Figure 36. Time series of daily, global and all vertical level NRT dataset L2B Rayleigh-clear O-B statistics via operational monitoring from 7 September 2018 until 5 July 2023. The average range-bin thickness (purple) is the mean over all observations which pass the QC.

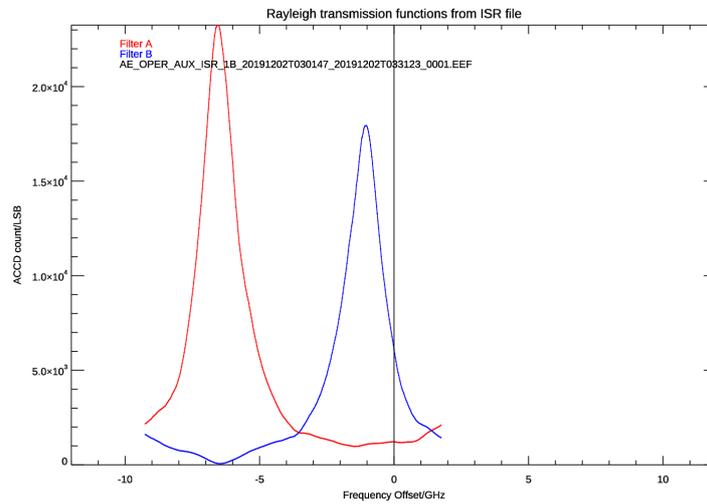
Features of the Figure 36 time-series are discussed below:

#### Systematic error statistics:

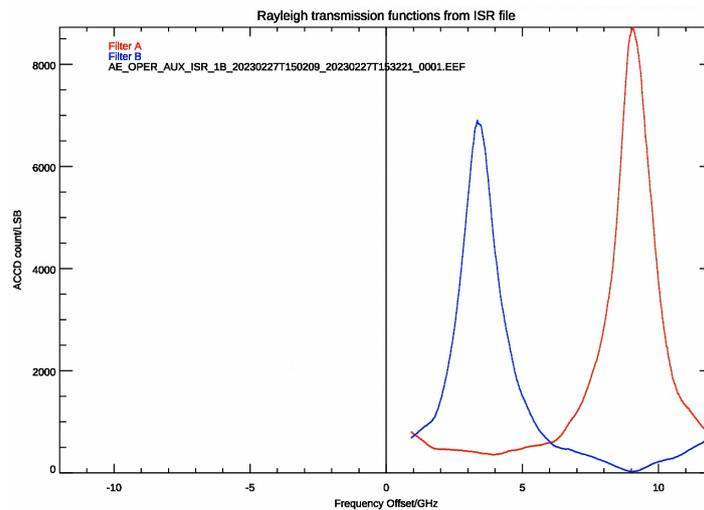
- With the early mission with the FM-A laser NRT dataset, the global average HLOS wind bias (blue line) had a positive trend (of around 0.5 m/s per month) which apparently continued during the temporary data outage in February 2019 (as discussed in section 3.1). The bias then seemed to stabilise in March 2019. A new calibration file was implemented in May 2019, reducing the positive bias. The bias then remained steady until the FM-A laser was switched off in June 2019.
- When the FM-B laser was switched-on on 28 June 2019 the bias was initially very negative (off the scale) due to the initial calibration files being inappropriate for FM-B (still using FM-A calibration files). The Rayleigh transmission curve ( $T_{A,B}$ ) peaks were in very different frequency offset positions for FM-A and FM-B e.g. see the Instrument Spectral Registrations (ISRs) in Figure 37 ( $T_A$  is shifted by about -5 GHz in FM-B relative to FM-A). Also note that with FM-A data since November 2022, the alternative filter cross-point has been used with channel A peak transmission on the right and channel B on the left (which changes the sign of the Rayleigh response).



a) FM-A ISR from 9 January 2019



b) FM-B ISR from 2 December 2019



c) FM-A ISR from 27 February 2023

Figure 37. Example ISRs during the mission showing the Rayleigh spectrometer internal path transmission curves for channels A (red) and B (blue) as a function of frequency offset for a) FM-A from 9 January 2019, b) FM-B from 2 December 2019 and c) FM-A ISR from 27 February 2023.

- The bias for FM-B data jumped to near zero after 2 August 2019 when FM-B derived calibration files were implemented operationally. However, the bias drifted quickly to negative values, particularly after mid-September 2019. The drift rate varied being around -1 m/s per month from August to December 2019 and even larger from January to April 2020. The positive bias jumps on 16 December 2019 and 7 January 2020 were due to the application of manual offset bias correction to the L2B data (via an option in the AUX\_PAR\_2B file; see appendix 9.11) as a crude attempt to mitigate this. The jump around 22 January 2020 was due to an internal reference anomaly, which has been investigated by the DISC, but not yet fully explained. DISC investigations have shown that the negative bias drift for FM-B is mostly driven by the drifting internal reference Rayleigh response whereas the atmospheric path Rayleigh response was relatively stable; this could be caused by a drift in the internal path laser pointing or divergence properties. The bias drift also has some correspondence with the decrease in internal and atmospheric signal levels with time possibly linked to a geometrical effect e.g. Laser Induced Contamination affecting the beam (early 2022 results) — as indicated by DLR, ESA and industry investigations.
- The significant improvement in bias since 20 April 2020 (close to zero and steady with time) was due to the operational implementation of the M1-temperature-dependent bias correction scheme (see section 3.2), which additionally corrects the global offset bias, using the ECMWF model as a reference. It is not possible to subtract off the M1-temperature-dependent bias correction from the ECMWF ODB archived O-B statistics to see how the underlying instrumental bias drift continued, because the L2B BUFR does not contain this information. However, this could be assessed from the L2B EE format products (via comparisons to AUX\_MET), which does contain the M1 T correction value.
- Several gaps in the statistics occurred in late May 2020, early July 2020 (due to ITR (Instrument Telescope Refocus) and M1 thermal control tests respectively) and in late 2020 and 2021 due to further instrument tests (Special Operations Requests).
- There were several instances during the mission of automatic switching of the laser to Survival Mode due to the automatic detection of anomalies, triggering FDIRs (Failure Detection, Isolation and Recovery). It takes a while to safely switch the laser back on again, leading to data gaps with no wind measurements. The first data gap was for FM-A between 14 January and 15 February 2019. There was a large data gap from 23 March to 5 May 2021 for the FM-B laser. A similar incident occurred from 22 October to 2 November 2021 (with FM-B) and from 21 February to 2 March 2023 (with FM-A). There were also large data gaps during the shutdown of one laser and switching to the other laser: e.g. for the switch from FM-A to FM-B from 16 to 28 June 2019 and for the switch back to FM-A between 1 October 2022 and 1 December 2022 (with considerable tuning of the FM-A laser settings).
- Larger biases occurred in December 2020 due to large step changes in the internal Rayleigh response; therefore a patch L2Bp processor was implemented in operations (v3.40.3) which effectively stopped using the internal Rayleigh response in the calculation of L2B Rayleigh winds, by setting the internal RR to a fixed value. This improved the biases without any detected degradation in wind quality.
- NRT data biases remained well controlled from then until the final operational data on 30 April 2023; apart from hot-pixel induced biases affecting specific range-bins, which are not evident in the global average, all-range-bin plots. The End-of-life (EOL) testing of the laser(s) from 1

May 2023 until the final wind measurements. (5 July 2023) led to some periods of large wind biases.

- Range-bin 15 (counting from 0) was rejected due to a large bias since 22 December 2022 related to improper DCMZ correction values. This reduces the data count. This has been mostly explained by DLR investigations as part of the DISC (not discussed here) as being due to a pixel damaged by cosmic rays leading to relatively inefficient photon detection.

#### Random error statistics:

- The Rayleigh standard deviation of O-B (red line) has varied during the mission due to many factors, for example:
  - Changes in the emitted laser energy and atmospheric path radiometric performance i.e. instrumental changes.
  - Seasonal changes in solar background noise: larger in a global average sense during summer months, and larger for the NH polar summer than in the SH polar summer due to the asymmetry in the line-of-sight pointing of the telescope. The solar background noise influence can be reduced by focusing on data where Aeolus is pointing to the dark side of the terminator; hence the tropics-only plot in Figure 38 (note that the solar background is small, but not zero at the equator and Aeolus' LTAN can be  $\pm 10$  minutes from the terminator).
  - Seasonal changes in the M1 mirror temperature-gradient-dependent wind biases (larger sub-orbital variability in NH summer).
  - Changes in range-bin settings and laser N/P (see Appendix 9.10) settings.
  - Ground processing software baseline changes for NRT datasets.
- From the tropics-only result we see the L2B Rayleigh-clear random error increased with time for the 2018-2019 period with the FM-A laser. This was due to: an increasing percentage of uncorrected hot pixels (in the NRT dataset); a significant reduction in atmospheric signal levels due to decreasing laser energy with time; and seasonal changes in the M1 mirror temperature-dependent-bias variations (became notably large in April 2019), large sub-orbital M1 temperatures biases increase the apparent large temporal-spatial scale random error statistics.
- For the August 2019 FM-B data the Rayleigh O-B standard deviation in the tropics was significantly smaller than the 2018-2019 FM-A data (partly due to the high wind noise of 250 m thick RBS for early FM-A). Unfortunately, the FM-B random errors then increased steadily with time, as shown in Figure 38 (apart from occasional laser energy increases) due to the FM-B atmospheric path transmission loss issue.
- The standard deviation of O-B was  $\sim 33\%$  larger in August 2021 compared to August 2019 (the absolute values here are not so important due to the relaxed QC).
- There was an apparent improvement in the random error from 20 April 2020 due to the inclusion of the M1 temperature bias correction (see section 3.2), due to the correction of sub-orbital bias variations.
- There was an instrument-based improvement in the random error in early December 2020 due to an increase in FM-B laser energy per pulse from about 60 to 70 mJ. There was also a significant boost in November 2021 ( $> 80$  mJ) and in September 2022 (100 mJ) to mitigate the ongoing FM-B laser transmission loss.

- Since 17 June 2021, moon-blinding affected measurements were rejected and not used in the L2Bp wind retrieval (an AUX\_PAR\_2B setting). This was to avoid the occasional  $\sim 4$  m/s biases that are associated to mispointing during moon-blinding of the star trackers (switch to inertial navigation system). This QC threw out a significant fraction of data in the tropics which causes the oscillations in the data counts every 2 weeks (grey line).
- There was an instrument-based improvement in random error from 13 December 2021 due to changes to the instrument's on-board N/P settings (N=15, P=38). Doubling the measurement horizontal scale improved the SNR with benefits in L2B processing classification, but there was also a surprisingly large influence from the reduction in read-out noise (DLR).
- Significant changes to the tropical RBS were implemented in late January 2022 which affected the random error trend. The range-bins were raised to 30 km in the  $-35^\circ$  to  $10^\circ$  latitude band to help capture the Hunga-Tonga volcanic eruption plume (which provided a good sample of Mie-cloudy winds). This improved the noise of winds in the troposphere due to thicker range-bins but added extra very noisy Rayleigh-clear winds at higher altitudes (20-30 km). Therefore, it is perhaps better to focus on e.g. NH extratropics to assess instrument-related data quality trends in that period.
- NRT processed data noise levels peaked for the FM-B laser in July 2022. This was despite a significant improvement from 4 April 2022 due to N=5, P=114 being applied on the detector. This reduced the read-out noise significantly, without any significant effect on the resolution of the L2B wind "observation" results (still up to  $\sim 83$  km for the Rayleigh-clear and  $\sim 17$  km for Mie-cloudy).
- The random errors improved by a large amount after switching back to FM-A (reduced by 30%) – data operational after 1 December 2022. It provided the smallest NRT Rayleigh-clear random errors of the mission; thanks to reasonable signal levels combined with B15, B16 processing algorithms. The errors remained at similar levels until the end of the operational NRT dataset.
- The random errors increased with the switch-back to the FM-B laser for the EOL tests in May 2023, due to the atmospheric path transmission issues persisting. This was despite the laser energy reaching record highs in the final days of wind measurements (182 mJ on 5 July 2023) and RBS been more favourable (and fixed).

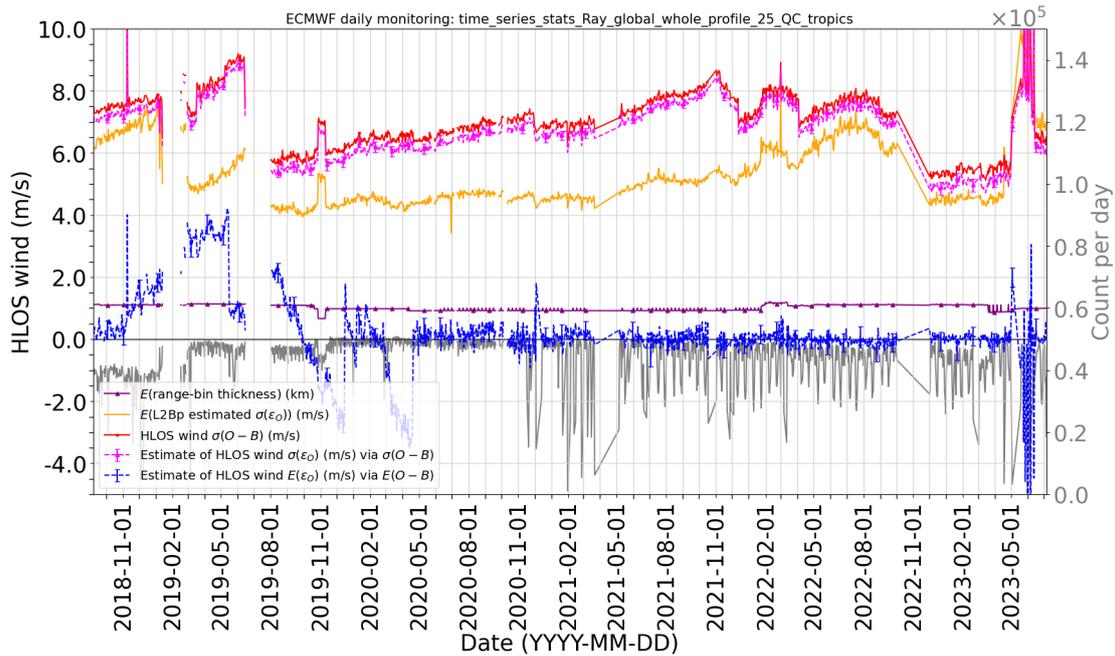


Figure 38. Daily time series for the “tropics” only ( $\pm 30^\circ$  latitude band) and all pressure level NRT dataset L2B Rayleigh-clear O-B statistics via operational monitoring from 7 September 2018 until 5 July 2023. The range-bin thickness is the mean over all range-bins. The same QC criteria as for the global plot, Figure 36, is applied.

### 3.3.2 Trends in L2B Mie-cloudy statistics for NRT data

Monitoring of NRT L2B Mie-cloudy HLOS wind O-B statistics for the whole mission is shown in Figure 39. The QC applied is reasonably relaxed (reject data if  $|O - B| > 10 \text{ m/s}$ ), given the typical noise levels of most Mie-cloudy winds ( $\sim 2\text{--}4 \text{ m/s}$ ). Only valid L2B winds are used. The Mie-cloudy winds are not strongly influenced by solar background noise; hence a global statistic is sufficient.

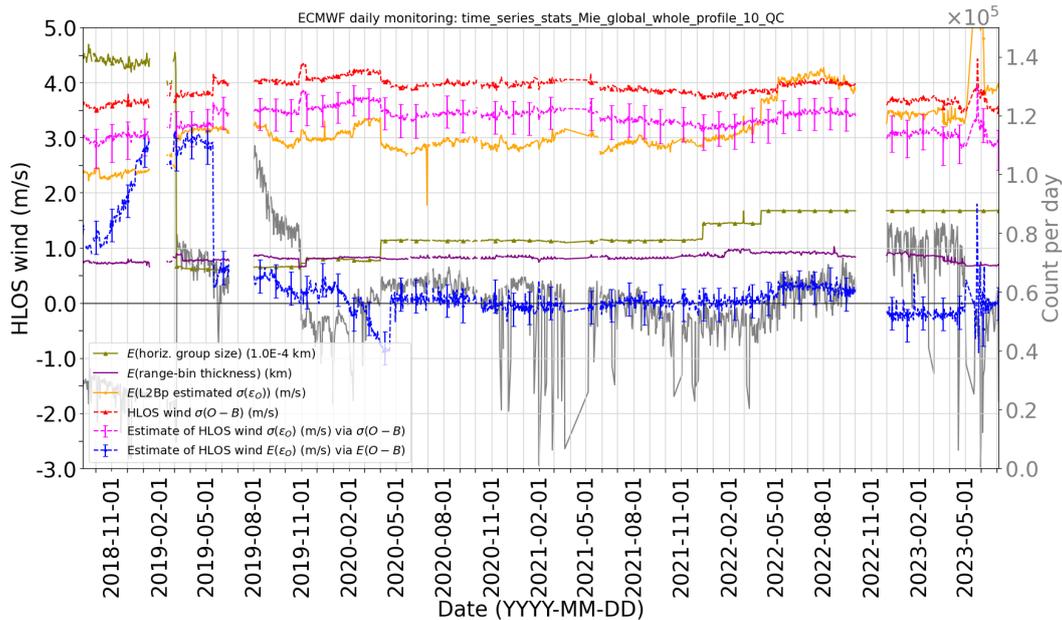


Figure 39. Daily time series of global and all pressure level NRT dataset L2B Mie-cloudy O-B statistics via operational monitoring from 7 September 2018 until 5 July 2023.

Features in the statistics of Figure 39 are discussed below:

- Global average bias had a very strong positive trend for FM-A (as discussed in section 3.1), although it settled in April 2019 onwards (in a similar manner to the Rayleigh bias; it is unclear what instrumental effects drove this). The bias then dropped by 2.5 m/s when a new AUX\_MRC\_1B file was applied in May 2019.
- The global bias for FM-B data (after appropriate calibration files were implemented on 2 August 2019) was reasonably stable and  $< 1$  m/s until January 2020, after which it drifted to negative values (as did the Rayleigh-clear winds).
- The global average bias improved significantly from 20 April 2020 due to the operational implementation of the M1 temperature-dependent-bias correction, which also corrects for the global offset bias with respect to the ECMWF background.
- There are some data gaps in the time-series in 2020 and 2021, for the same reasons outlined for the Rayleigh-clear winds in section 3.3.1.
- The Mie random error had a slight increase with time for FM-A (estimated  $\sigma_O$  increased from  $\sim 3$  m/s to 3.5 m/s); there was a step increase when a new MRC was applied in May 2019 — suggesting the updated MRC was suboptimal in terms of wind-speed dependent bias (atmospheric path Mie non-linearities).
- For the early FM-A period the L2Bp horizontal grouping of measurements was set to be up-to 1 BRC, which resulted in the mean average Mie-cloudy horizontal group size being  $\sim 44$  km (green line). This was modified to 10 km maximum group size on 5 March 2019, which resulted in the average group size being  $\sim 7$  km. This change was chosen because it increased the number of winds by a factor 2-3, with only a modest increase in noise as shown via stdev(O-B).
- For FM-B the random error was reasonably steady with time (compared to the Rayleigh-clear) with estimated  $\sigma_O$  of  $\sim 3.5$  m/s, which was larger than for FM-A (which is thought to be related to the smaller FWHM of the Fizeau fringe on FM-A versus FM-B; perhaps related to pointing differences).

- There was a small improvement in the random error around April 2020 associated with ground processing software updates and the M1 T bias correction.
- There was also an improvement with the implementation of processing baseline 12 on 26 May 2021; due to an improvement in the L1B processing Mie useful signal formatting (allowing negative values).
- There was perhaps a slight improvement evident since 1 July 2021 due to a better atmospheric path Mie-nonlinearity look-up table (using NWP winds as a reference to infer the fine-scale Mie non-linearity function; (Marseille et al. (2022)).
- There was a small increase in random error on 5 May 2022 due to changes in the AUX\_PAR\_2B Mie core algorithm QC thresholds (see Appendix 9.11), but with a concomitant increase in the data count, thus justifying the change. The L2Bp estimated error jumped up with the N=5, P=114 change on 4 April 2022 and with the change in AUX\_PAR\_2B on 5 May 2022; it is unclear why and should be investigated.
- The initial data count for the early FM-B period (2019) was much larger than the end of the early FM-A period (2019). This was due to a combination of higher transmitted signal and therefore more sensitivity to small backscatter (e.g. aerosols) and a temporary large source of Mie-cloudy winds from lower stratospheric wildfire aerosols in NH high latitudes which peaked in August 2019 (see also section 3).
- There was a large decrease in the data count in late 2019 due to a mistake in the L2Bp classification with a new processor baseline; this was later rectified.
- The average horizontal group size has steps to larger values with FM-B, due to intentional settings changes and to accommodate the modified N/P settings; it was ~17 km since 4 April 2022.
- The random errors improved with the switch back to FM-A from 1 December 2022, with estimated  $\sigma_o$  of ~3.0 m/s and the data counts increased by ~20% compared to the end of FM-B (due to more signal). There is a high sensitivity of the number of “valid” L2B winds to the atmospheric path signal levels.

### 3.4 First reprocessed dataset (early FM-B) monitoring

To help understand the NWP impact results using the first reprocessed L2B dataset (2B10), the data quality via the O-B departure statistics from 29 June to 31 December 2021 is assessed. The O-B departures are calculated from the control run of the OSE in which Aeolus was not assimilated (passive) followed by the application of the same monitoring tools used for the NRT operational dataset (as in section 3.3).

Figure 40 shows the global, all altitudes, daily statistics for the Rayleigh-clear winds, whilst Figure 41 restricts this to the tropics to reduce the effect of solar background noise. The O-B derived standard error for the Rayleigh-clear winds increases from 4.0 to 4.8 m/s during this period in the tropics, mostly due to a decrease in useful signal with time. However, there was a reduction in average range-bin thickness from mid-November 2019 onwards<sup>25</sup> increasing the random error (after a temporary, large reduction in range-bin thickness due to the CAL/VAL AMV RBS, with an associated reduction in the data count passing the QC).

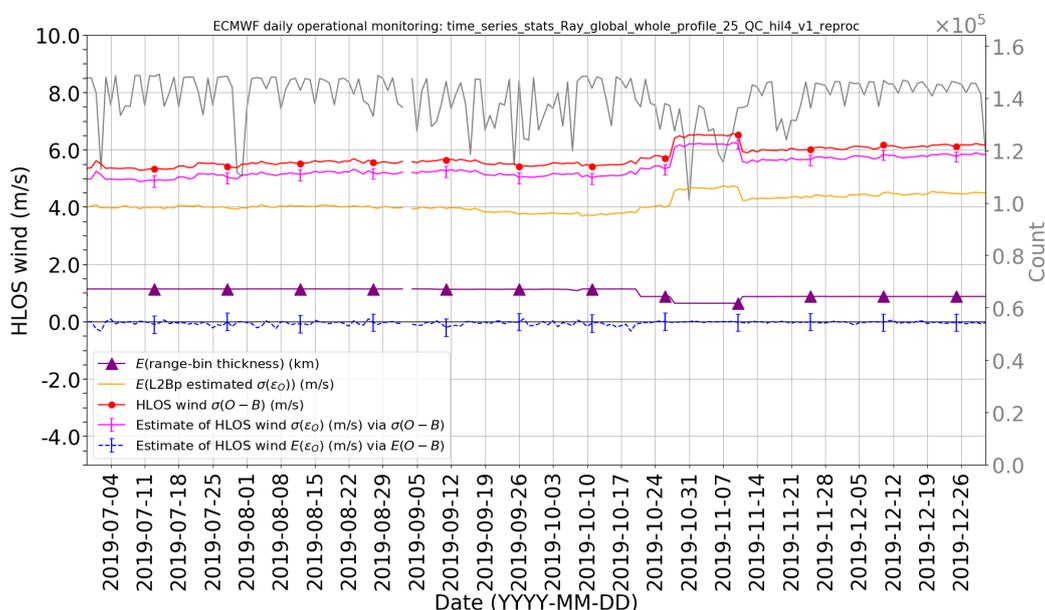


Figure 40. Time series of daily, global and all pressure level L2B Rayleigh-clear O-B statistics via the control run of the first reprocessed dataset (2B10) OSE from 29 June to 31 December 2019.

<sup>25</sup> The change of RBS increased the vertical resolution in the UTLS in the tropics since this zone has large vertical wind shear and gives the largest NWP impact from Aeolus.

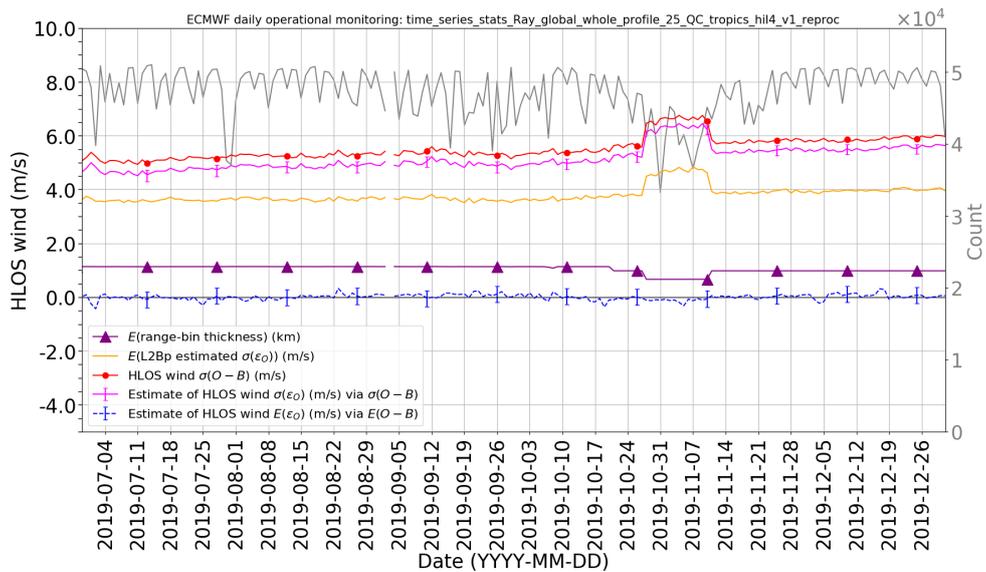


Figure 41. Daily time series for the “tropics” only ( $\pm 30^\circ$  latitude band) and all pressure level L2B Rayleigh-clear O-B statistics via the control run of the first reprocessed dataset OSE from 29 June to 31 December 2019.

The Mie-cloudy data count was significantly larger in late July and August 2019 than in December 2019; see the explanation in section 3.5. The Mie-cloudy HLOS wind noise was quite stable with time; being typically around 3.3-3.5 m/s. An increase in noise associated with the reduced range-bin thicknesses of the CAL/VAL AMV RBS is clearly seen. It is interesting that the AMV RBS increased the standard deviation of O-B but not the L2B processor estimated error (instrument noise only); perhaps suggesting Aeolus captured some small-scale vertical wind shear that the ECMWF background could not represent. A positive bias (reaching 0.3 m/s) occurred during late July and early August and then decreased steadily — see the explanation in section 3.5.

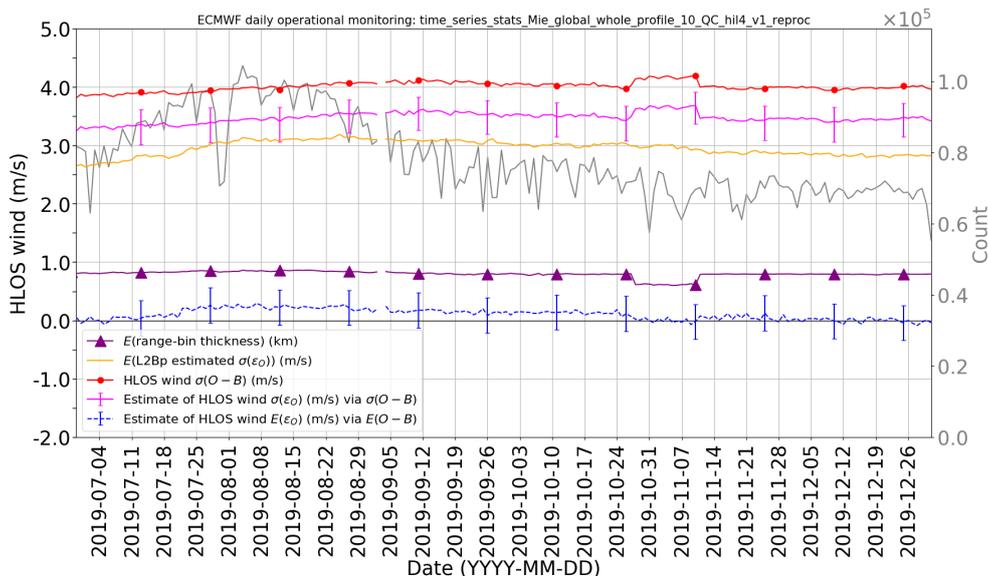


Figure 42. Daily time series of global and all pressure level L2B Mie-cloudy O-B statistics via the control run of the first reprocessed dataset OSE from 29 June to 31 December 2019.

### 3.5 Second reprocessed dataset (FM-B) monitoring

Monitoring of the second reprocessed dataset (2B11) O-B departure statistics (and L2Bp estimated error) versus time via the control of the OSE (hlpv), shows a gradual increase of the L2B Rayleigh-clear wind noise by 40% over the period from late June 2019 to October 2020 (whilst the data count is roughly constant); see Figure 43. The increasing noise trend is clearer in the tropics (Figure 44) due to the reduced influence of solar background noise. The bias remains very small in global, daily averages.

The L2B Mie-cloudy statistics show reasonably stable random errors; however, the data counts show a significant decrease with time following a peak in August 2019 (a 37% decrease in counts); see Figure 45. The August 2019 peak count was found to be due to an unusually high aerosol load in the lower stratosphere at high northern latitudes – from Siberian wildfire smoke and a smaller influence from the Raikoke volcanic eruption, according to Engelmann et al. (2021). The data counts rose sharply in July 2019 and returned to normal levels by December 2019. This is confirmed by the NH extratropics statistics of Figure 46 a), as compared to the tropics and SH extratropics in b) and c) respectively and by the zonal average data counts for August 2019 compared to August 2020 in Figure 47, which shows a peak in Mie-cloudy data towards the North Pole at  $\sim 130$  hPa ( $\sim 15$  km) in 2019 which was absent in 2020. Figure 48 shows maps of data counts in August 2019 and 2020 at the peak pressure range of the aerosol load, highlighting the remarkable anomalously high coverage of Mie winds in the NH extratropics in 2019; in areas usually with no Mie winds during the summer. There is also a gradual decline in data counts in the tropics, which may be due to aerosol scenes, e.g. desert dust, gradually not producing valid winds or being rejected by QC due to large noise because of ongoing FM-B atmospheric path signal loss.

The Mie-cloudy global average, daily, bias is slightly positive (+0.2 m/s) from 22 July to end of October 2019; this is correlated with the extra winds in the NH extratropics from the wildfire smoke. There is an altitude dependence to the bias (not shown), with more positive bias at higher altitudes.

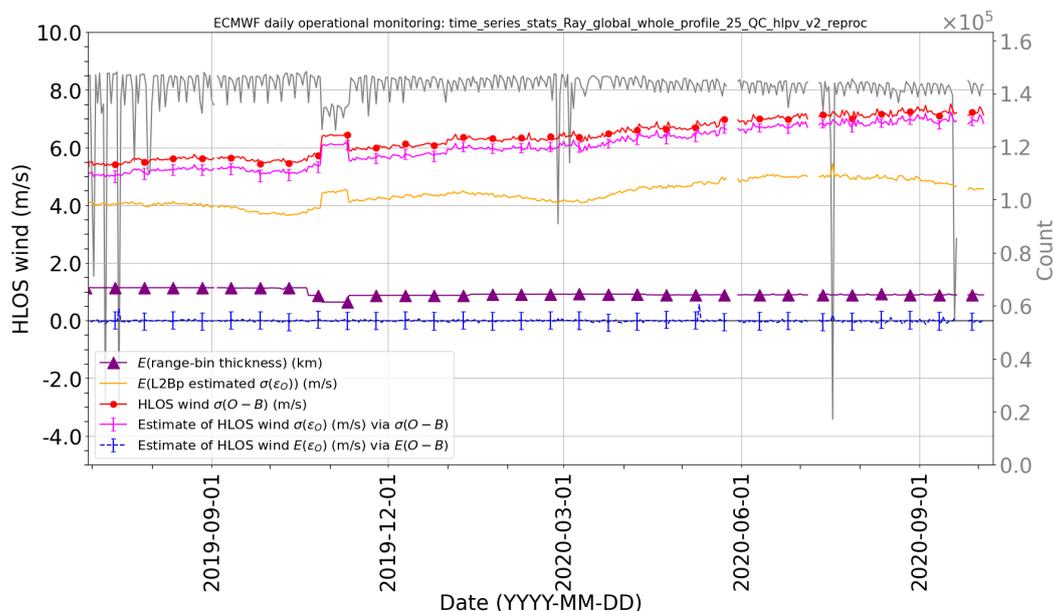


Figure 43. Time series of daily, global and all pressure level L2B Rayleigh-clear O-B statistics via the control run of the second reprocessed dataset OSE from 29 June 2019 to 9 October 2020.

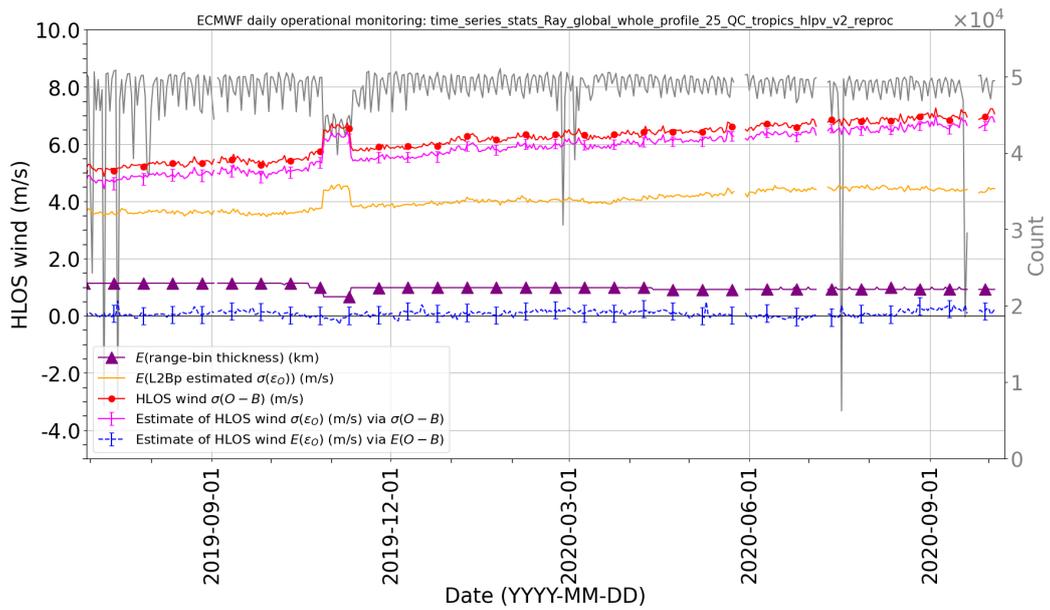


Figure 44. Daily time series for the “tropics” only ( $\pm 30^\circ$  latitude band) and all pressure level L2B Rayleigh-clear O-B statistics via the control run of the second reprocessed dataset OSE from 29 June 2019 to 9 October 2020.

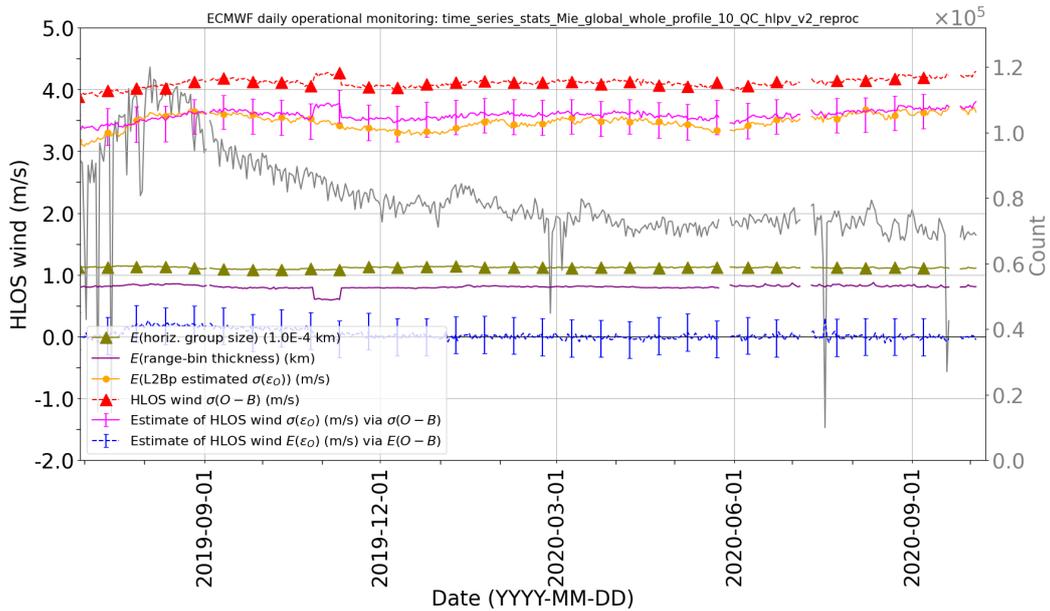
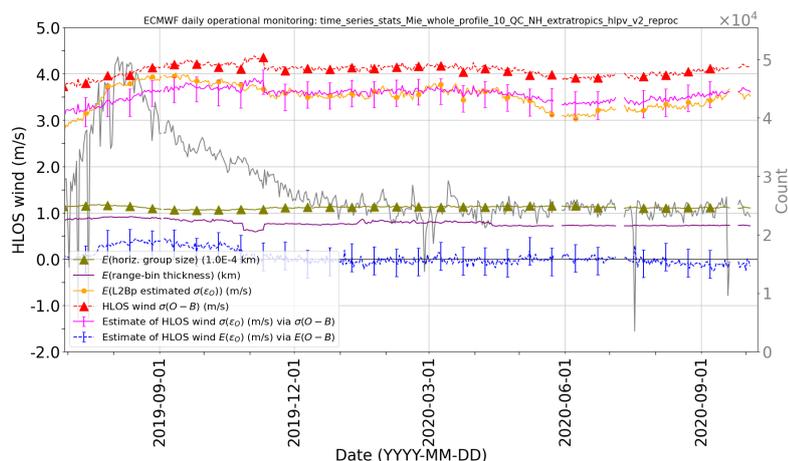
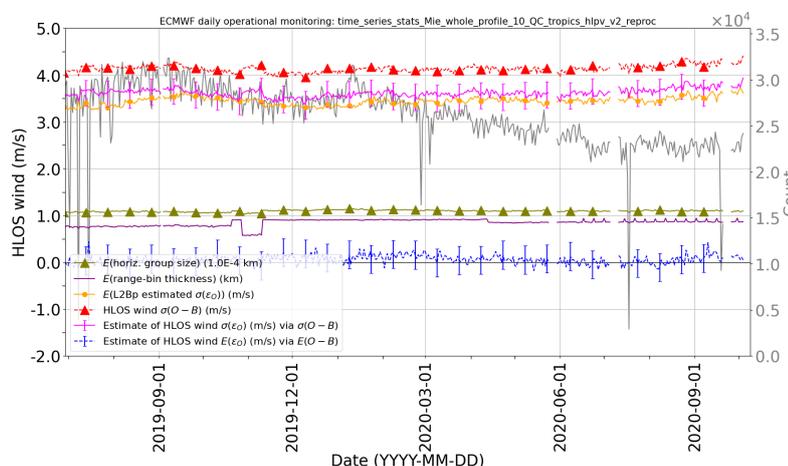


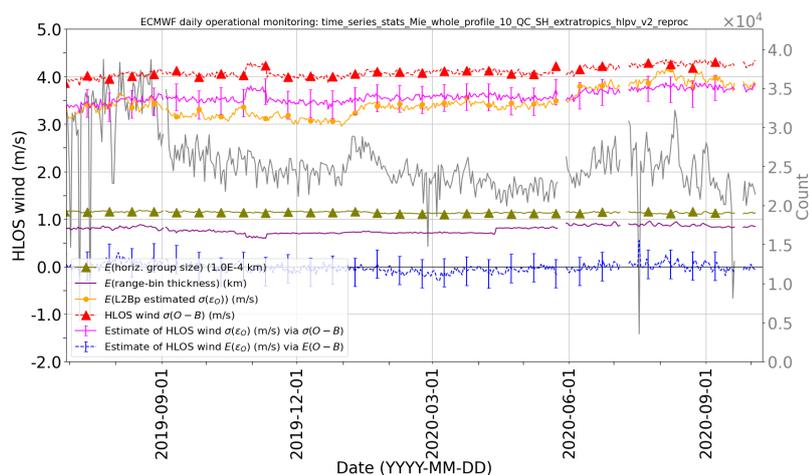
Figure 45. Daily time series of global and all pressure level L2B Mie-cloudy O-B statistics via the control run of the second reprocessed dataset OSE from 29 June 2019 to 9 October 2020.



a) NH extratropics

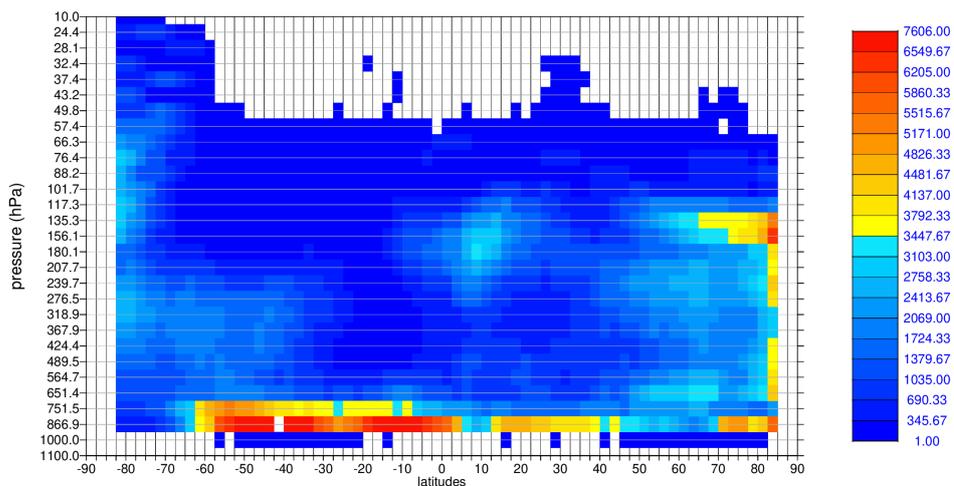


b) Tropics

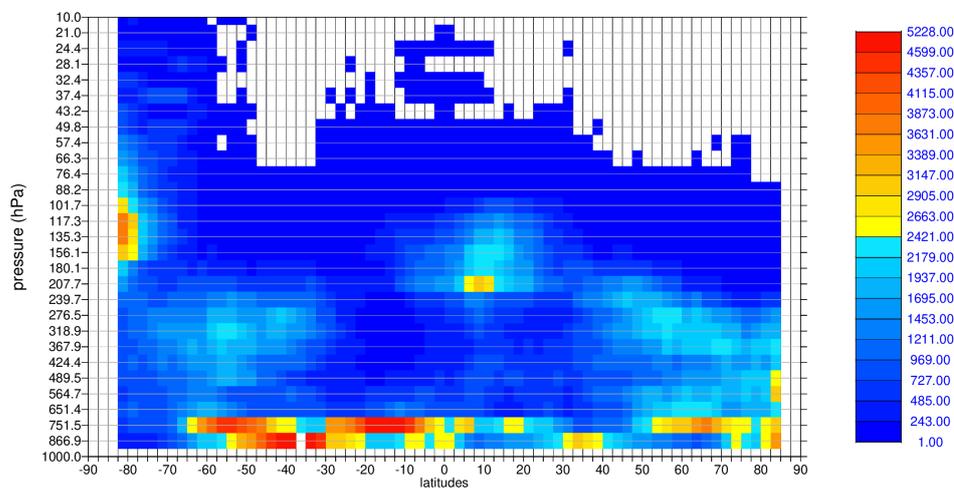


c) SH extratropics

Figure 46. Daily time series of all pressure level L2B Mie-cloudy O-B statistics via the control run of the second reprocessed dataset OSE from 29 June 2019 to 9 October 2020 for a) NH extratropics b) the tropics and c) SH extratropics.

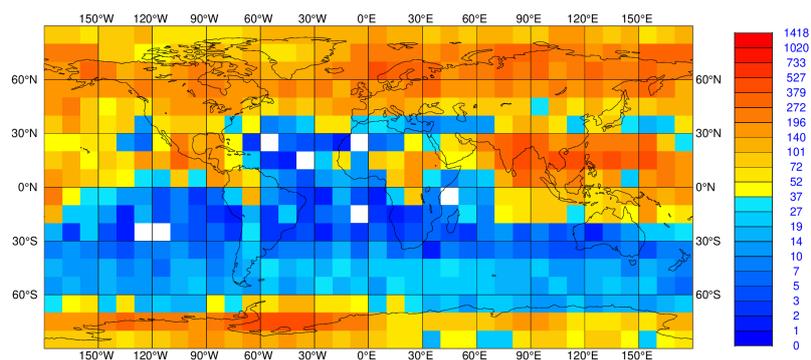


a) August 2019

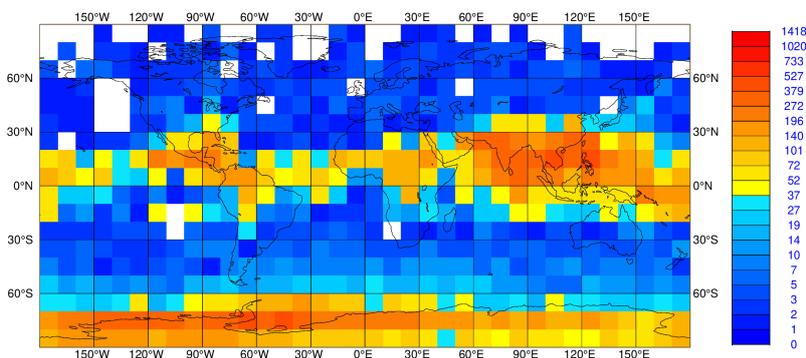


b) August 2020

Figure 47. Zonal average L2B Mie-cloudy wind data counts for the 2<sup>nd</sup> reprocessed dataset for a) August 2019 and b) August 2020. Note the scales differ.



a) August 2019



b) August 2020

Figure 48. Map of L2B Mie-cloudy wind data counts for the pressure range 117-135 hPa for the 2<sup>nd</sup> reprocessed dataset for a) August 2019 and b) August 2020.

### 3.6 Third reprocessed dataset (FM-A) monitoring

Time-series monitoring of the third reprocessed dataset (B14) using the control (no Aeolus assimilation; expid: i0jg) of the OSE was performed.

The global and profile average Rayleigh-clear random error remained reasonably stable (at around 6 m/s) from September 2018 to January 2019 and the global, daily bias was close to zero (Figure 49). The biases were also close to zero when partitioned into ascending and descending orbits (not shown). Following the FM-A laser shutdown (14 January 2019) and re-switch-on (15 February 2019), the random errors improved from 26 February 2019 due a significant change in the RBS (more suited for measuring winds) and the number of winds passing the QC increased. This was because the RBS of the CP (250 m thick range-bins near surface) were very noisy and mostly rejected by QC. This was followed by a gradual increase in noise until June 2019, as the FM-A laser energy decreased.

The Rayleigh-clear wind monitoring in the tropics (Figure 50) shows a biweekly drop in data counts which is due to a L2Bp switch to flag the winds invalid if the L1B moon-blinding flag is set. This avoids assimilating biases of order 4 m/s that have been found to occur with moon-blinding of the star-trackers – due to temporary mispointing. However, this QC is too conservative and could be improved in future to avoid rejections when the winds are not biased using O-B statistics. In the tropics the noise does not increase as sharply with time as in the global statistics, presumably due to the smaller influence of solar background noise during boreal summer.

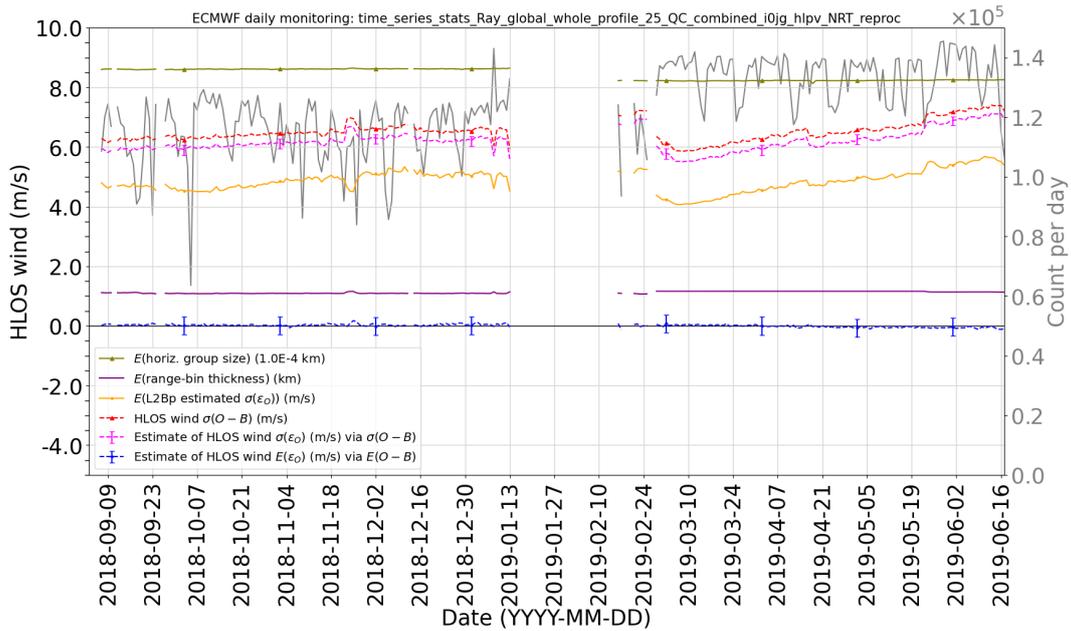


Figure 49. Time series of daily, global and all pressure level L2B Rayleigh-clear O-B statistics via the control run of the third reprocessed dataset OSE from 3 September 2018 until 16 June 2019.

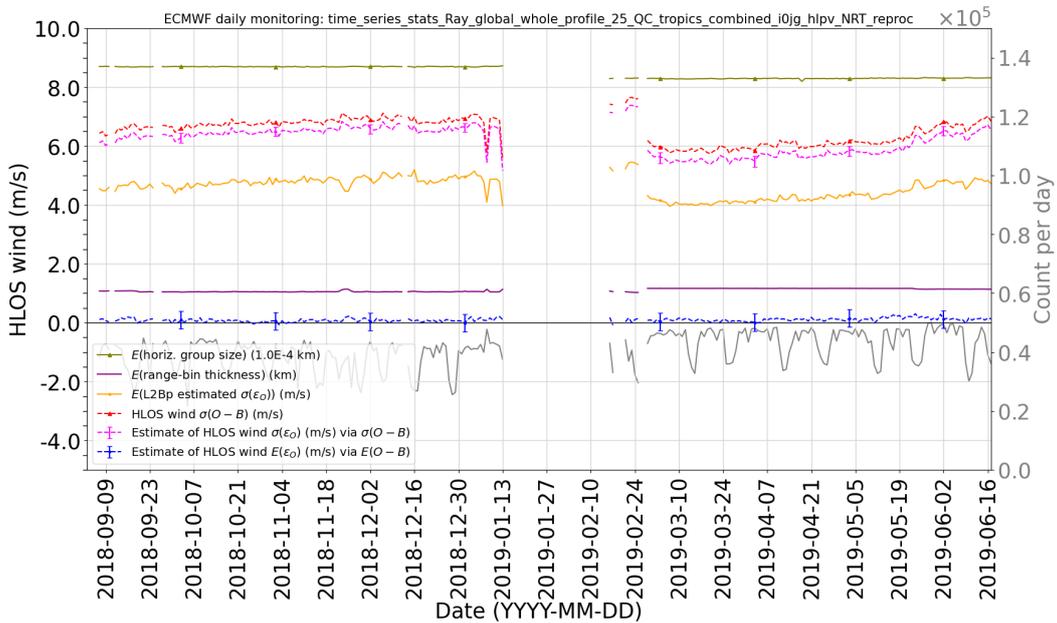


Figure 50. Daily time series for the “tropics” only ( $\pm 30^\circ$  latitude band) and all pressure level L2B Rayleigh-clear O-B statistics via the control run of the second reprocessed dataset OSE from 3 September 2018 until 16 June 2019.

The Mie-cloudy daily, global, whole profile monitoring (Figure 51) shows that the random error only increased a little during the period with a random error of typically 3 m/s (which is quite similar to FM-A in late 2022/early 2023; see Figure 52). However, the data counts decreased with time by about 20% from the peak in September 2018. This is thought to be caused by the reduction in SNR causing low backscatter Mie winds to be increasingly rejected given the fixed QC thresholds.

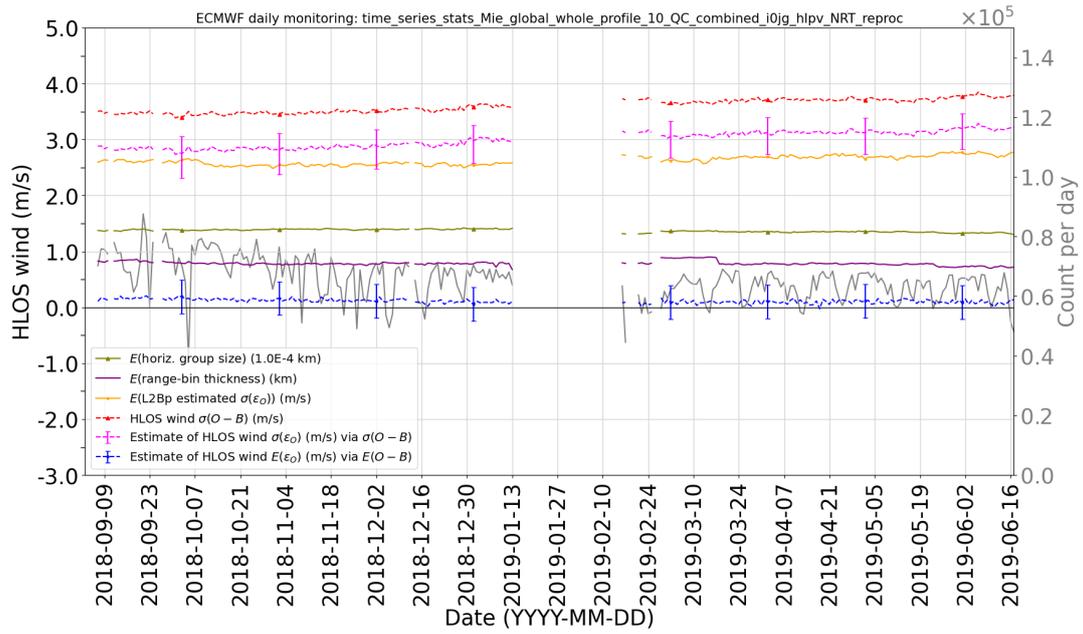
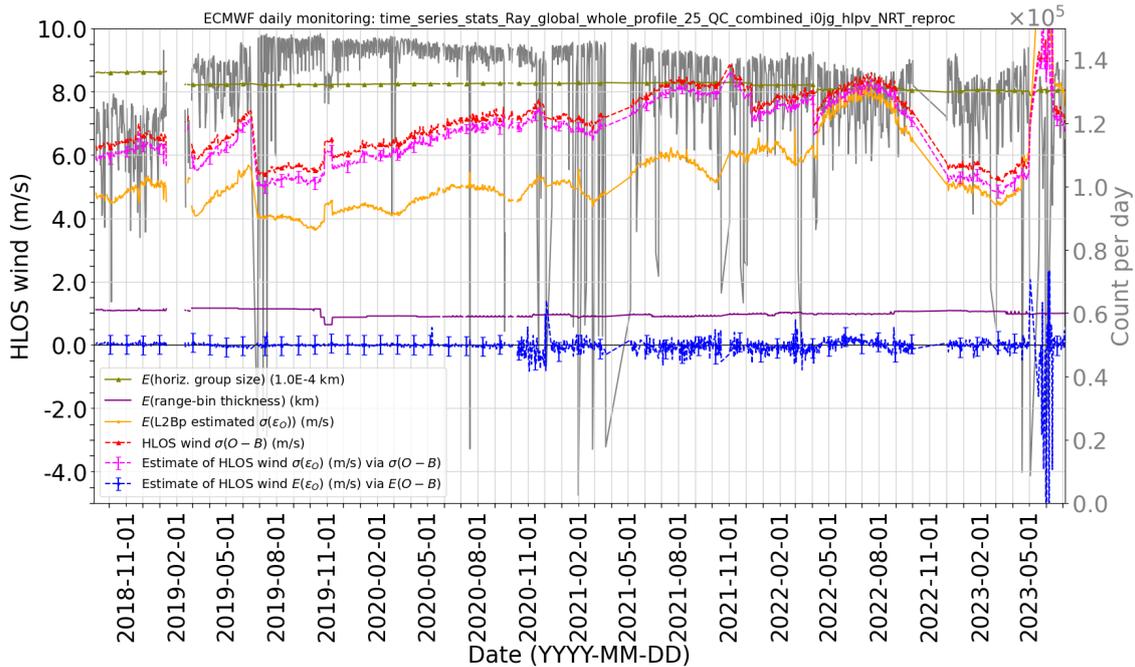


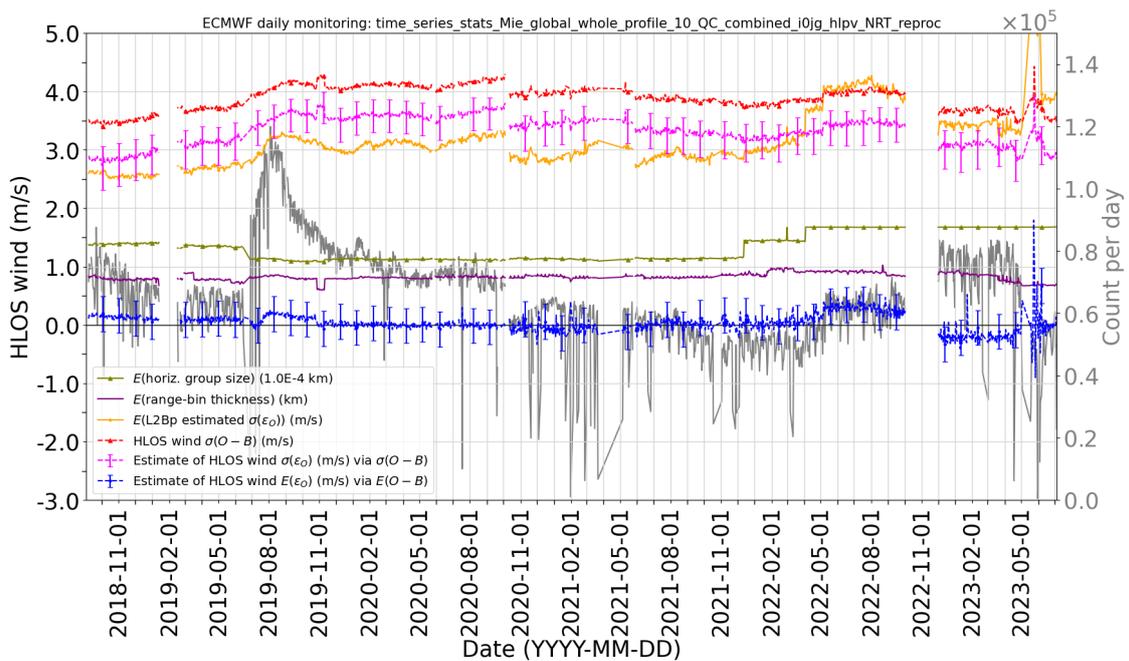
Figure 51. Daily time series of global and all pressure level L2B Mie-cloudy O-B statistics via the control run of the second reprocessed dataset OSE from 3 September 2018 until 16 June 2019.

### 3.7 Whole mission monitoring: combining the third and second reprocessing with the NRT dataset monitoring

To see a more consistent processing baseline throughout the whole mission, Figure 52 shows the concatenation of “no-Aeolus” control OSE-based statistics from the third reprocessing (early FM-A), the second reprocessing (early FM-B) and the NRT dataset operational monitoring after that until the last wind measurements of the mission, made on 5 July 2023. A noticeable feature is that the FM-A data from December 2022 until end of April 2023 were of rather good quality compared to the rest of the mission.



a) Rayleigh-clear, global



b) Mie-cloudy, global

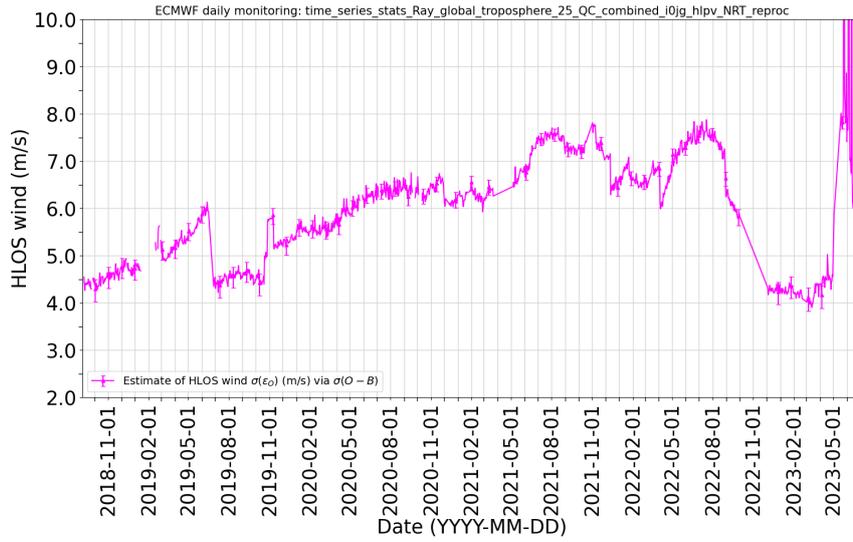
Figure 52. Time series of daily, global and all pressure level O-B statistics from the controls of the combined third (FM-A) and second reprocessing (FM-B) (no-Aeolus control OSE runs) followed by NRT (operational) dataset monitoring. A) Rayleigh-clear and b) Mie-cloudy. For the period 3 September 2018 until 5 July 2023.

### 3.7.1 Summary of Aeolus L2B wind random error statistics

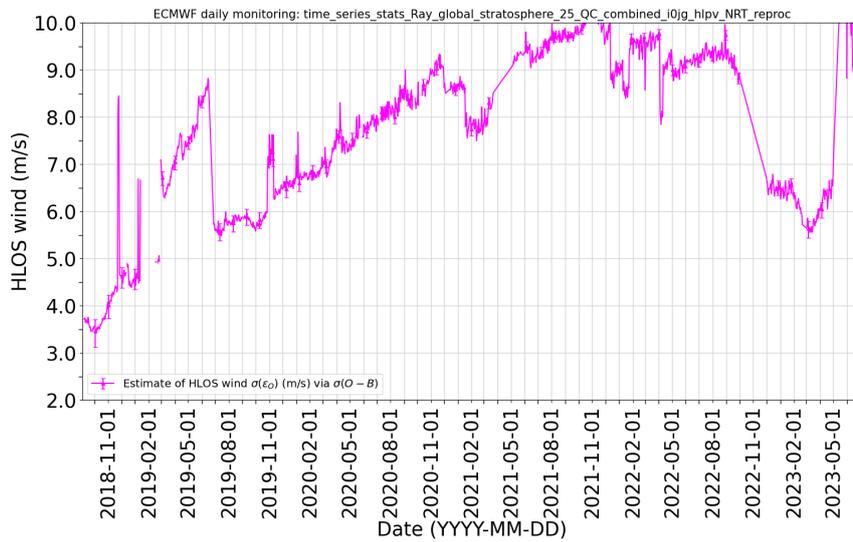
Given global O-B standard deviation statistics from the previous section (using a combination of best available reprocessing and NRT data) and an estimate of the background forecast error, an approximate global average level of the HLOS wind retrieval random error (one standard deviation) is calculated with the formula shown in Appendix 0. This assumes that the observation and background errors are uncorrelated. The observation error statistic estimate also includes the representativeness error due to mismatch between what the wind retrieval represents and what the model can represent, given, for example, discrepancies due to the point-like observation operator, and the limited spatial and temporal resolution of the ECMWF model. One should note that the magnitude of the background forecast random error statistic  $\sigma(\epsilon_B)$  varies geographically. Typically,  $\sigma(\epsilon_B)$  is largest at the tropical tropopause and smallest in the NH mid-latitudes based on Ensemble Data Assimilation spread statistics (see Figure 11).

An estimate for  $\sigma(\epsilon_B)$  was derived using “Desroziers diagnostics” (Desroziers et al. (2005)) for the radiosonde zonal wind component and Aeolus HLOS wind retrievals. This led to a typical range for  $\sigma(\epsilon_B)$  from 1.5 to 2.5 m/s in HLOS wind space, varying with atmospheric pressure (altitude) and geolocation. Due to the variability of  $\sigma(\epsilon_B)$ , we have used this range to define confidence intervals for the  $\sigma(\epsilon_O)$  estimates.

Plots of the estimated global average random observation error,  $\sigma(\epsilon_O)$ , for Aeolus L2B HLOS winds throughout the mission are shown in Figure 53 and Figure 54. The plots are restricted to the defined pressure ranges: 800-1100 hPa (surface to about 2 km) to roughly represent the planetary boundary layer (PBL); 100-800 hPa (~2-16 km) to approximately represent the free troposphere; and above 100 hPa to represent the lower stratosphere.

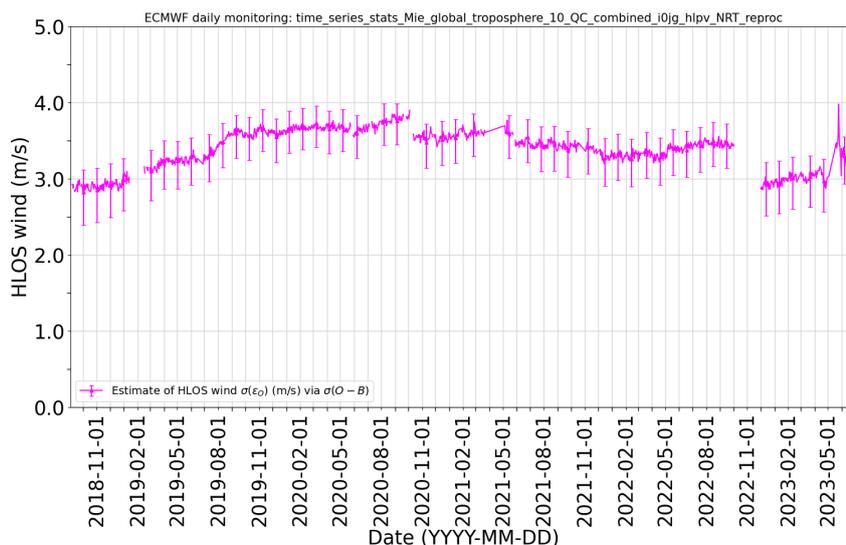


a) Free troposphere

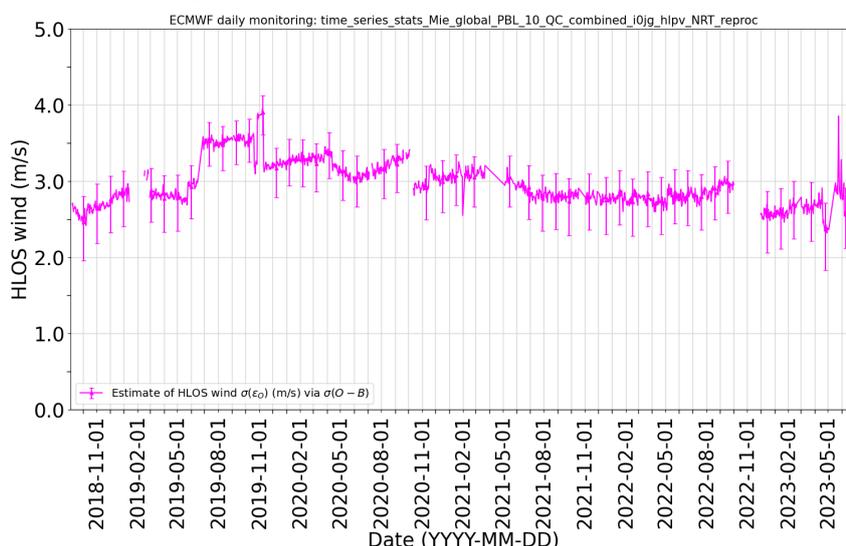


b) Lower stratosphere

Figure 53. Estimated global average random error ( $\sigma(\epsilon_0)$ ) for Aeolus L2B Rayleigh-clear HLOS winds during the mission for a) Free troposphere and b) Lower stratosphere.



a) Free troposphere



b) PBL

Figure 54. Estimated global average random error ( $\sigma(\epsilon_0)$ ) for Aeolus L2B Mie-cloudy HLOS winds during the mission for a) Free troposphere and b) Planetary Boundary Layer.

The following tables summarise the best and worst estimated global average random error ( $\sigma(\epsilon_0)$ ) found during the mission for the different wind types and pressure ranges.

Table 2. A summary of the estimated global average random error ( $\sigma(\epsilon_0)$ ) for Aeolus L2B Rayleigh-clear HLOS winds in different parts of the atmosphere.

Pressure range	1- $\sigma$ observation error estimate (m/s)	
	Best	Worst
Free troposphere	4.0	7.5
Lower stratosphere	3.5	10.0

Table 3. A summary of the estimated global average random error ( $\sigma(\epsilon_0)$ ) for Aeolus L2B Mie-cloudy HLOS winds in different parts of the atmosphere (PBL is planetary boundary layer).

Pressure range	1- $\sigma$ observation error estimate (m/s)	
	Best	Worst
Free troposphere	3.0	3.6
PBL	2.5	3.5

The Mie-cloudy winds are significantly less noisy than the Rayleigh-clear, with the best case for Mie-cloudy random error (1- $\sigma$ ) of 2.5 m/s in the PBL and 3.0 m/s in the free troposphere (both found with FM-A data), whereas the best performance for the Rayleigh-clear is 3.5 m/s in the lower stratosphere (due to 2 km thick range-bins for the Commissioning Phase) and 4.0 m/s in the free troposphere (for FM-A in late 2022/early 2023). The Aeolus L2B HLOS Rayleigh wind random error statistics are larger than the pre-launch mission requirements (Ingmann and Straume, 2016) which agrees with the atmospheric path signal levels being poorer than pre-launch expectations (see the Discussion in Section 0).

## 4 Results of Aeolus NWP impact assessment at ECMWF

### 4.1 Commissioning Phase NRT FM-A OSEs

The Commissioning Phase data assimilation experiment using Aeolus L2B Rayleigh-clear and Mie-cloudy HLOS winds (referred to as “Rayleigh+Mie”, see section 2.4.1.1) showed the expected 4D-Var analysis behaviour, with the analysis being pulled closer to Aeolus, as is shown by the smaller standard deviation and mean of O-A relative to O-B statistics, see Figure 55. The analysis pulls much closer (relatively) to the Mie-cloudy winds compared to the Rayleigh-clear due to the Mie’s significantly smaller assigned observation errors (less than half the Rayleigh). However, there are about five times as many Rayleigh winds as Mie assimilated and by sampling the clear air the Rayleigh cover a much larger fraction of the atmosphere. Global average statistics such as in Figure 55 show relatively small biases of less than 1 m/s (as expected during this period chosen for its small biases, see section 2.4.1.2). However, there does appear to be some positive global average bias for upper levels for both the Mie-cloudy and the Rayleigh-clear winds, and at lower levels for the Mie-cloudy winds. There is a negative bias for the Rayleigh at around 500 hPa, which is also seen in Figure 13, which is a range-bin affected by hot pixels that was not totally removed by quality control.

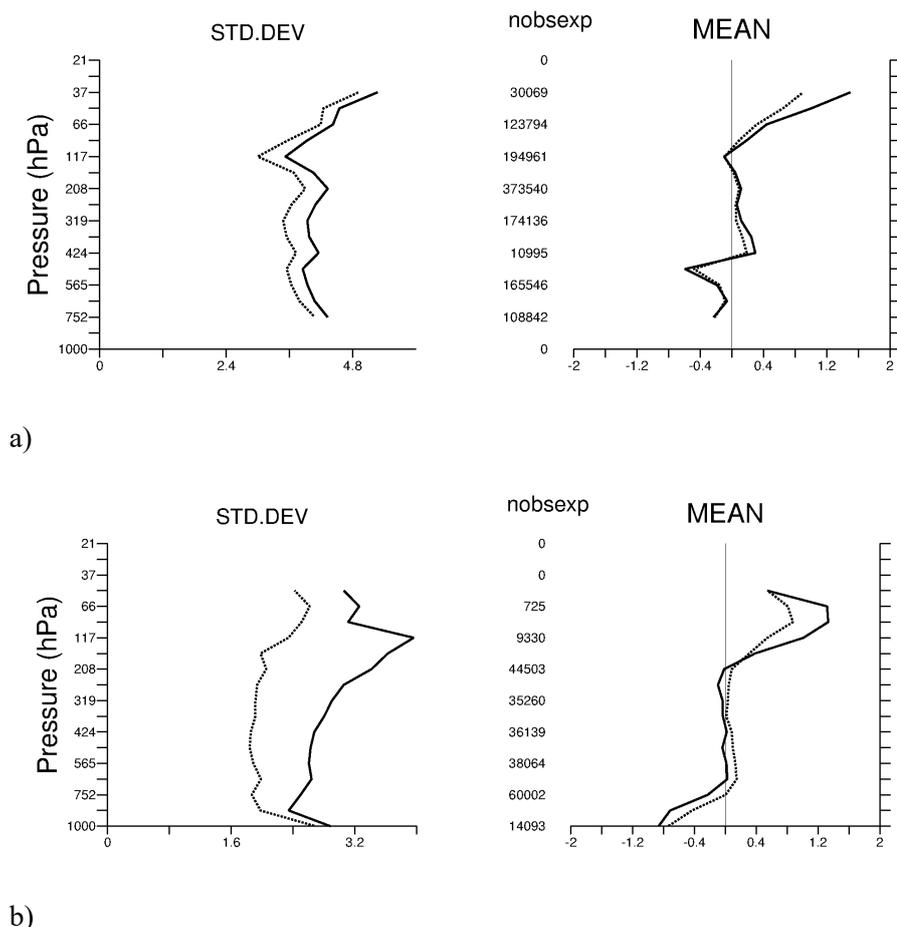


Figure 55. Aeolus HLOS wind global O-B (solid) and O-A (dotted) departures statistics for “used” data as a function of pressure in a 4D-Var experiment in which L2B Rayleigh-clear plus Mie-cloudy

*HLOS winds are assimilated. The variable on the x-axis is HLOS wind (m/s). The standard deviation is shown on the left, and the mean on the right and the sample size for the pressure bins is written in the middle. A) Shows the L2B Rayleigh-clear statistics and in b) the L2B Mie-cloudy statistics. The period is from 12 September 2018 to 16 October 2018.*

Assimilating Rayleigh-clear winds influences the mean analysis u-wind component, as shown in Figure 56 (in both a) Rayleigh+Mie and b) Mie only experiments). Note that the effect of Mie-cloudy winds on the mean state is much smaller than the Rayleigh-clear. The effect is strongest in the tropics for Rayleigh+Mie, where the generally easterly winds are increased by 0.5 m/s in the zonal average (not shown). This change leads to an increase in bias with respect to tropical radiosonde u-wind mean(O-B) (not shown), hence it is assumed to be caused by Aeolus observation bias and not by a truthful correction of an ECMWF model bias. However, the bias is improved with respect to radiosonde v-wind. The systematic u-wind changes persist in the longer forecast ranges (up to 8 days at least).

The wind speed dependence of the Rayleigh-clear bias (as discussed in reference to Figure 21 c) shows a slow bias of -2.5% for most of September 2018. Such a bias should reduce the tropical easterly wind speed and not increase it. On closer inspection, it was found that for the small absolute wind speeds present in the tropics, the descending orbit phase is positively biased relative to the ascending phase by about 2 m/s (not shown; but in hindsight likely to be an M1 temperature-dependent-bias; see section 3.2). It appears that the descending phase bias is causing the tropical easterlies to become stronger. The slight decrease by 0.03 m/s in the stratospheric polar vortex (polar night jet) in the SH, is however thought to be a result of the Rayleigh-clear slow bias, as shown in Figure 21 c).

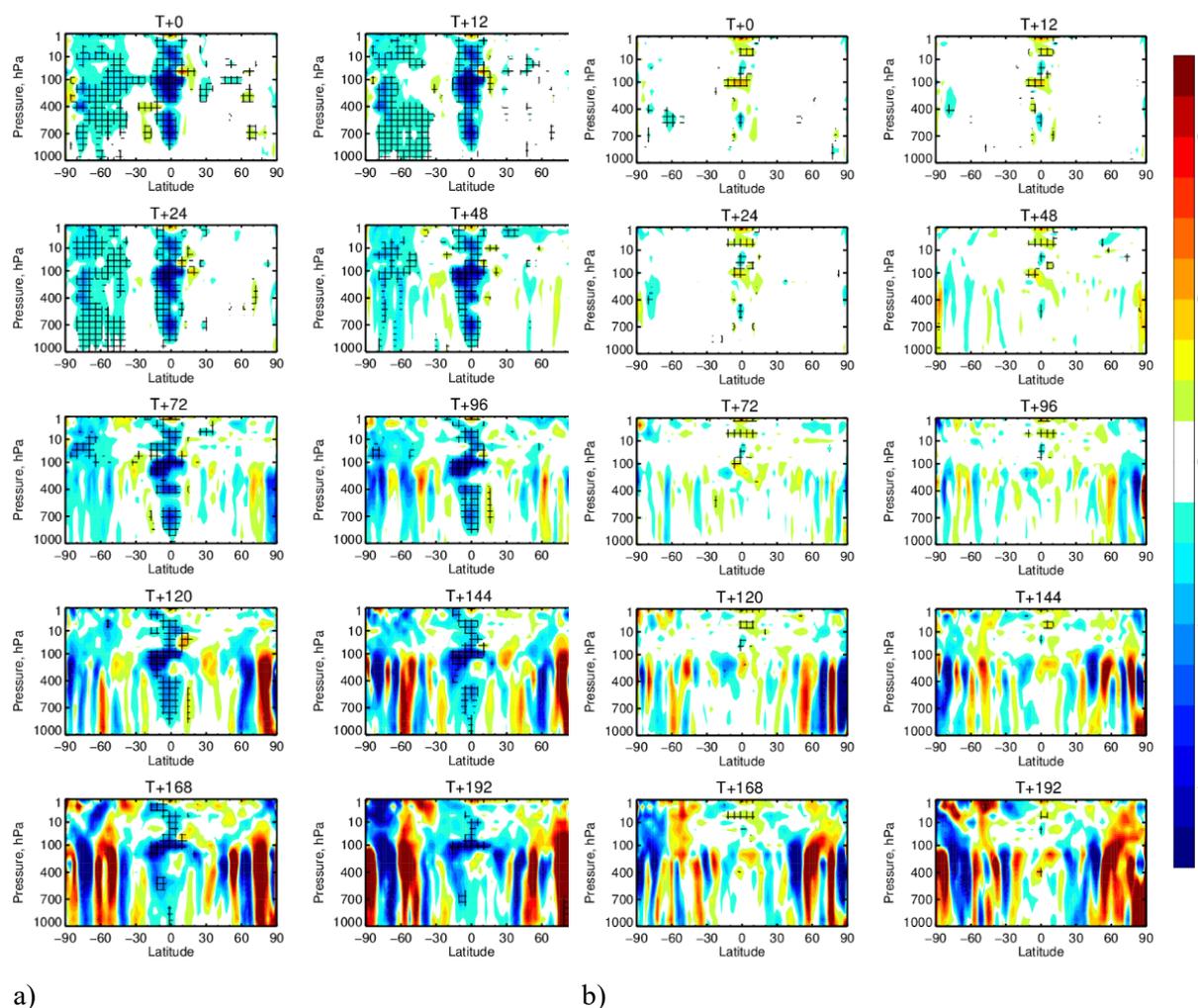


Figure 56. The zonal mean change in analysis ( $T+0$  hours) and forecasted ( $T+12$  to  $T+192$  hours)  $u$ -wind field resulting from assimilating a) both L2B Rayleigh-clear and Mie-cloudy HLOS winds and b) L2B Mie-cloudy winds only. The scale covers  $\pm 0.25$  m/s.

The impact of Aeolus in this relatively short OSE was most robustly demonstrated via the short-range forecasts (up to 12-hour forecast) fit to other assimilated observation types, see Figure 57. The other observation types shown consist of conventional wind observations (radiosondes, pilots, aircraft and radar wind profilers); AMSR-2 (Advanced Microwave Scanning Radiometer 2) all sky radiances; radiosonde temperature; aircraft temperature; geostationary satellite radiances; GNSSRO (Global Navigation Satellite System radio occultation) and AMSU-A (Advanced Microwave Sounding Unit-A) radiances (mostly temperature information). For the Rayleigh+Mie experiment, the fit to conventional wind observations (which is dominated by the very large sample of aircraft wind observations) shows an improvement of around 1% in the SH at 300 hPa ( $\sim 8$  km). The improvement in the tropics is roughly 0.7%, with only a small improvement evident in the NH. Note that the fit to  $u$  and  $v$  radiosonde wind components per se (not shown) is larger e.g. up to 1% in the NH, 2% in the tropics and 1% in the SH; perhaps because aircraft follow flight paths restricted in spatial coverage meaning that the short-range forecast winds are much better in such areas. The assimilation of only the Mie-cloudy winds (red lines) seems to provide a reasonable fraction of the Rayleigh+Mie impact (black lines). Wind improvements of order 1% in SH and tropics are comparable to the impact of other satellite observation types such as

infrared sounders, AMV (atmospheric motion vectors) or GNSSRO in troposphere as shown in ECMWF OSE denial experiments (Bormann et al., 2019).

Globally it is seen that the fit to radiosonde and aircraft temperatures is improved by 0.5% at 300-400 hPa. The fit to GPSRO is improved by 0.8% at 11 km. In terms of humidity and cloud information there is an improvement of 0.8% against some channels of AMSR-2 and improvements in other humidity sensitive observations such as e.g. MHS (Microwave Humidity Sounder) and geostationary infrared imagery. The only observation type to show a clear degradation are the microwave temperature sounders in the lower stratosphere i.e. AMSU-A and ATMS (Advanced Technology Microwave Sounder). AMSU-A channel 11, which peaks around 25 km, shows the largest degradation, however Aeolus only provided winds only up to around 20 km altitude for this period, therefore the degradation must be an indirect effect of its assimilation. The Rayleigh-clear winds lead to this degradation, because the Mie-cloudy only experiment does not show the negative impact against AMSU-A.

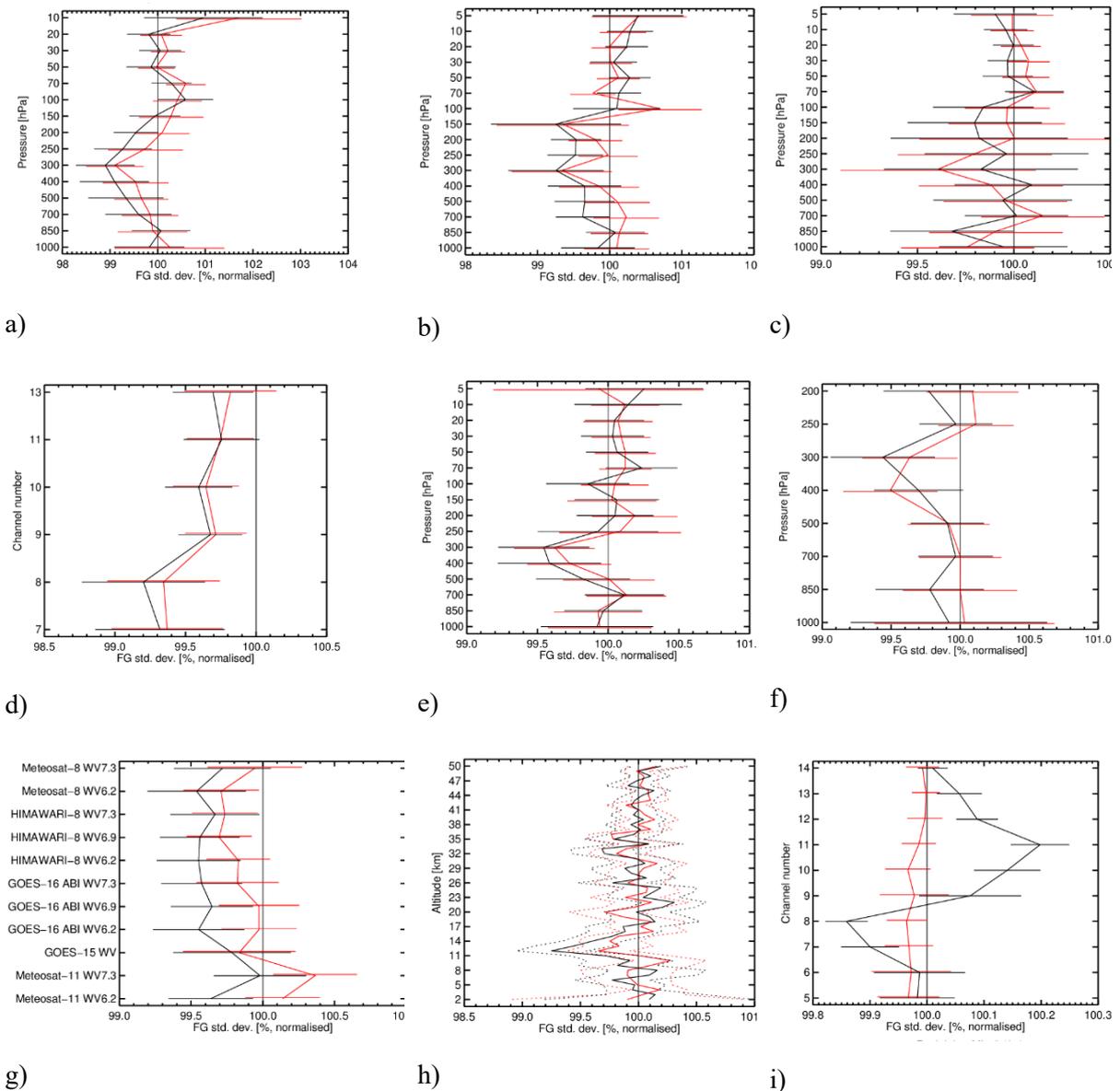


Figure 57. Change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are: a) conventional vector wind in the SH extratropics, b) conventional vector wind in the tropics, c) conventional vector wind in the NH extratropics, d) global AMSR-2 all sky radiances, e) global radiosonde temperature, f) global aircraft temperature, g) global geostationary satellite radiances, h) global GNSSRO bending angle and i) global AMSU-A radiances. The black lines are the Rayleigh+Mie experiment and the red lines are the Mie only experiment. Horizontal bars show the 95% confidence range. The period is from 12 September 2018 to 16 October 2018.

The impact of Aeolus on the longer forecast ranges as verified against own analyses is shown in Figure 58 (both experiments) and Figure 59 (the Rayleigh+Mie experiment only). There is a tendency towards positive impact for the Rayleigh+Mie experiment in the SH at longer forecast ranges of order 2%; and in the tropics at higher altitudes. The impact on the 500 hPa geopotential heights in the SH is also positive (not shown). However, the confidence intervals suggest the impact is mostly lacking in statistical significance due to the sample being too small. A longer reprocessed FM-A L2B dataset without bias drift (e.g. six months) should allow for a more robust assessment of the longer-range forecast impact (longer experiments were since then performed for different time periods, as shown in the following sections). The impact of the Mie-cloudy only experiment is neutral, as can be seen in Figure 58, in contrast to the short-range forecast fit to other observation types; this lack of Mie-cloudy impact for this period is not yet understood. The apparent degradation in the shorter range forecasts (red areas in Figure 59) is a typical feature of verification against own analyses when adding new observations which add new structures to the forecast fields, and given the improved fit of short-range forecasts to other observation types it is not of great concern. However, some of the red persists to T+72 hours in the lower stratosphere which may indicate an overweighting of the Aeolus winds.

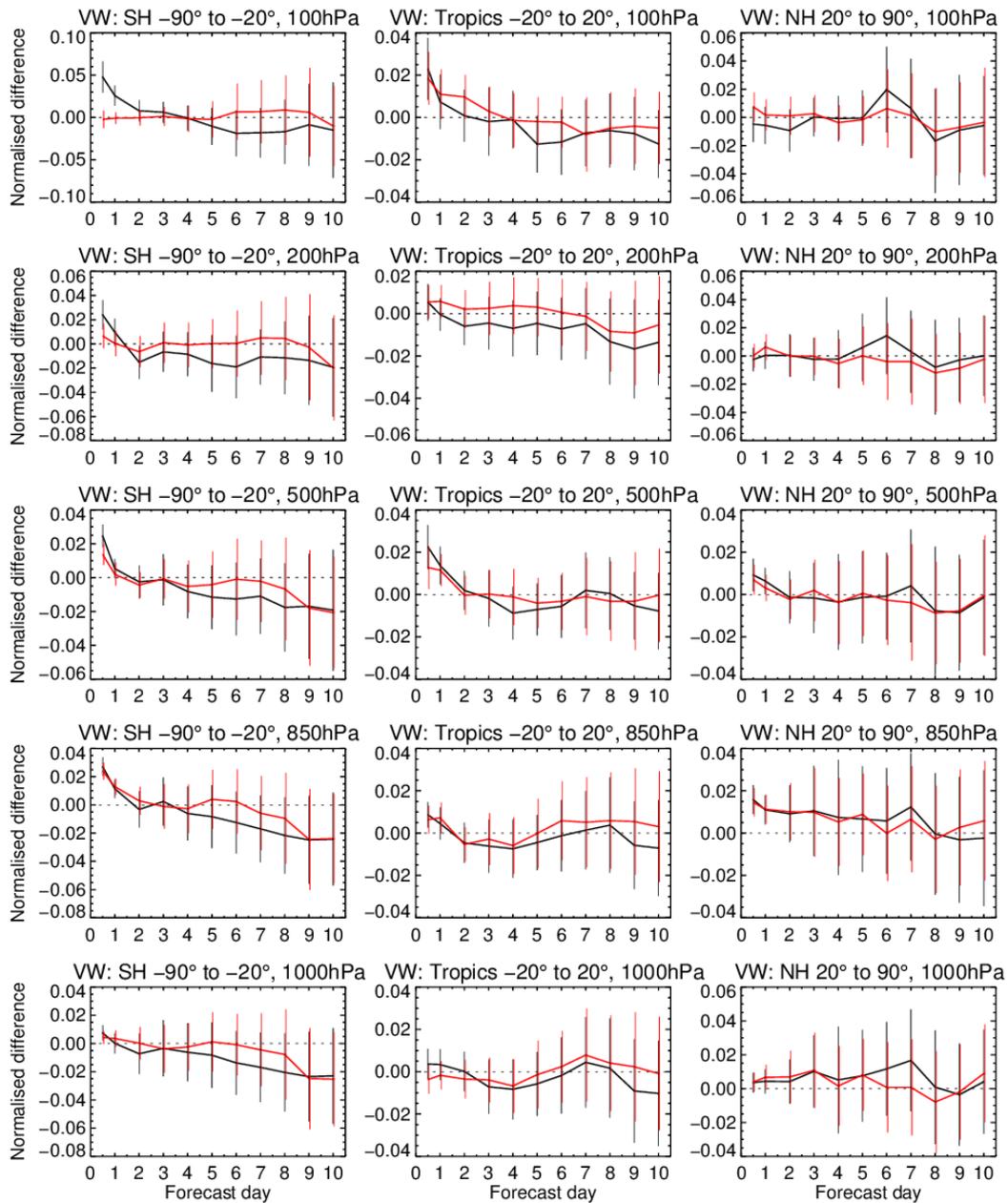


Figure 58. Normalised change in the RMS (root mean square) error in wind vector for different vertical levels (from 100 to 1000 hPa, top to bottom) and for SH, tropics and NH (left to right). The black line is for the Rayleigh-clear plus Mie-cloudy experiment and the red line is for the Mie-cloudy only experiment. The period is from 12 September 2018 to 16 October 2018. Confidence ranges are 95%. Negative values indicate a reduction in error from assimilating Aeolus. Verified against own analyses.

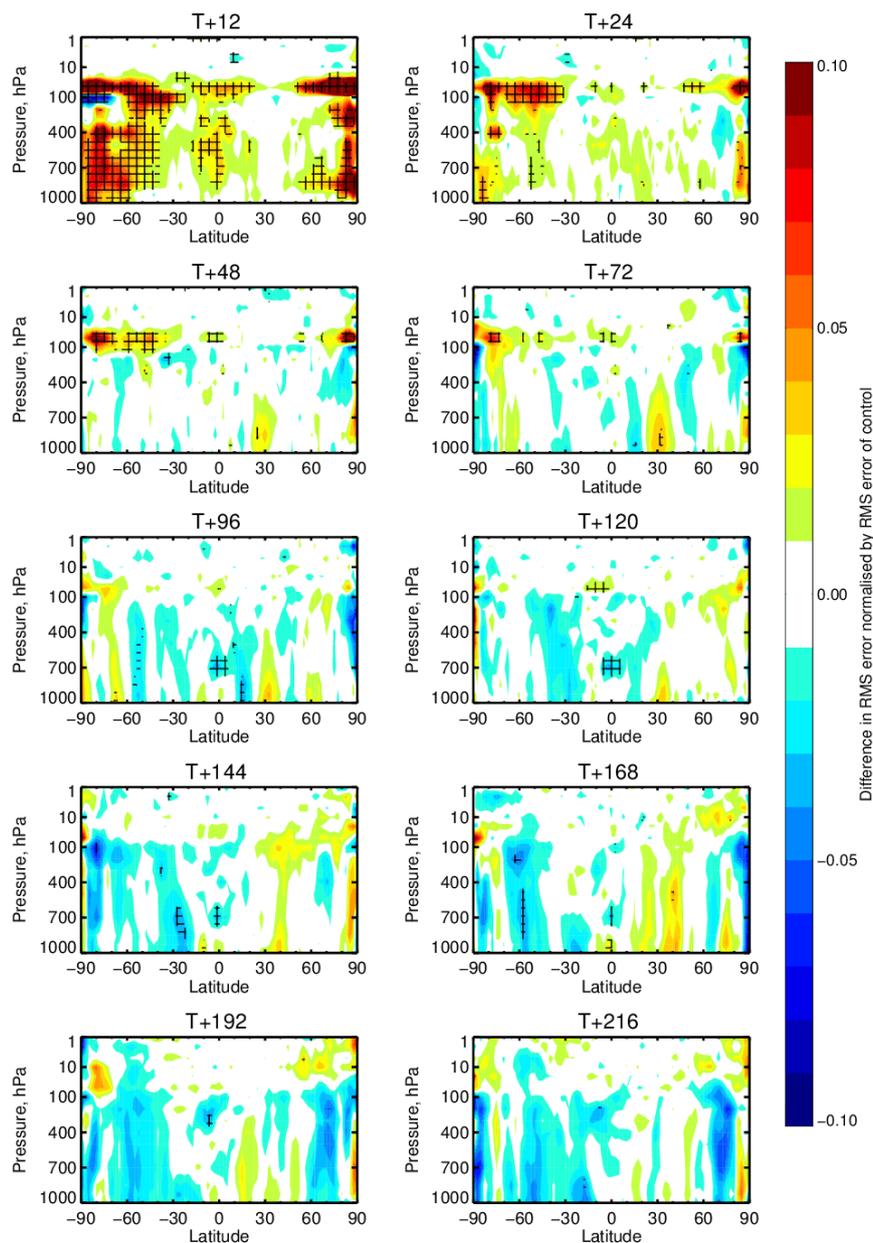


Figure 59. Zonal average normalised change in the RMS (root mean square) error of wind vector for the Rayleigh-clear plus Mie-cloudy experiment for the period 12 September 2018 to 16 October 2018. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against own analyses. Note that the scale is large at  $\pm 10\%$  which usually happens when an experiment has not run for long enough period to settle down.

## 4.2 2019 NRT FM-A OSEs

This section provides the verification results for Aeolus experiments performed for the 2019 FM-A period as described in sections 2.1.2 and 2.4.2.

### 4.2.1 Experiments without bias correction

Firstly, the results obtained before a bias correction scheme was considered are presented.

Figure 60 is a map of the standard deviation of the differences of 200 hPa zonal wind analyses with and without Aeolus assimilation. The largest effects of Aeolus L2B winds on the analysis are in tropical convectively active areas (along the intertropical convergence zone) and further south over the Pacific Ocean in the extratropical convergence zones (as shown by the satellite imagery of Figure 61, a snapshot from early April 2019). There are also large differences in the south-eastern United States, which was convectively active for this period, and in the extratropical convergence zones in the Northern Hemisphere oceans.

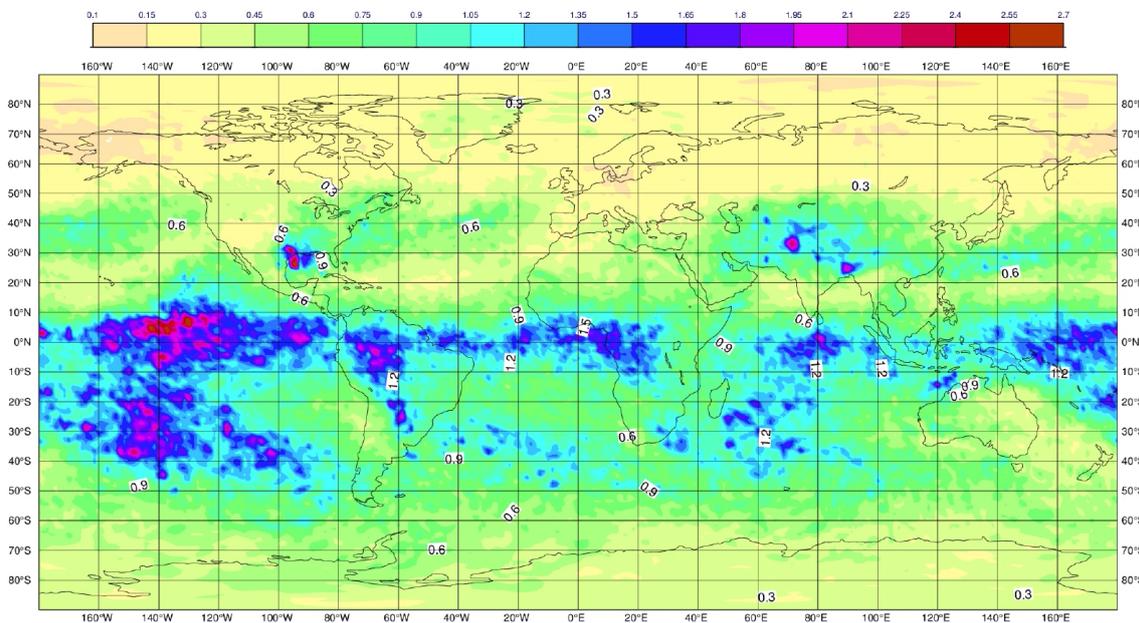


Figure 60. Standard deviation of the differences of 200 hPa u-wind component between analyses assimilating Aeolus and the control analyses without Aeolus for the period 4-23 April 2019. Unit of the colour-scale is m/s. The scale is from 0.1 to 2.7 m/s.

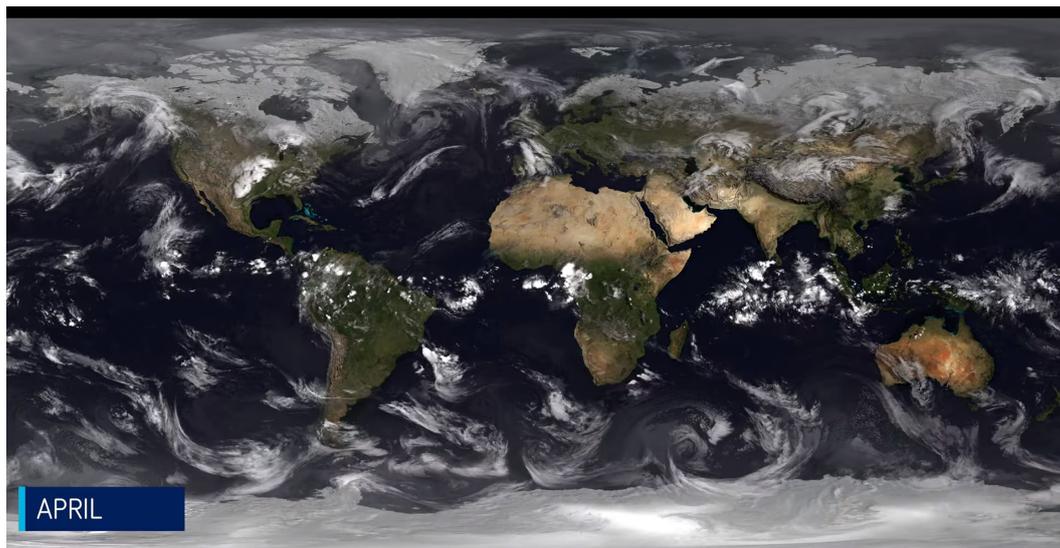


Figure 61. Composite of geostationary satellite imagery in early April 2019 (courtesy of EUMETSAT).

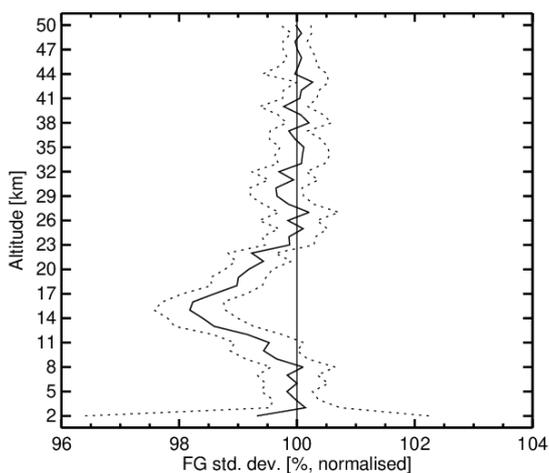
The impact of Aeolus late FM-A period data without bias correction on short-range forecasts is shown in Figure 62, as assessed by the background fit to other observation types. The positive impact is greatest in the tropics. In a), the strong (2%) improvement in the fit to GNSS radio occultation bending angles in tropical upper troposphere indicates an improvement in background temperature; this peaks at about 15 km altitude. The tropical impact is also seen via the improved background fit to in situ wind observations in b), for which the positive impact peaks at 150 hPa (upper troposphere). In c) the fit to AMSU-A radiances in the global sense is improved most for channel 9, which has a peak sensitivity to temperatures at roughly 17 km.

The pattern of degraded fit to AMSU-A radiances which occurred in the early FM-A experiments (Figure 57) is absent, which is perhaps explained by thinner range-bin settings in upper troposphere and lower stratosphere for the late FM-A period. Thinner range-bins reduces the observation operator error (in which we assume no vertical averaging of the winds, see section 2.2). Alternatively, it could be because of the improved Rayleigh-clear calibration in terms of wind-speed dependent error for this period.

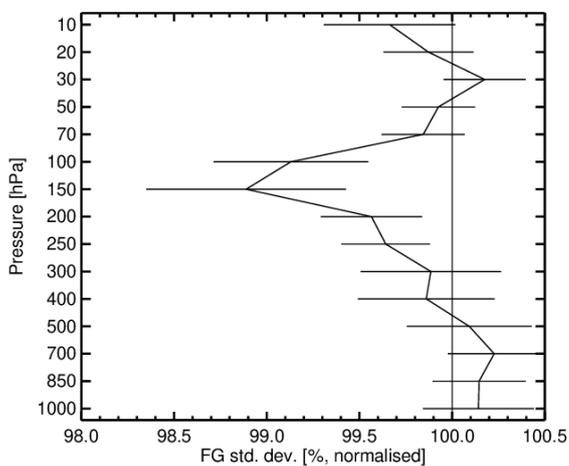
The statistical significance is improved compared to the early FM-A period due to the longer test period. The results show that Aeolus is improving short-range temperature and wind forecasts in the upper troposphere. However, there appears to be less impact in mid-troposphere than the early FM-A period and not so much impact on humidity. This may be due to overweighting the much more numerous  $\sim 10$  km horizontal resolution Mie winds<sup>26</sup> or possibly due to thicker Mie range-bins at lower altitudes. These results were obtained with the scaling of observation error of 1.4 for the Rayleigh and 1.2 for the Mie.

It is reassuring that positive NWP impact was obtained from Aeolus for this period in which the FM-A laser energy was rather low (40-50 mJ) and hence the atmospheric path signal was weaker; there were also many hot-pixels reducing the assimilated data count; and there were considerable uncorrected biases. This impact should be significantly larger with reprocessed data for this period.

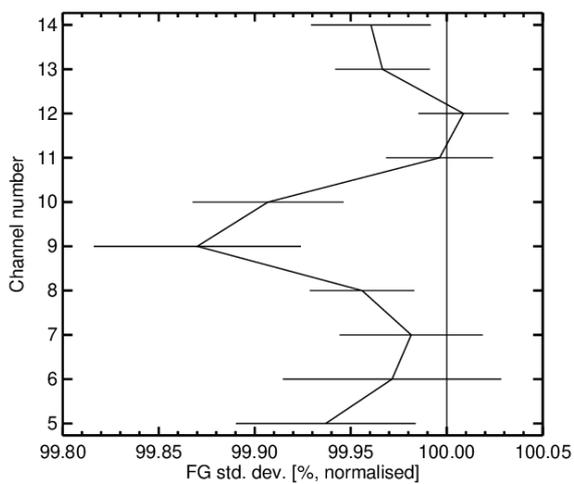
<sup>26</sup> An experiment should be considered (from reprocessed datasets) to assess if we can obtain more NWP impact from a larger sample of higher resolution Mie-cloudy winds or from longer horizontal averaging and fewer winds.



a)



b)



c)

Figure 62. Change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations (Rayleigh-clear and Mie-cloudy), normalised so that the control is 100%. Values

*below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) GNSS radio occultation in the tropics, b) conventional vector wind in the tropics and c) global AMSU-A radiances<sup>27</sup>. Horizontal bars show the 95% confidence range. For the period 2 April 2019 to 14 June 2019.*

The impact of Aeolus FM-A data in 2019 with no bias correction on longer range forecasts as verified against operational ECMWF analyses is shown in Figure 63. The results are made up from 128 to 147 samples (with more samples for short-range forecasts). The figure shows more positive (blue) than negative (red) impact for days two to four, particularly in the tropics, but it lacks statistical significance. There seems to be some negative impact in the NH for days five to seven.

---

<sup>27</sup> An internet search for AMSU-A weighting functions will show vertically which vertical layers in the atmosphere the channels are typically sensitive to.

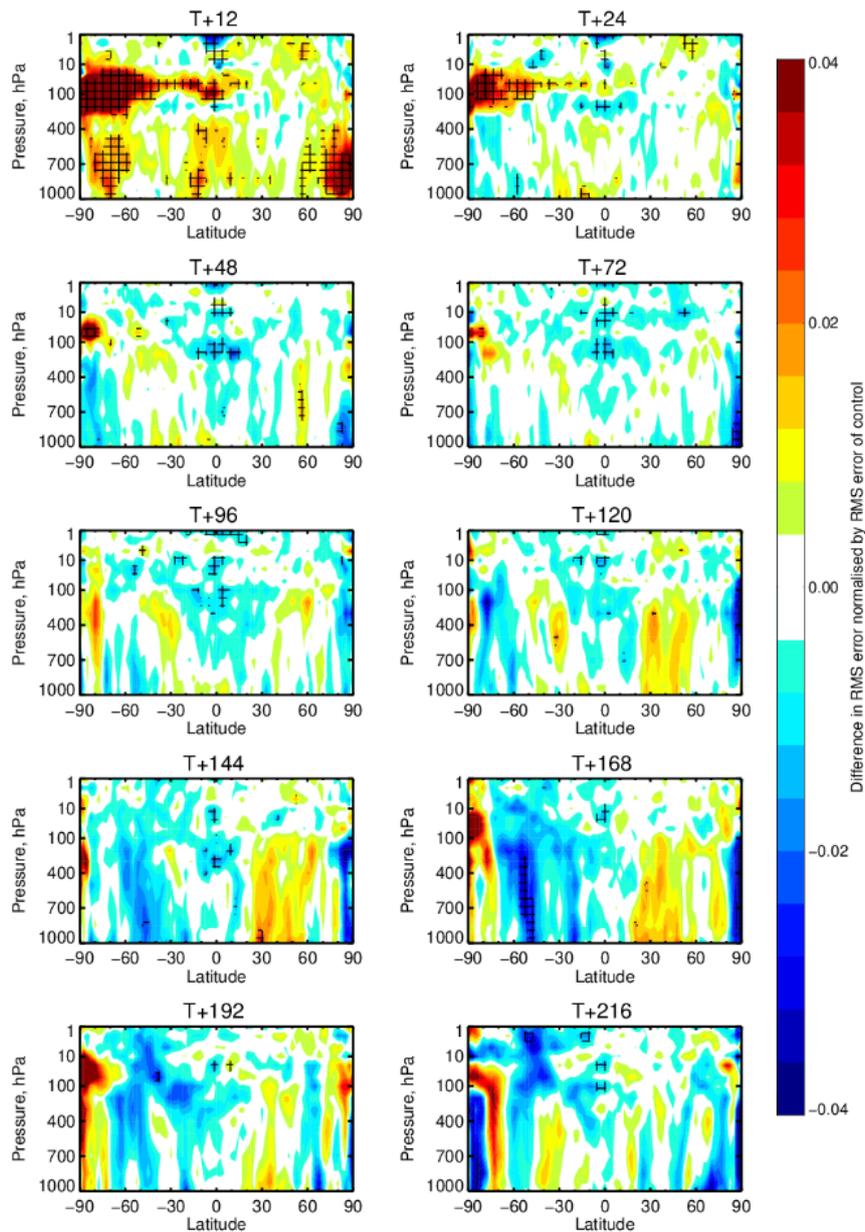


Figure 63. Zonal average normalised change in the RMS (root mean square) error of wind vector for the Rayleigh-clear plus Mie-cloudy experiment (with no bias correction applied) for the period 2 April 2019 to 14 June 2019. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicates 95% confidence.

#### 4.2.2 Experiments with a bias correction scheme

As shown with O-B departure statistics in section 3.2, the L2B Rayleigh-clear bias was found to strongly vary with argument of latitude for this late FM-A period and the magnitude of the bias became large in the April-June 2019 period, changing significantly over a matter of weeks in the NH. To try to correct for this, a bias correction scheme was developed in the ECMWF IFS code using the ECMWF model as a reference. It was implemented via look-up tables of mean(O-B) binned by argument of latitude, which

at the time the bias was known to have strong dependence on<sup>28</sup>. The bias correction look-up table is applied in the BUFR to ODB (Observation Database) IFS code step. Further information on the bias correction is provided in section 2.4.3.2 and section 4.3.1. The bias correction look-up tables were calculated from operational O-B departure statistics (i.e. bias with respect to the operational short-range forecasts); typically using a week-long sample. It was found that weekly updates of the bias correction seemed to be a reasonable update rate for the Rayleigh channel. The Mie biases were found to be much more stable with time so less frequent updates were required, however some dependence on argument of latitude existed.

The ECMWF bias correction scheme improved the impact of Aeolus on short-range forecast skill, as can be seen in the fit of short-range forecasts to important observation types sensitive to temperature, humidity and winds (ATMS and in situ vector wind observations) in Figure 64. The black line (with bias correction) being more to the left than the red line (without bias correction) demonstrates an improvement in Aeolus impact (statistically significant in the troposphere with respect to ATMS). It is interesting that the negative impact on ATMS channel 15 (peaking at around 42 km altitude) without bias correction becomes neutral with a bias correction. This perhaps relates to the propagation of gravity waves from the troposphere to the stratosphere and how biased wind data is negatively affecting that.

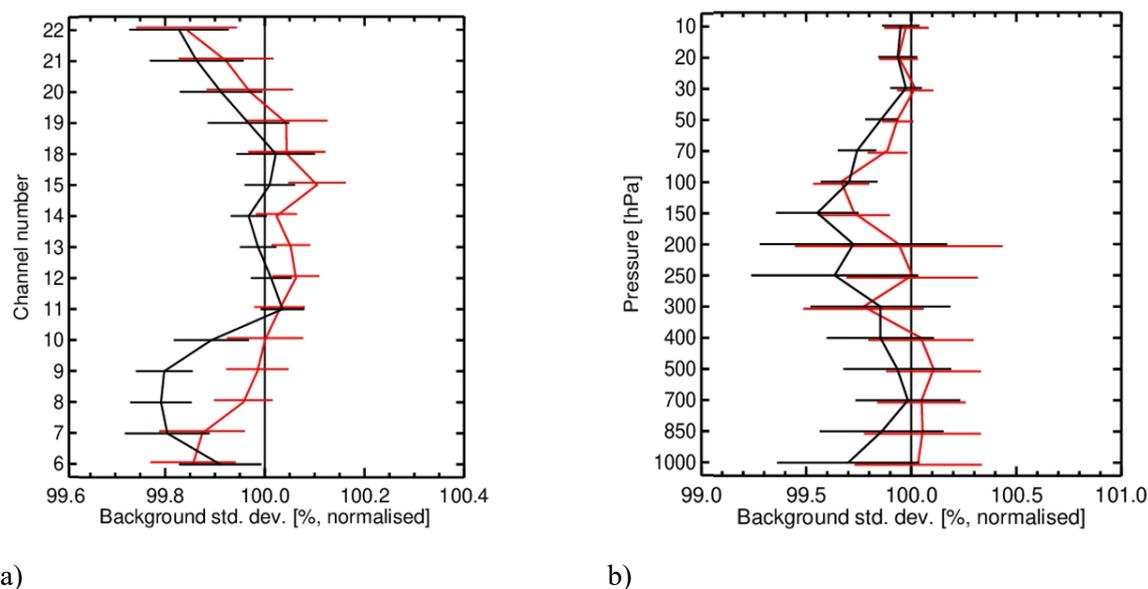


Figure 64. Change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS winds with the bias correction scheme (black line) and without the bias correction (red line) normalised so that the control (without Aeolus) is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) global ATMS radiances b) global conventional vector winds. Horizontal bars show the 95% confidence range. For the period 2 April 2019 to 14 June 2019.

The ECMWF bias correction look-up table was also found to improve the impact of Aeolus on longer range forecast skill, as can be seen by comparing the impact using bias correction (Figure 65) to that without (Figure 63); for vector wind forecast skill as verified against operational analyses. The bias correction leads to larger patches of statistically significant impact in the tropics (cross-hatched areas), particularly in the upper troposphere at days 2-5. The negative impact at day 5 without bias correction

<sup>28</sup> We later (with FM-B data) noticed that the Rayleigh biases can also vary significantly with longitude in certain latitude bands. The bias correction scheme does not vary with altitude.

(although not statistically significant) is reduced to be more neutral with the bias correction. This improvement in NH skill is likely because of the correction of large changes with space and time in Rayleigh-clear bias that occurred in the NH during the period of the experiment (see section 3.2). The positive impact of bias correction is demonstrated more quantitatively in Figure 66; since the black lines (with bias correction) are more negative (i.e. positive impact) than the red lines (without bias correction). Again, the largest benefit of the bias correction is in the tropical upper troposphere.

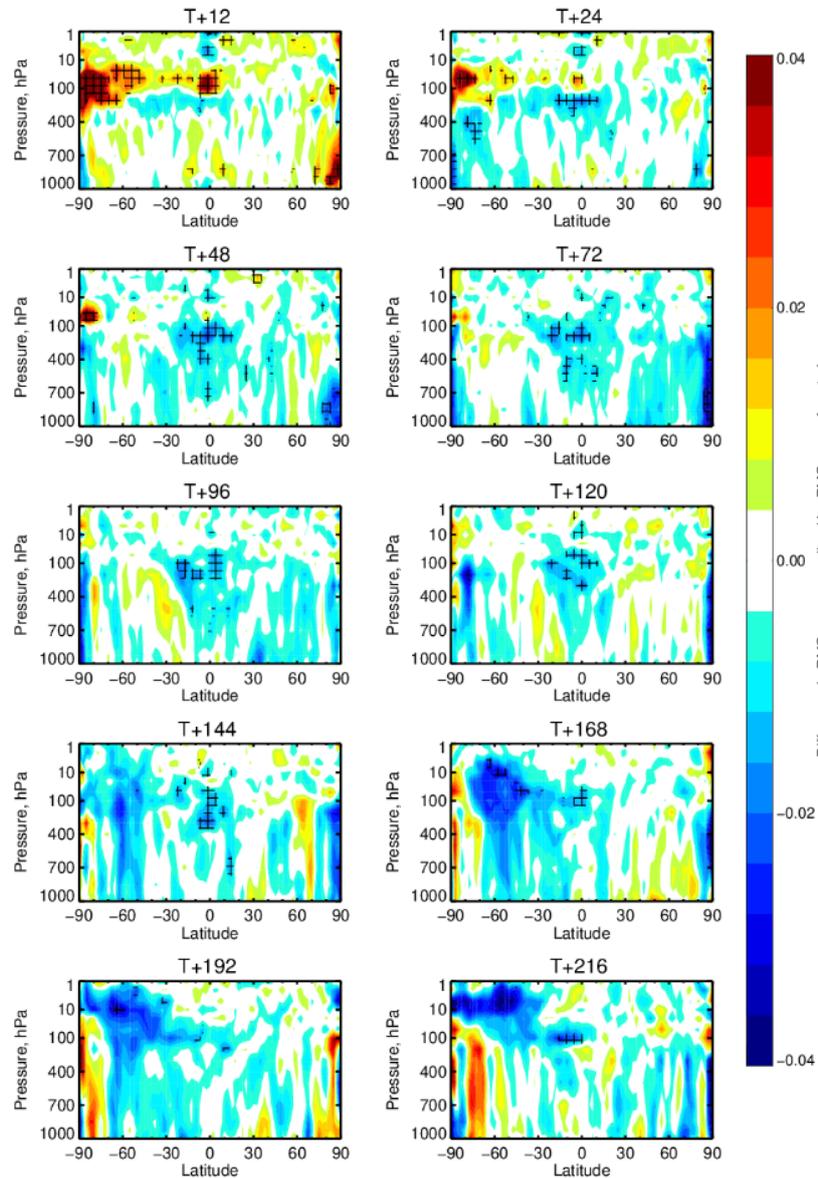


Figure 65. Zonal average normalised change in the RMS (root mean square) error of wind vector for the Rayleigh-clear plus Mie-cloudy experiment for the period 2 April 2019 to 14 June 2019 with the ECMWF bias correction scheme applied (dependence of argument of latitude only). Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicates 95% confidence.

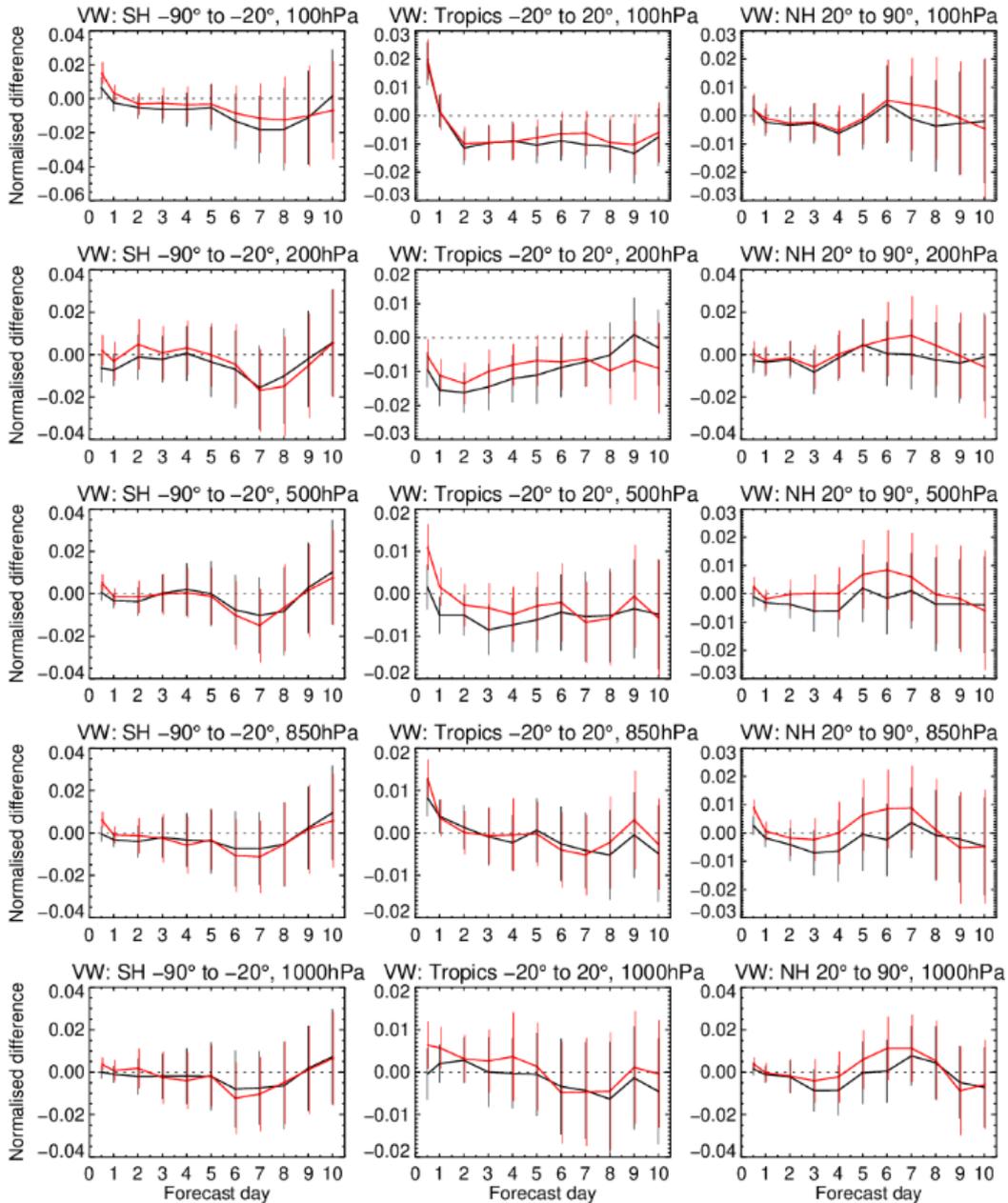


Figure 66. Normalised change in the RMS (root mean square) error in wind vector for different vertical levels (from 100 to 1000 hPa, top to bottom) and for SH, tropics and NH (left to right). The black line using the bias correction scheme and the red line is without the bias correction scheme. The period is from 2 April 2019 to 14 June 2019. Confidence ranges are 95%. Negative values indicate a reduction in error from assimilating Aeolus relative to a control without Aeolus. Verified against operational analyses.

Despite having to reject a quarter of the range-bins (due to blacklisting of hot pixel affected range-bins) and given a low laser energy (and hence low atmospheric path signal), the late FM-A period data was, when bias corrected, able to provide some useful improvement in forecast skill. Also, the short-range forecasts were statistically significantly improved throughout the troposphere when verified against other observation types. The tropical wind forecasts were statistically significantly improved by 1.5% at day two dropping to 1% by day six at 200 hPa when verified against operational analyses.

However, the test period was not long enough, or the impact was not large enough, for the promising improvements in the SH and NH to have achieved statistical significance.

### 4.3 2019 NRT FM-B dataset OSEs

This section provides the results of the Aeolus assimilation experiments performed for the early FM-B period using the NRT (operational PDGS) dataset as described in sections 2.1.3 and 2.4.3.

#### 4.3.1 Bias correction scheme

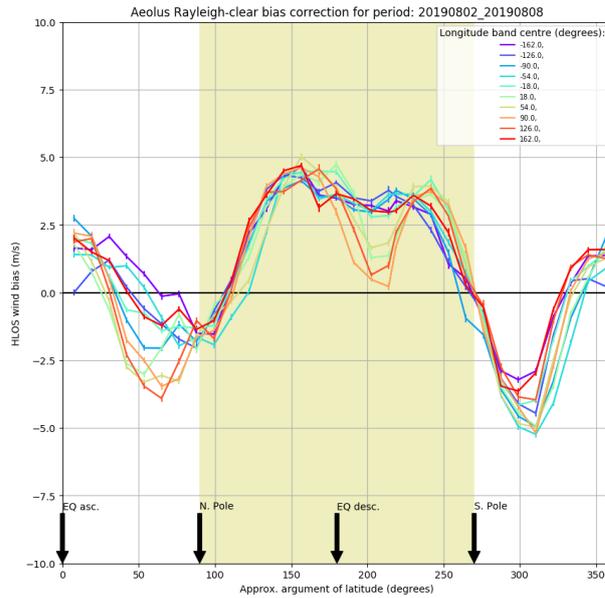
Following from the positive impact of ECMWF's data assimilation Aeolus bias correction scheme when applied to the 2019 FM-A period (see section 4.2.2) it was decided to continue testing it for the early FM-B period. Initially the bias correction scheme continued to be applied as a function of argument of latitude only, however as evidence mounted from FM-B monitoring results (see section 3.2), a longitudinal dependence to the bias correction was also introduced<sup>29</sup>.

Figure 67 (Rayleigh-clear) and Figure 68 (Mie-cloudy) show examples of how the bias varied with argument of latitude (and geodetic longitude for Rayleigh only) for two periods: one week in August 2019 and one week in December 2019. The Rayleigh bias was particularly large in August 2019 with a range (from negative to positive) of 9 m/s; in mid-December 2019 the bias range was smaller at 6 m/s. The varying biases in space and time result from varying top of atmosphere radiation and telescope thermal control laws affecting the M1 mirror temperature gradients (as explained in section 3.2). Despite the range of bias being smaller in December 2019, there is still quite a large longitudinal variation in the bias. The longitudinal bias variations were large in the N. Hemisphere in August 2019, as evident in the figures of section 3.2.

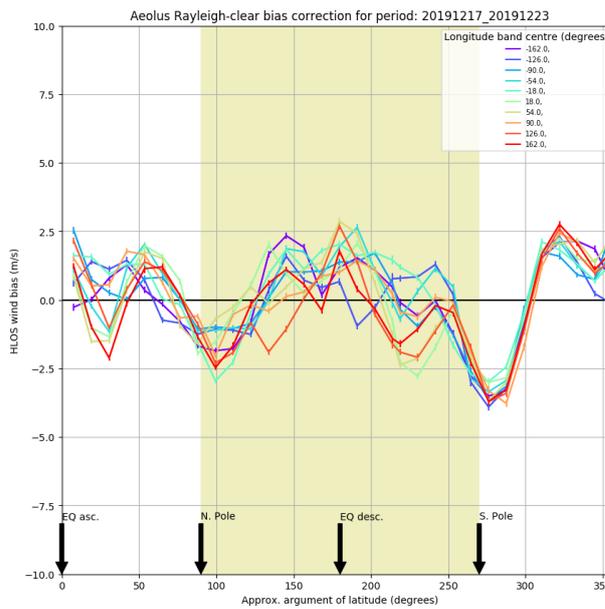
The Mie wind biases are much smaller than the Rayleigh and did not show a clear longitudinal variation, but the structure of Mie bias with argument of latitude has changed with the seasons. Therefore, we corrected the Mie bias using the mean(O-B) as a function of argument of latitude only; see the examples in Figure 68.

---

<sup>29</sup> Note that this was investigated before the ground processing M1 temperature-based bias correction was possible, because the M1 temperatures were then not available in Level-1 nor Level-2 products (nor were the algorithms).

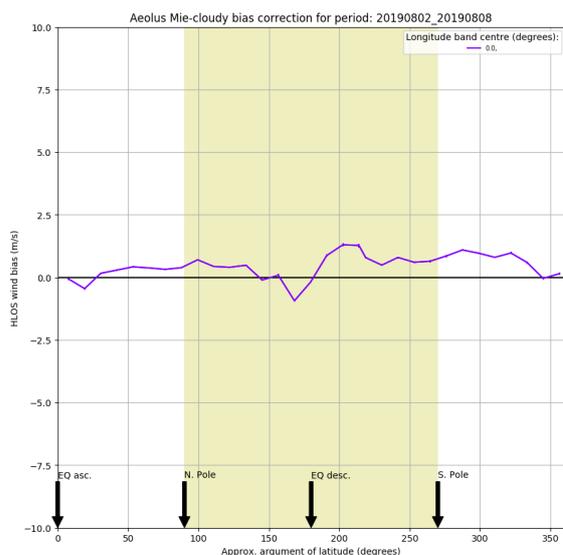


a)

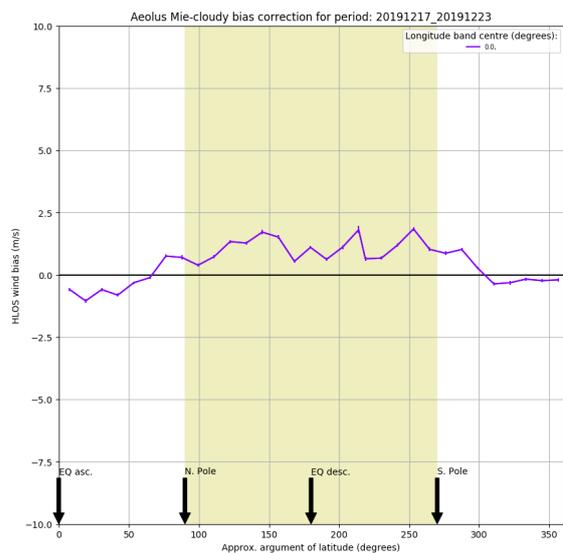


b)

Figure 67. The bias corrections determined for the Rayleigh-clear HLOS winds for the argument of latitude and longitude dependent bias correction; a) for the week from 2-8 August 2019 and b) for the week from 17-23 December 2019. Different longitude bands are shown by different coloured lines (as labelled).



a)



b)

Figure 68. The bias corrections determined for the Mie-cloudy HLOS winds for the argument of latitude dependent bias correction; a) for the week from 2-8 August 2019 and b) for the week from 17-23 December 2019. A longitudinal dependence to the bias correction was deemed unnecessary for the Mie-cloudy winds.

#### 4.3.2 Analysis changes from assimilating Aeolus HLOS winds

Figure 69 shows a map of the standard deviation of the differences of 250 hPa zonal wind analyses with and without Aeolus assimilation (with both Rayleigh-clear and Mie-cloudy). The largest magnitude differences occur in convectively active areas: for example, along the Inter Tropical Convergence Zone (ITCZ), the South Pacific Convergence Zone (SPCZ), the Indian Ocean and to a lesser extent in the South Atlantic Convergence Zone (SACZ). The pattern has similarities with the results for the 2019 FM-A period in Figure 60. Perhaps this is because of the short-range forecasts larger wind errors in convectively active areas, which was demonstrated to be the case in the well-observed (aircraft,

radiosondes) meso-scale convection systems in the USA (Rodwell et al. (2013)). At 700 hPa, as shown in Figure 70, the analysis differences also reach polar regions, and the differences are again larger in convergence zones. In the upper troposphere, the largest magnitude changes to u-wind occur at 150 hPa in the East Pacific ITCZ just north of the equator, with standard deviation reaching 4 m/s, see Figure 71. This area consistently has the largest background zonal and meridional wind errors via EDA spread. Interestingly the wind field is changed a lot at very high altitudes (1 hPa, ~48 km) far above the direct influence of Aeolus. This may be due to changes in gravity wave propagation due to Aeolus' influence on mean winds in the troposphere, see Figure 72. The patterns in v-wind component have similar geographical patterns and magnitude, and hence are not shown.

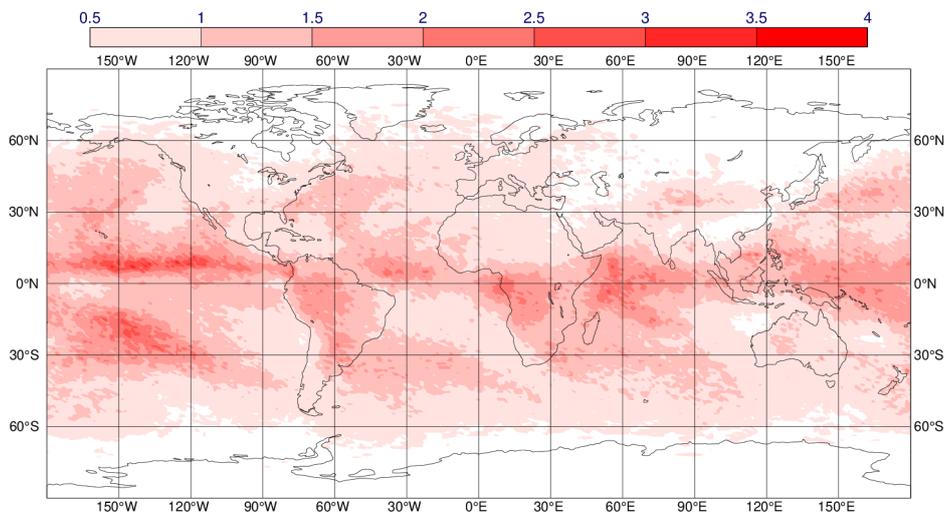


Figure 69. Standard deviation of the differences of u-wind component (m/s) at 250 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 2 August to 31 December 2019.

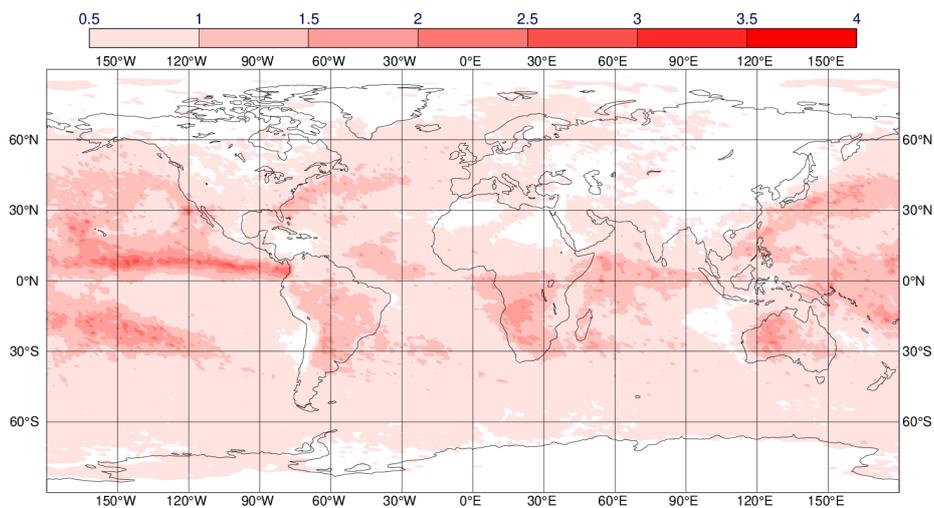


Figure 70. Standard deviation of the differences of u-wind component (m/s) at 700 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 2 August to 31 December 2019.

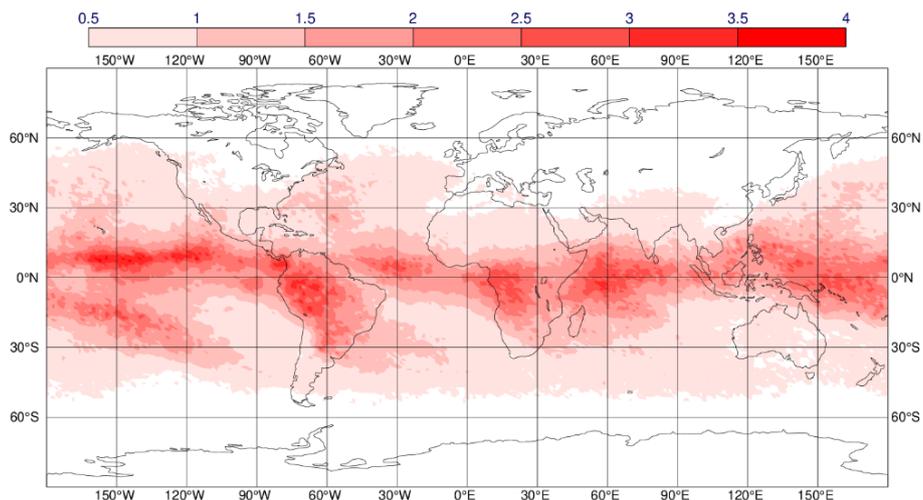


Figure 71. Standard deviation of the differences of u-wind component (m/s) at 150 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 2 August to 27 December 2019.

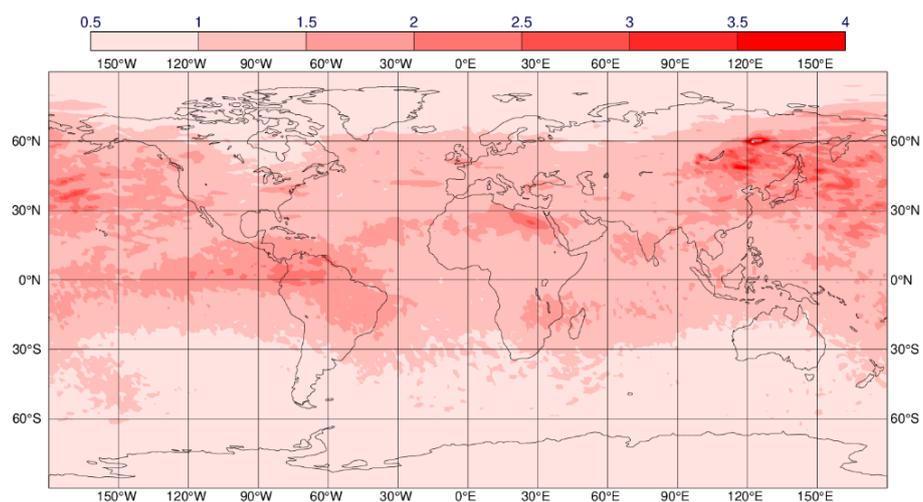


Figure 72. Standard deviation of the differences of u-wind component (m/s) at 1 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 2 August to 27 December 2019.

Figure 73 shows the zonal average mean changes in the wind field after assimilating Aeolus. There is a complicated structure and larger magnitude to the mean changes in the tropics and an apparent slowing down of the westerly flow in the Southern Hemisphere at -60 to -70 degrees latitude. The tropical changes are largest at 100 hPa, whereas for the extratropics they are largest at 400 hPa. We cannot tell from these plots if these changes are bringing the mean analysis state closer to the truth. The zonal average magnitude of 0.1 m/s is not particularly large, however.

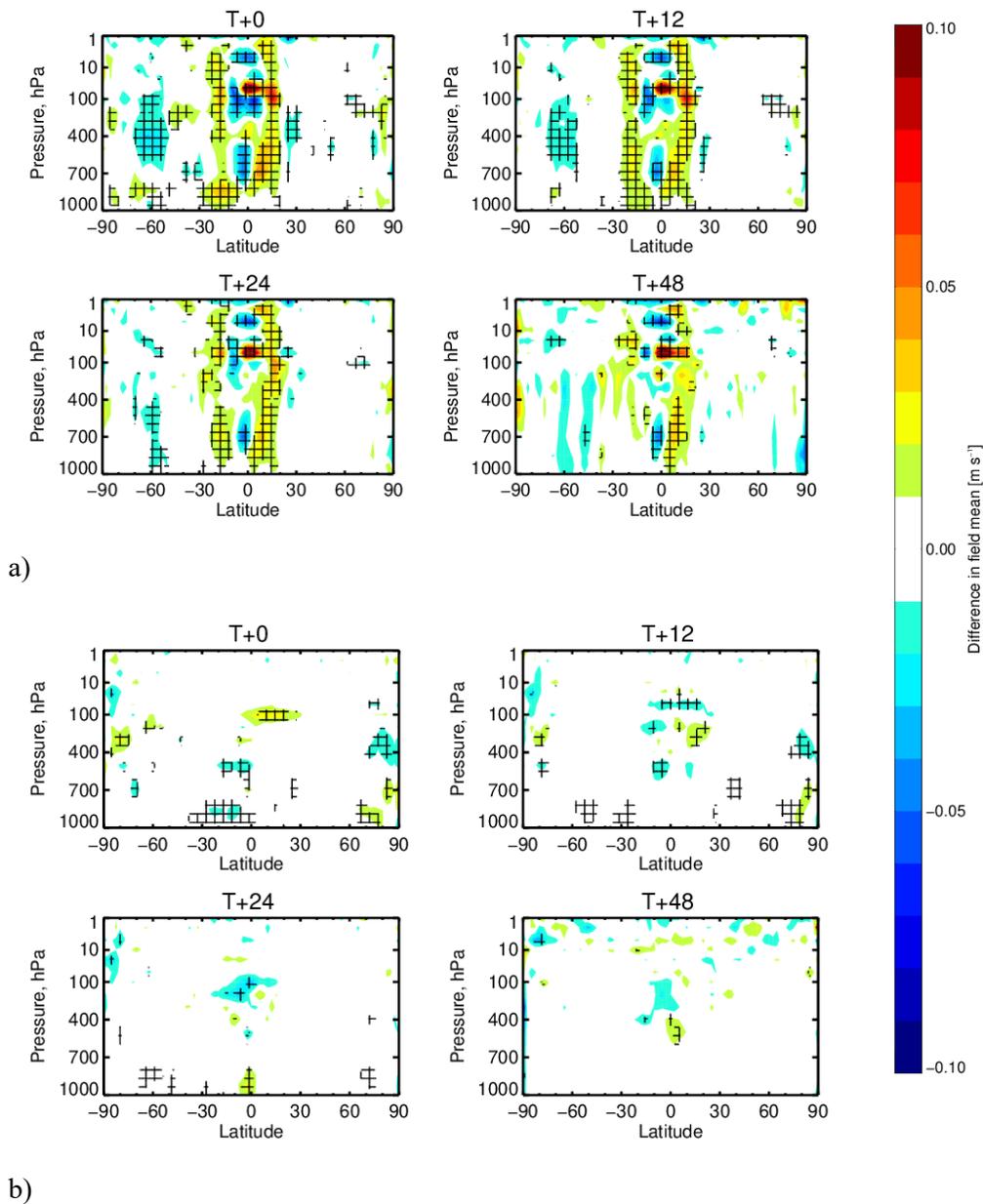


Figure 73. The zonal mean change in the analysis (T+0 hours) and the forecasts (T+12 to T+48 hours) a) u-wind and b) v-wind fields because of assimilating Aeolus (Rayleigh-clear and Mie-cloudy). For the period from 12/8/19 to 31/12/19.

Figure 74 shows that the mean change in the tropical zonal wind at 100 hPa occurs in complicated longitudinal bands with opposite signs. The changes tend to persist throughout the forecast range. There is also an interesting pattern of mean changes with alternating sign over Antarctica, which is evident in the v-wind component also, which is not understood.

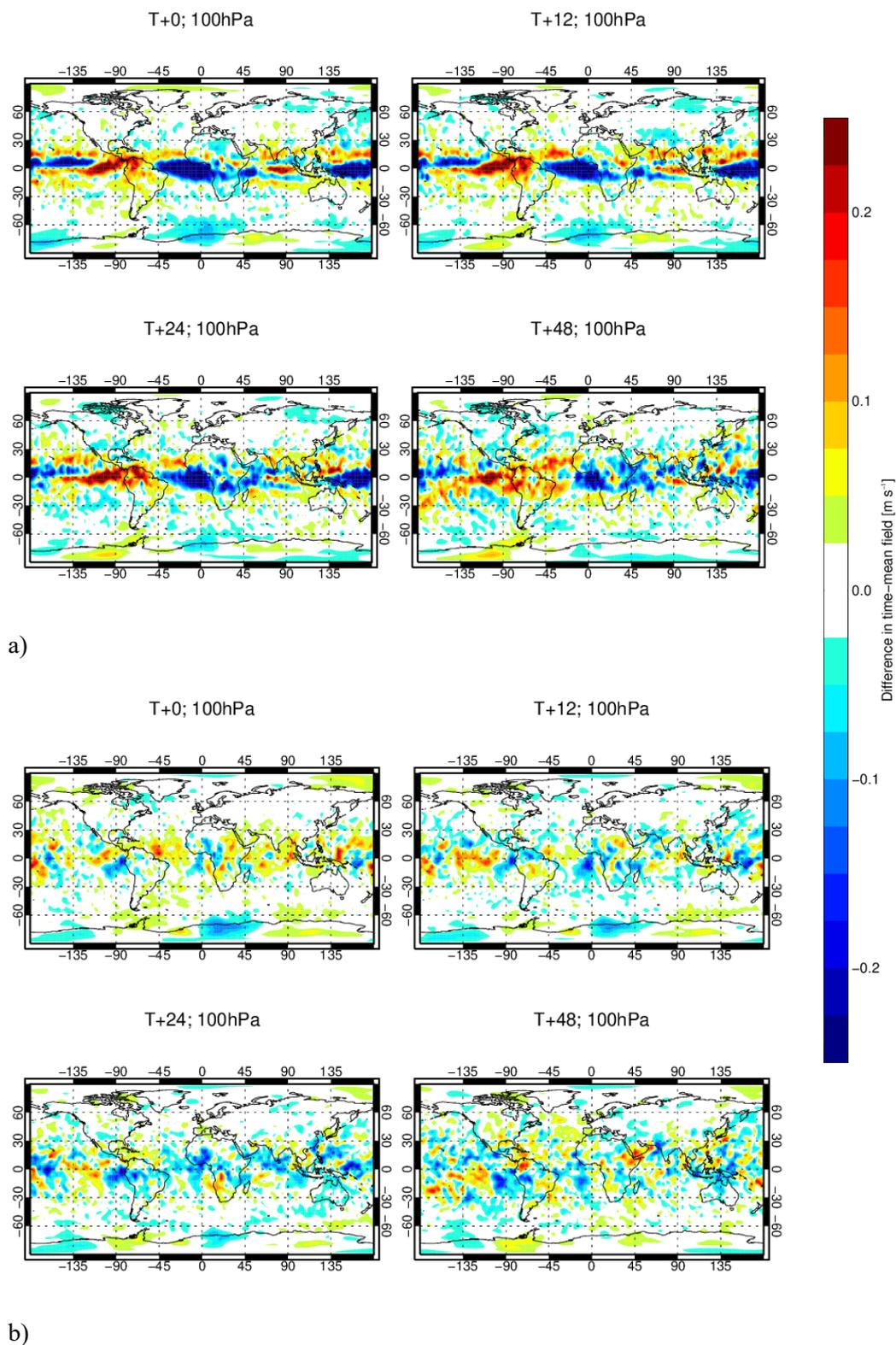
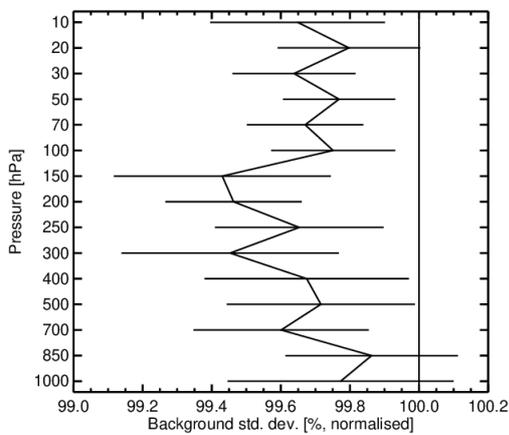


Figure 74. Maps of the mean change in the analysis ( $T+0$  hours) and the forecasts ( $T+12$  to  $T+48$  hours) at 100 hPa for a)  $u$ -wind and b)  $v$ -wind resulting from assimilating Aeolus. For the period from 12 August 2019 to 31 December 2019.

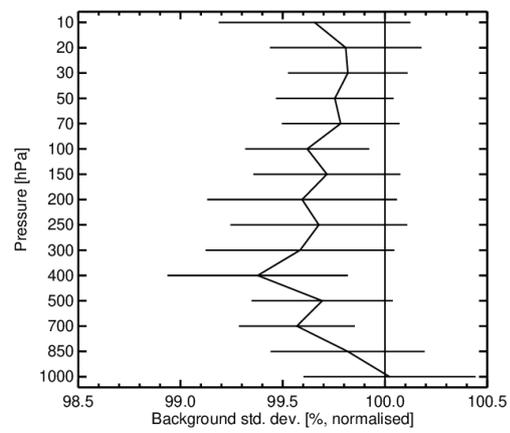
### 4.3.3 Short-range forecast impact

#### 4.3.3.1 Rayleigh-clear and Mie-cloudy

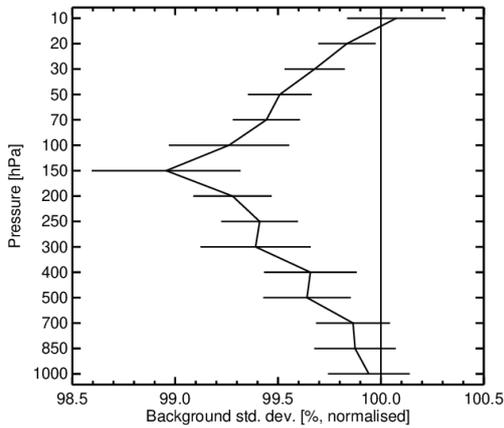
The impact of Rayleigh-clear and Mie-cloudy data on the short-range forecasts (12 hours) as verified by the fit to other observation types is shown in Figure 75 and Figure 76. Figure 75 shows in situ (known as *conventional* at ECMWF) wind and radiosonde temperature observations. Aeolus improves the short-range forecast fit to vector wind and temperatures typically by less than 1%. The largest impact is in the tropics, peaking at 150 hPa (around 1%), but positive impact is present throughout the troposphere and lower stratosphere. The impact throughout the troposphere and lower stratosphere in the SH does not vary much with pressure, with a peak impact of 0.5%. NH impact is positive above 150 hPa for wind, but otherwise neutral.



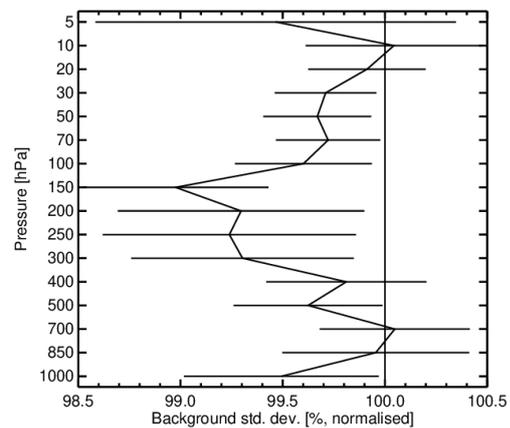
a)



b)



c)



d)

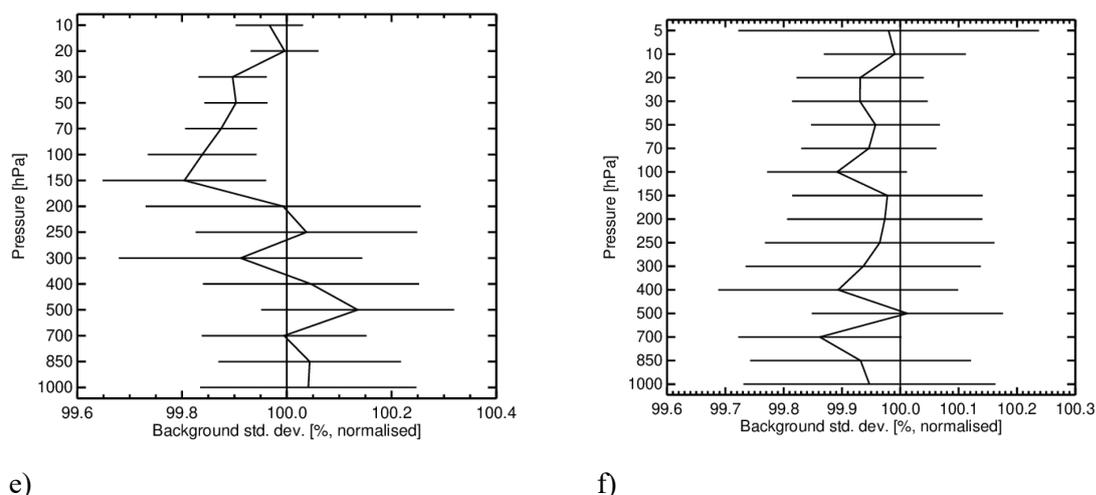
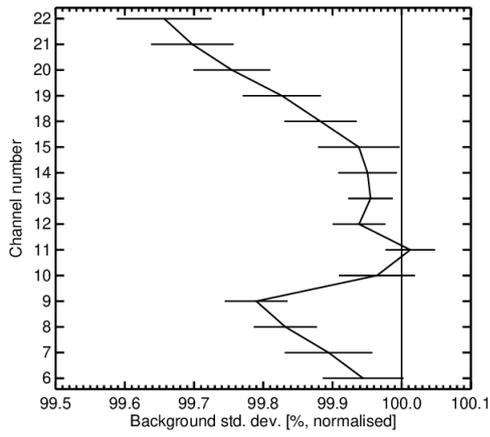
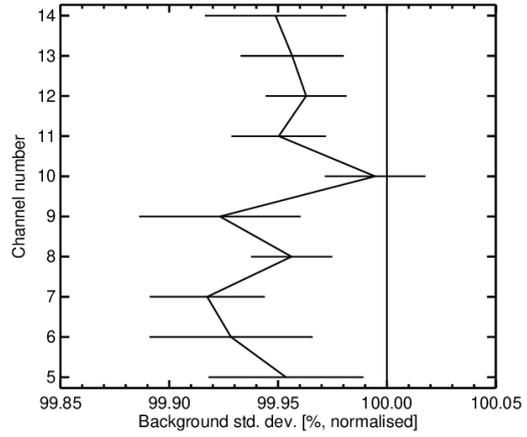


Figure 75. Change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations (Rayleigh-clear and Mie-cloudy), normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) conventional vector wind in the SH extratropics, b) radiosonde temperatures in the SH extratropics, c) conventional vector wind in the tropics, d) radiosonde temperatures in the tropics, e) conventional data vector wind in the NH extratropics and f) radiosonde temperatures in the NH extratropics. Horizontal bars show the 95% confidence range (Student's *t*-test). For the period from 12 August 2019 to 31 December 2019.

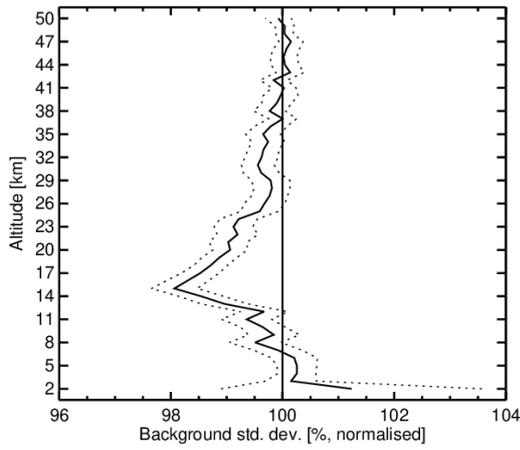
Figure 76 shows the short-range forecast fit to other satellite observation types. Aeolus improves the fit to global microwave sounding temperature sensitive channels (for ATMS and AMSU-A) and to ATMS humidity sensitive channels (18-22). GNSS radio occultation fits are improved by 2% at ~14-15 km in the tropics (upper troposphere), but the positive impact is seen throughout much of the troposphere and lower stratosphere; this is similar to the equivalent result for the late FM-A data (see Figure 62a). Relative to the infrared sounding instrument CrIS, Aeolus shows improvements in temperature sensitive channels in the upper troposphere and lower stratosphere and in channels sensitive to the surface temperature and humidity. The change relative to Atmospheric Motion Vectors (AMVs) is positive at 200-250 hPa and near the surface; but there is a small negative impact at 700 hPa (perhaps associated with the top of boundary layer clouds). The fit to geostationary water vapour imagery is improved, as is the fit to the FY-3B microwave humidity sounder. Finally, in the NH, there are improvements relative to scatterometer winds (ASCAT) and wave height altimeter data (RALT).



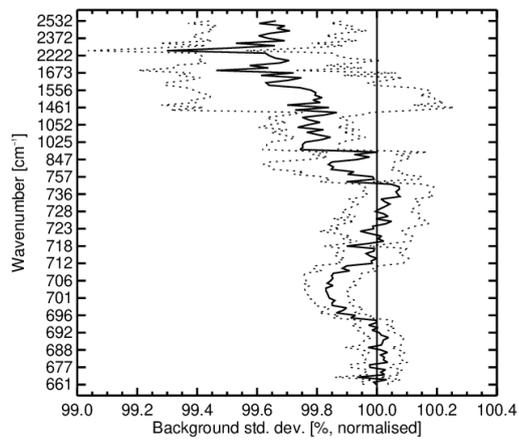
a)



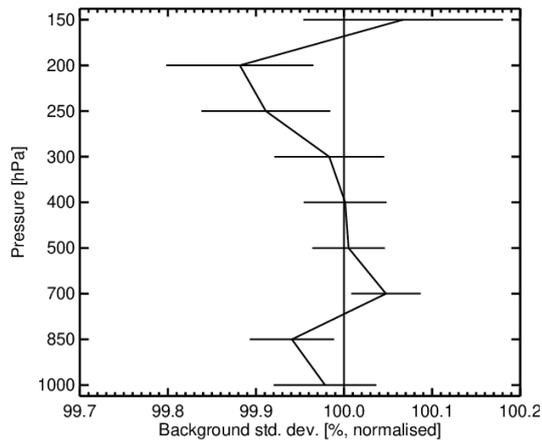
b)



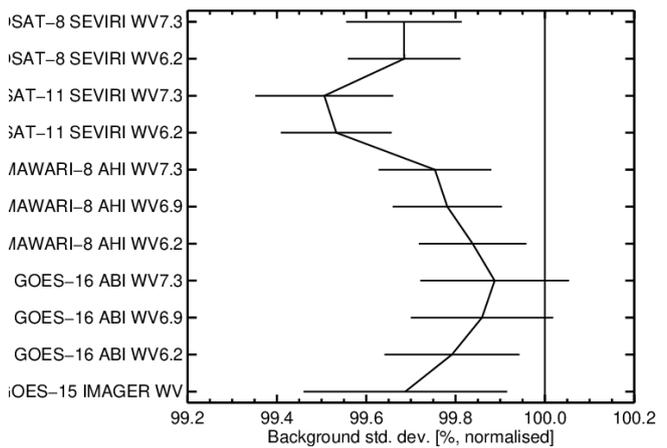
c)



d)



e)



f)

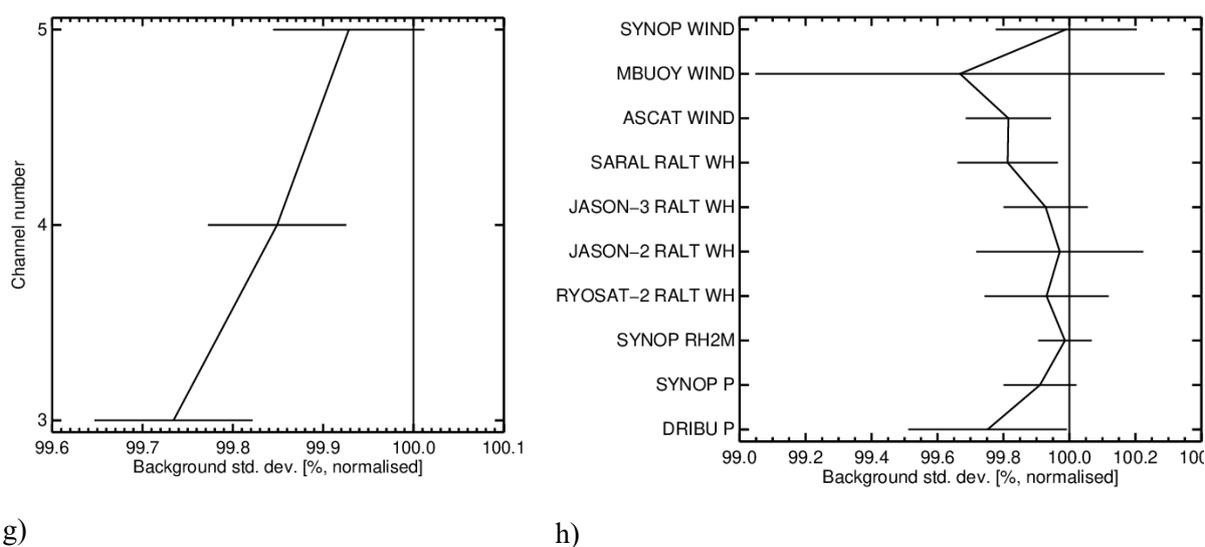


Figure 76. Change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations (Rayleigh-clear and Mie-cloudy), normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) ATMS global statistics, b) AMSU-A global statistics, c) GNSS radio occultation in the tropics, d) CrIS global statistics, e) AMVs global statistics, f) geostationary imagery global statistics, g) FY-3B MWHS global statistics and h) various surface observation types in the NH. Horizontal bars show the 95% confidence range (Student's t-test). For period from 12 August 2019 to 31 December 2019.

Aeolus early FM-B NRT data improved the short-range forecast wind, temperature and humidity. The magnitude of the impact is consistently of a good magnitude for one satellite, when compared to other satellite data in ECMWF OSEs e.g. Bormann et al. 2019.

Assimilating Aeolus leads to small changes in the mean fit to tropical conventional wind data (mostly radiosondes sampling the Maritime Continent) as shown in Figure 77. These changes mostly bring the forecast closer to radiosondes, which is a good indication that the mean analysis changes from Aeolus are moving it closer to the truth (as shown in Figure 73 and Figure 74); at least at geolocations near radiosondes. This is reassuring given that the bias correction scheme relied on the model which has biases in the tropical upper troposphere<sup>30</sup>, but perhaps these biases average out in a global sense.

<sup>30</sup> In previous years we have seen biases between the ECMWF and Met Office model wind fields in tropical upper troposphere.

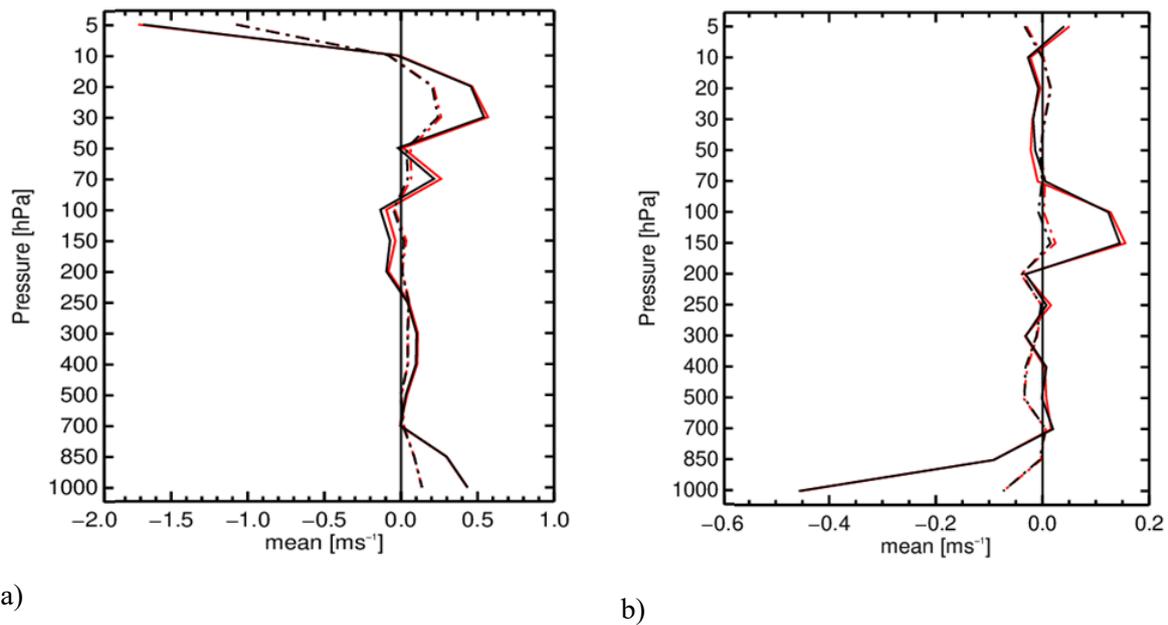


Figure 77. The mean(O-B) fit to conventional wind data (mostly aircraft and radiosondes) in the tropics for a) u-wind component and b) v-wind component. The red line is the control (without Aeolus) and the black line is with Aeolus assimilation. For the period from 2 August 2019 to 31 December 2019.

#### 4.3.3.2 Mie-cloudy only

The short-range forecast impact relative to ATMS microwave radiances from assimilating only Mie-cloudy winds in comparison to using both Rayleigh-clear and Mie-cloudy is shown in Figure 78. The black line is when assimilating only Mie-cloudy and the red line is when assimilating both Rayleigh-clear and Mie-cloudy. The impact of Mie-cloudy on its own is mixed relative to ATMS microwave radiances (and AMSU-A, not shown). The addition of Rayleigh-clear winds changes the negative impact of the Mie winds to a positive impact. The Mie winds are degrading the fit to the stratospheric temperature sensitive channels (channels 12-15, with weighting functions peaking around 25-42 km) and for the humidity sensitive channel near the surface (channels 18). However, the impact of the Mie-cloudy on tropospheric temperature i.e. channels 6 to 9 (below 10-12 km) is positive. Further inspection showed that the stratospheric temperature degradation is strongest in the SH. With Mie winds in the SH typically only reaching 16 km altitude this must be an indirect effect, which could be due to gravity wave propagation from the troposphere to the stratosphere being affected by Mie winds altering the flow at low levels. The high peaking microwave radiances are very sensitive to the temperature variability associated with the large amplitude orographically driven gravity waves.

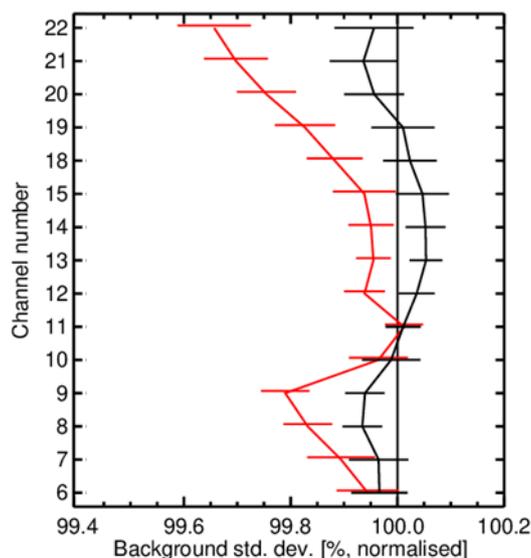


Figure 78. Change in standard deviation of O-B departures relative to global ATMS microwave radiances resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. The red line is with both Rayleigh-clear and Mie-cloudy used and the black line is with only Mie-cloudy. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. For the period 2 August 2019 to 31 December 2019.

#### 4.3.3.3 Rayleigh-cloudy only

Despite the L2B Rayleigh-cloudy winds naturally having the same spatial coverage as the Mie-cloudy winds but being significantly noisier and having larger biases (at least for the early FM-B NRT dataset), it is interesting to see the impact of assimilating Rayleigh-cloudy winds on their own. There are around 1800 Rayleigh-cloudy observations per 12 hours assimilated (strict QC); which is relatively few compared to Rayleigh-clear (~70000) and Mie-cloudy (~20000) (Mie-cloudy winds are produced at higher horizontal resolution).

An OSE using Rayleigh-cloudy winds only showed mostly neutral impact on short-range forecast O-B departures relative to other observation types. The only positive impact is seen for the humidity and cloud sensitive radiances in the tropics e.g. Figure 79.

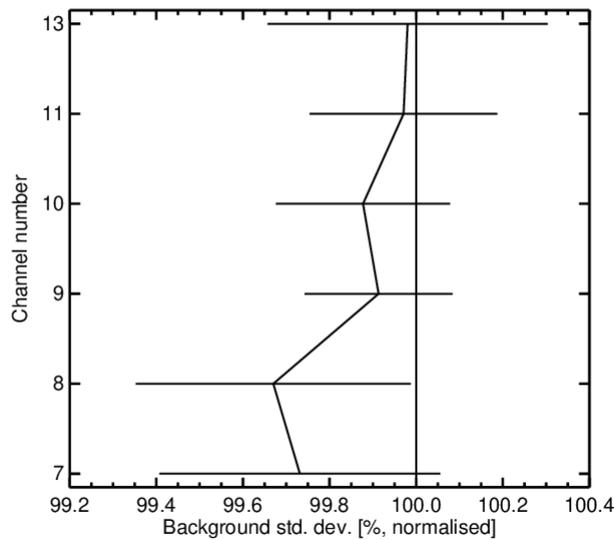


Figure 79. Change in standard deviation of O-B departures relative to tropical AMSR-2 (Advanced Microwave Scanning Radiometer 2) all-sky microwave radiances resulting from assimilating Aeolus Rayleigh-cloudy HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. For the period 2 August 2019 to 31 December 2019.

There are now methods available to improve the biases for Rayleigh-cloudy winds by applying improved Rayleigh response calibration curves for high scattering ratios scenes (see work by Marseille et al. 2023) with the aid of NWP based calibration procedures.

#### 4.3.4 Medium-range forecast impact

##### 4.3.4.1 Rayleigh-clear and Mie-cloudy

Figure 80 and Figure 82 show the combined impact of Aeolus Rayleigh-clear and Mie-cloudy winds for the period from 2 August 2019 until 31 December 2019 on vector wind, verified against ECMWF operational analyses (for which Aeolus was not assimilated). The impact on 2 to 4-day forecasts is mostly positive with statistical significance in the tropical troposphere. The percentage improvements are on the order of 1-2% in the tropics. In the extratropics the positive impact up to the lower stratosphere is less than 1% for the whole area, particularly poleward of 60 degrees. The positive impact has similar patterns to that shown for the 2019 FM-A period (Figure 65) but has a larger magnitude at two to four days. This larger magnitude impact is likely to be due to the smaller Rayleigh-clear winds random error and greater wind count for the early FM-B compared to late FM-A period; or perhaps a seasonal effect.

The apparent negative impact at 12-hour forecasts around 100 hPa and in the lower troposphere and at 24 hours around 100 hPa is thought to be an artefact of the extra variability that Aeolus adds to wind forecasts in otherwise poorly observed areas (for winds), when verifying against the operational analysis which did not assimilate Aeolus. This is corroborated by the short-range forecast fit to mass observations which is improved in such areas. There may also however be an element of overweighting or improper use of the data causing some part of this e.g. in the lower stratosphere.

There is negative impact in the SH for five-day forecasts and beyond, particularly in the stratosphere e.g. 10 hPa (~30 km). It is unclear what is causing this, but it may be a manifestation of the possible Mie wind induced gravity wave issue as discussed in section 4.3.3.2. Another possible explanation is that there are large gradients in the Rayleigh-clear wind bias for this region, which were

perhaps not corrected well enough via the bias correction look-up table method. It will be interesting to see if the M1 temperature dependent bias correction method can improve upon this (using reprocessed datasets, see section 4.5). The 2019 FM-A period showed hints of positive impact in the same region (Figure 65).

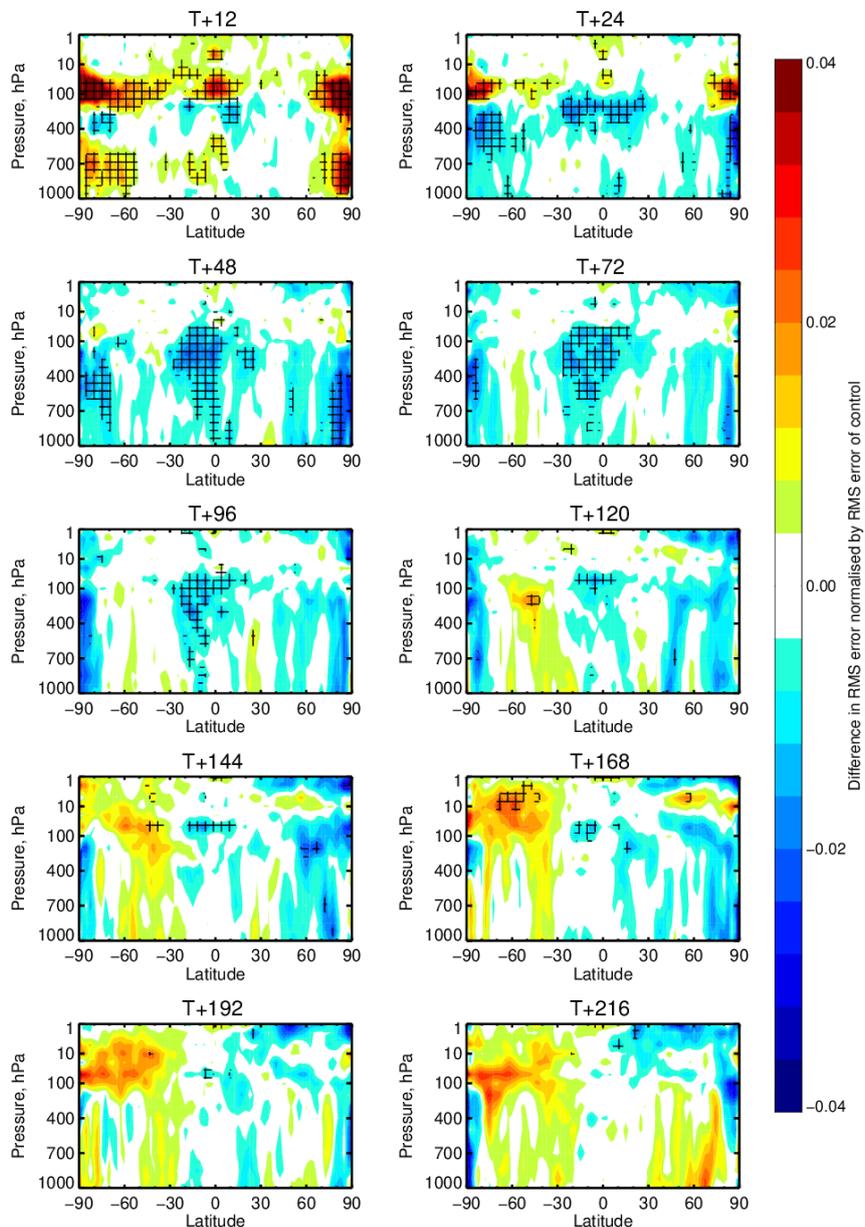


Figure 80. Zonal average normalised change in the RMS (root mean square) error of wind vector from assimilating Aeolus (Rayleigh-clear and Mie-cloudy (nominal assigned observation error) for the period 2 August 2019 to 31 December 2019. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence.

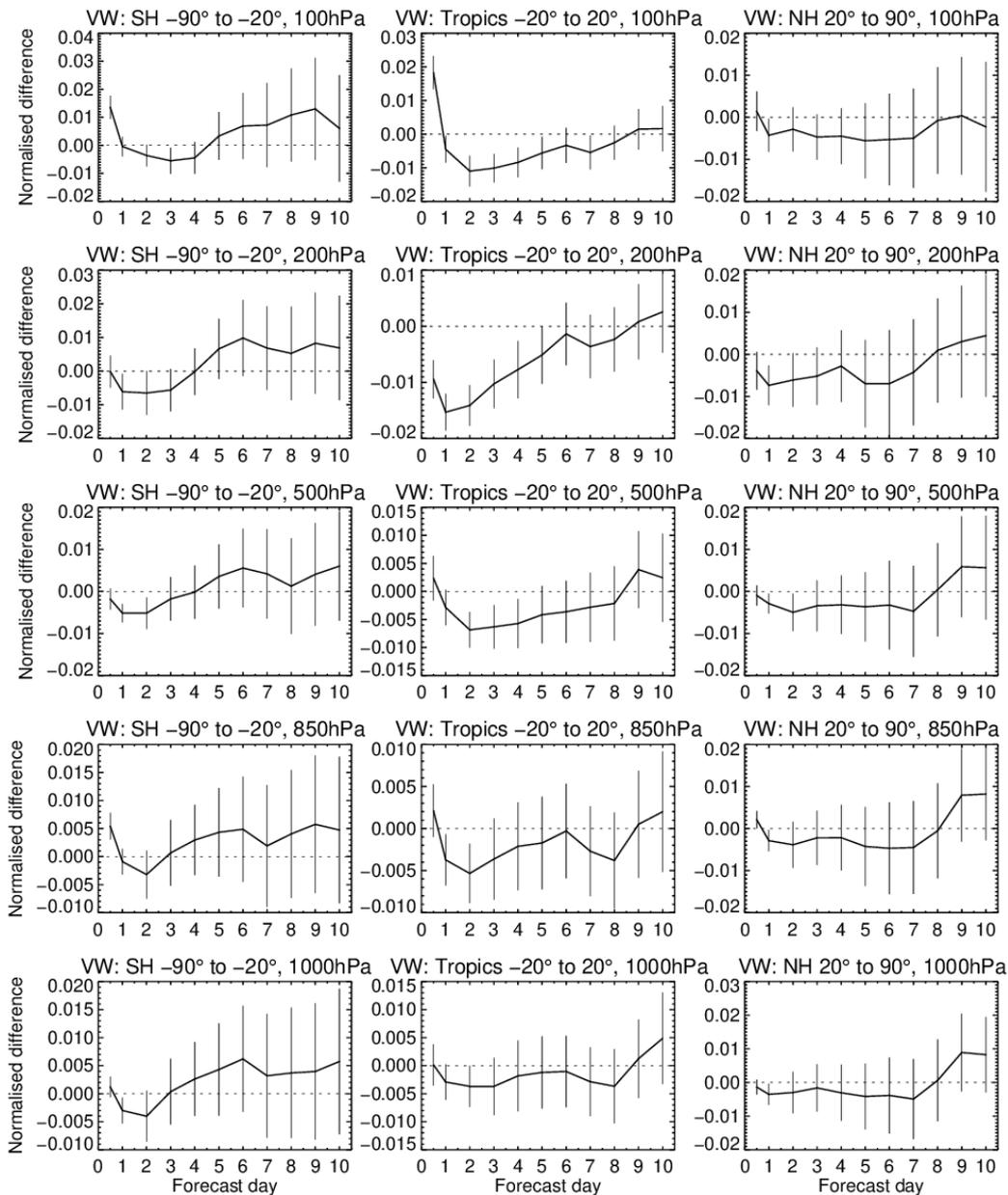


Figure 81. Normalised change in the RMS (root mean square) error in vector wind for different vertical levels (from 100 to 1000 hPa, top to bottom) and for SH, tropics and NH (left to right) from assimilating Aeolus (Rayleigh-clear and Mie-cloudy). The period is from 2 August 2019 to 31 December 2019. Confidence ranges are 95%. Verified against operational analyses. Negative values indicate a reduction in error from assimilating Aeolus.

A map of 200 hPa vector wind impact is shown in Figure 82. The positive impact (blue colours) tends to exist in the tropics, over oceans and other areas that a relatively lacking in wind profile observations. The negative impact in the South Pacific seems to grow and propagate westward with time. Other variables such as temperature, relative humidity and geopotential are affected in a similar pattern to the wind impact as shown in Figure 53 to Figure 55.

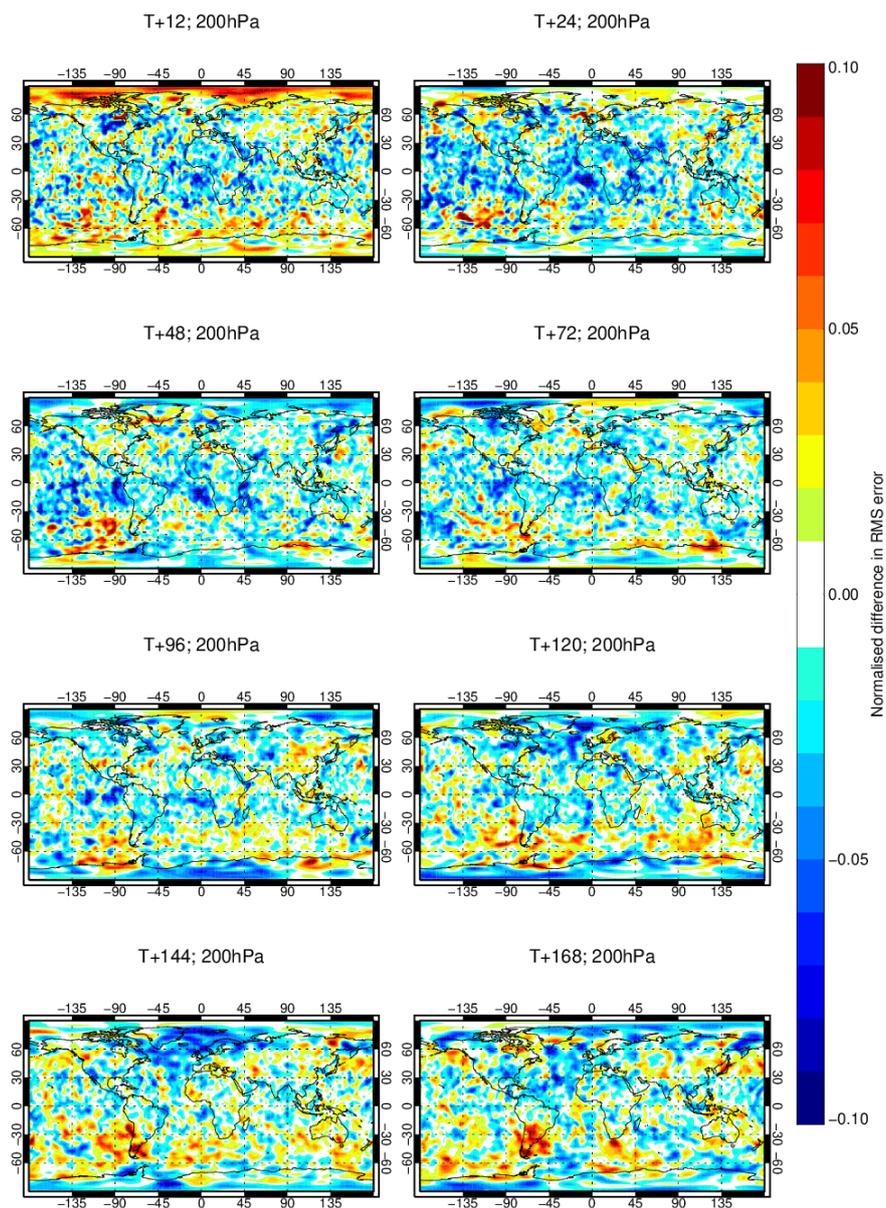


Figure 82. Maps of normalised change in the RMS (root mean square) error of wind vector at 200 hPa from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 2 August 2019 to 31 December 2019. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses.

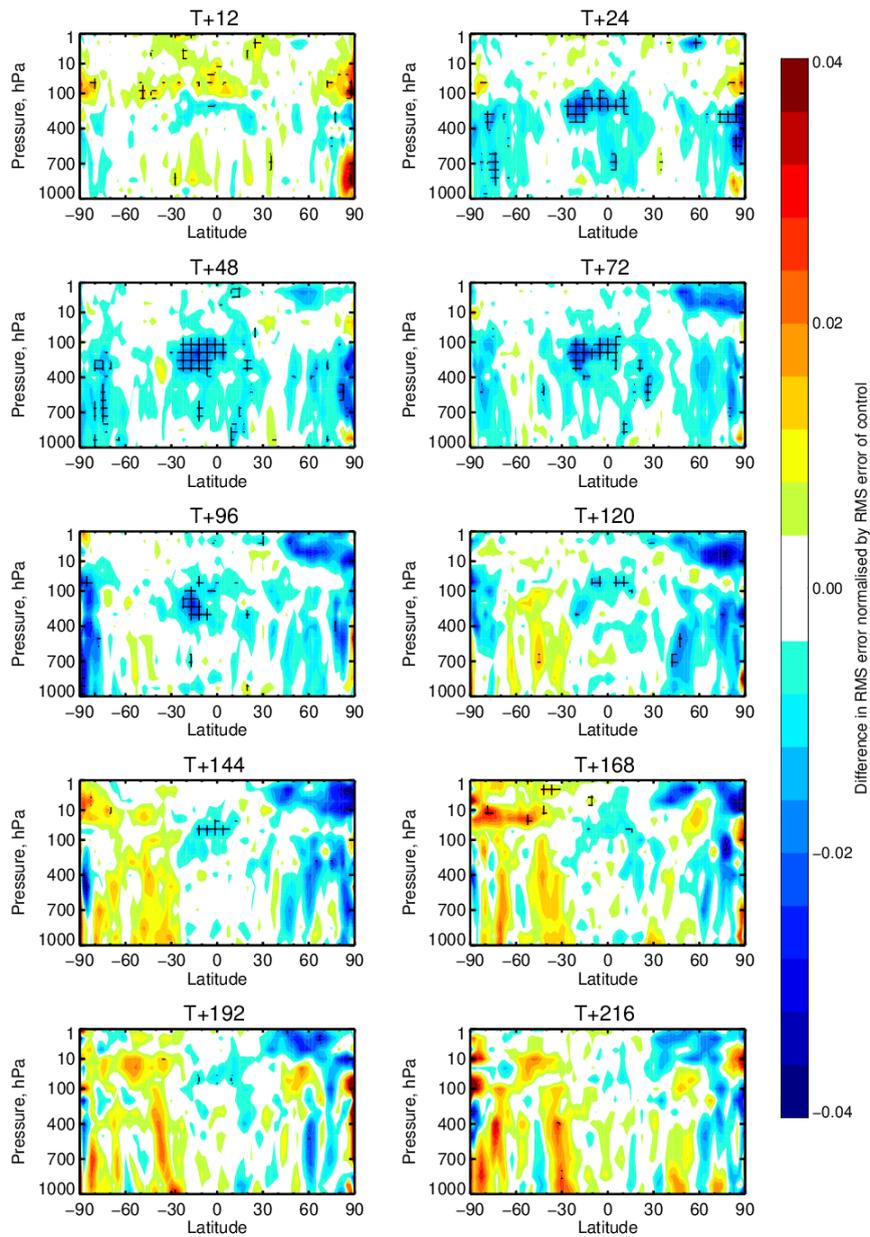


Figure 83. Zonal average normalised change in the RMS (root mean square) error of temperature from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 2 August 2019 to 31 December 2019. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Cross-hatching indicating 95% confidence. Verified against operational analyses.

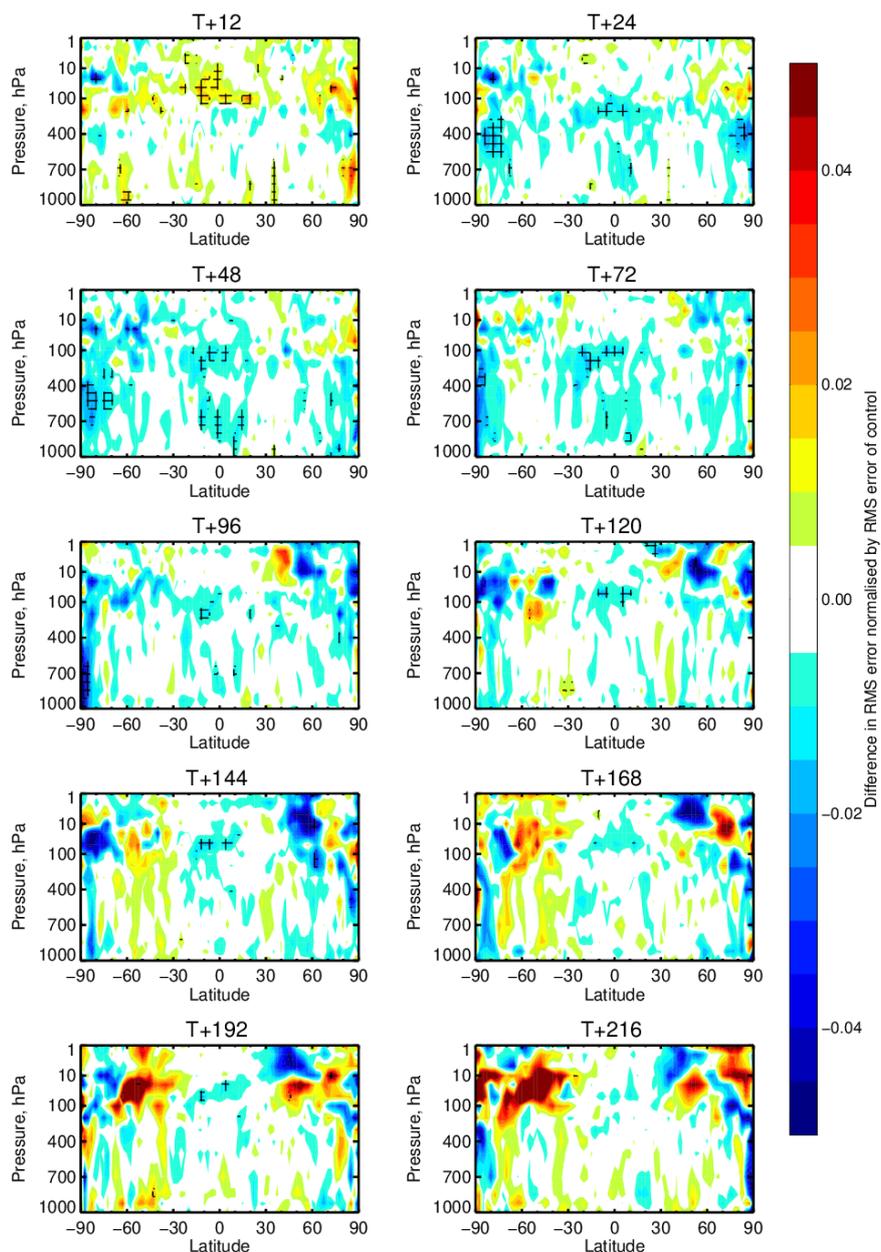


Figure 84. Zonal average normalised change in the RMS (root mean square) error of relative humidity from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 2 August 2019 to 31 December 2019. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Cross-hatching indicating 95% confidence. Verified against operational analyses.

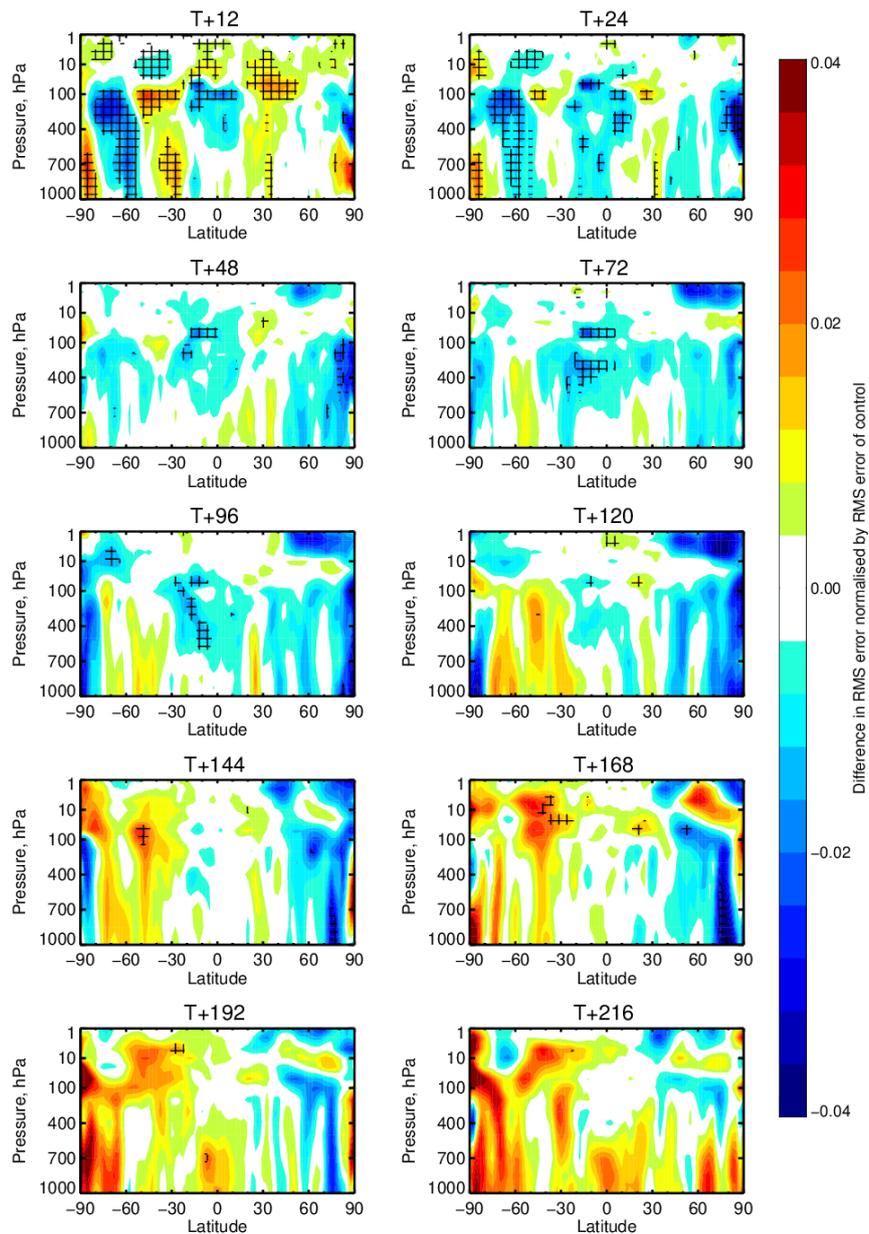


Figure 85. Zonal average normalised change in the RMS (root mean square) error of geopotential height from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 2 August 2019 to 31 December 2019. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Cross-hatching indicating 95% confidence. Verified against operational analyses.

#### 4.3.4.2 Impact of Aeolus in the stratosphere

The typically used Aeolus vertical range-bin settings provide clear-air winds (Rayleigh scattering) reaching up to 20-25 km (50-20 hPa) altitude and therefore sampling the lower stratosphere. Because Aeolus is limited to 24 range-bins, to keep the vertical resolution reasonable in the troposphere, the maximum possible altitude of 30 km is not nominally applied<sup>31</sup>. The Rayleigh winds become increasingly noisy with altitude due to the exponential drop in the molecular backscatter (proportional

<sup>31</sup> However, 30 km maximum altitude range-bins have been tested in special cases e.g. the TONGA RBS applied to capture the Hunga-Tonga volcanic plume in early 2022.

to density). These noise aspects and the limited number of range-bins mean the stratospheric range-bin resolution is typically set to 1-2 km. The highest altitude cloudy-air winds (Mie scattering) are due to scattering from ice clouds at around the tropopause level. An exception is Polar Stratospheric Cloud (PSC) and volcanic/wildfire aerosol plumes for which decent quality Mie-cloudy winds can reach 20-30 km altitude. PSCs are very common over Antarctica in the SH winter and less common over the NH winter pole. The flexible vertical sampling of Aeolus has typically been commanded (on several occasions) to be raised over the winter poles to better sample the dynamics of the Polar Night Jet.

The strongest stratospheric impact from Aeolus (as demonstrated in ECMWF OSEs) occurs in the tropical lower stratosphere; see Figure 86, which shows the impact for vector winds at 50 hPa (~20 km). The positive impact on wind and temperature persists well into the medium range (statistically significant at day 8). Positive impact in the extratropical stratosphere relative to conventional wind observations is also evident particularly above 150 hPa (see Figure 75a and e). The tropical impact is corroborated by short-range forecast verification against GNSS radio occultation (see Figure 76c) — with positive impact reaching as high as 35 km altitude. The impact is also positive relative to ATMS (microwave radiances) channels 13 to 15 which peak between 20-40 km (Figure 76a). The positive impact occurs at significantly higher altitudes than are directly observed by Aeolus winds. The mechanism for this remote impact could be via lower altitude changes to the mean wind field, leading to an improved propagation gravity waves from the troposphere to stratosphere (however we have seen that the Mie-cloudy winds *per se* can degrade upper levels via the same speculated mechanism).

For the same time period, Figure 74 shows the mean change in the zonal wind due to Aeolus at 100 hPa and Figure 71 shows the standard deviation of changes in zonal wind due to Aeolus at 150 hPa (where the effect is largest). The changes are strongest in the ITCZ; with a pattern perhaps suggesting a modification to the Walker Circulation. Large magnitude analysis changes are also present at 1 hPa (Figure 72) which must be associated with gravity wave propagation. Research continues to understand these changes. Some evidence for the importance of the phase of the QBO on Aeolus NWP impact was found in Žagar et al. (2021).

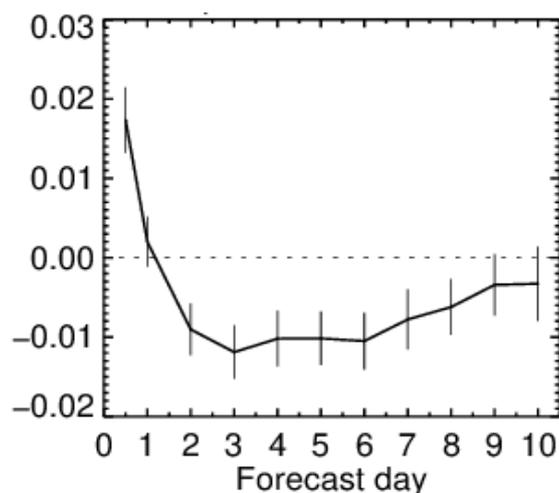


Figure 86. Normalised change in the root mean square error in vector wind at 50 hPa in the tropics from assimilating Aeolus (Rayleigh-clear and Mie-cloudy). The period is from 2 August 2019 to 31 December 2019. Confidence ranges are 95%. Verified against operational analyses. Negative values indicate a reduction in error from assimilating Aeolus.

### 4.3.4.3 Mie-cloudy only

The impact of assimilating Mie-cloudy winds, with the nominal assigned observation error (as described in 2.4.3) on vector wind forecasts is shown in Figure 87; this should be compared to the impact of the combined Rayleigh-clear and Mie-cloudy winds in Figure 80. Mie winds show hints of positive impact for days 2-4 but it lacks statistical significance. Therefore, it was concluded that the Rayleigh-clear winds are providing the majority of Aeolus' positive impact. This concurs with the short-range forecast fit to observations in Figure 78. Despite the much larger noise of a single Rayleigh-clear wind compared to a single Mie-cloudy wind, the Rayleigh-clear's much greater spatial coverage and smaller representativeness error (83 km along-track averaging and broader vertical averaging in clear air) leads to a greater positive impact than the Mie-cloudy.

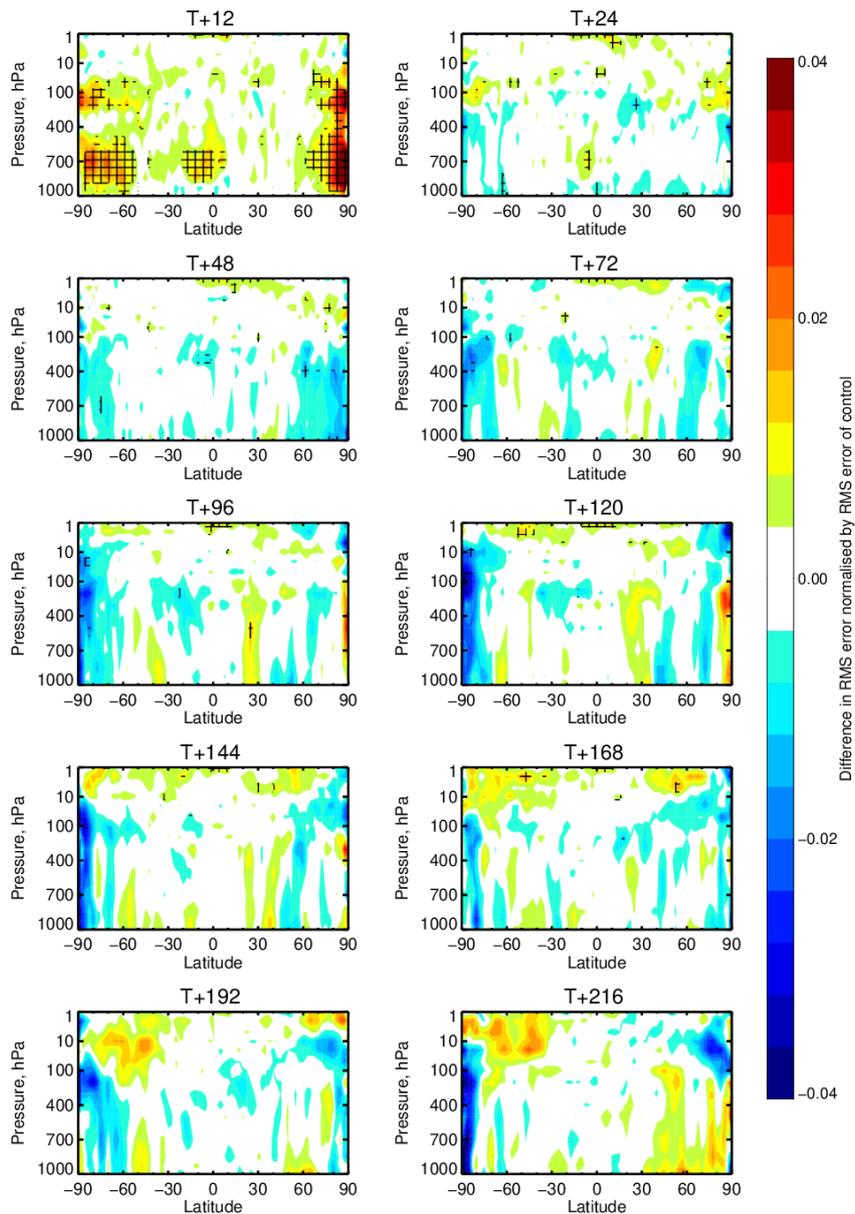


Figure 87. Zonal average normalised change in the RMS (root mean square) error of wind vector from assimilating only Mie-cloudy for the period 2 August 2019 to 31 December 2019. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence.

#### 4.3.4.4 Improving Mie-cloudy impact via the inclusion of representativeness error

However, it is demonstrated that the Mie-cloudy impact can be significantly improved by accounting for representativeness error in the assigned observation error. The formula for the assigned observation error for data assimilation including representativeness error is:

$$\sigma_{O,assign} = \sqrt{\sigma_{rep}^2 + \alpha^2 \sigma_{L2Bp,instr}^2} \quad (4)$$

Where  $\sigma_{rep}$  is the representativeness error standard deviation,  $\sigma_{L2Bp,instr}$  is the L2B processor reported instrument error standard deviation and  $\alpha$  is the scaling factor applied to the generally underestimated instrument error estimate.

With the help of “Desroziers diagnostics”; using the expectation for the standard deviation of O-B as a function of the assigned observation error and with knowledge of the assigned observation errors for radiosondes (2-3 m/s, in which instrument error is negligible, hence it is all representativeness error) it was found that  $\sigma_{rep} = 2 \text{ m/s}$  and  $\alpha_{Mie} = 1.25$  seemed to be reasonable parameters for the global average. Representativeness error is scene-dependent, so it may be possible to refine this formulation to have a variable  $\sigma_{rep}$  varying with e.g. geolocation.

The short-range forecast impact of the updated assigned error model compared to the previous weighting of  $\sigma_{rep} = 0 \text{ m/s}$  and  $\alpha_{Mie} = 1.75$  is shown in Figure 88 via the fit to ATMS microwave radiances. The impact on the tropospheric humidity channels (18-22) and stratospheric temperature channels (14 and 15) is improved the most. There is also some improvement in the tropospheric temperature channels (6-9).

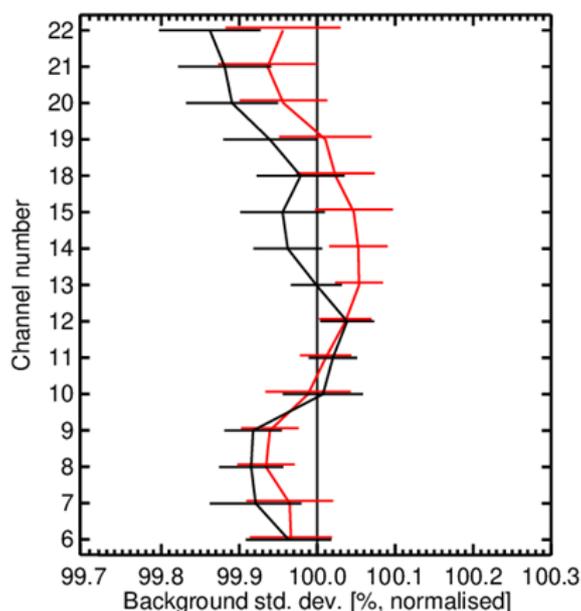


Figure 88. Change in standard deviation of O-B departures relative to global ATMS microwave radiances resulting from assimilating Aeolus Mie-cloudy HLOS wind observations, normalised so that the control is 100%. The red line is with Mie-cloudy only using the original assigned error model and the black line is with the updated error model including representativeness error. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. For the period 2 August 2019 to 31 December 2019.

Mie-cloudy forecast impact verified against operational analyses is also improved with the new assigned observation error model, as shown in Figure 89. This should be compared to the impact

of the original error model as shown in Figure 87; there is considerably more blue (positive impact) when using the refined error model, and some statistical significance in the NH high latitudes at day two. Although statistical significance is still generally lacking.

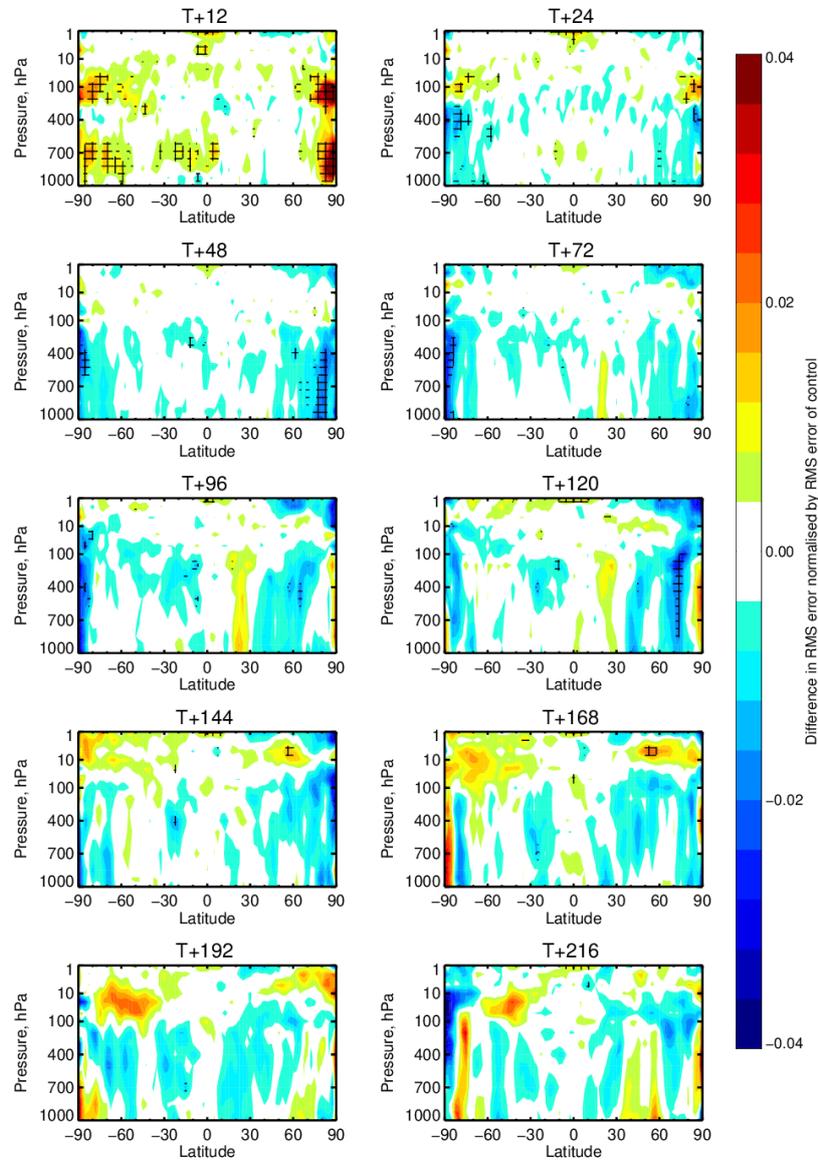


Figure 89. Zonal average normalised change in the RMS (root mean square) error of wind vector from assimilating only Mie-cloudy with an improved assigned error model for the period 2 August 2019 to 31 December 2019. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence.

The combined impact of Rayleigh-clear and Mie-cloudy with the new representativeness error observation error is shown in Figure 90. This is a useful improvement in impact relative to the case with the nominal observation weighting for the Mie-cloudy winds in Figure 80.

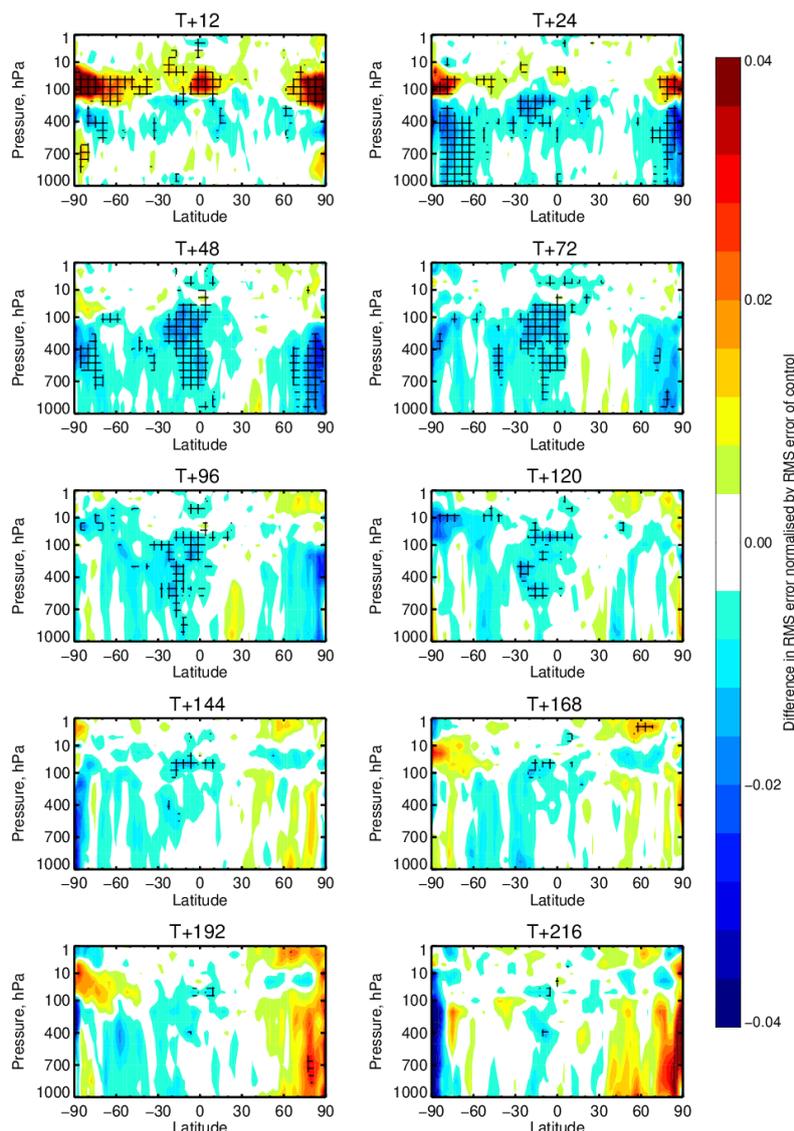


Figure 90. Zonal average normalised change in the RMS (root mean square) error of wind vector from assimilating only Mie-cloudy with an improved assigned error model in combination with Rayleigh-clear winds for the period 2 August 2019 to 31 December 2019. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence.

#### 4.3.4.5 Rayleigh-cloudy only

The medium-range forecast impact from assimilating L2B Rayleigh-cloudy winds only was found to be generally neutral, however there is suggestion of a small 0.3% improvement in the tropical vector wind at 500 hPa into the medium range. This is supported by the positive impact on short-range forecast tropical humidity as shown in section 4.3.3.3. This is promising given the expected future reductions in the biases of Rayleigh-cloudy winds.

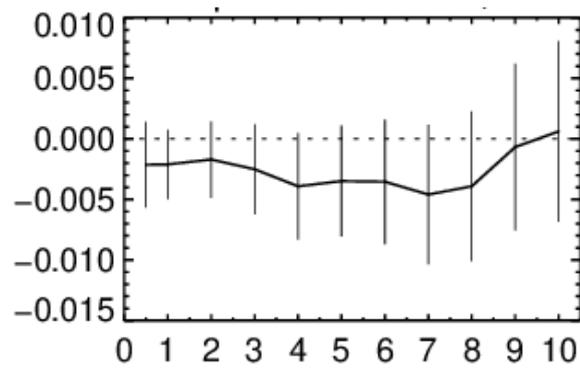


Figure 91. Normalised change in the root mean square error in vector wind at 500 hPa in the tropics from assimilating Rayleigh-cloudy winds only (from a full observing system baseline). The period is from 2 August 2019 to 31 December 2019. Confidence ranges are 95%. Verified against operational analyses. Negative values indicate a reduction in error from assimilating Aeolus.

#### 4.4 Mid-2020 NRT FM-B dataset OSEs

In this section we provide the results of the Aeolus HLOS wind assimilation experiments performed for the mid-2020 FM-B NRT dataset as described in sections 2.1.4 and 2.4.4.

##### 4.4.1 Bias correction scheme

The ECMWF's IFS (look-up table) bias correction of Aeolus HLOS winds was not required for experiments where the M1 temperature-bias-correction was applied to the L2B data (see sections 2.1.4 and 3.2). However, one experiment was performed to compare the impact of the M1 temperature-bias-correction to the ECMWF geolocation based bias correction look-up table method. The bias correction look-up table method (as a function of argument of latitude and longitude) is the same as described in section 4.3.1 for application on early FM-B NRT data.

##### 4.4.2 Analysis changes due to assimilating Aeolus HLOS winds

The standard deviation of differences between analyses using Rayleigh-clear and Mie-cloudy winds and the control not using Aeolus for this period are shown in Figure 92 to Figure 95. The spatial patterns resemble those of the early FM-B period. The largest random differences (as seen previously) occur in convergence zones, peaking in the ITCZ. There are some more subtle discrepancies relative to the early FM-B period which are probably associated with different seasons. The magnitude of the standard deviation is smaller in the extratropics for the mid-2020 period compared to early FM-B, which may be explained by larger random errors for Aeolus Rayleigh-clear winds during the more recent period (see section 3.3.1), and hence Aeolus has less weight in the analysis and smaller increments.

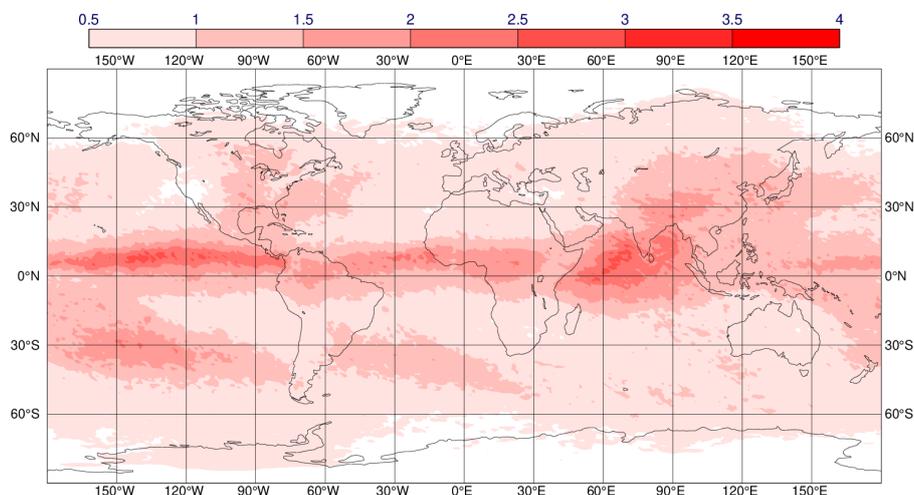


Figure 92. Standard deviation of the differences of  $u$ -wind component (m/s) at 250 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 4 April to 19 August 2020.

Some enhanced variance is present over the southern USA at 700 hPa, which is thought to be associated with large background forecast wind errors associated with mesoscale convective systems for this time of the year (in situ wind observations i.e. aircraft and radiosondes confirm large background errors in such cases). Again, this concurs with the hypothesis that Aeolus is mostly correcting the short-range forecast winds in the areas of high wind uncertainty associated with convection. The largest standard deviation of analysis differences occurs in the tropical upper troposphere, e.g. at 150 hPa,

particularly in the East Pacific as shown in Figure 94. Again, this is another area which tends to have the largest EDA spread for wind and hence background forecast errors.

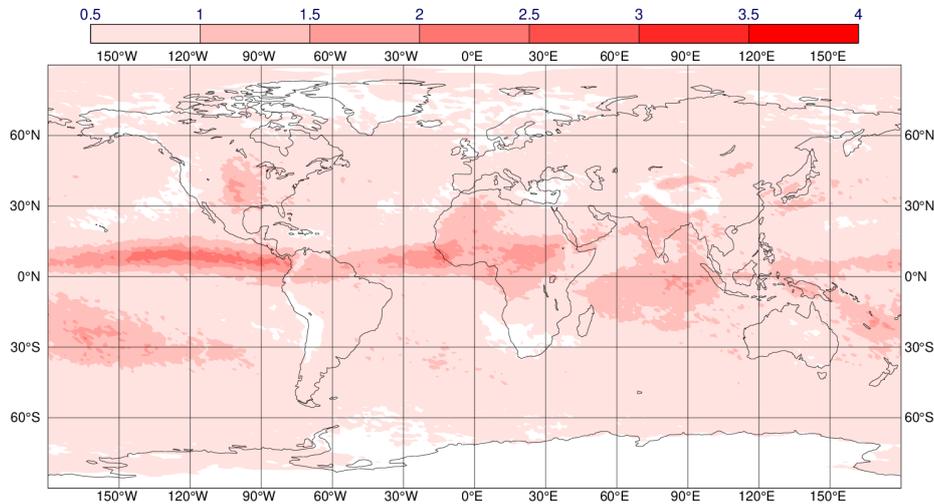


Figure 93. Standard deviation of the differences of u-wind component (m/s) at 700 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 4 April to 19 August 2020.

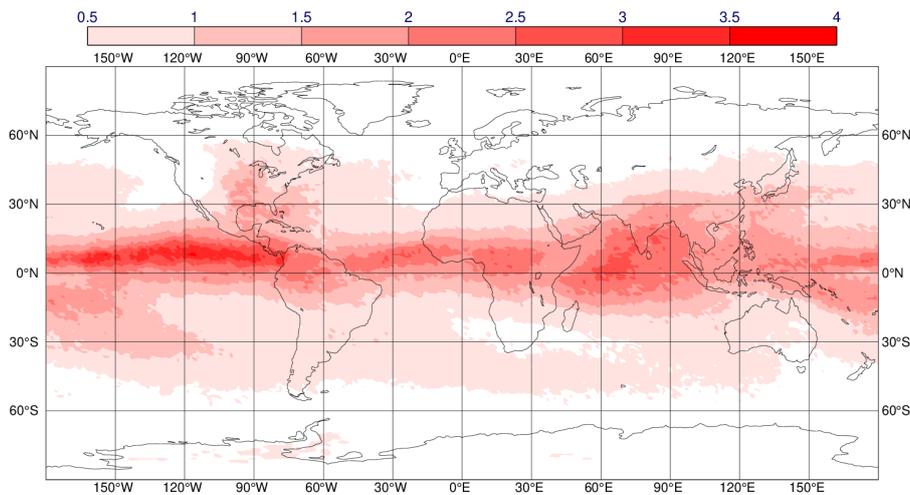


Figure 94. Standard deviation of the differences of u-wind component (m/s) at 150 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 4 April to 19 August 2020.

There are again large magnitude changes in the winds very high in the model (1 hPa) because of assimilating Aeolus at lower altitudes; as was found for the early FM-B NRT dataset period. The pattern in the SH suggests orographic gravity wave related activity downwind of Patagonia, which encircles the latitude band near 60°S. Also, pronounced differences in the tropics at the top of the model exist, which are likely to be associated with Aeolus-induced changes in the propagation of inertia-gravity waves; as shown in previous OSEs.

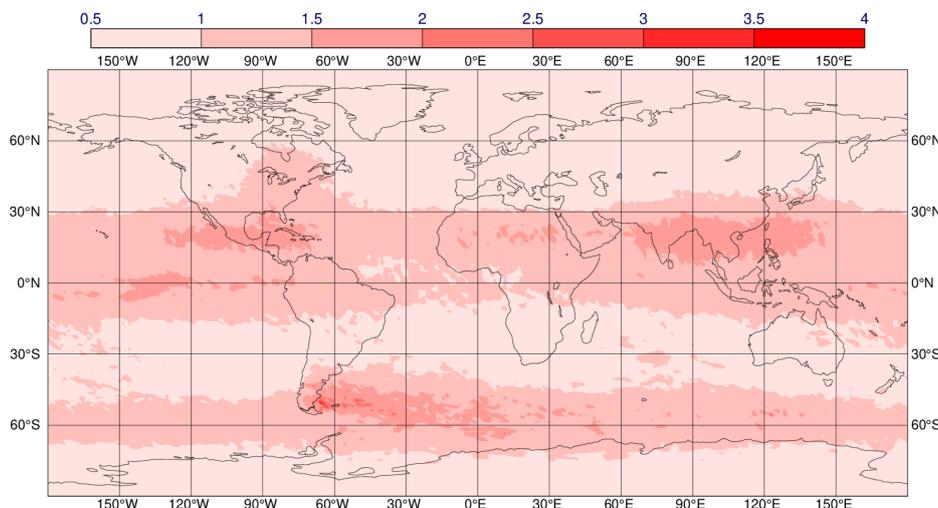


Figure 95. Standard deviation of the differences of u-wind component (m/s) at 1 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 4 April to 19 August 2020.

A similar geographical pattern to the wind differences is evident in the temperature analysis differences; Figure 96.

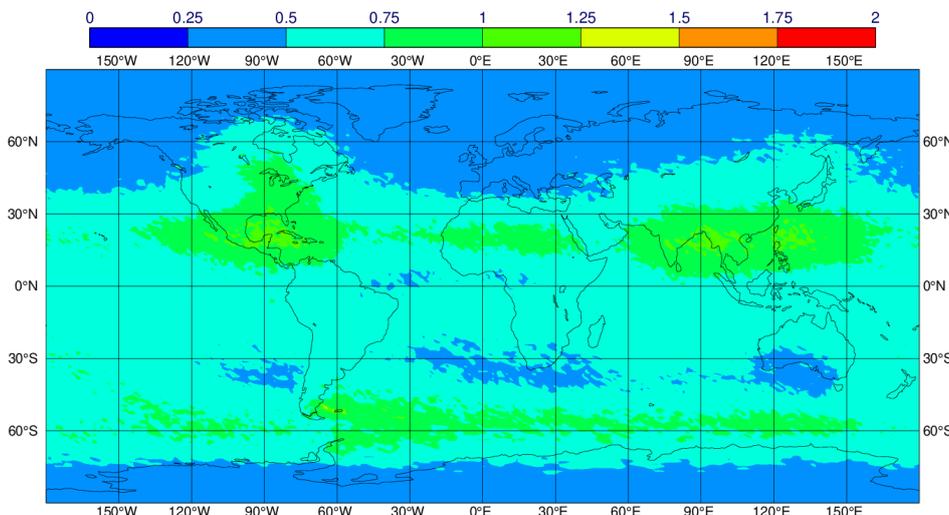


Figure 96. Standard deviation of the differences of temperature (K) at 1 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 4 April to 19 August 2020.

The temperature differences at 900 hPa (see Figure 97) are largest near the N. Pole, which is not seen in the zonal and meridional wind component differences (not shown). The reason for this pattern is not understood.

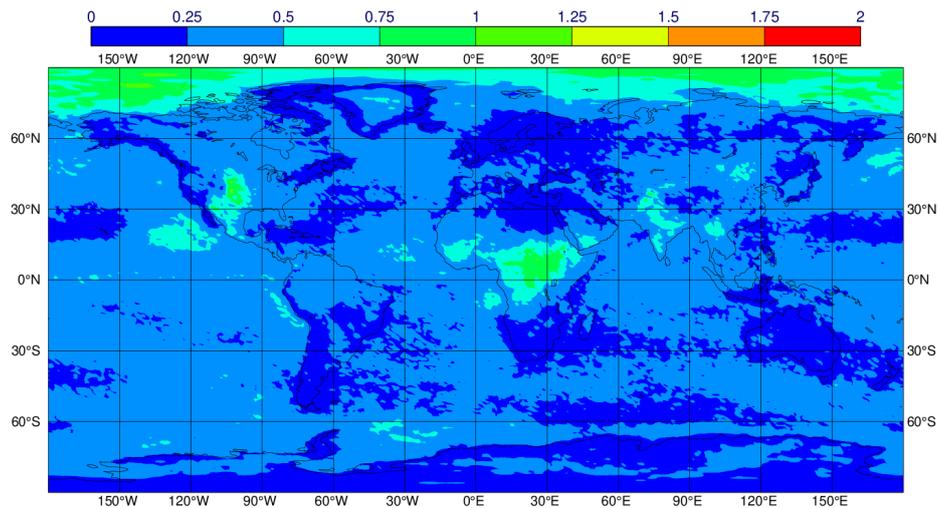


Figure 97. Standard deviation of the differences of temperature (K) at 900 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 4 April to 19 August 2020.

The standard deviation of the temperature differences at 70 hPa (Figure 98) is particularly large over the Indian subcontinent and surrounding ocean, which may be associated with uncertainties from convection during the Indian Monsoon, which became active from around June 2020.

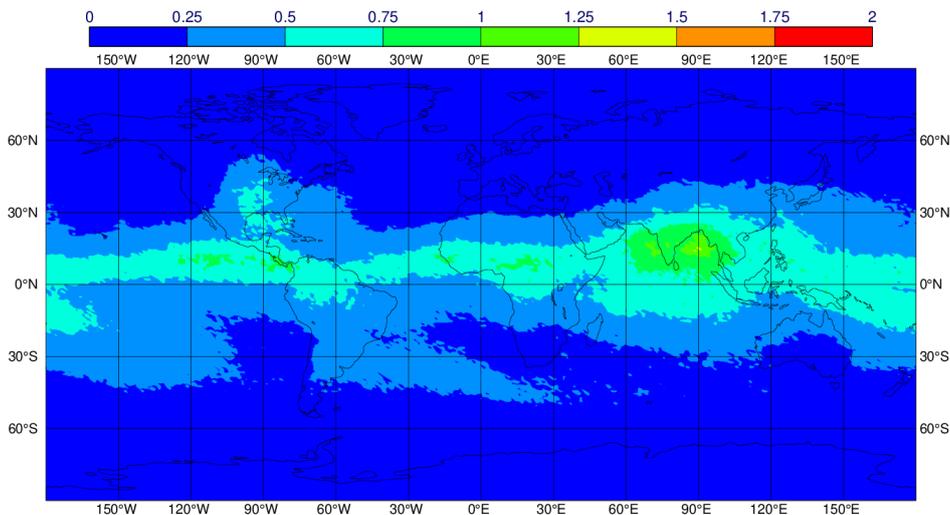


Figure 98. Standard deviation of the differences of temperature (K) at 70 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 4 April to 19 August 2020.

The largest systematic (mean) difference in the analysis with and without Aeolus occurs at around 100 hPa for zonal wind in the tropics (see Figure 99) which is consistent with early periods (typically less than 0.5 m/s differences). Aeolus tends to make the u-wind component larger at the equator (i.e. more westerly) and more easterly at latitudes either side (near Africa and India). Looking at the mean u-wind component from the control analysis (Figure 100), then perhaps Aeolus is making systematic adjustments to the position and strength of the Tropical Easterly Jet associated with the Indian Monsoon.

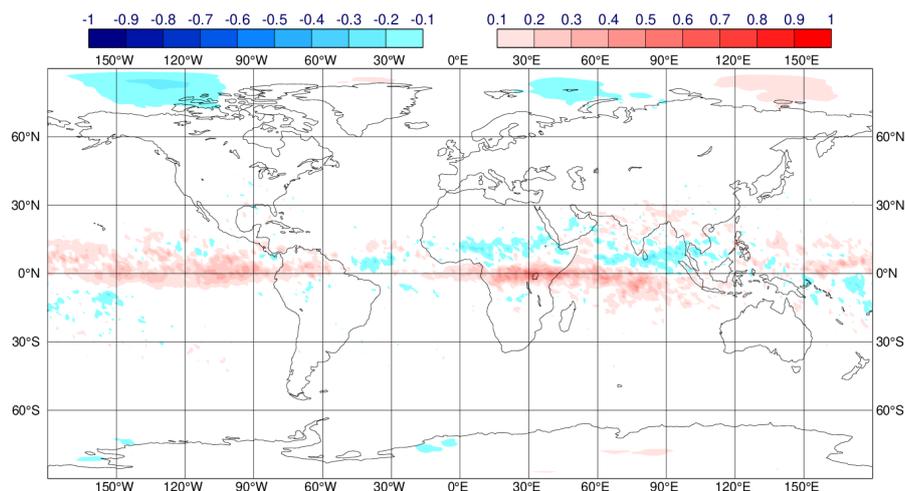


Figure 99. Mean of the differences of  $u$ -wind component (m/s) at 100 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 4 April to 19 August 2020.

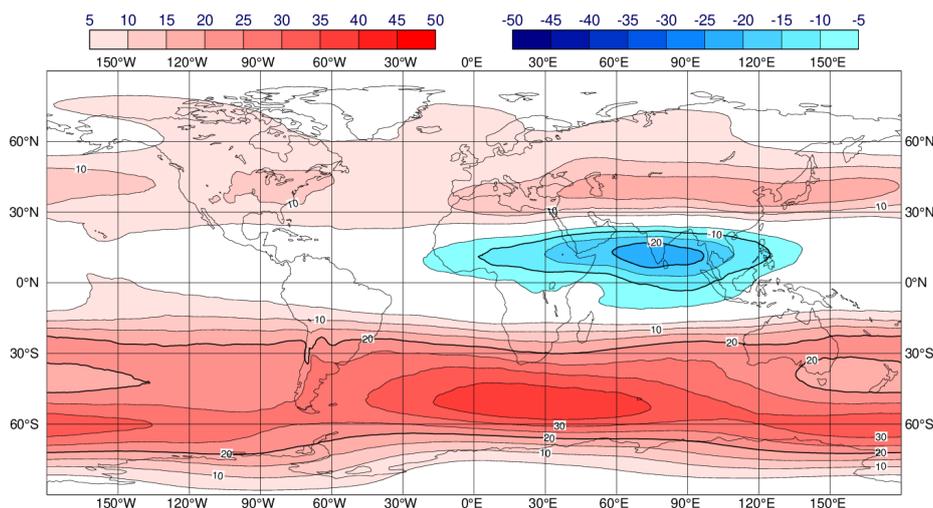


Figure 100. Mean of  $u$ -wind component (m/s) at 100 hPa of the control (not using Aeolus HLOS winds) for the period 4 April to 19 August 2020.

There are also systematic changes in the NH polar wind due to Aeolus in Figure 99, which are also seen in the  $v$ -wind component (Figure 101). But they are small mean changes at typically 0.1-0.2 m/s. Such patterns have been seen in early test periods. The tropical mean differences in  $v$ -wind are much less systematic in spatial extent than for  $u$ -wind.

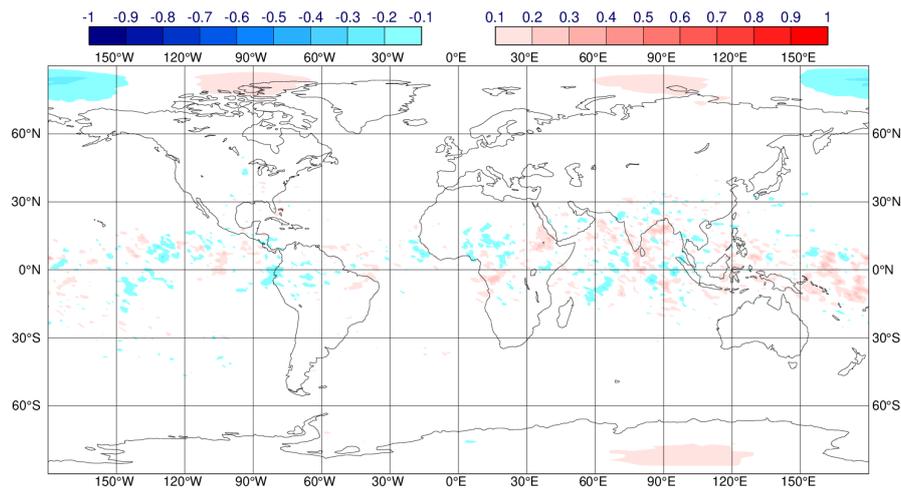


Figure 101. Mean of the differences of v-wind component (m/s) at 100 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 4 April to 19 August 2020.

#### 4.4.3 Short-range forecast impact

##### 4.4.3.1 A comparison of bias correction methods: geolocation-based look-up table versus the M1 temperature-dependent

A short OSE was performed for the period 4-19 April 2020 to compare the impact of L2B data (Rayleigh-clear and Mie-cloudy winds) assimilated with the ECMWF geolocation-based look-up table bias correction method (as a function of argument of latitude and longitude) to the equivalent L2B data using the M1 temperature dependent bias correction (see section 3.2). It was a short test period during which the ECMWF (DAMI contract) back-up server ran the L2B processor with the M1 T bias correction whilst the operational L2B products were still not bias corrected in the L2B processing. The M1 T bias correction was applied in operations from 20 April 2020 onwards.

Because this is a short OSE the forecast verification using analyses is not statistically robust. However, the short-range forecast verification with respect to other observation types has statistical significance, therefore we only show those results. Figure 102 shows that the impact with the M1 T bias correction (black line) is better than with the ECMWF geolocation-based look-up table bias correction (red line) with respect to ATMS microwave radiances in the tropics. Similar improvements were shown with the M1 T bias correction relative to other observation types. It is unclear why the Aeolus impact with the geolocation look-up table bias correction was rather neutral for this period, whereas it was shown to be positive in the early FM-B NRT dataset experiments. It could be because April is a month with fast changes in the M1 T bias structure versus geolocation and the look-up table based method lagged the real biases.

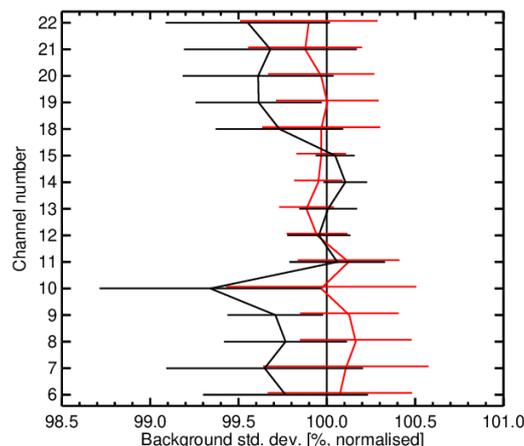
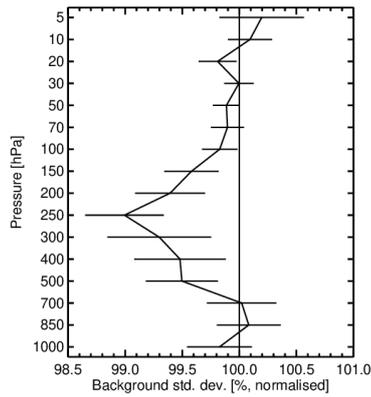


Figure 102. The change in standard deviation of O-B departures relative to ATMS microwave radiances in the tropics resulting from assimilating Aeolus HLOS wind observations with different bias correction methods. The plots are normalised so that the control is 100%. The red line is with the ECMWF geolocation based bias correction look-up table method and the black line is with the M1 temperature-dependent-bias correction (as part of the L2B processing). Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. For the period 4 to 19 April 2020.

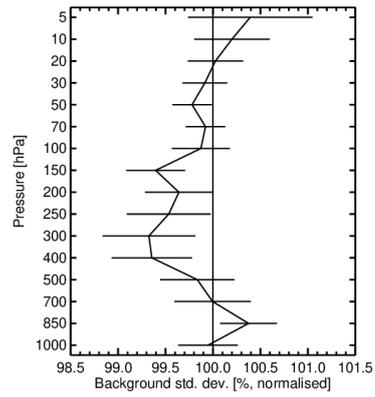
#### 4.4.3.2 Rayleigh-clear and Mie-cloudy with M1 temperature-dependent-bias correction combined with improved Mie assigned observation error

The impact of Rayleigh-clear and Mie-cloudy data on the short-range forecasts (12 hours) as verified by the fit to other observation types is shown in Figure 103 and Figure 104. The L2B data had the M1 T bias correction applied (see section 3.2) and the Mie-cloudy winds were assimilated with the improved assigned observation error model (see section 4.3.4.4).

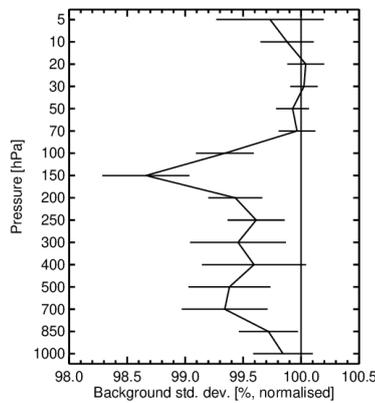
Figure 103 shows in situ (conventional) wind and radiosonde temperature observations — this should be compared to the equivalent plots for early FM-B data: Figure 75. Aeolus improves the short-range forecast fit to vector wind and temperatures by typically less than 1%, which is consistent with the early FM-B impact NRT dataset experiment. The impact is largest for tropical winds, peaking at 150 hPa (around 1.3%), agreeing with the earlier OSEs. The positive impact is present throughout the troposphere and lower stratosphere. The impact on wind is stronger than for temperature. The impact throughout the troposphere and lower stratosphere in the SH peaks at 250 hPa and is stronger than in the early FM-B period (1% rather than 0.5% for vector wind); perhaps because of the longer stretch of SH winter. NH impact is mostly neutral but there is a slight positive impact in the troposphere (around 500 hPa) for temperature.



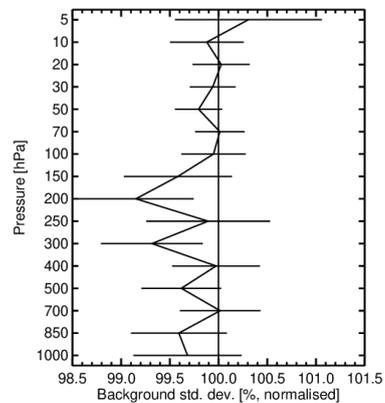
a)



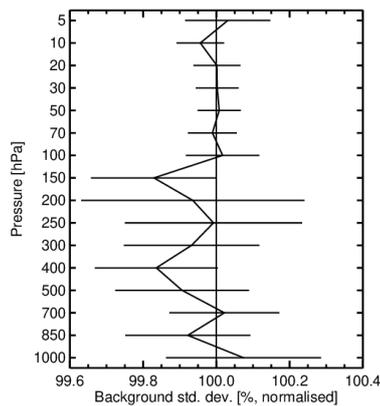
b)



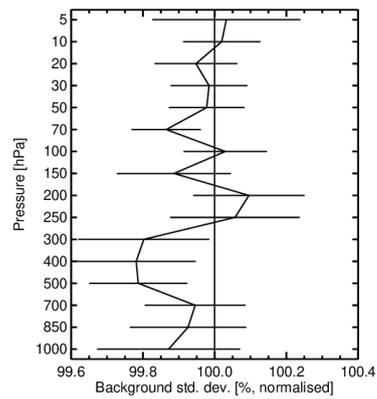
c)



d)



e)

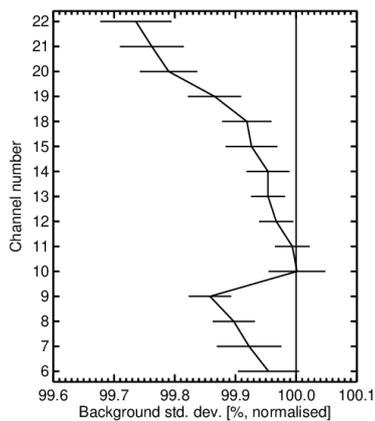


f)

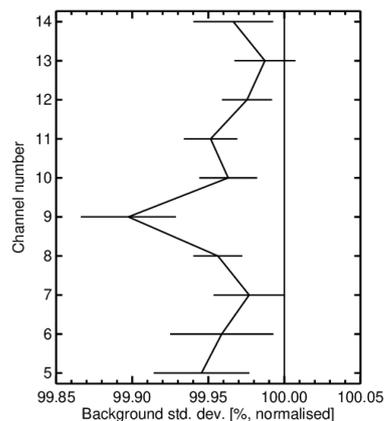
Figure 103. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) conventional data vector wind in the SH extratropics, b) radiosonde temperatures in the SH extratropics, c) conventional data vector wind in the tropics, d) radiosonde temperatures in the tropics, e) conventional data vector wind in the NH extratropics and f) radiosonde temperatures in the

*NH extratropics. Horizontal bars show the 95% confidence range (Student's t-test). For the period from 4 April 2020 to 20 September 2020.*

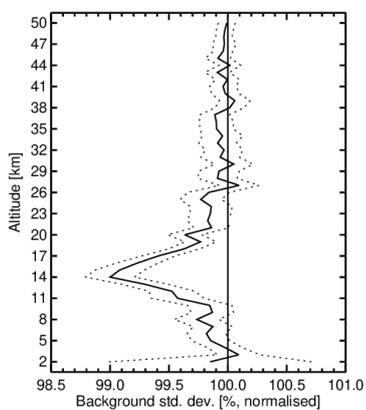
Figure 104 shows the short-range forecast fit to other satellite observations (which should be compared to Figure 76 for the early FM-B period). Aeolus improves the fit to global microwave sounding temperature sensitive channels of ATMS and AMSU-A and also ATMS humidity sensitive channels (18-22) in a similar manner to early period OSEs. GNSS radio occultation fits are improved by 1% at ~14 km in the tropics (upper troposphere), which is a similar pattern to earlier periods, but of smaller magnitude. There was a significant increase in the number of radio occultation measurements assimilated in 2020 (from the COSMIC-2 and Spire constellations), which results in smaller background errors for Aeolus to improve upon in the tropical upper troposphere temperatures and geostrophic winds — also the Rayleigh winds are noisier relative to early FM-B. Relative to the infrared sounding instrument CrIS, Aeolus shows improvements in temperature sensitive channels in the upper troposphere and lower stratosphere and in channels sensitive to the surface temperature and humidity; which again is like the early FM-B period impact. The change relative to Atmospheric Motion Vectors (AMVs) is more consistently positive than shown in earlier periods. The fit to geostationary water vapour imagery is improved, as is the fit to the FY-3B microwave humidity sounder; in a similar manner to early OSEs. Finally, in the NH, there are useful improvements in the fit to surface pressure observations, with more significance than earlier periods.



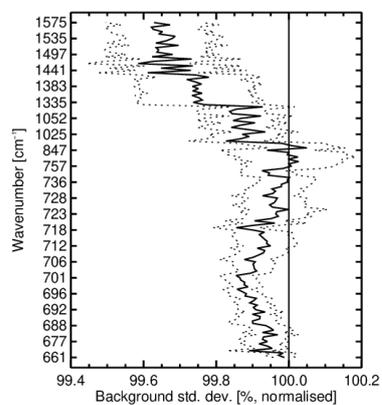
a)



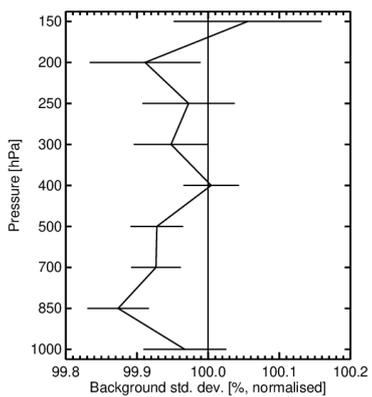
b)



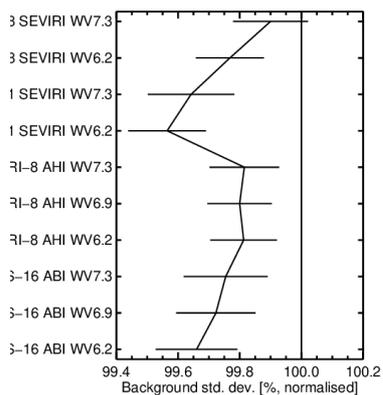
c)



d)



e)



f)

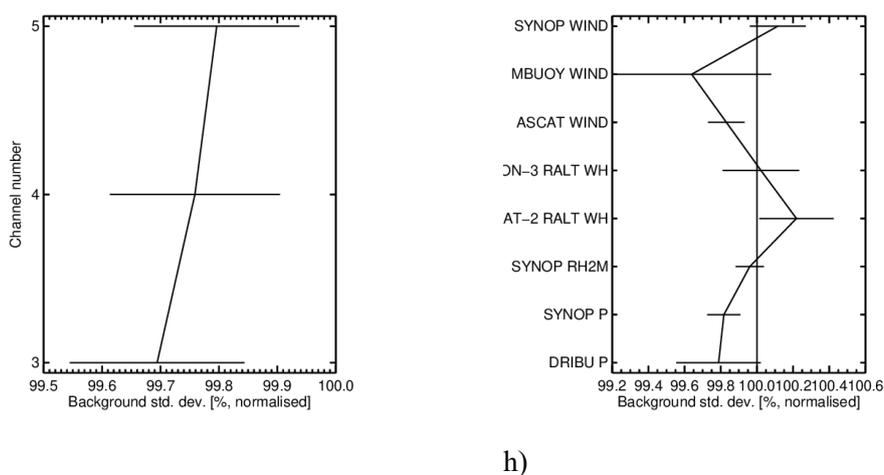


Figure 104. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) ATMS global statistics, b) AMSU-A global statistics, c) GNSS radio occultation in the tropics, d) CrIS global statistics, e) AMVs global statistics, f) geostationary imagery global statistics, g) FY-3B MWHS global statistics and h) various surface observation types in the NH. Horizontal bars show the 95% confidence range (Student's t-test). For period from 4 April 2020 to 20 September 2020.

#### 4.4.4 Medium-range forecast impact

##### Rayleigh-clear and Mie-cloudy with M1 T-dependent-bias correction combined with improved Mie assigned observation error

Figure 105 to Figure 107 show the impact of Aeolus Rayleigh-clear and Mie-cloudy winds with the M1 T-bias-correction applied (see section 3.2) and the improved Mie error assignment for the period from 4 April 2020 until 20 September 2020 on vector wind. The improved Mie error assignment (including representativeness error) was successfully tested in the early FM-B period; see section 4.3.4.4. The same assigned observation error parameters are used here. Because Aeolus was operationally assimilated during this period, the verification against own analyses is also shown in Figure 106 to compare to the impact against operational analyses in Figure 105 to help assess the real impact of Aeolus at shorter forecast ranges. The figures should be compared Figure 90 showing the equivalent verification for the early FM-B NRT dataset OSE. It appears that the impact in the tropical UTLS does not persist so well into the medium range for the mid-2020 period OSE compared to the early FM-B period OSE.

The impact on 2 to 4-day forecasts is mostly positive with statistical significance in the tropical troposphere and in polar areas, with some agreement for verification against operational analyses and own analyses. The percentage improvements are on the order of 1-2% in the tropics. In the extratropics the positive impact up to the lower stratosphere is typically less than 1% for the whole area. The very large magnitude of the positive impact at 12 hours relative to operational analysis is probably overestimated, due to the operational assimilation of Aeolus; it may reflect where Aeolus is similarly changing the experiment analysis and operational analysis. The positive impact has similar geographical patterns to that shown for the early period OSEs.

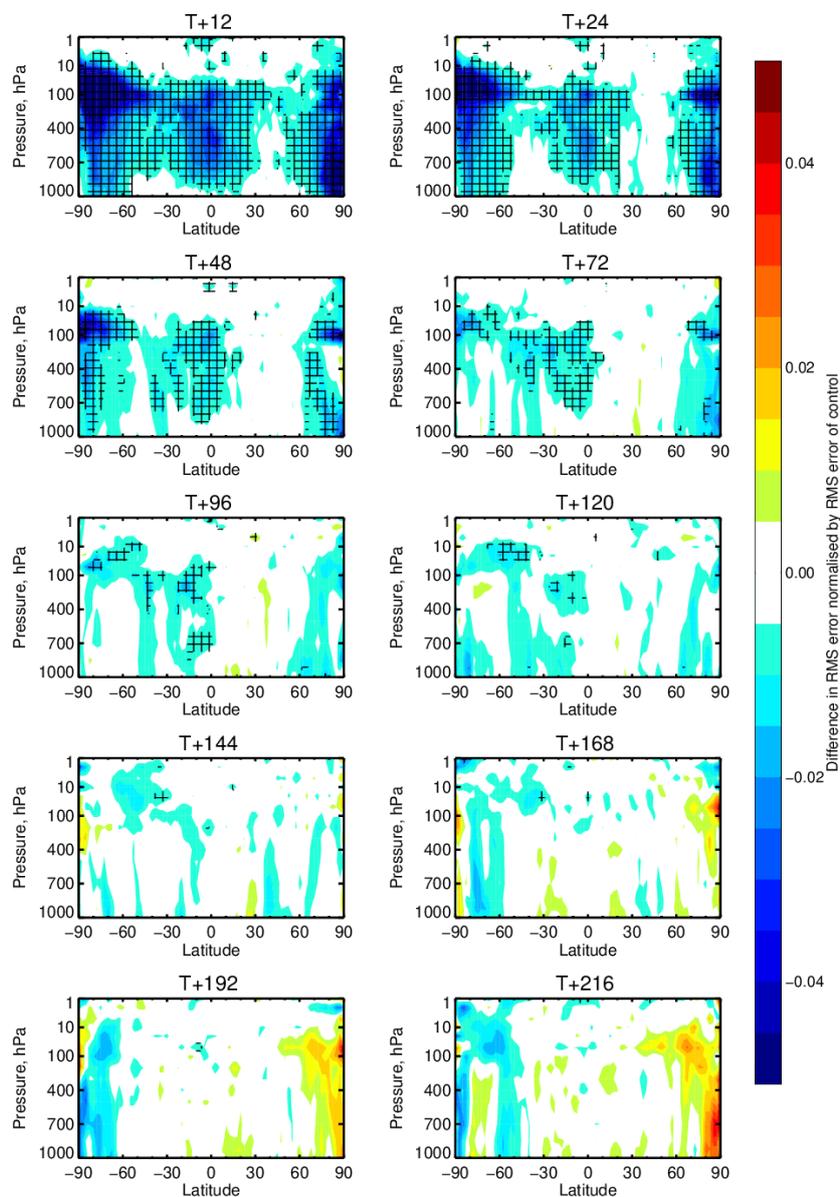


Figure 105. Zonal average normalised change in the RMS (root mean square) error of wind vector from assimilating Aeolus for the period 4 April 2020 to 20 September 2020. Rayleigh-clear and Mie-cloudy winds were used with the MI T bias correction. The Mie-cloudy winds used the improved formulation for assigned observation error. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence. Note the scale is  $\pm 0.05$  ( $\pm 5\%$ ).

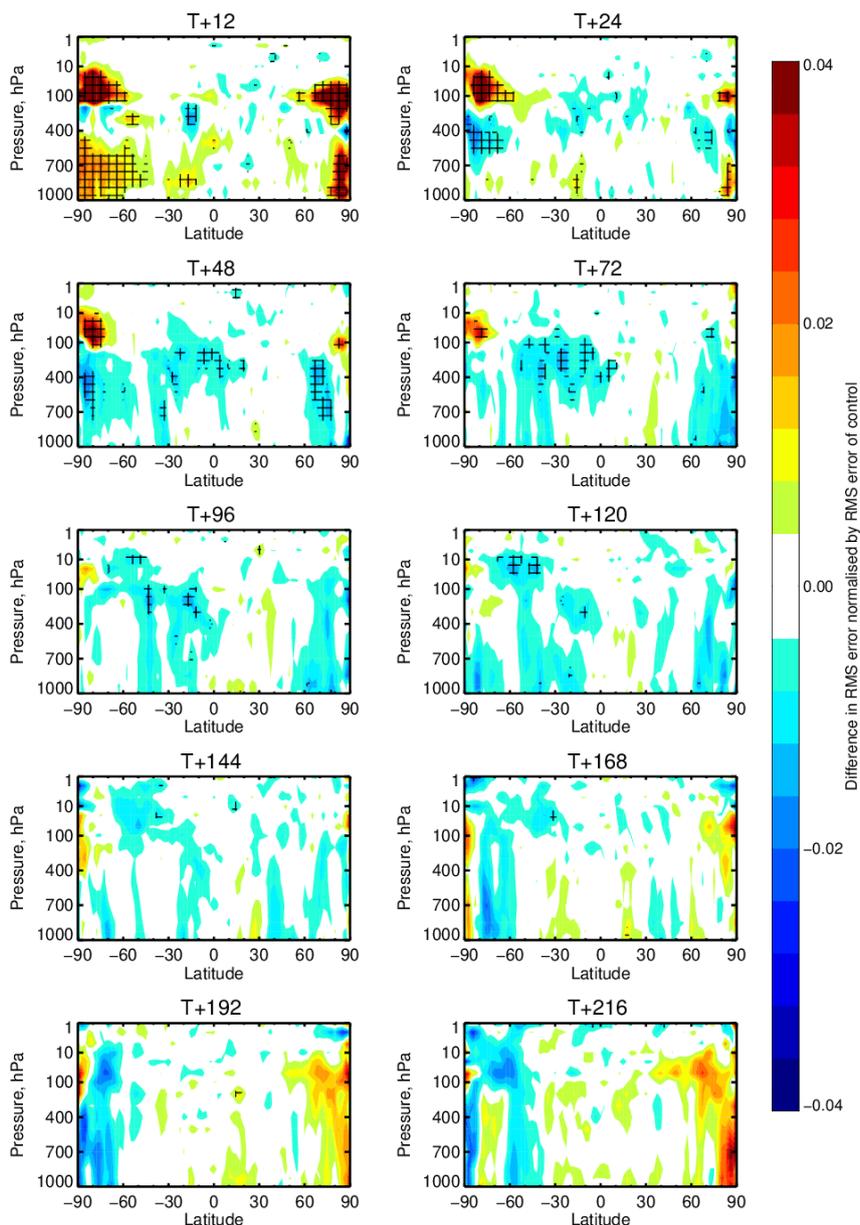


Figure 106. Equivalent to Figure 105, but with verification against own analyses. Note the scale is  $\pm 0.04$  ( $\pm 4\%$ ).

The improvements in other fields such as temperature, geopotential and relative humidity follows similar geographical patterns to the wind vector improvement (not shown).

Figure 107 demonstrates an improvement of impact, particularly in the extratropics, from the application of the representativeness error formulation for the Mie-cloudy winds error assignment.

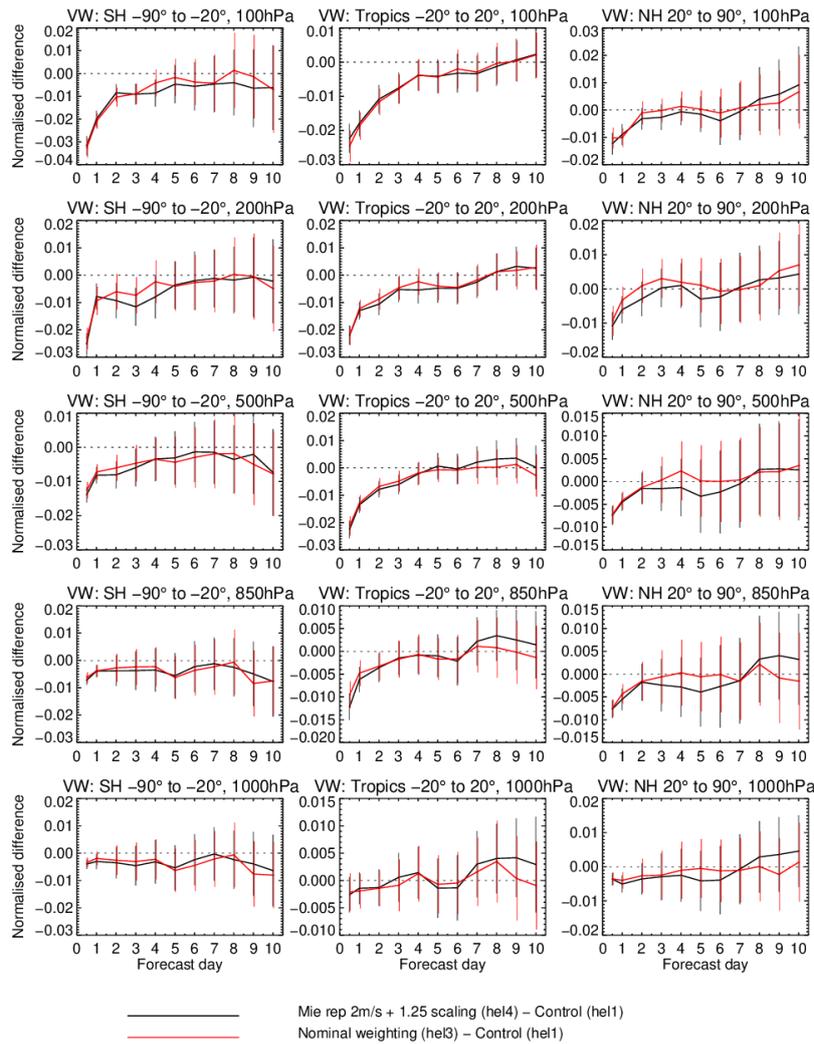


Figure 107. The normalised change in the RMS (root mean square) error in vector wind for different vertical levels (from 100 to 1000 hPa, top to bottom) and for SH, tropics and NH (left to right) from assimilating Aeolus. Rayleigh-clear and Mie-cloudy winds with the MI T bias correction were used. The red line shows the impact for the nominal observation weighting and the black line when using the improved Mie-cloudy assigned observation error (including representativeness error). The period is from 4 April 2020 to 20 September 2020. Confidence ranges are 95%. Verified against operational analyses. Negative values indicate a reduction in error from assimilating Aeolus.

## 4.5 First reprocessed dataset (FM-B) OSEs

This section provides the verification results from OSEs using the first reprocessed L2B dataset (the first official reprocessing campaign, 2B10) covering FM-B in 2019, as discussed in sections 2.1.5, 1.1.1 and 0. Several experiments with Aeolus “on” have been run: one using both the Rayleigh-clear and the Mie-cloudy winds (with the improved Mie-cloudy error assignment as discussed in section 4.3.4.4); another with increased Mie-cloudy representativeness error to 3 m/s instead of 2 m/s; one testing a method of atmospheric-temperature-dependent bias correction (look-up table) for the Rayleigh-clear winds; and one testing modifications to the assumed VarQC distribution.

### 4.5.1 Atmospheric temperature-dependent-bias correction for Rayleigh-clear winds

Monitoring of the first reprocessed dataset and of the operational L2B Rayleigh-clear winds in 2020-2021 revealed a relatively small, but persistent bias, with an apparent dependence on altitude (typically within  $\pm 1$  m/s HLOS). The bias is generally more negative at lower altitudes and positive at higher altitudes. There is no known physical reason for why a bias would causally depend on altitude and it does not have a clear dependence on satellite-to-target range and is therefore inconsistent with a range-dependent bias<sup>32</sup>. The ECMWF Advanced Monitoring showed a reasonably robust variation of the bias with atmospheric temperature, being positively biased at cold temperature and negatively biased at warmer temperatures, which due to the climatological temperature profile could explain the variation with altitude. A dependence on temperature is more plausible because of the application of, possibly imperfect, Rayleigh response curves as a function of atmospheric temperature (provided by the AUX\_RBC\_L2 file). However, the true causation is still being investigated, with the help of the IRONICs off-nadir calibrations.

As a sensitivity test to see how important the atmospheric-temperature-dependent bias is for NWP impact, it was decided to develop in the IFS (bufr2odb) a bias correction of the L2B Rayleigh-clear winds as a function of atmospheric temperature. The atmospheric temperature for each wind result is provided via the AUX\_MET data used in the L2B processing (used in the selection of Rayleigh response curves (from AUX\_RBC\_L2)) and which is conveniently stored in the L2B BUFR. The bias correction was done using the same EMCWF infrastructure that was developed for the bias correction as a function of geolocation (argument of latitude and longitude). This method uses a Fortran namelist storing the mean(O-B) binned per atmospheric temperature bin. The O-B statistics versus temperature are provided via the control run, which does not assimilate Aeolus, but does have Aeolus departures available for monitoring purposes.

For the first reprocessed dataset, the structure of atmospheric-temperature-dependent-bias changed significantly on 22 July 2019 — the dynamic range of the bias increased. This is associated with a laser frequency adjustment (LFA) and Rayleigh Cover Temperature (RCT) change on that date, which supports the hypothesis that the bias is somehow related to imperfect calibration issues. Due to this time-variation in bias characteristics, it was decided to also take advantage of the ECMWF bias correction scheme’s ability to vary with date and hence to have a different set of coefficients before and after the 22 July 2019. Examples of the bias (mean(O-B)) binned as a function of atmospheric temperature before and after the 22 July 2019 is shown in Figure 108, which shows how the bias

---

<sup>32</sup> Range-dependent-bias (RDB) was considered before launch as a possible source of wind bias due to the change in angle of incidence of the light upon the spectrometers associated with the changing angular position of the satellite along the orbit for the different time-gates (range-bins).

dependence on temperature increased after 22 July. Temperature-bin resolution was selected to be sufficient capture the function’s variation.

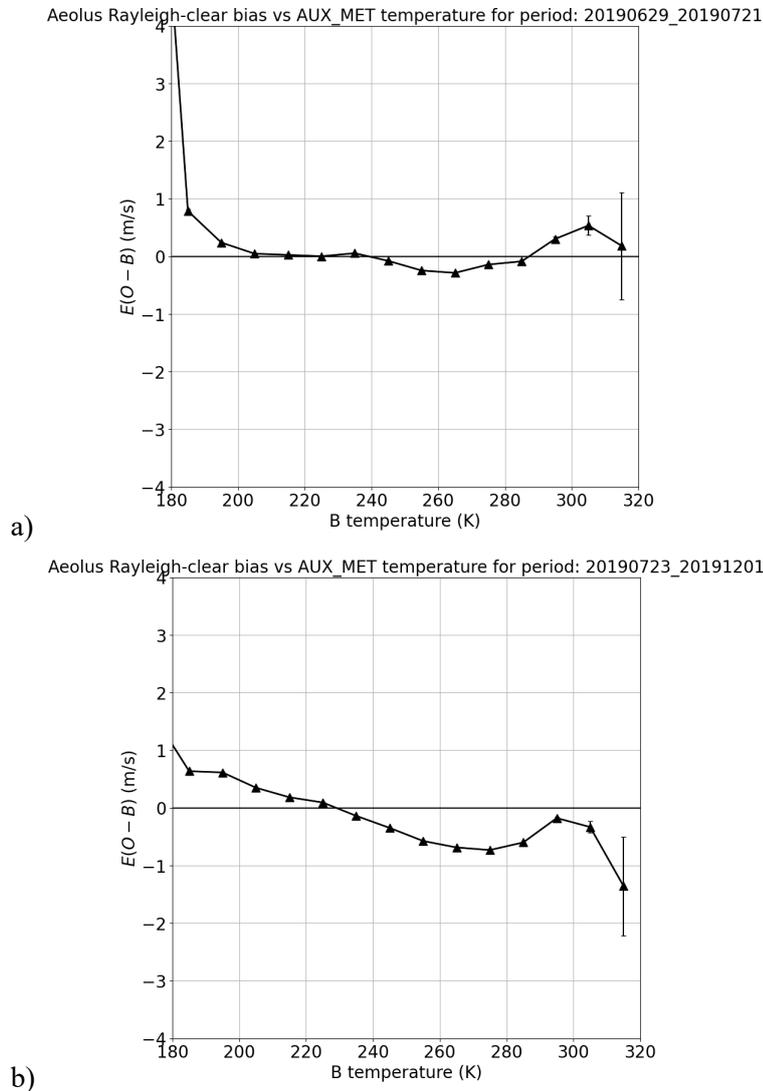
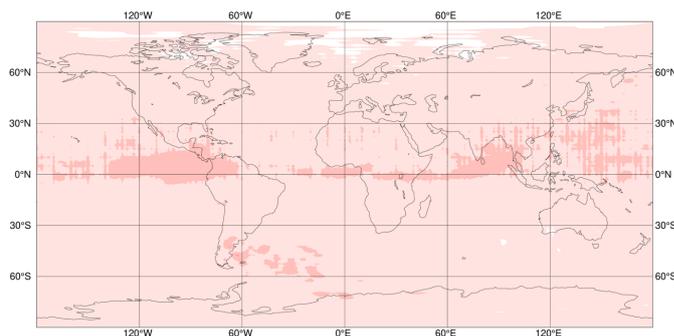


Figure 108. Dependence of the L2B Rayleigh-clear bias (mean(O-B), m/s) as a function of ECMWF background forecast temperature (K, via AUX\_MET) for the period a) 29 June to 21 July 2019 and b) 23 July to 1 December 2019. These values were used in an OSE testing the bias correction of L2B Rayleigh-clear winds as a function of atmospheric temperature.

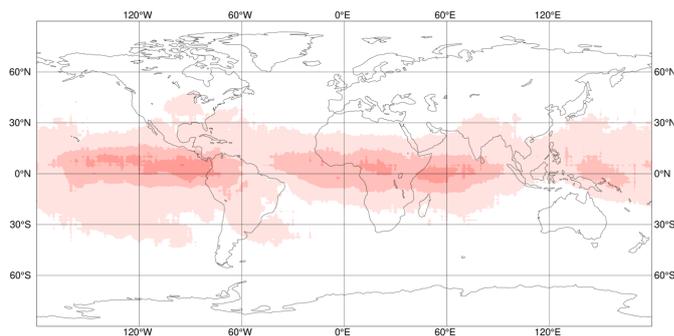
#### 4.5.2 Analysis changes due to assimilating Aeolus HLOS winds

To save on the MARS archiving costs (due to ECMWF’s data centre moving to Bologna), all 47R2 experiments have been by default archiving model fields on pressure levels at lower horizontal resolution of 2.5-degree latitude/longitude grid (~277 km grid spacing). This low resolution makes it hard to compare directly to earlier OSEs with the full resolution T<sub>CO399</sub> grid (~29 km grid spacing) archived. The lower resolution removes some smaller horizontal-scale variance from the analysis differences and affects the magnitude of the differences.

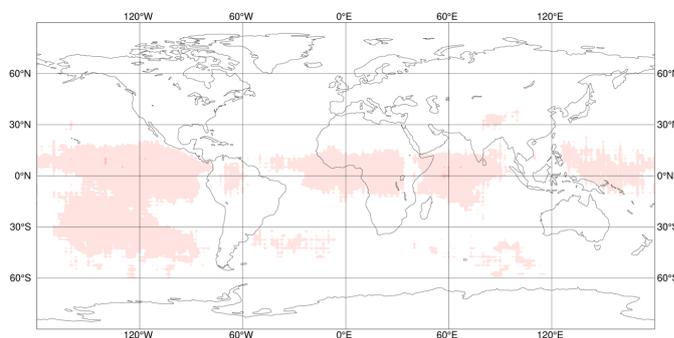
Despite the smaller magnitudes due to lower resolution, Figure 109 still shows familiar geographical patterns as seen in previous OSEs, in where Aeolus L2B winds modify in a random-sense the zonal wind analysis fields i.e. largest changes in the convergence zones.



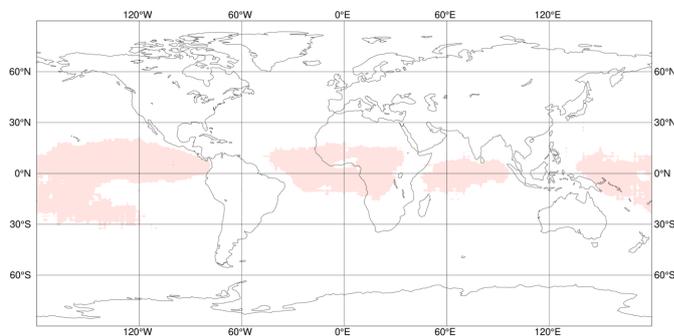
a) 1 hPa



b) 150 hPa



c) 250 hPa



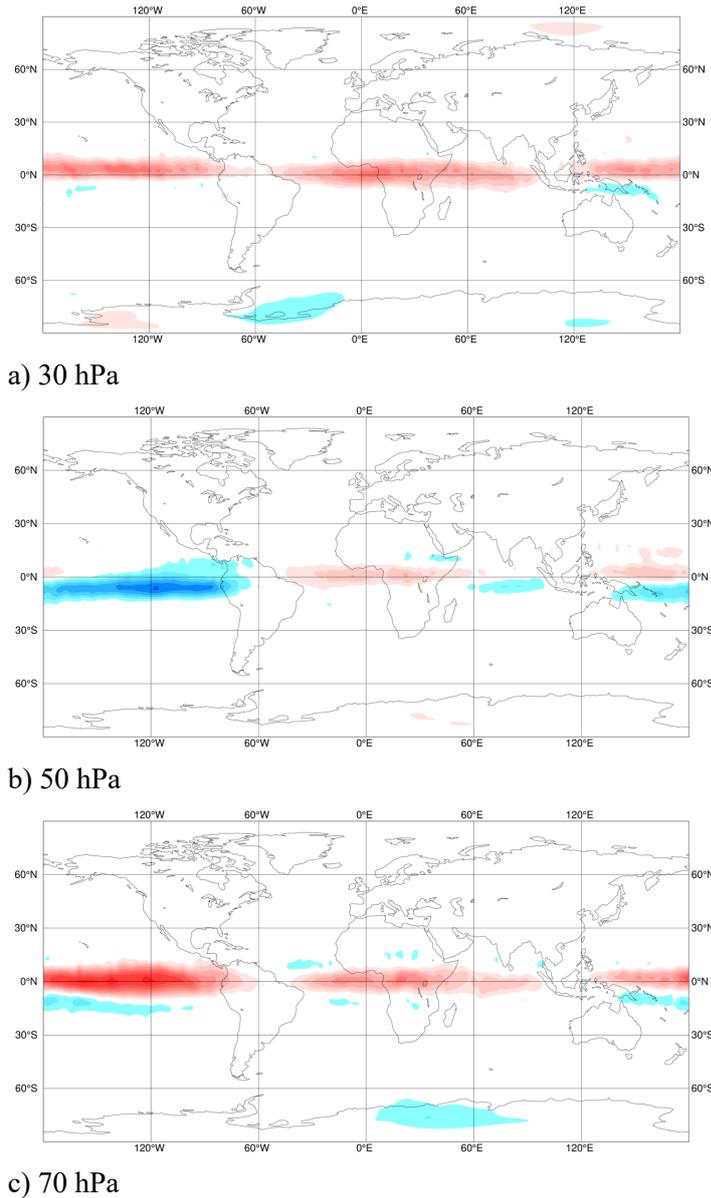
d) 700 hPa



m/s

Figure 109. Maps of the standard deviation of the differences of u-wind component (m/s) at a) 1 hPa, b) 150 hPa, c) 250 hPa and d) 700 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 29 June 2019 to 31 December 2019.

Figure 110 shows that Aeolus L2B winds systematically modify the zonal wind analysis fields most strongly near the equator and to a lesser extent in polar regions, with similar patterns seen in previous OSEs. The strongest changes are in the tropical UTLS.



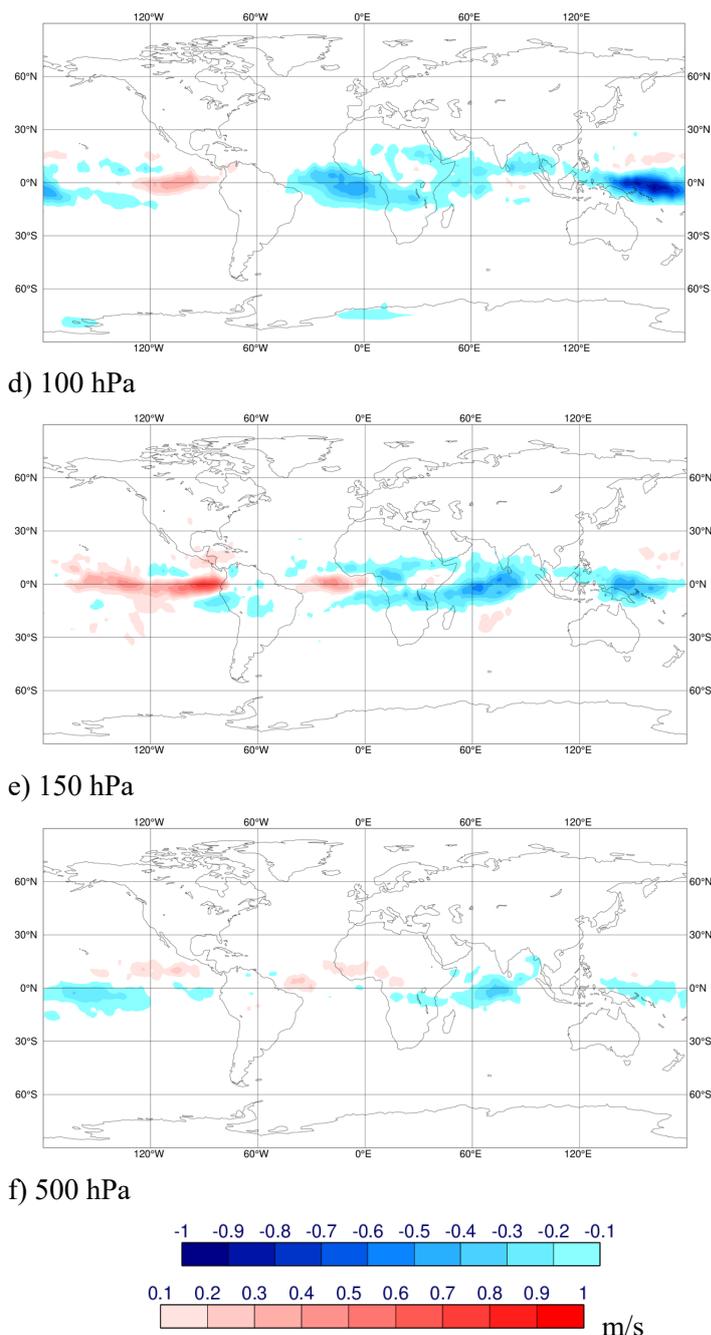


Figure 110. Maps of the mean of the differences of  $u$ -wind component (m/s) at a) 30 hPa, b) 50 hPa, c) 70 hPa, d) 100 hPa, e) 150 hPa and f) 500 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 29 June 2019 to 31 December 2019.

#### 4.5.3 Short-range forecast impact

Figure 111 and Figure 112 show the short-range forecast verification relative to other observation types for the nominal Aeolus “on” OSE i.e. Rayleigh-clear plus Mie-cloudy (with 2 m/s representativeness error). The patterns of impact is similar to the NRT dataset early FM-B OSE (Figure 75 and Figure 76), however the magnitude of positive impact is larger, particularly in the tropics; but note that these OSEs do not cover the exact same period (they overlap from 2 August to 31 December 2019).

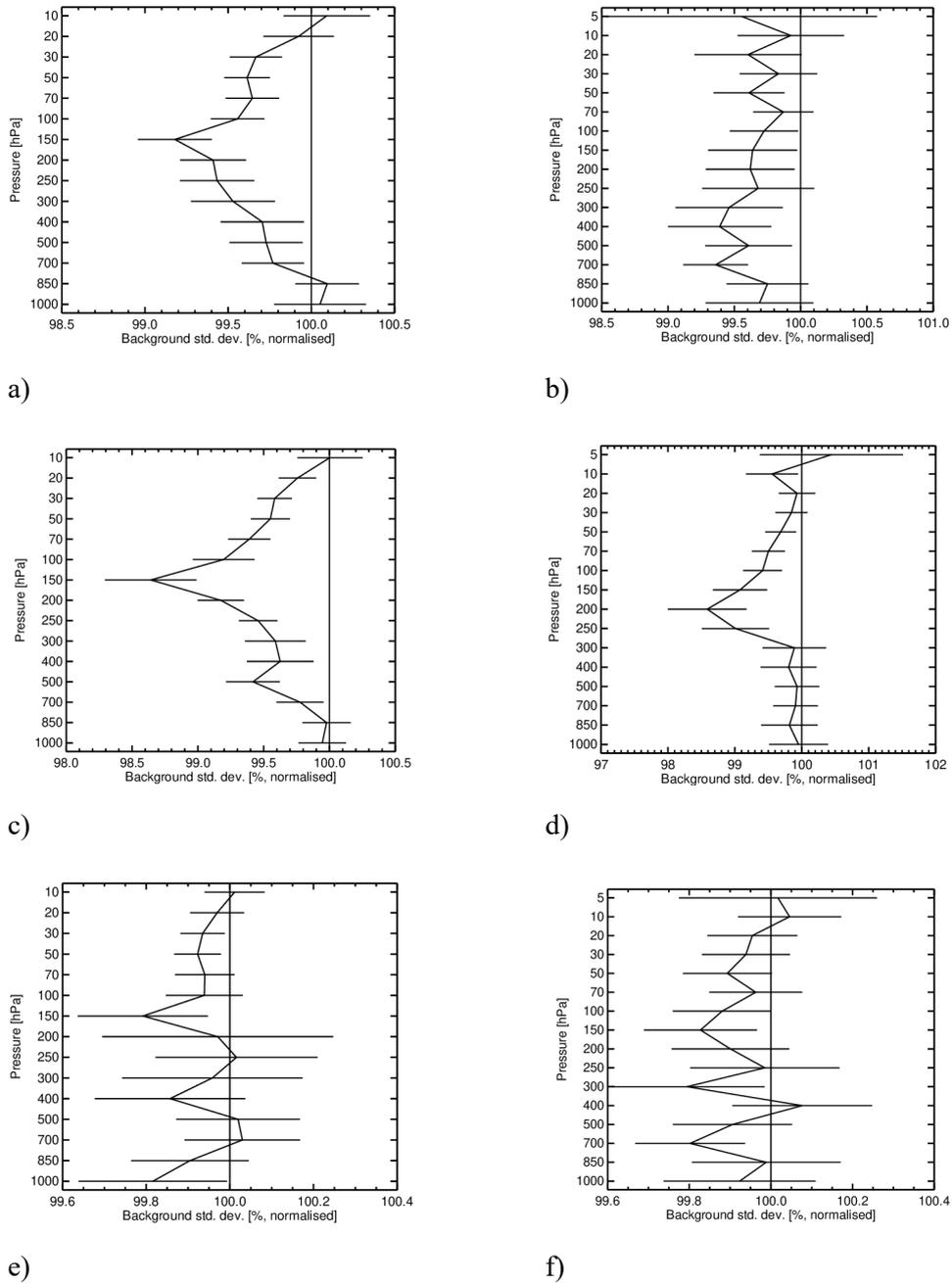
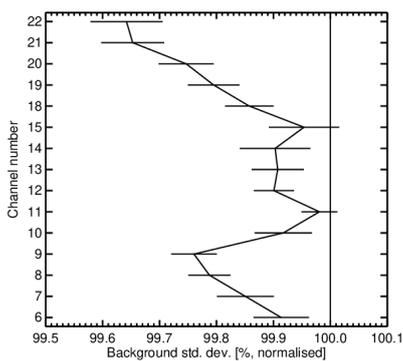
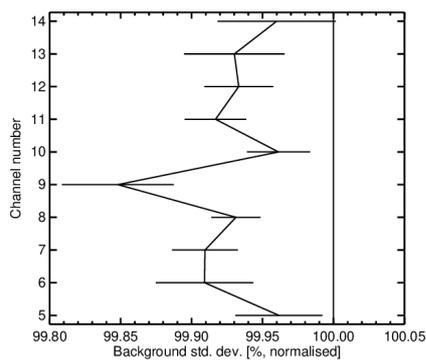


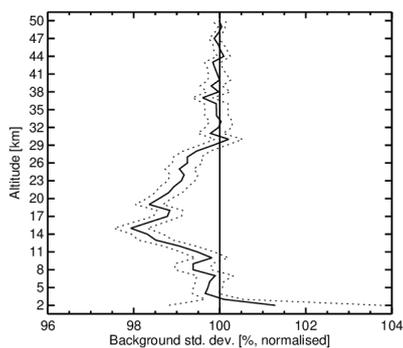
Figure 111. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) conventional data vector wind in the SH extratropics, b) radiosonde temperatures in the SH extratropics, c) conventional data vector wind in the tropics, d) radiosonde temperatures in the tropics, e) conventional data vector wind in the NH extratropics and f) radiosonde temperatures in the NH extratropics. Horizontal bars show the 95% confidence range (Student's t-test). For the period from 29 June 2019 to 31 December 2019.



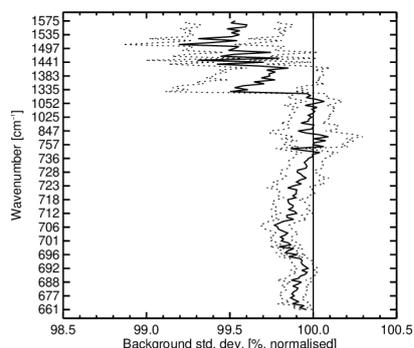
a)



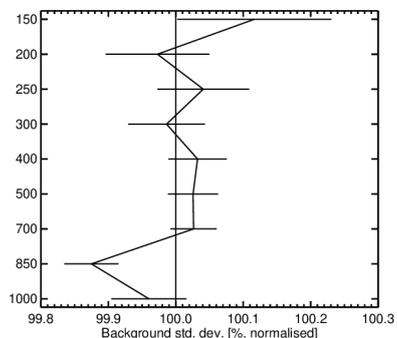
b)



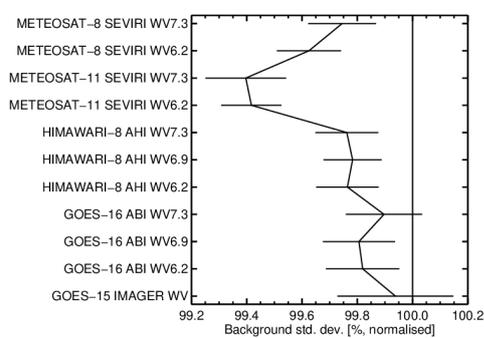
c)



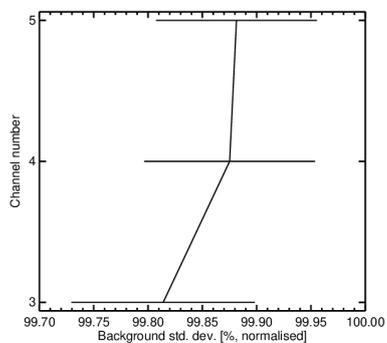
d)



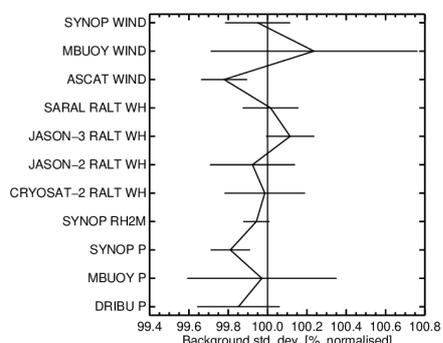
e)



f)



g)



h)

Figure 112. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) ATMS global statistics, b) AMSU-A global statistics, c) GNSS radio occultation in the tropics, d) CrIS global statistics, e) AMVs global statistics, f) geostationary imagery global statistics, g) FY-3B MWHS global statistics and h) various surface observation types in the NH. Horizontal bars show the 95% confidence range (Student's *t*-test). For period from 29 June 2019 to 31 December 2019.

The short-range fit to ATMS, due to the Rayleigh-clear atmospheric-temperature-dependent bias correction, shows stronger positive impact (black line) than without the bias correction (red line); this experiment is covers from 29 June to 31 December 2019; see Figure 113. However, some other observation-based verification is more mixed e.g. positive in NH, but not so in SH extratropics i.e. not consistently positive.

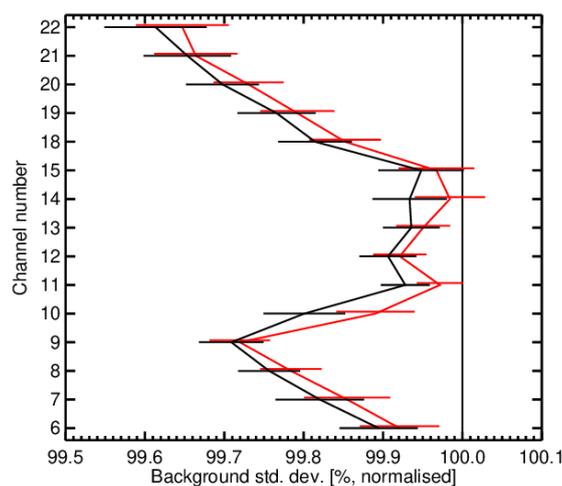


Figure 113. The change in standard deviation of O-B departures of ATMS resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The red-line is the nominal assimilation, and the black-line is with the atmospheric-temperature-dependent bias correction of L2B Rayleigh-clear winds. For period from 29 June 2019 to 31 December 2019.

The impact of applying a modified Variational QC via the Huber norm with parameter  $c=1.5$  (see Appendix 9.9) plus a very relaxed first-guess check (20-sigma) is a negative one (when tested on top of the atmospheric-temperature-dependent bias correction experiment). It is thought that the very relaxed first-guess check may have spoiled the impact somewhat due to a finite proportion of genuine Aeolus gross-errors which will get some weight. This should be revisited.

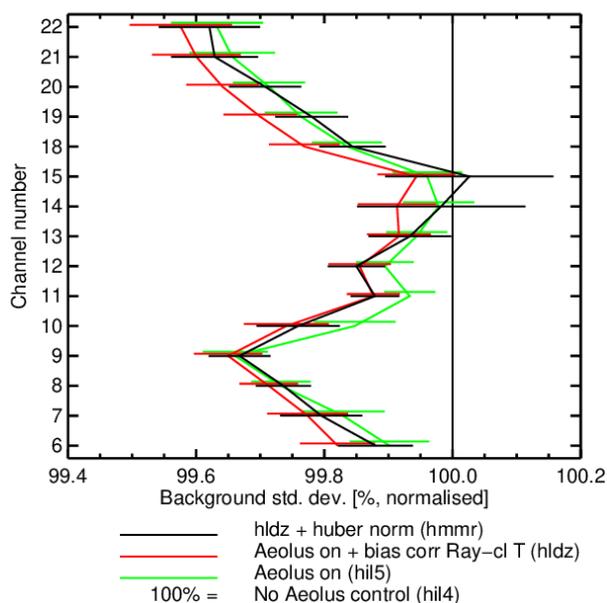


Figure 114. The change in standard deviation of O-B departures of ATMS resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The green-line is the nominal assimilation, the red-line is with the atmospheric temperature-dependent Rayleigh-clear bias correction, and the black-line is with an additional relaxation of the variational QC using a Huber norm and relaxed first-guess check. For period from 29 June 2019 to 19 October 2019.

#### 4.5.4 Medium-range forecast impact

Medium-range forecast verification with respect to operational ECMWF analyses for the nominal Aeolus “on” OSE for the first reprocessed dataset is shown for vector wind and temperature RMS error in Figure 115 and Figure 116 respectively. The impact is positive in the tropics (troposphere and stratosphere) and in the SH extratropical troposphere up to four days. However, there is a tendency towards more negative impact (not statistically significant) in the NH extratropics at longer forecast ranges (e.g. day five). The medium-range tropical lower stratospheric positive impact remains statistically significant throughout the forecast range (up to 9 days shown in this plot). This positive medium-range tropical impact is stronger than in the NRT early FM-B period OSE equivalent plot in Figure 90. Note that the first reprocessed dataset starts on 29 June compared to 2 August for the NRT data OSE and so had a longer period with higher atmospheric path signal and lower random errors, which could partially explain the differences.

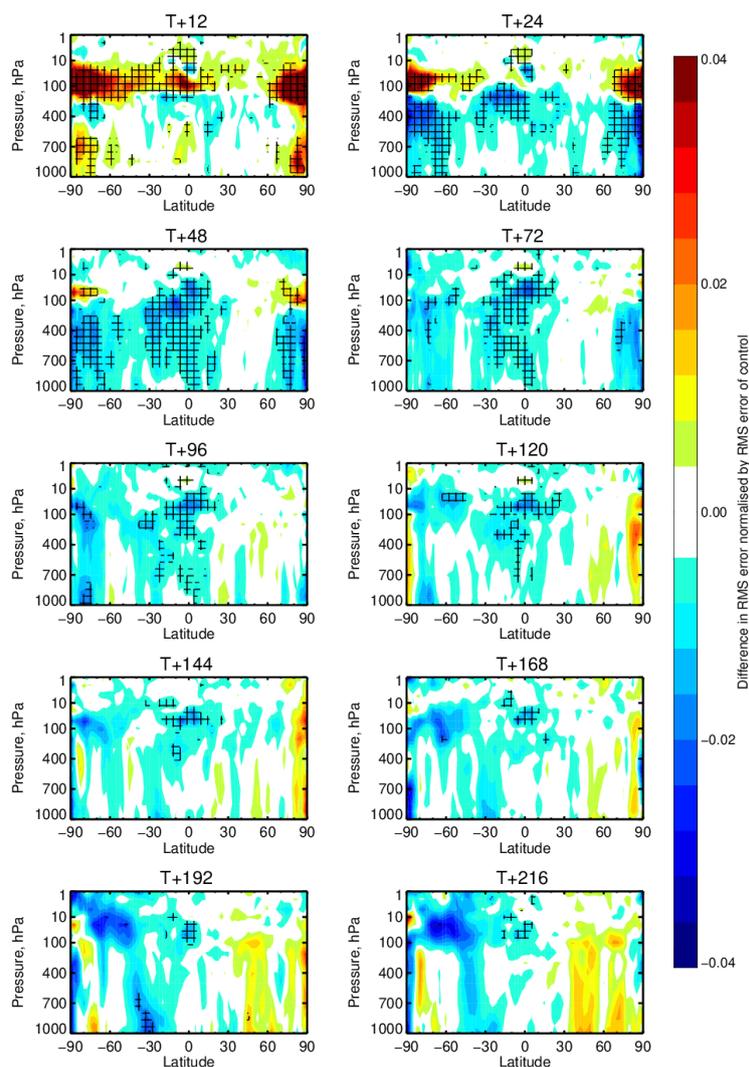


Figure 115. Zonal average normalised change in the RMS (root mean square) error of wind vector from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 29 June 2019 to 31 December 2019. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence.

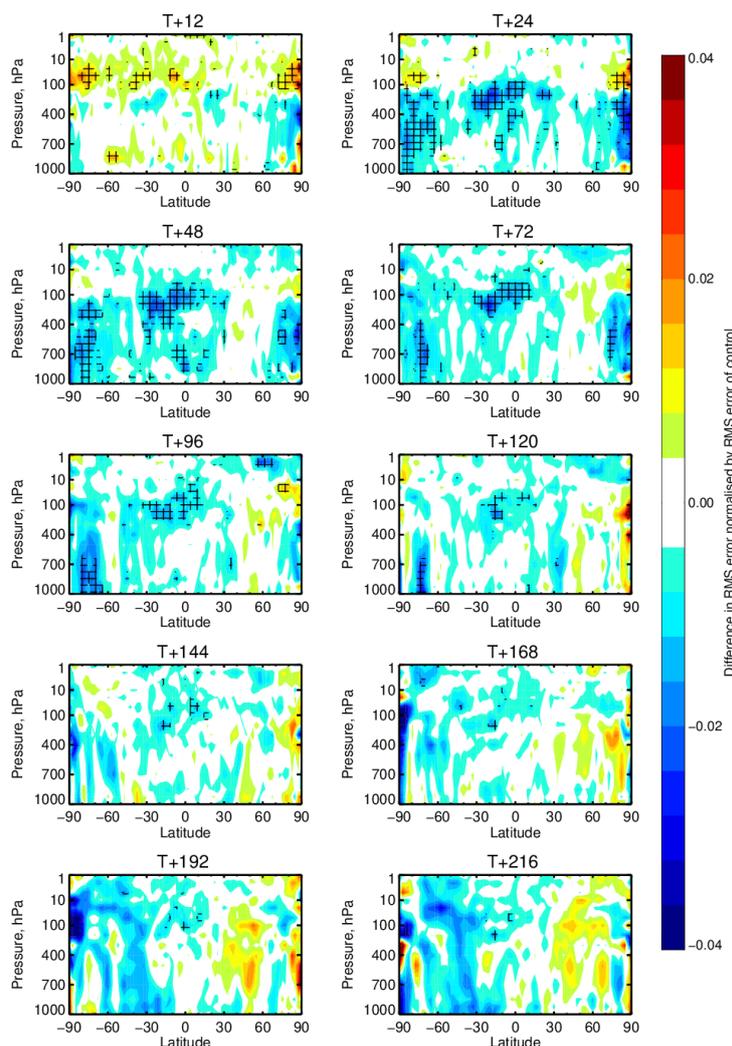


Figure 116. Zonal average normalised change in the RMS (root mean square) error of temperature from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 29 June 2019 to 31 December 2019. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence.

An experiment with reduced the weight on Mie-cloudy winds was run to see if the NH medium range tendency to negative impact is at all mitigated. The negative impact may have arisen from large changes made to the background towards the N. Pole (T+12) which may be due to overweighting the large sample of Mie winds near the poles. One hypothesis is that known wind-speed-dependent Mie-cloudy biases<sup>33</sup> and perhaps also the height-dependent Rayleigh biases could cause this. However, the impact was not improved by a larger representativeness error of 3 m/s compared to 2 m/s; see Figure 117 and so the experiment was stopped early, 3 October 2019, due to the lack of positive impact.

<sup>33</sup> A method to improve the L2B Mie-cloudy wind-speed dependent biases has been developed as part of the DISC and was operationally implemented on 1 July 2021 (see Marseille et al. 2022). It uses NWP HLOS winds as a reference to determine to atmospheric path Mie non-linearity correction more accurately i.e. an improvement in calibration look-up tables.

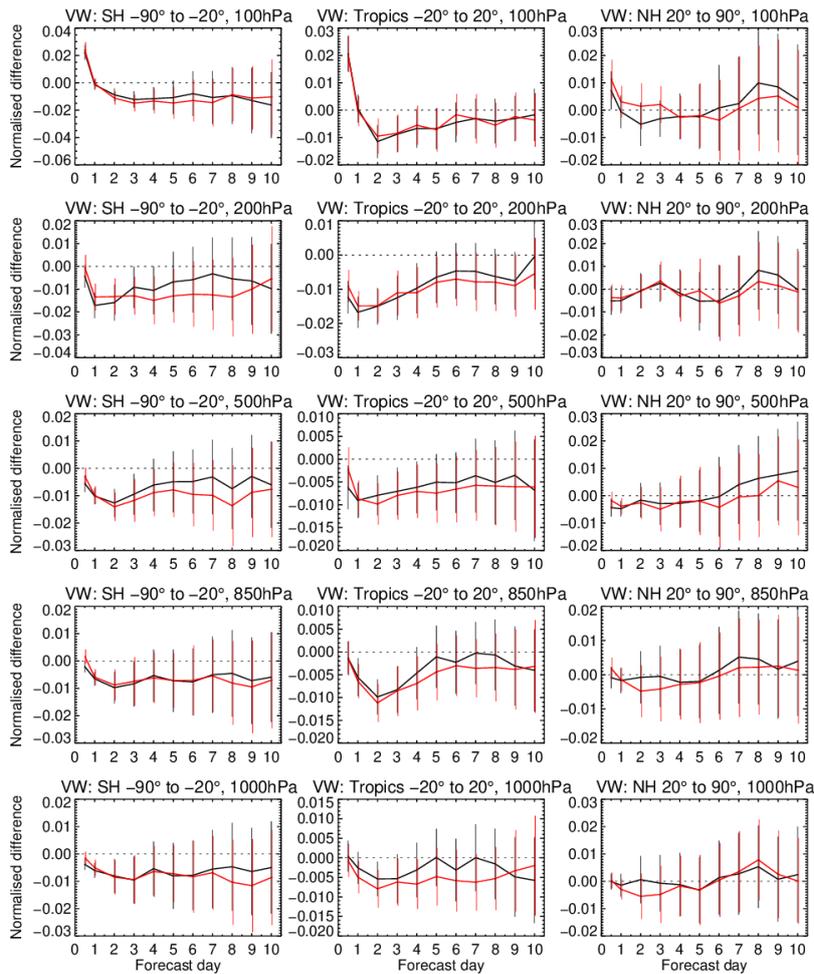


Figure 117. Normalised change in the RMS (root mean square) error of vector wind from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 29 June 2019 to 3 October 2019. Negative values indicate a reduction in error from assimilating Aeolus and positive values an increase in error. Verified against operational analyses. The red-line is with the 2 m/s representativeness error for the Mie-cloudy and the black line is with the 3 m/s.

The experiment with the atmospheric-temperature-dependent-bias correction of the Rayleigh-clear winds is compared to that without the bias correction in Figure 118. The impact of the bias correction seems somewhat neutral, if not slightly negative, on average. This points to the likely scenario that the bias correction is correcting some of the bias, but not completely. It is believed that correcting the bias at source in the Rayleigh calibration file (AUX\_RBC\_L2), i.e. by finding a more correct calibration Rayleigh response versus frequency curves for different temperatures and pressures, is the best way to resolve this. Research is ongoing in the DISC to that end.

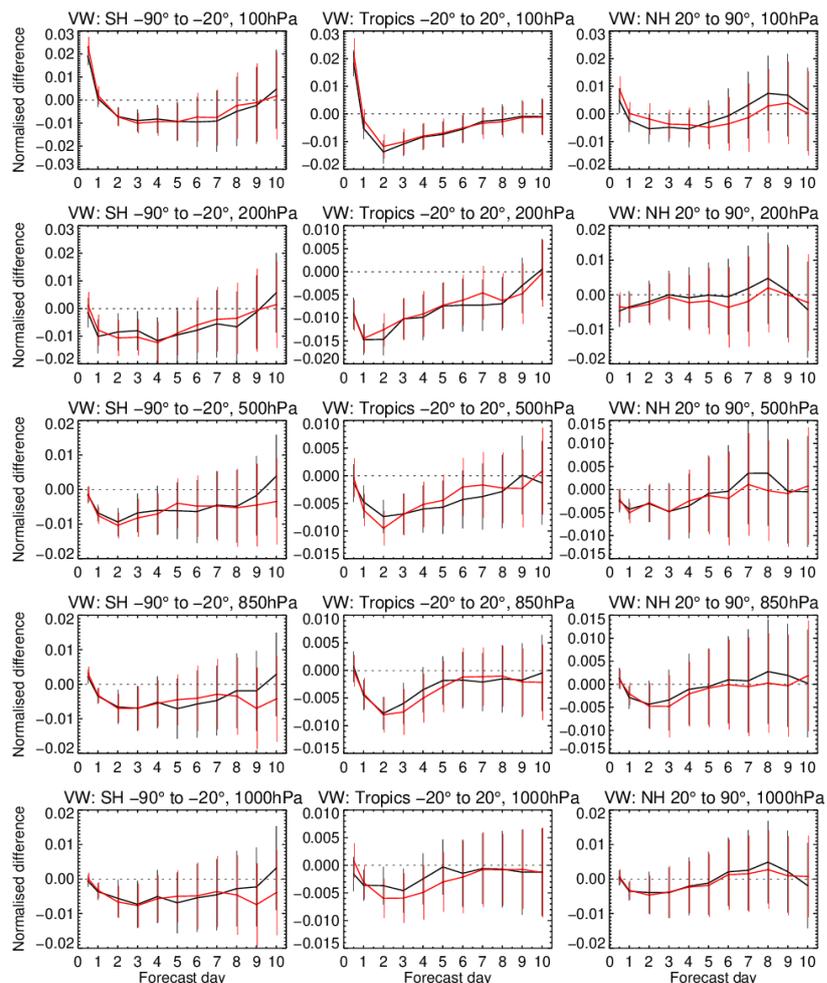


Figure 118. Normalised change in the RMS (root mean square) error of vector wind from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 29 June 2019 to 31 December 2019. Negative values indicate a reduction in error from assimilating Aeolus and positive values an increase in error. Verified against operational analyses. The red-line is the experiment without, and the black-line is with the atmospheric-temperature-dependent bias correction for the Rayleigh-clear winds.

The impact of the Huber norm VarQC and relaxed first-guess check was overall negative according to the results of Figure 119, and so the experiment was stopped early (19 October 2019).

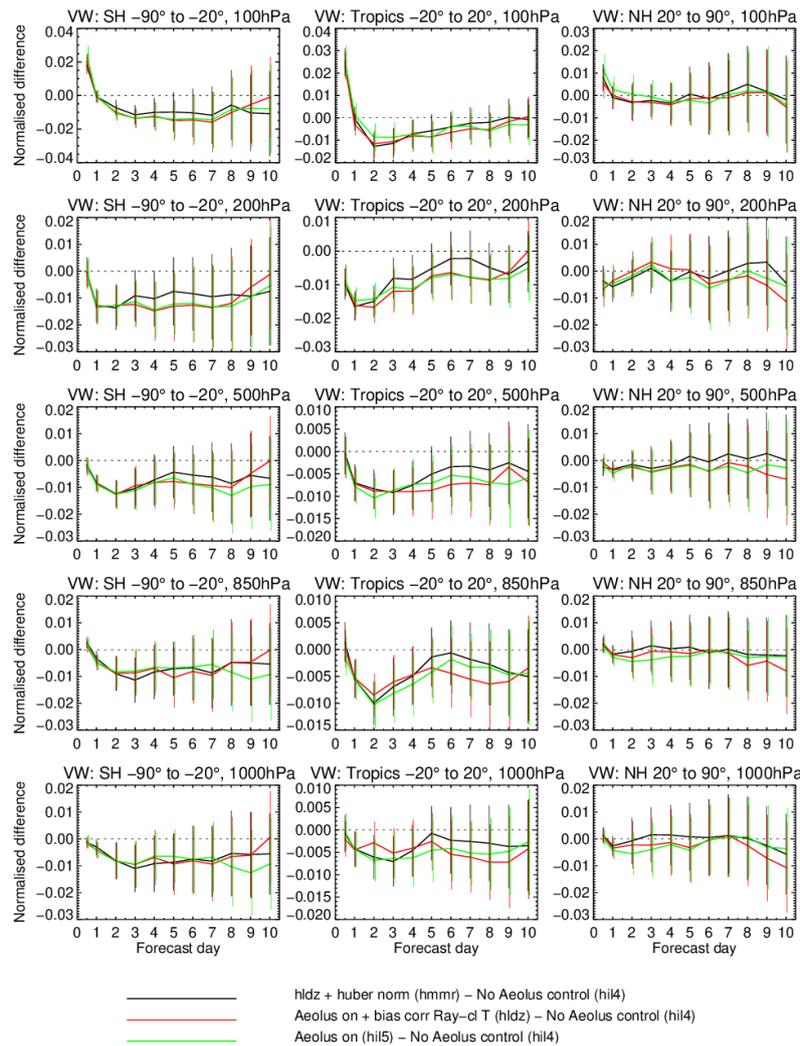


Figure 119. Normalised change in the RMS (root mean square) error of vector wind from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 29 June 2019 to 19 October 2019. Negative values indicate a reduction in error from assimilating Aeolus and positive values an increase in error. Verified against operational analyses. The green-line is the experiment with the nominal Aeolus set-up; the red-line is with the atmospheric-temperature-dependent bias correction for the Rayleigh-clear winds; and the black-line is with an additional modification of VarQC to a Huber norm with a relaxed first-guess check.

### 4.5.5 Tropical cyclone impact

For this first reprocessed dataset Aeolus has a statistically significant positive impact on tropical cyclone central pressure mean errors at the 4 and 5 day forecast range by 0.5-1.5 hPa (200 and 136 samples). Other metrics are neutral i.e. the confidence intervals cross the zero line. However, it is not far from being a positive result on position errors e.g. at day 2 by 3% — perhaps needs a longer OSE period (e.g. 2<sup>nd</sup> reprocessed dataset). This topic was investigated in a lot more depth by the parallel ESA funded Aeolus Extreme Weather NWP impact study (see De Chiara et al. 2023).

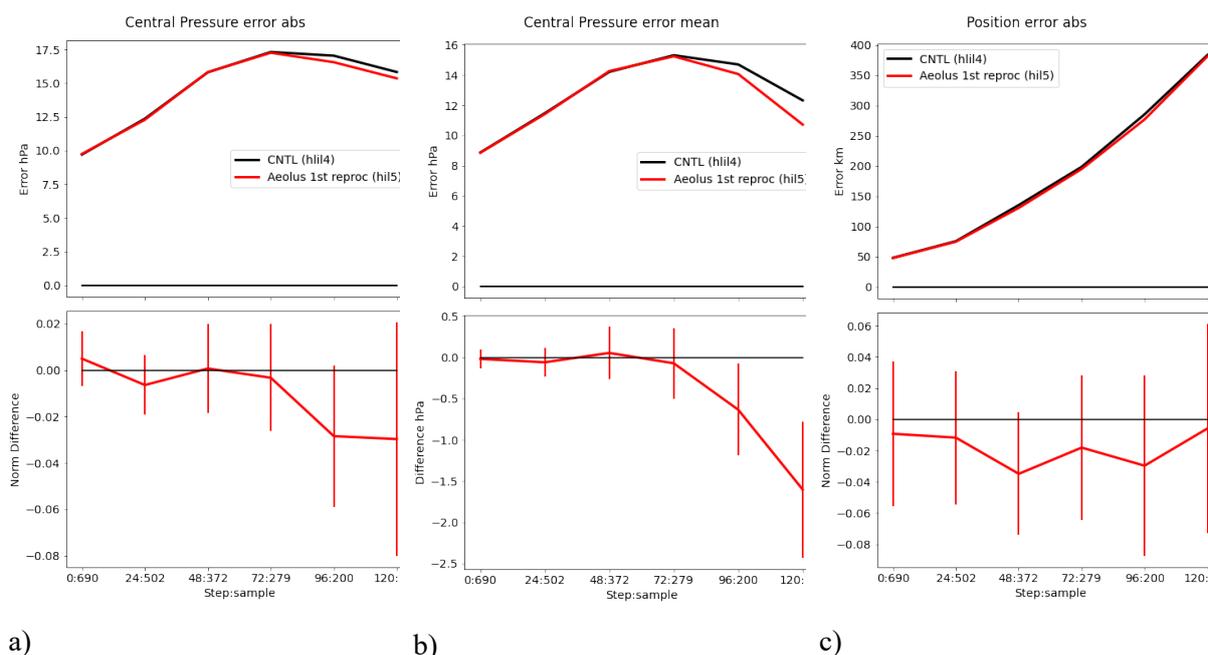


Figure 120. The impact on tropical cyclones from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 29 June 2019 to 31 December 2019. On the upper plots, black lines are the control without Aeolus and red lines are with Aeolus. Negative values on the lower plots indicate a reduction in error from assimilating Aeolus. Verified against satellite observation derived best estimates of TC position and central pressure (which are not perfect). a) central pressure absolute error (hPa) and normalised difference b) central pressure mean error (hPa) and difference (hPa) c) absolute position error (km) and normalised difference. Plots produced by a tool of Linus Magnusson (ECMWF).

## 4.6 Second reprocessed dataset (FM-B) OSEs

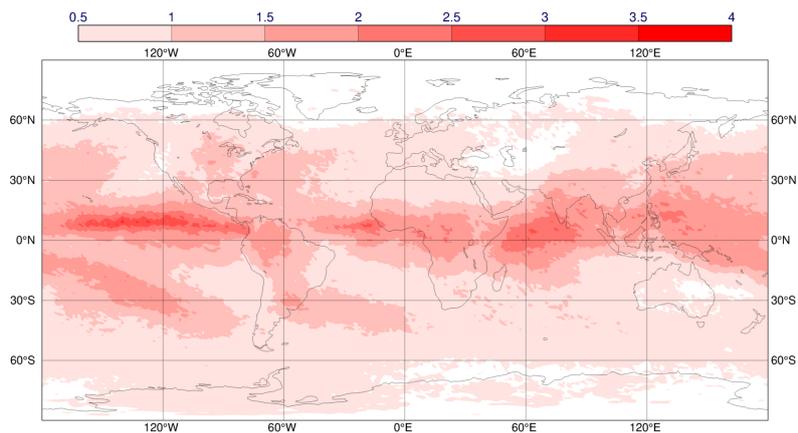
This section provides the verification results from OSEs using the second reprocessed L2B dataset (2B11), covering the period of the FM-B laser until October 2020, as discussed in sections 2.1.6, 2.4.5 and 0.

The main OSE for this dataset is an experiment with Aeolus “on” (Rayleigh-clear plus Mie-cloudy, with nominal operational settings), and a control without Aeolus. It is the longest OSE that has been run for Aeolus. Also, for a fair comparison of the NWP impact from the first and second reprocessing, an experiment with the same higher model resolution setup used for the second reprocessing OSE (T<sub>CO639</sub> outer loop) is done for the first reprocessing.

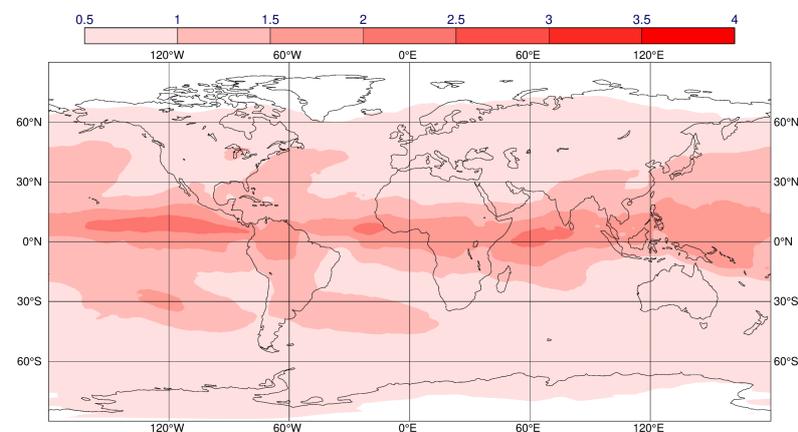
### 4.6.1 Analysis changes due to assimilating Aeolus HLOS winds

The standard deviation of the differences between the experiment and control analyses, e.g. Figure 121 a) u-wind at 200 hPa, shows very similar spatial patterns to those observed in previous OSEs, e.g. see Figure 92. It was noticed that the spatial pattern of increments is similar to the climatological map of mesoscale convective systems as shown in Figure 10 of Feng et al. (2021); thus, supporting the hypothesis of Aeolus correcting the larger background errors for wind in convective areas. Also, the spatial pattern is very similar to the ECMWF operational EDA spread for u-wind at 200 hPa the same period, Figure 121 b). The operational EDA is an input to the background error covariances used in the OSE; which have a significant flow dependence. Therefore it is not surprising that Aeolus is changing the background most in areas where the B matrix has largest wind errors. Note this result is not based on the whole OSE period, but up to 30 December 2019.

Figure 122 shows the standard deviation of differences between ECMWF operational u-wind analyses (before Aeolus was assimilated) and the equivalents from other NWP centres (Met Office and NCEP) – for the same period, but at 250 rather than 200 hPa (due to a 200 hPa being unavailable on MARS from other NWP centres). The variance between the ECMWF and other models has peaks in tropical ocean areas and convergence zones in similar areas to where Aeolus is changing the ECMWF analysis and ECMWF EDA spread is highest. This supports the hypothesis that NWP forecasts of upper tropospheric wind are fundamentally uncertain in these areas.

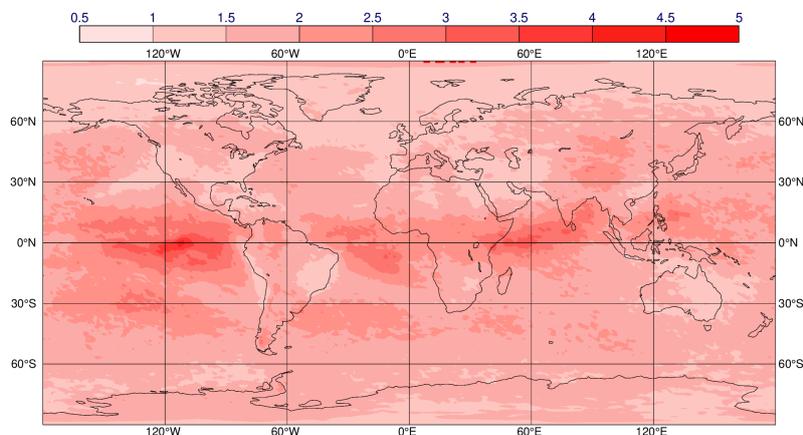


a)

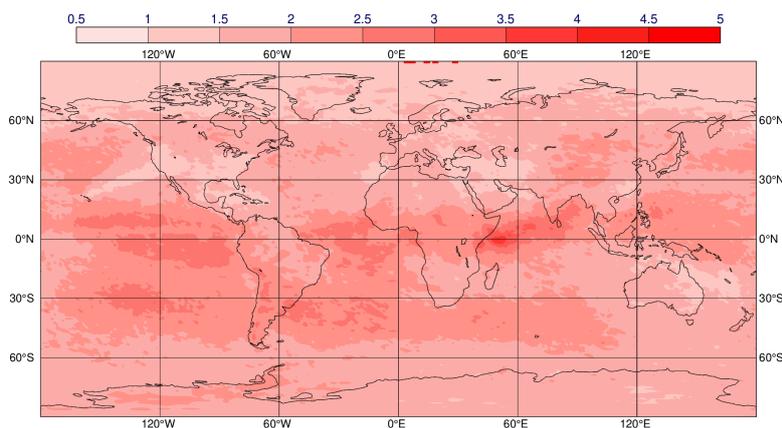


b)

Figure 121. a) Standard deviation of the differences of  $u$ -wind component (m/s) at 200 hPa between the analysis assimilating Aeolus (Rayleigh-clear and Mie-cloudy) and the control not assimilating Aeolus HLOS winds for the period 29 June to 30 December 2019. b) Mean of the EDA spread (a proxy for background forecast error) for  $u$ -wind component (m/s) for model level 75 (~200 hPa) from ECMWF operational EDA for the period 29 June to 30 December 2019.



a)



b)

Figure 122. Standard deviation of analysis differences of 250 hPa u-wind component (m/s) between the operational high-resolution ECMWF analysis and other global NWP model analyses: a) Met Office b) NCEP GFS, for the period 29 June to 30 December 2019. Note the scale goes up to 5 m/s compared in 4 m/s in Figure 121.

As in previous OSEs, some notable changes (of up to 1 m/s) in the mean difference of the analysis zonal wind component are evident due to Aeolus data in the tropical UTLS; see Figure 123.

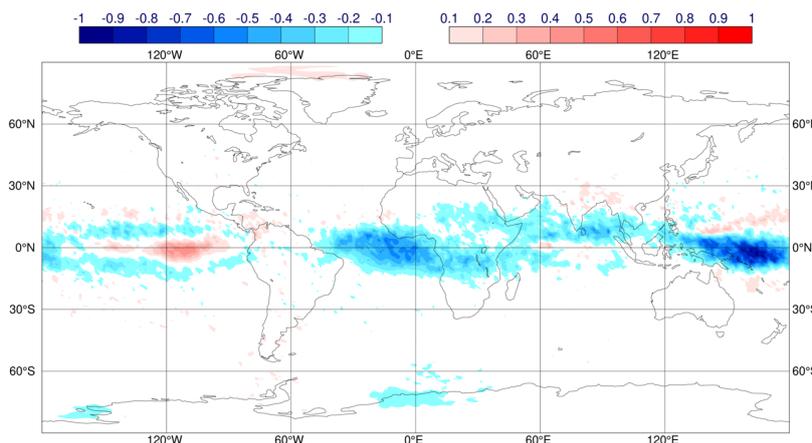


Figure 123. Mean of the differences of u-wind component (m/s) at 100 hPa between the analysis using Aeolus (Rayleigh-clear and Mie-cloudy) and the control not using Aeolus HLOS winds for the period 29 June to 30 December 2019. Colour-scale range is  $\pm 1$  m/s.

Figure 124 shows mean analysis differences between ECMWF operations (before Aeolus was assimilated) and other NWP centres (Met Office and NCEP) for 100 hPa u-wind. The largest systematic differences (often over 1 m/s) occur in the tropics and sub-tropics; however, differences are also rather large for high latitudes with respect to the Met Office. The ECMWF model is in closer agreement with the NCEP model than the Met Office in terms of 100 hPa winds. Systematic differences between the analyses suggest NWP winds are not well constrained at 100 hPa, which suggests Aeolus is trying to correct some systematic errors in the ECMWF background. Zonal wind biases are known to be present in the ECMWF model around the tropical lower stratosphere due to too much vertical diffusion (Polichtchouk et al., 2021), leading to too little vertical wind shear. This may partly explain the biases.

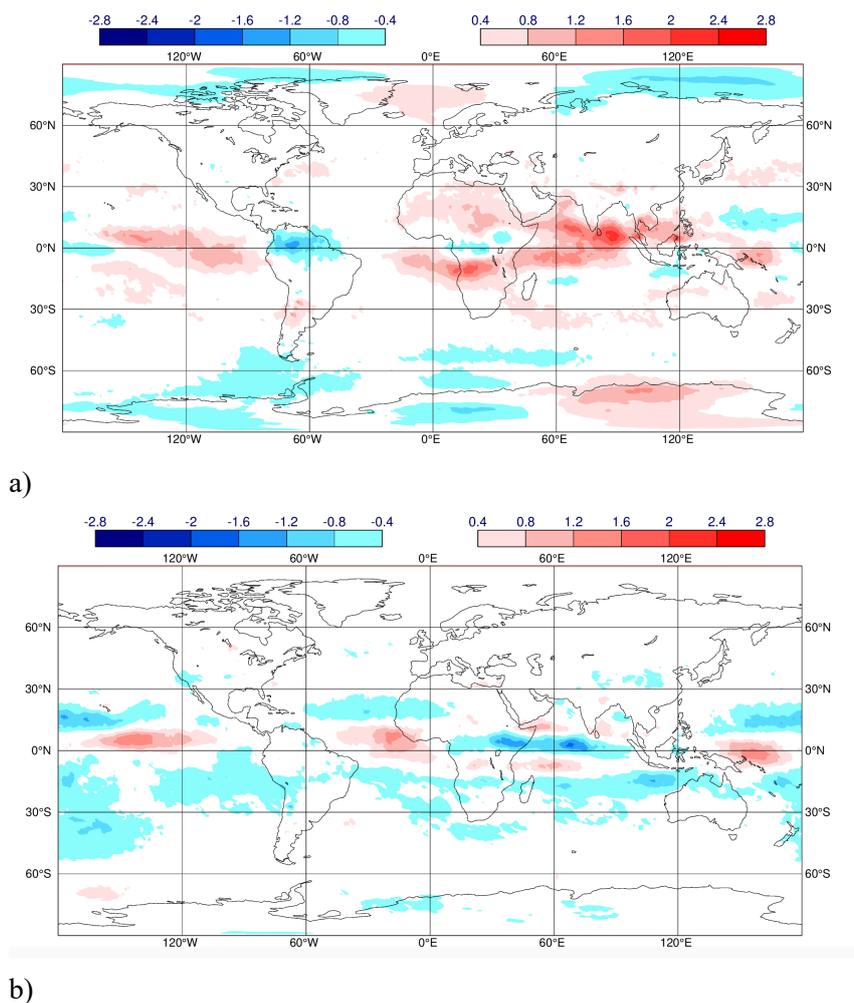


Figure 124. Mean of analysis differences of 100 hPa u-wind component (m/s) between the operational high-resolution ECMWF analysis and other global NWP model analyses: a) Met Office b) NCEP GFS, for the period 29 June to 30 December 2019. Colour-scale range is  $\pm 2.8$  m/s. The differences have the sign: ECMWF minus NWP centre.

#### 4.6.2 Short-range forecast impact

The verification of the 2<sup>nd</sup> reprocessed dataset OSE with conventional wind vectors and radiosonde temperature observations (Figure 125) shows a similar pattern of impact to the 1<sup>st</sup> reprocessed dataset OSE (Figure 111), except there is greater statistical significance to the NH extratropics positive impact, which is probably due to the larger sample.

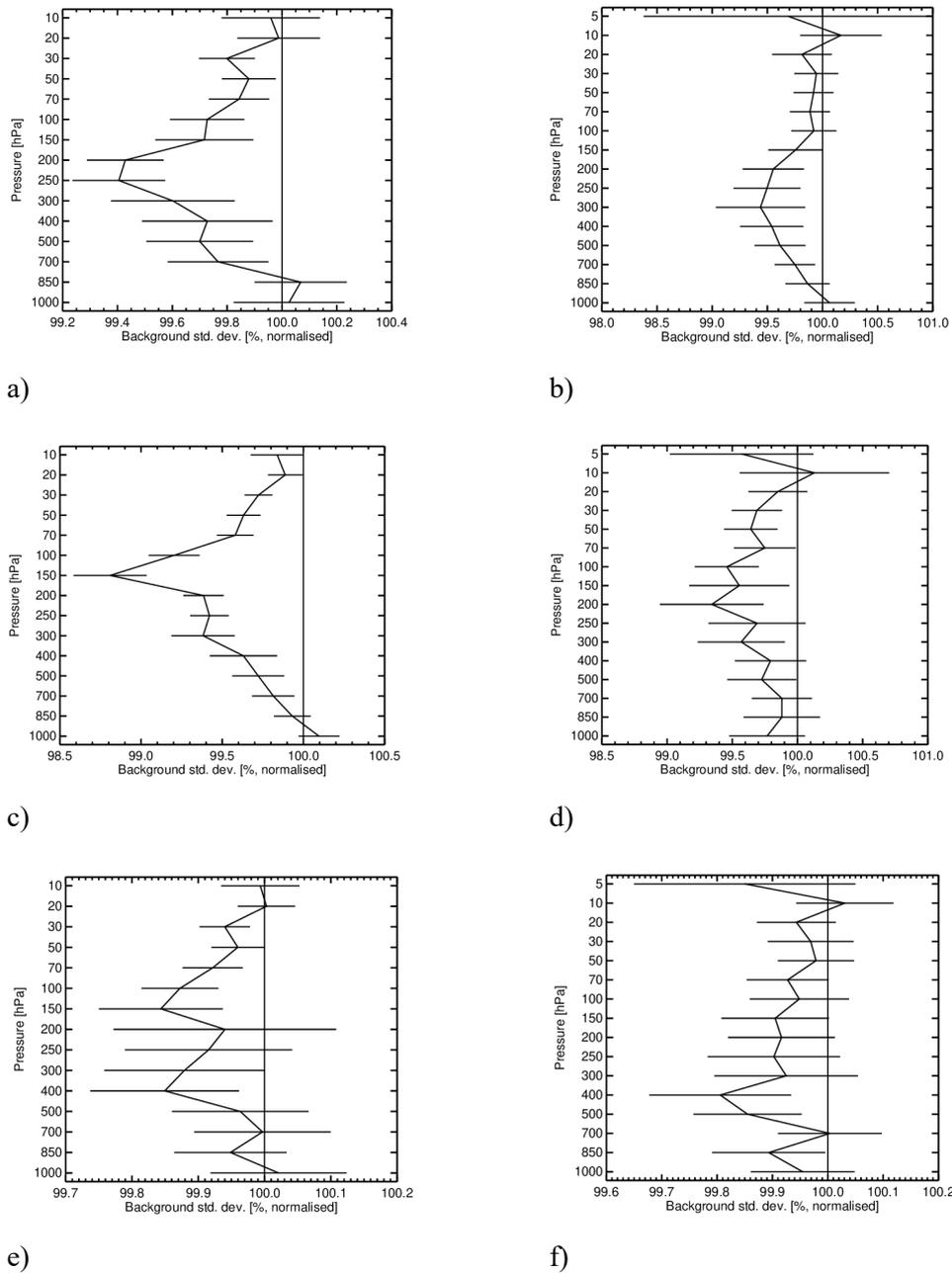
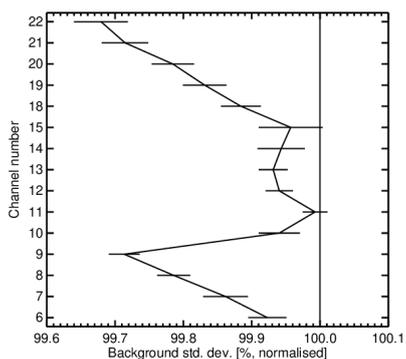


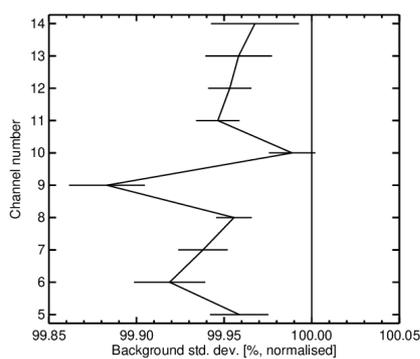
Figure 125. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) conventional data vector wind in the SH extratropics, b) radiosonde temperatures in the SH extratropics, c) conventional data vector wind in the tropics, d) radiosonde temperatures in the tropics, e) conventional data vector wind in the NH extratropics and f) radiosonde temperatures in the NH extratropics. Horizontal bars show the 95% confidence range (Student's t-test). For the period from 29 June 2019 to 9 October 2020.

The verification against other satellite observation types (Figure 126) also looks to be similar to that shown for the 1<sup>st</sup> reprocessed dataset OSE (Figure 112), apart from slightly larger impact in some cases e.g. with respect to the ATMS radiances (channel 9). The slight degradation with respect to AMVs in the 400-700 hPa range is evident again. This degradation is strongest in the SH extratropics (not shown);

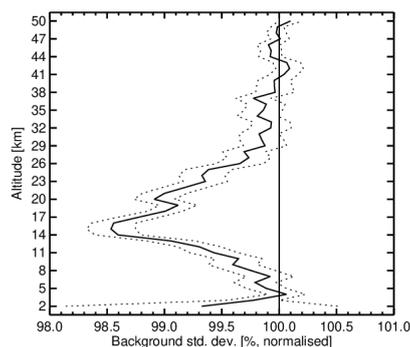
in the tropics the impact is positive at the same pressure range. Also, there is a slight degradation against altimeter wave-heights which is not understood. There is strong positive impact for CrIS wavenumber  $\sim 1500\text{ cm}^{-1}$ , which is sensitive to water vapour at 300 hPa, and wavenumber  $700\text{ cm}^{-1}$  which is sensitive to temperature at 100 hPa. There is a slight degradation with respect to near surface temperatures for CrIS wavenumber  $\sim 757\text{ cm}^{-1}$ . Otherwise, the impact is generally a consistent positive one. There is a strong positive impact on buoy surface pressures in the SH extatropics (not shown), so Aeolus is providing positive impact is very remote areas.



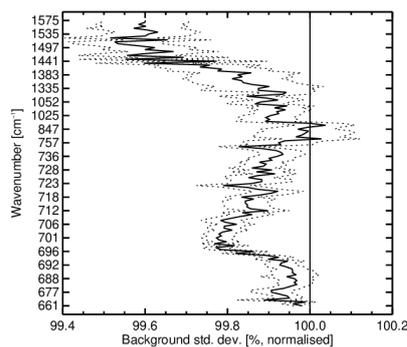
a)



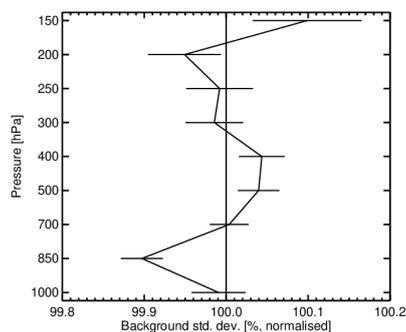
b)



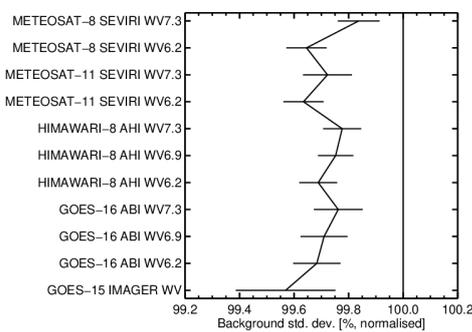
c)



d)



e)



f)

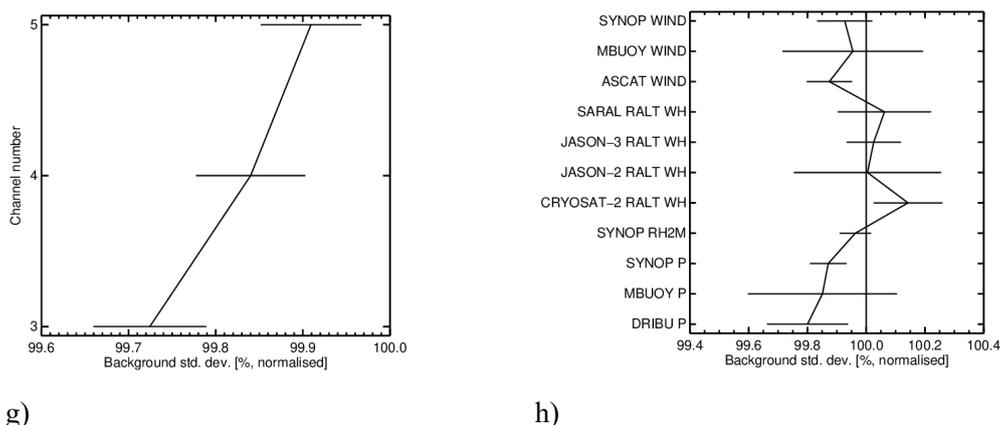


Figure 126. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) ATMS global statistics, b) AMSU-A global statistics, c) GNSS radio occultation in the tropics, d) CrIS global statistics, e) AMVs global statistics, f) geostationary imagery global statistics, g) FY-3B MWHS global statistics and h) various surface observation types in the NH. Horizontal bars show the 95% confidence range (Student’s t-test). For period from 29 June 2019 to 9 October 2020.

#### 4.6.3 Medium-range forecast impact

The forecast verification into the medium-range with operational analysis as the reference (Figure 127, Figure 128) demonstrates the most positive impact found with Aeolus at ECMWF. The usual patterns of stronger impact in the  $\pm 30^\circ$  latitude band and poleward of  $60^\circ$  degrees is evident. The main difference compared to previous OSEs is that the positive impact in the tropics and lower stratosphere is clearer at day nine forecasts, when compared to e.g. Figure 115; the NH extratropics negative impact is resolved. This could be due to the longer period (so greater statistical significance), the higher model resolution of this OSE compared to previous ones and the improved data quality in the 2<sup>nd</sup> reprocessing compared to the 1<sup>st</sup>. It was noticed that if the OSE verification end date is set earlier e.g. February 2020, then the magnitude of positive impact is larger. This may be due to the signal loss with time and hence noisier winds (particularly Rayleigh-clear) and decreasing Mie-cloudy winds by spring 2020, as shown in section 3.5.

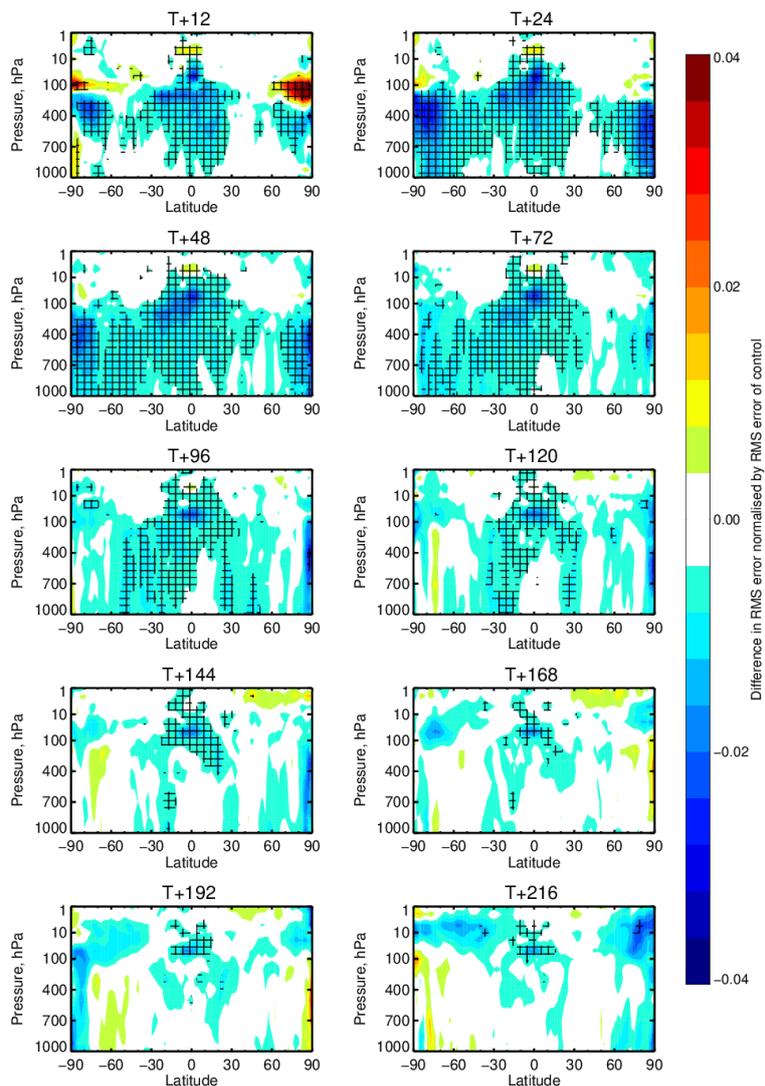


Figure 127. Zonal average normalised change in the RMS (root mean square) error of wind vector from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 29 June 2019 to 9 October 2020. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence.

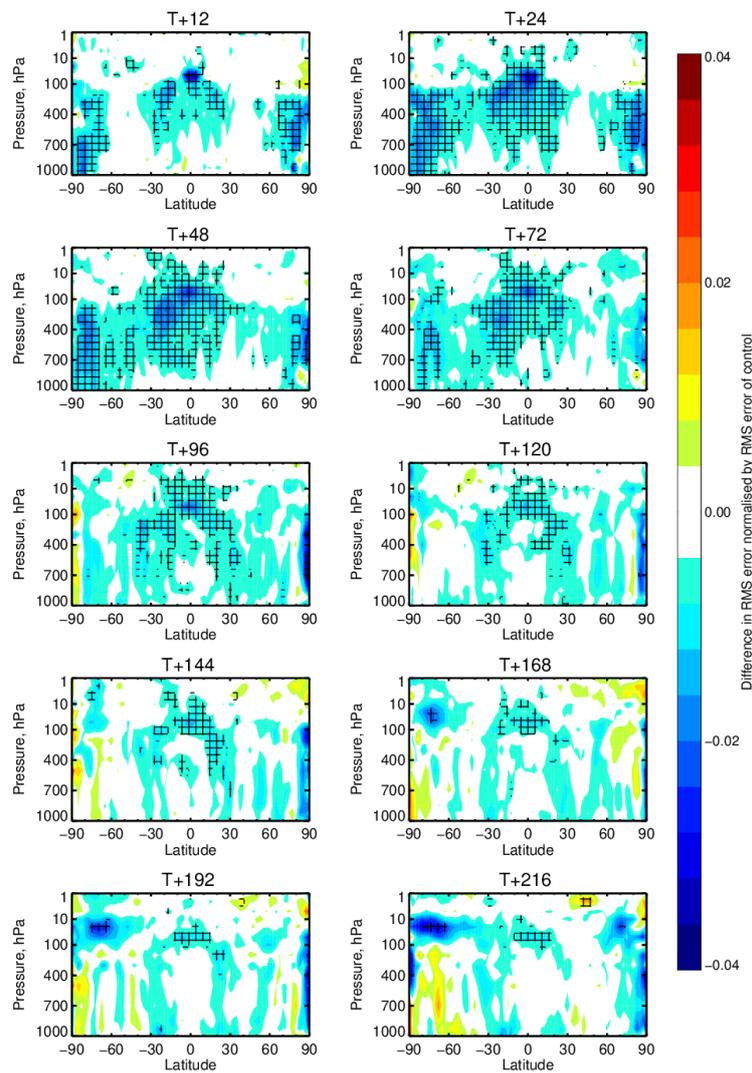


Figure 128. Zonal average normalised change in the RMS (root mean square) error of temperature from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 29 June 2019 to 9 October 2020. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence.

Maps of Aeolus impact at 100 hPa in Figure 129, shows that the largest impact for this upper-troposphere lower-stratosphere height is in the tropical east Pacific – and is particularly strong at two day forecast range with greater than 5% improvement in RMSE. This area has the largest u-wind errors according to EDA spread and comparisons between different NWP centre analyses.

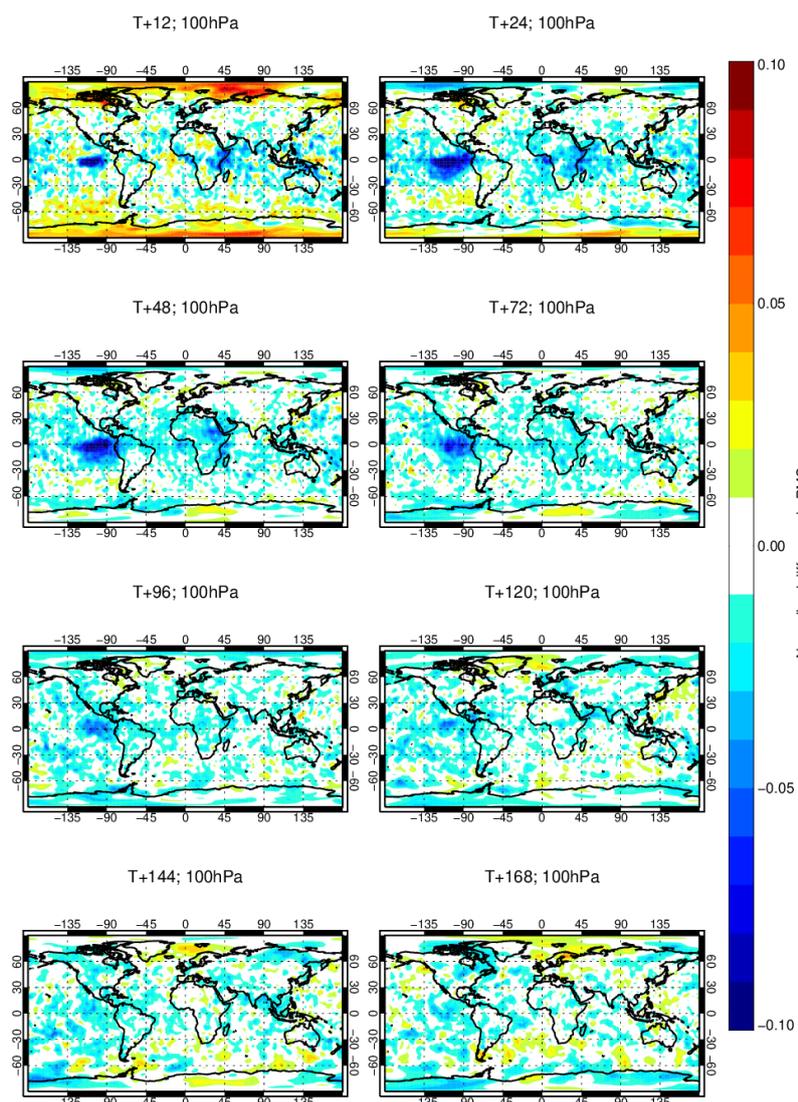


Figure 129. Maps of normalised change in the RMS (root mean square) error of wind vector at 100 hPa from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 29 June 2019 to 9 October 2020. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error.

To directly assess which of the potential factors may have contributed more to the apparent improvement in impact of the 2<sup>nd</sup> reprocessing, the 1<sup>st</sup> reprocessed dataset was reassessed in an OSE with the same higher model resolution (T<sub>CO</sub>639) as used in the 2<sup>nd</sup> reprocessing. The verification of the two reprocessed datasets is compared in Figure 130. The impact of both reprocessed datasets is reasonably similar; however, the 2<sup>nd</sup> reprocessing has a larger positive impact in the tropics and in the NH extratropics near the surface. This may be partly explained by there being more data assimilated between 850 and 150 hPa for the 2<sup>nd</sup> reprocessing (11% more data at 850 hPa); see Figure 131, however the extra data must be relatively noisy since the stdev(O-B) is increased.

Tropical cyclone impact for the 2<sup>nd</sup> reprocessing was covered in great depth in the separate ESA contract on extreme weather; see the report: De Chiara et al. 2023.

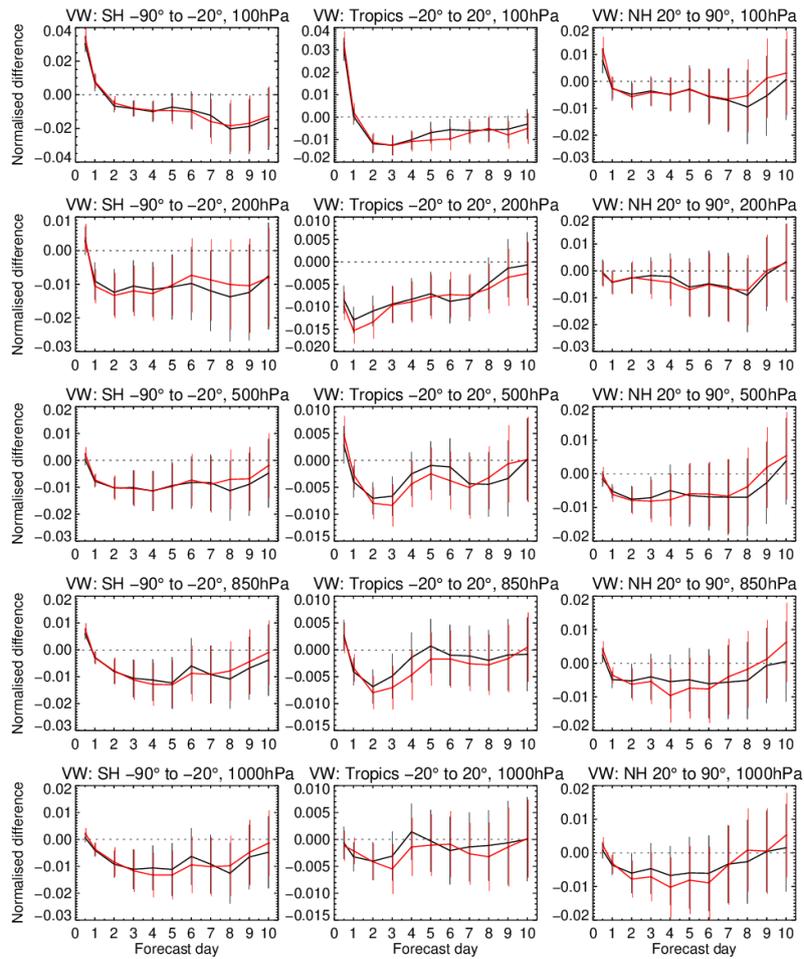


Figure 130. Normalised change in the RMS (root mean square) error of vector wind from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 29 June 2019 to 31 December 2019. Negative values indicate a reduction in error from assimilating Aeolus and positive values an increase in error. Verified against operational analyses. The red-line is the experiment with the 2<sup>nd</sup> reprocessed dataset, and the black-line is with the 1<sup>st</sup> reprocessed dataset. Both OSEs were run with a 4D-Var outer loop resolution of Tco639.

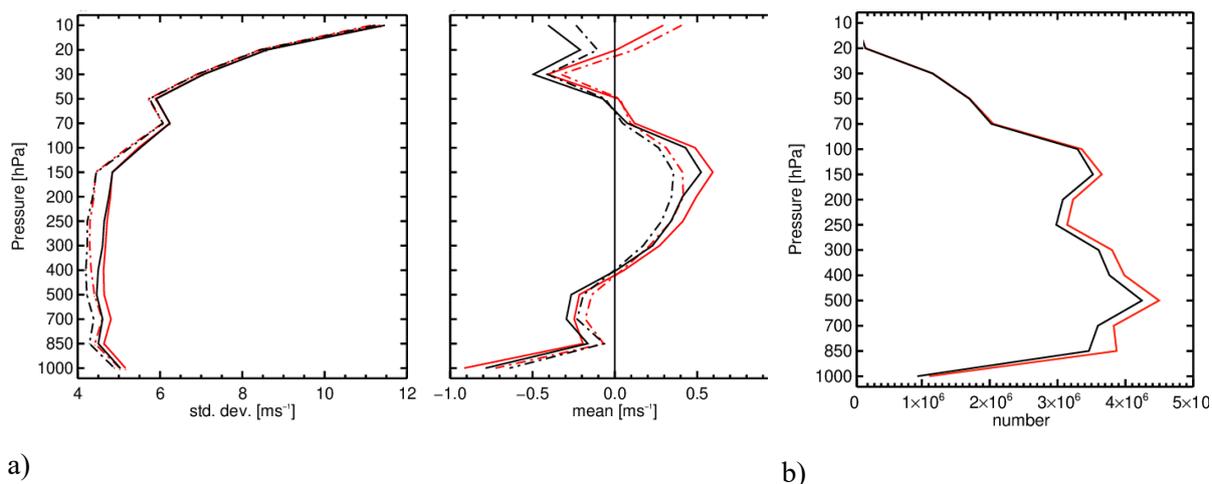


Figure 131. Combined Rayleigh-clear and Mie-cloudy O-B statistics. a) Global standard deviation and mean O-B and b) the number of Aeolus HLOS wind observations assimilated, as function of pressure, for the period 29 June to 31 December 2019. The red-line is the experiment with the 2<sup>nd</sup> reprocessed dataset, and the black-line is with the 1<sup>st</sup> reprocessed dataset.

#### 4.7 Late 2021/2022 NRT FM-B dataset OSE

This section provides the verification results from an OSE using the NRT PDGS produced L2B dataset, covering from Dec 13 2021 until 28 September 2022, as discussed in sections 2.1.7, 2.1.7 and 0.

The OSE for this dataset is an experiment with Aeolus “on” (Rayleigh-clear plus Mie-cloudy) and a control without Aeolus, using the nominal operational set-up for observations. A high model resolution for a research experiment is used (T<sub>CO639</sub> outer loop) since it is closer to the operational resolution and the impression was that impact seems to be greater from Aeolus winds when using higher resolution, as discussed in previous section.

##### 4.7.1 Short-range forecast impact

Aeolus’ impact on short-range forecasts with respect to conventional winds for this period is rather small in the SH extratropics and tropics compared to the 2<sup>nd</sup> reprocessed dataset OSE (even some degradation in tropical temperature at 250 hPa, despite the wind being improved). This must be due to the very large noise of Rayleigh-clear winds associated with the atmospheric path signal decrease. The global Rayleigh-clear stdev(O-B) statistic was 55% larger in July 2022 (NRT dataset) compared to July 2019 (2<sup>nd</sup> reprocessed dataset). According to the formulation of appendix 9.3, this increase in noise should reduce the NWP impact of the Rayleigh-clear by a factor 0.45. However, the relatively small NH extratropics looks similar in magnitude to what has been seen in previous OSEs; perhaps the Mie-cloudy winds (for which the random error has not increased, but coverage has reduced) is maintaining that impact.

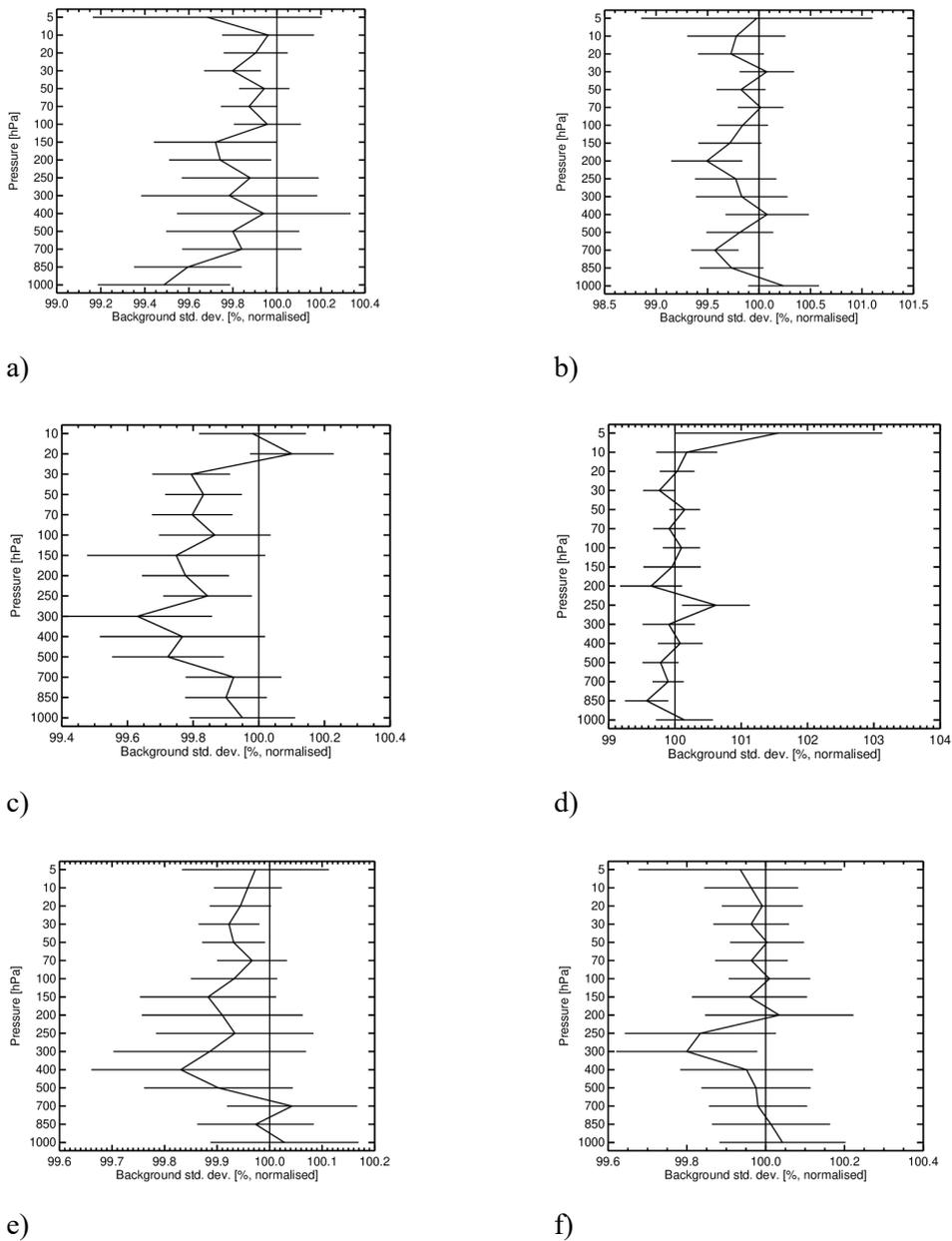
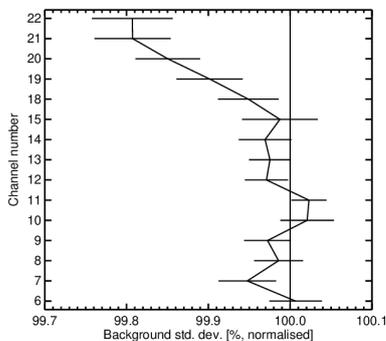


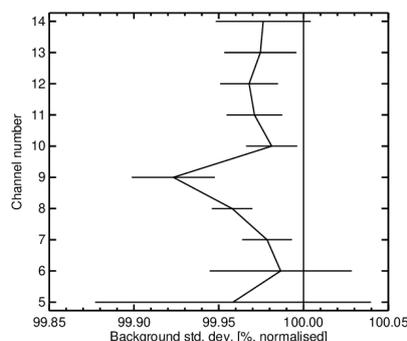
Figure 132. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) conventional data vector wind in the SH extratropics, b) radiosonde temperatures in the SH extratropics, c) conventional data vector wind in the tropics, d) radiosonde temperatures in the tropics, e) conventional data vector wind in the NH extratropics and f) radiosonde temperatures in the NH extratropics. Horizontal bars show the 95% confidence range (Student’s t-test). For the period from 13 December 2021 to 28 September 2022.

Positive impact is still evident with short-range forecast verification with respect to other satellite observation types (Figure 133). However, it is rather small for temperatures sensitive observations, e.g. ATMS channels 6-9, GNSSRO and CrIS channels  $\sim 600-800\text{ cm}^{-1}$ , compared to the 2<sup>nd</sup> reprocessed dataset OSE. The positive impact for humidity sensitive observations, e.g. CrIS channels  $1028-1536\text{ cm}^{-1}$  and ATMS channels 18-22, is still reasonable, at about half what is used to be. The Rayleigh-clear

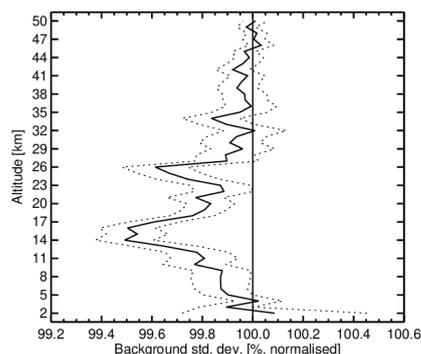
winds used to provide a large impact with respect to ATMS channel 9 (temperature sensitive, peaking at ~12 km); but this is no longer present. This suggests the Mie-cloudy winds are providing the bulk of the NWP impact in 2022. Interestingly the impact with respect to AMVs is positive in the 400-700 hPa range, whereas in previous OSEs the opposite was shown. No changes to AMV usage in this pressure range have been made. Perhaps the improved Mie-cloudy wind-speed dependent biases since 1 July 2021 have helped or the smaller influence of Rayleigh-clear on the analysis.



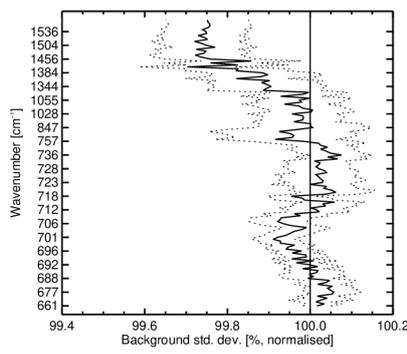
a)



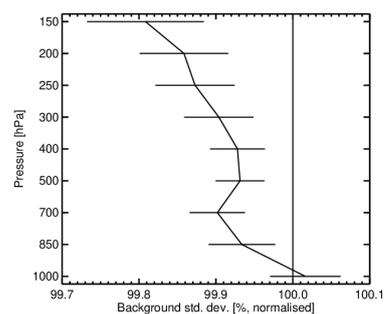
b)



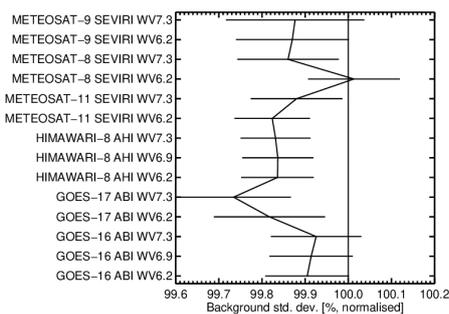
c)



d)



e)



f)

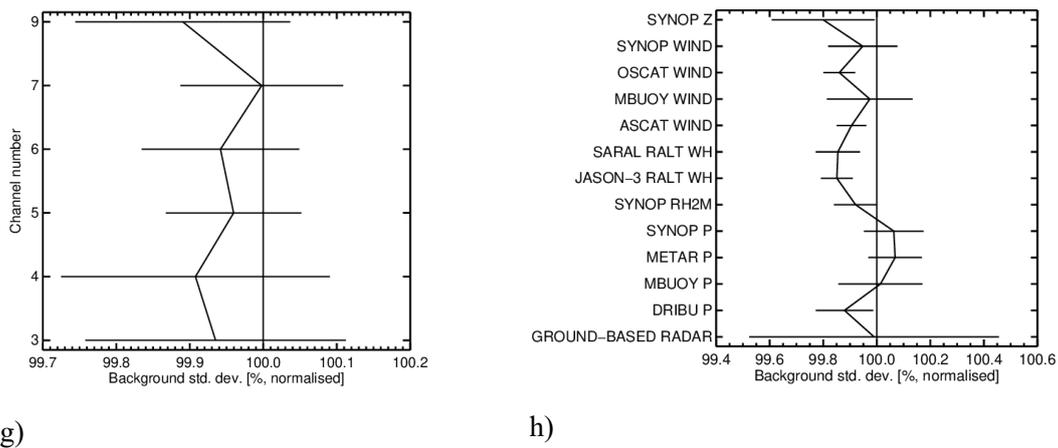


Figure 133. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) ATMS global statistics, b) AMSU-A (all-sky) global statistics, c) GNSS radio occultation in the tropics, d) CrIS global statistics, e) AMVs global statistics, f) geostationary imagery global statistics, g) FY-3D MWRI global statistics and h) various surface observation types global. Horizontal bars show the 95% confidence range (Student's *t*-test). For period from 13 December 2021 to 28 September 2022.

There is an unusual peak in positive impact with respect to tropical GNSSRO at ~26 km altitude which is assumed result from the assimilation of decent quality Mie-cloudy winds retrieved from the backscatter of the Hunga Tonga-Hunga Ha'apai volcanic eruption plume. The Aeolus RBS were raised to capture the plume in late January 2022. The plume was found to drift to the west with the easterly phase of the QBO and had circumnavigated the globe by the end of January 2022 (as could be seen from the distribution of the Mie winds). By Autumn 2022 (end of FM-B period) the aerosol backscatter of the plume was tenuous (scattering ratios < 1.2), and the Mie wind were becoming increasing noisy and so more were rejected by the QC leading to low data counts (see Figure 134). There was a resurgence in the data quality and data counts with the increased signal levels of FM-A since December 2022. The Mie winds of the plume were at altitudes of 23-28 km in late January 2022, but gradually descended to lower altitudes closer to 20 km. The peak in the data counts from the plume was in May 2022.

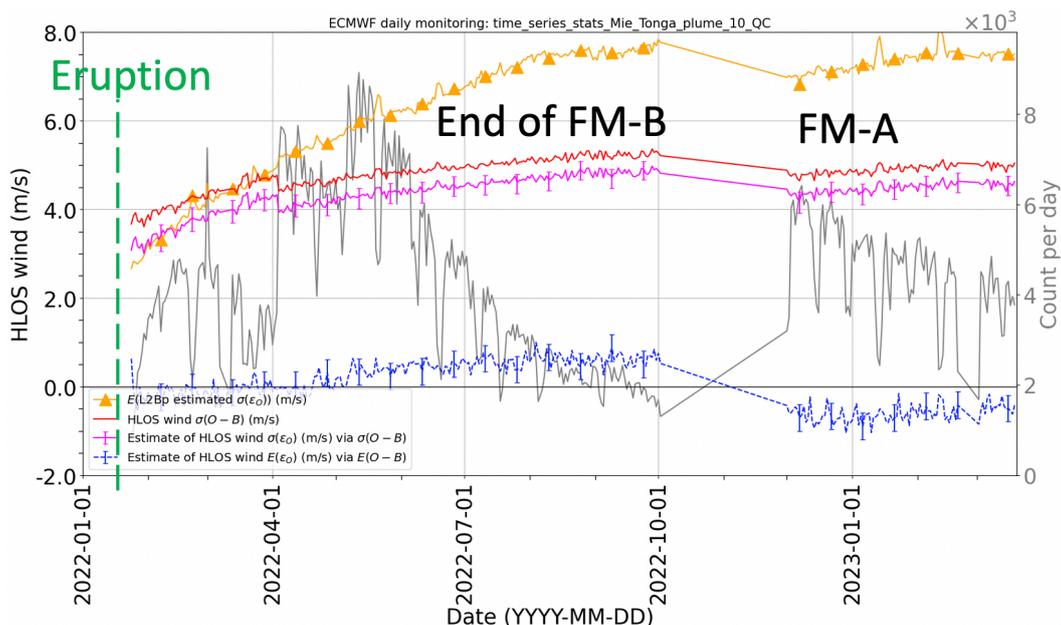


Figure 134. ECMWF monitoring of L2B Mie-cloudy winds from roughly the geographical area of the Hunga Tonga eruption plume i.e. latitude within  $\pm 40^\circ$  and pressures  $< 50$  hPa. QC rejects L2B winds flagged invalid and if  $\text{abs}(O-B) < 10$  m/s.

#### 4.7.2 Medium-range forecast impact

Figure 135 and Figure 136 shows that there is still some reasonable positive impact in the polar regions at day 2 and 3 and some positive impact in the tropics, however it is rather neutral impact in the medium range. Certainly, the impact for the 2<sup>nd</sup> reprocessed dataset OSE persisted much longer (Figure 127) and was a lot stronger in the tropics. This agrees with the hypothesis that the large noise on the Rayleigh-clear winds, which previously provided the bulk of the tropical impact, is causing the reduced impact. However, in addition there has been a significant increase in tropical GNSSRO data (COSMIC-2) which may also relatively reduce the impact of Rayleigh-clear winds, especially when very noisy.

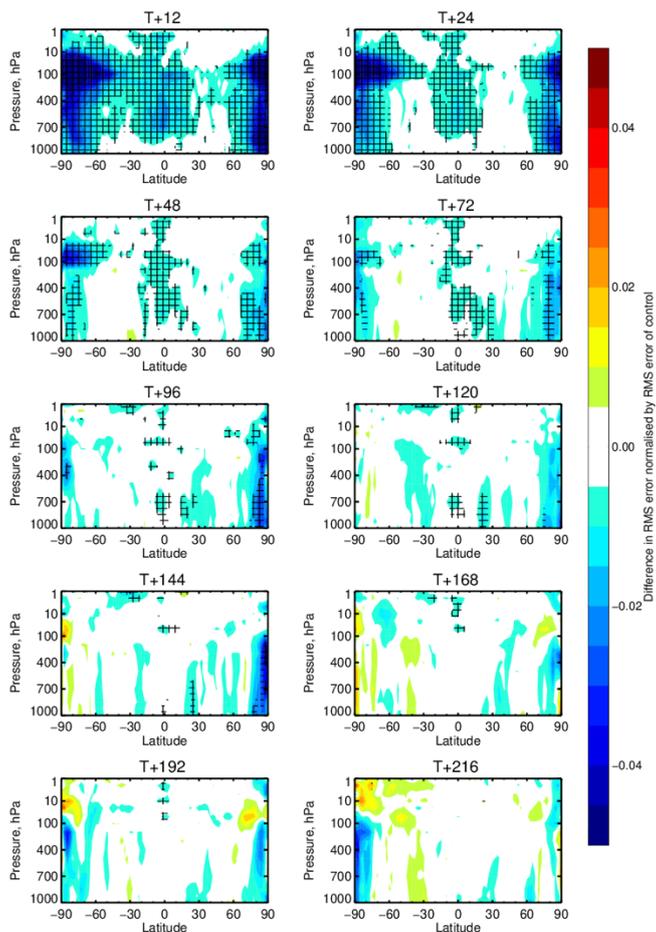


Figure 135. Zonal average normalised change in the RMS (root mean square) error of wind vector from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 13 December 2021 to 28 September 2022. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence. Note the scale is  $\pm 5\%$ .

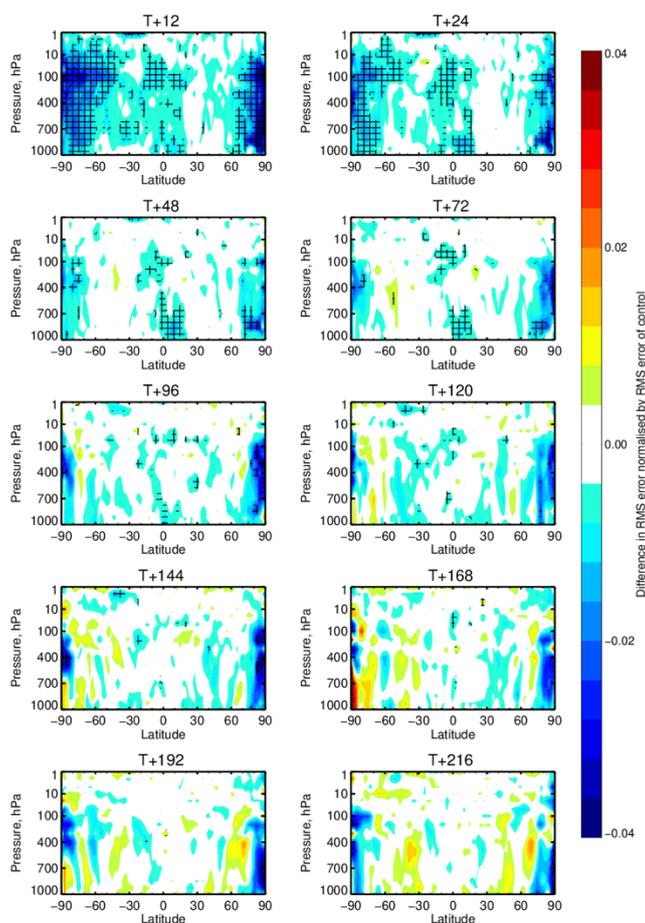


Figure 136. Zonal average normalised change in the RMS (root mean square) error of temperature from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period 13 December 2021 to 28 September 2022. Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence. Note the scale is  $\pm 5\%$ .

#### 4.8 Third reprocessed dataset (FM-A) OSEs

This section provides the verification results from an OSE using the third produced reprocessed L2B dataset, covering from 4 September 2018 until 16 June 2019, as discussed in sections 2.1.8, 2.4.7 and 3.6.

Several experiments were run for this dataset:

- A control without Aeolus, using the nominal operational set-up for observations.
- An experiment with additionally Aeolus “on” (Rayleigh-clear plus Mie-cloudy).
- An experiment with additionally Aeolus Mie-cloudy-only “on”.
- Three experiments testing different combinations of QC to reduce some degradation related to 2 km thick range-bin settings (particularly for Mie winds) in the upper troposphere applied during a large portion of the Commissioning Phase. One option was to not assimilate any Mie-cloudy winds derived from 2 km thick range-bins and the other was to tighten the first-guess QC to use a factor  $Z \sim 3.5\sigma$  rather than  $Z = 5\sigma$  (see explanation of the first-guess check in Appendix 9.9).

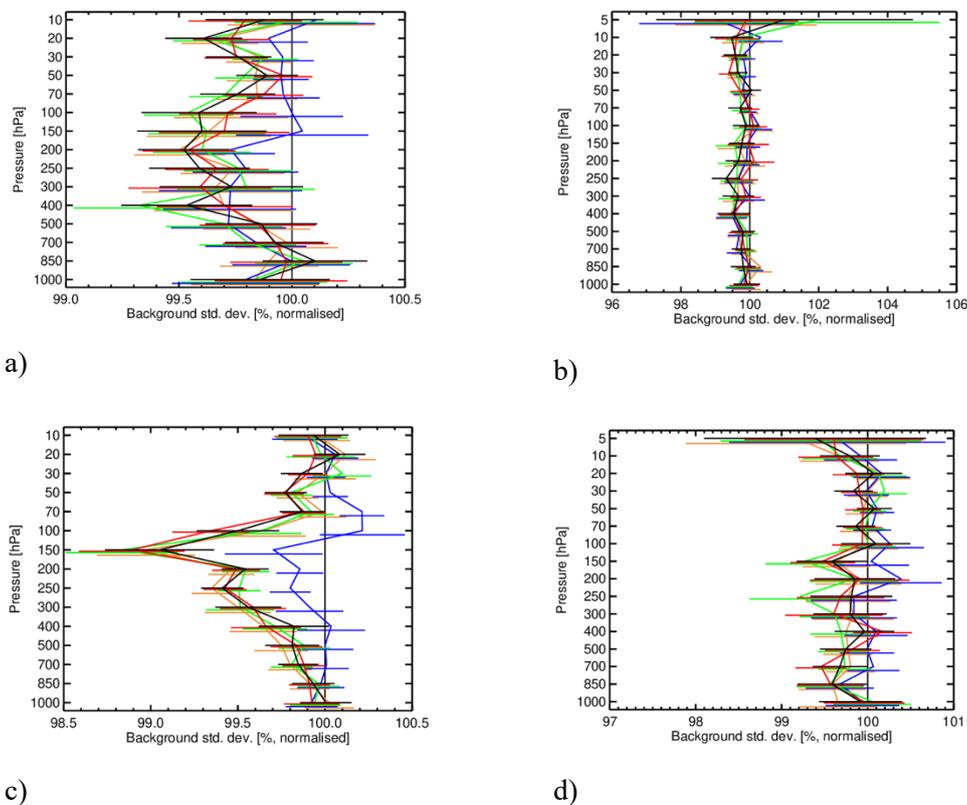
In all cases a high model resolution for a research experiment is used (T<sub>co639</sub> outer loop) since it is closer to the operational resolution and the impression was that impact seems to be greater from Aeolus winds when using higher resolution, as discussed in previous section.

The results that follow are based on concatenating the verification for the periods of 4 September 2018 to 13 January 2019 and 16 February 2019 until 16 June 2019, to avoid the significant data gap present when FM-A went into Survival Mode from 14 January to 15 February (see Figure 4).

#### 4.8.1 Short-range forecast impact

Some degradation in short-range forecast impact versus other observations was evident for the Mie-only experiment (blue) particularly temperature in the stratosphere as shown in Figure 137 and Figure 138. This is probably due to thick range-bins (2 km) used in Commissioning Phase until 26 Feb 2019 in upper troposphere – splitting the OSE into shorter periods (not shown) shows the negative impact was worse before this date. The thick range-bins are an issue when combined with the point-wind observation operator and large vertical wind shear (when we do not know where the clouds are within the range-bin). Evidence of some large erroneous increments in areas of strong vertical wind-shear were found for the top of tropical cyclones. Despite this Mie-only still provided positive wind impact in polar areas at 2 days.

To mitigate the Mie-cloudy degradation, extra QC to discard 2 km thick Mie range-bins and apply a stricter first-guess QC (for both Rayleigh-clear and Mie-cloudy) was tested and was shown (expid=i1b7, black line) to improve the impact significantly e.g. with respect to ATMS for the stratospheric weighting functions (channels 10-15).



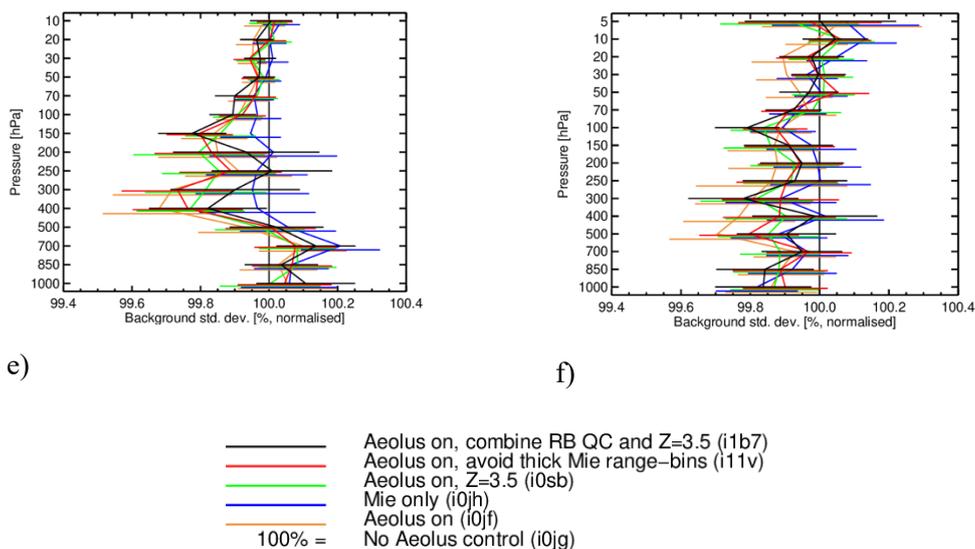
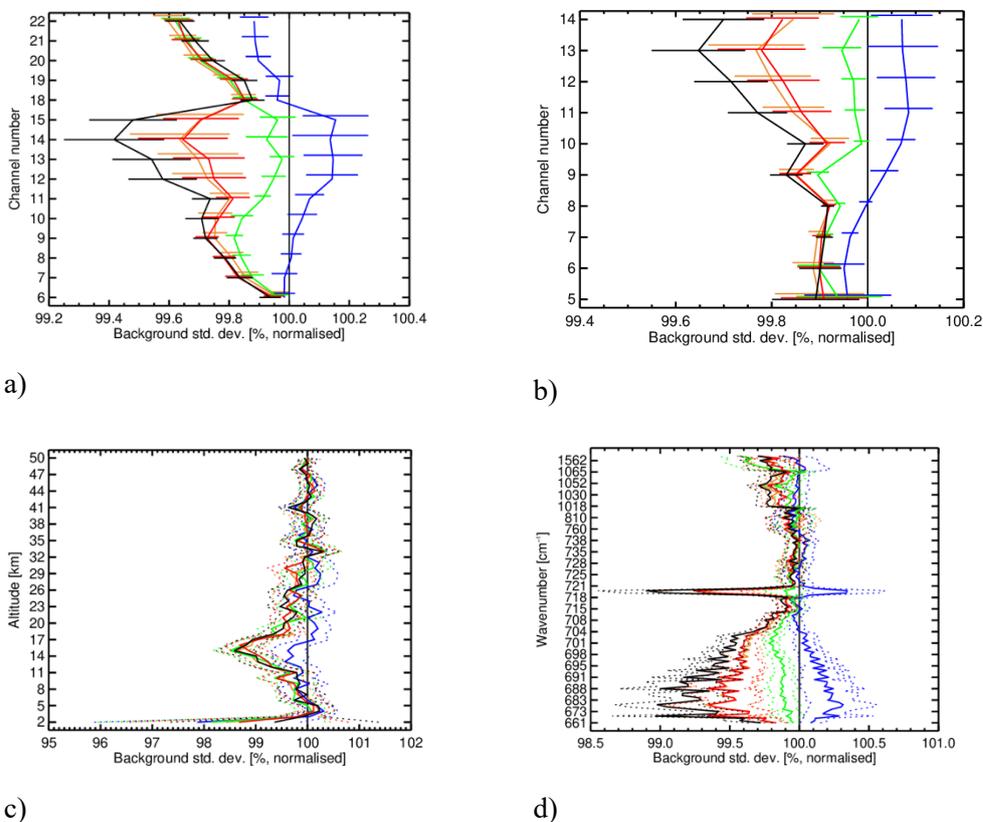


Figure 137. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) conventional data vector wind in the SH extratropics, b) radiosonde temperatures in the SH extratropics, c) conventional data vector wind in the tropics, d) radiosonde temperatures in the tropics, e) conventional data vector wind in the NH extratropics and f) radiosonde temperatures in the NH extratropics. Horizontal bars show the 95% confidence range (Student’s t-test). For the period from 4 September 2018 to 16 June 2019 (with a 32-day gap).



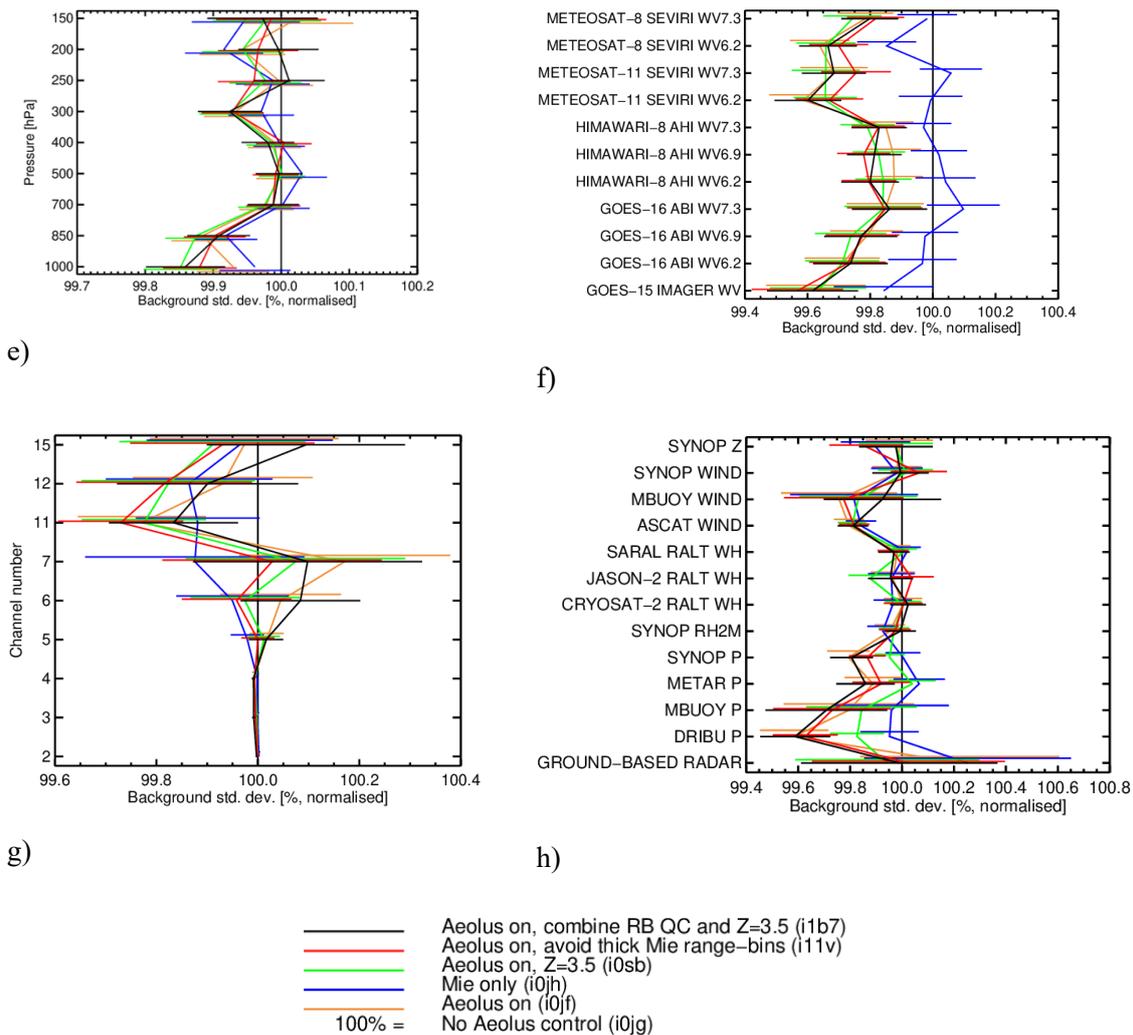


Figure 138. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) ATMS global statistics, b) AMSU-A (all-sky) global statistics, c) GNSS radio occultation in the tropics, d) CrIS global statistics, e) AMV global statistics, f) geostationary imagery global statistics, g) FY-3C MWHS2 global statistics and h) various surface observation types global. Horizontal bars show the 95% confidence range (Student's t-test). For the period from 4 September 2018 to 4 June 2019 (with a 32-day gap).

#### 4.8.2 Medium-range forecast impact

The operational analysis-based verification for vector wind, temperature and relative humidity is shown in Figure 139, Figure 140 and Figure 141 for the experiment with optimised QC settings (Z~3.5 and rejecting the 2 km thick Mie range-bins). Only the optimised QC setting results are shown as they gave resulted in the most positive impact.

Aeolus was not operationally assimilated during this period, which leads to an apparent degradation in the lower stratosphere at 12-hour and 1-day forecasts when verified against ECMWF's operational analysis, particularly over the SH extratropics at 100 hPa. The fit to conventional wind data at 12 hours shows an improvement in Figure 137 (black lines), which means it is likely to be an artefact from Aeolus adding variability to the wind field short-range forecasts which does not exist in the

operational analysis (due to limited wind observations assimilated at these levels). By 2-day forecasts the impact looks positive poleward of 45 degrees latitude and in the tropical upper troposphere – with similar patterns of impact as seen in other OSE periods e.g. the 2<sup>nd</sup> reprocessing (Figure 127) although the 2<sup>nd</sup> reprocessing impact was stronger and more persistent in the tropics, which could be due to the reduced random errors for the early FM-B period and the tuned range-bin settings for NWP impact. The impact on temperature follows similar spatial patterns to the wind impact. There is a very strong positive impact in the SH extratropics stratospheric relative humidity at T+12 to T+48 – which is not yet explained.

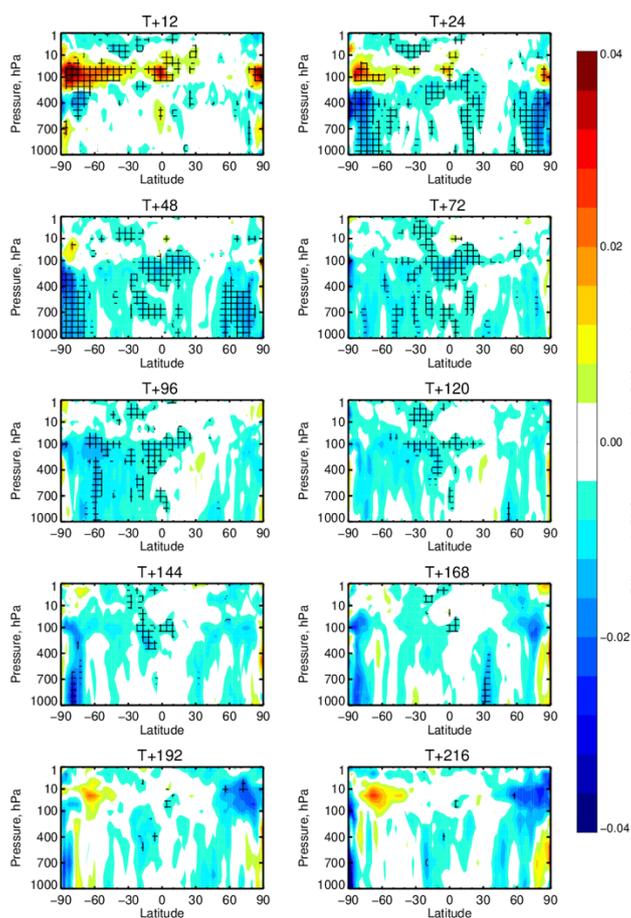


Figure 139. Zonal average normalised change in the RMS (root mean square) error of wind vector from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) with the optimised QC for the period from 4 September 2018 to 16 June 2019 (with a 32-day gap). Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence. Note the scale is  $\pm 4\%$ .

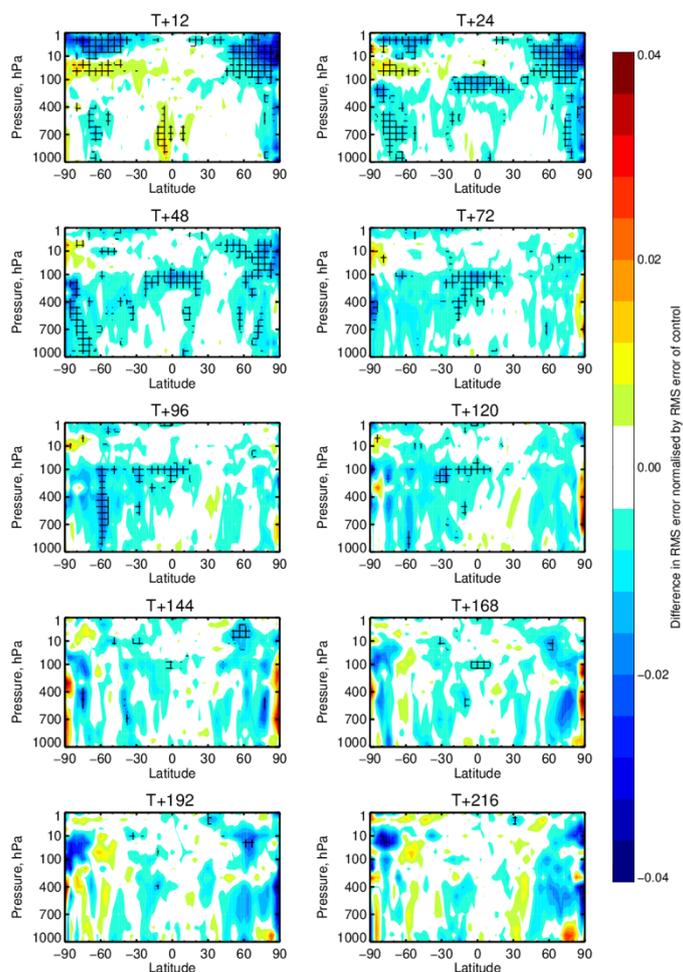


Figure 140. Zonal average normalised change in the RMS (root mean square) error of temperature from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) with the optimised QC for the period from 4 September 2018 to 16 June 2019 (with a 32-day gap). Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence. Note the scale is  $\pm 4\%$ .

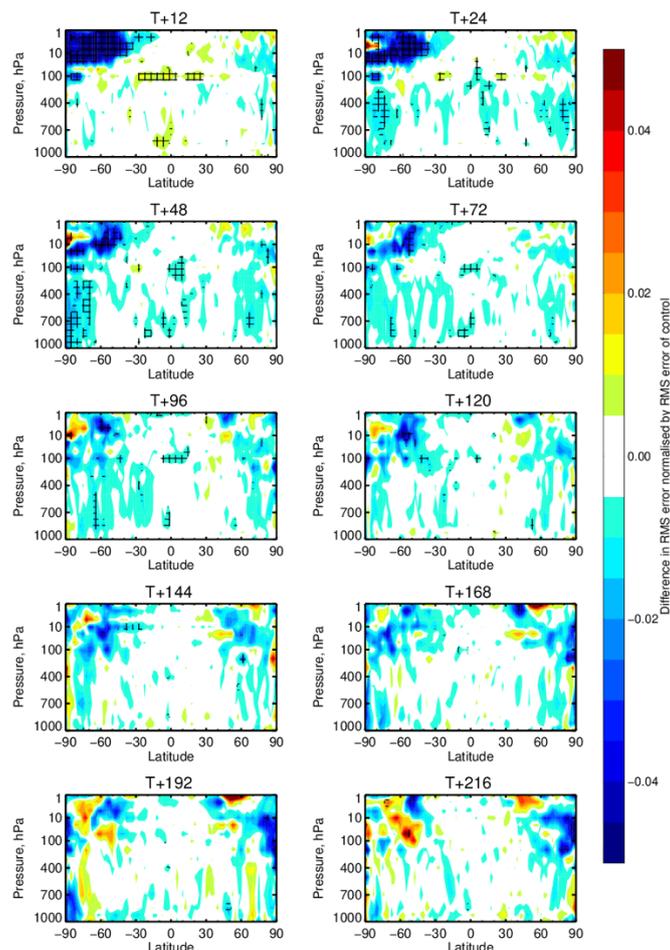


Figure 141. Zonal average normalised change in the RMS (root mean square) error of relative humidity from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) with the optimised QC for the period from 4 September 2018 to 16 June 2019 (with a 32-day gap). Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence. Note the scale is  $\pm 5\%$ .

The impact of Aeolus on extratropical 500 hPa geopotential RMSE was most positive with the optimised QC, giving improvements of 1.3% at day 4 in the SH extratropics and  $\sim 1\%$  at day 5-6 in the NH extratropics; see Figure 142 (black lines). The impact over Europe was good into the medium-range for 500 hPa geopotential height RMSE for this OSE with the optimised QC option, see Figure 143 (almost significant at 2% for day 6). The impact of Mie-cloudy only (prior to QC refinements) was rather mixed (sometimes positive, sometimes negative). A Mie-only experiment after the refinement of the QC settings was not performed due to computational cost but it is assumed to have improved the Mie-cloudy impact to achieve a combined best Rayleigh-clear and Mie-cloudy impact.

In summary, this first attempt at reprocessing the early FM-A period has shown a decent positive impact, particularly at days 2-3 forecast range in similar areas and metrics to other periods, however the impact is not as large in magnitude as found for the 2<sup>nd</sup> reprocessing with FM-B in 2019/2020.

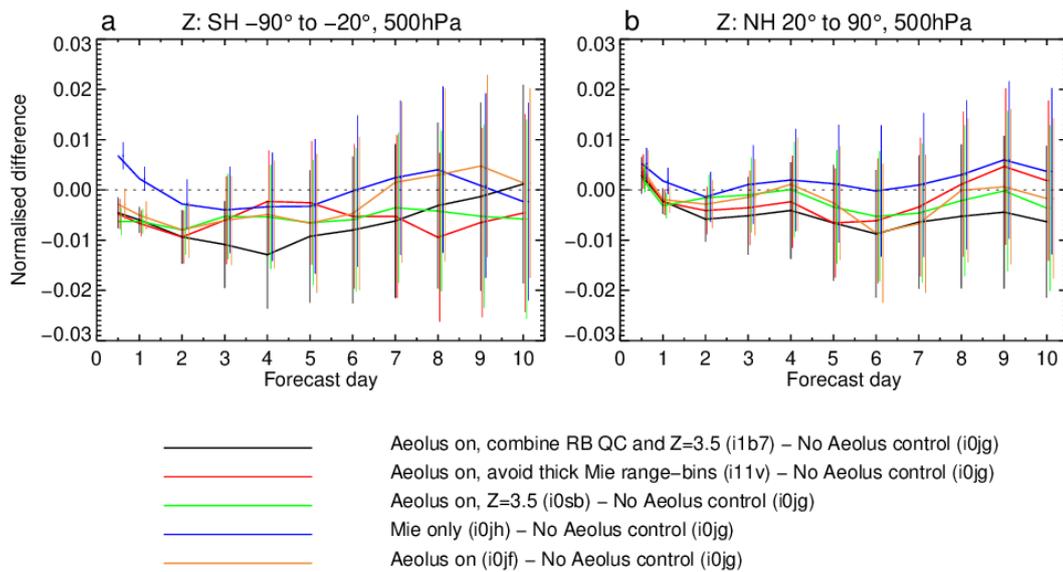


Figure 142. Normalised change in the RMS (root mean square) error of 500 hPa geopotential from assimilating Aeolus (various combinations shown) for the period from 4 September 2018 to 16 June 2019 (with a 32-day gap). Negative values indicate a reduction in error from assimilating Aeolus. Verified against operational analyses. a) SH extratropics and b) NH extratropics.

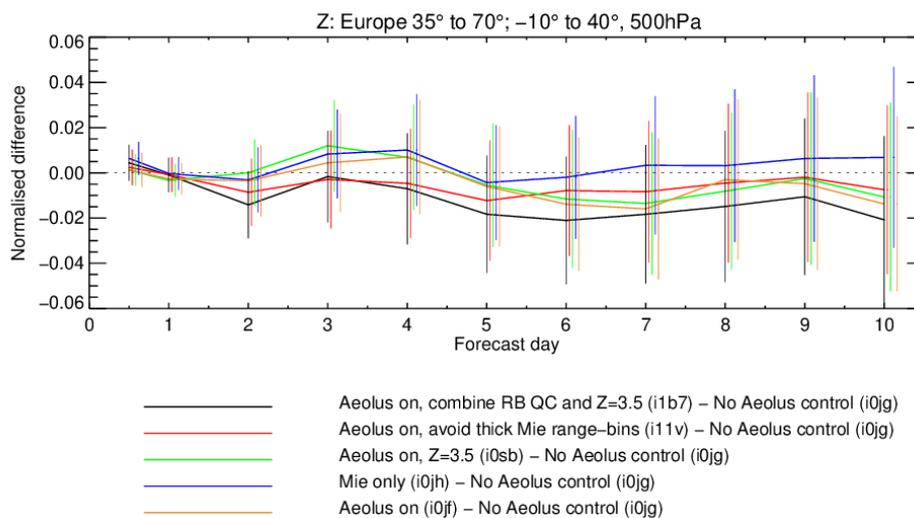


Figure 143. Normalised change in the RMS (root mean square) error of 500 hPa geopotential over Europe from assimilating Aeolus (various combinations shown) for the period from 4 September 2018 to 16 June 2019 (with a 32-day gap). Negative values indicate a reduction in error from assimilating Aeolus. Verified against operational analyses.

#### 4.8.3 Tropical cyclone impact

The tropical cyclone verification, Figure 144, shows an overall neutral impact. However, the impact is mixed with a tendency for position errors and central pressure errors to be improved by 5 day forecasts, whereas central pressure is degraded at analysis and 12 hour forecasts. The Mie-only experiment gives the best impact, particularly for central pressure (peaking at 72 hours). Mixed/neutral to slightly positive impact tendency was also found in the detailed study of the 2<sup>nd</sup> reprocessing: see De Chiara et al. 2023

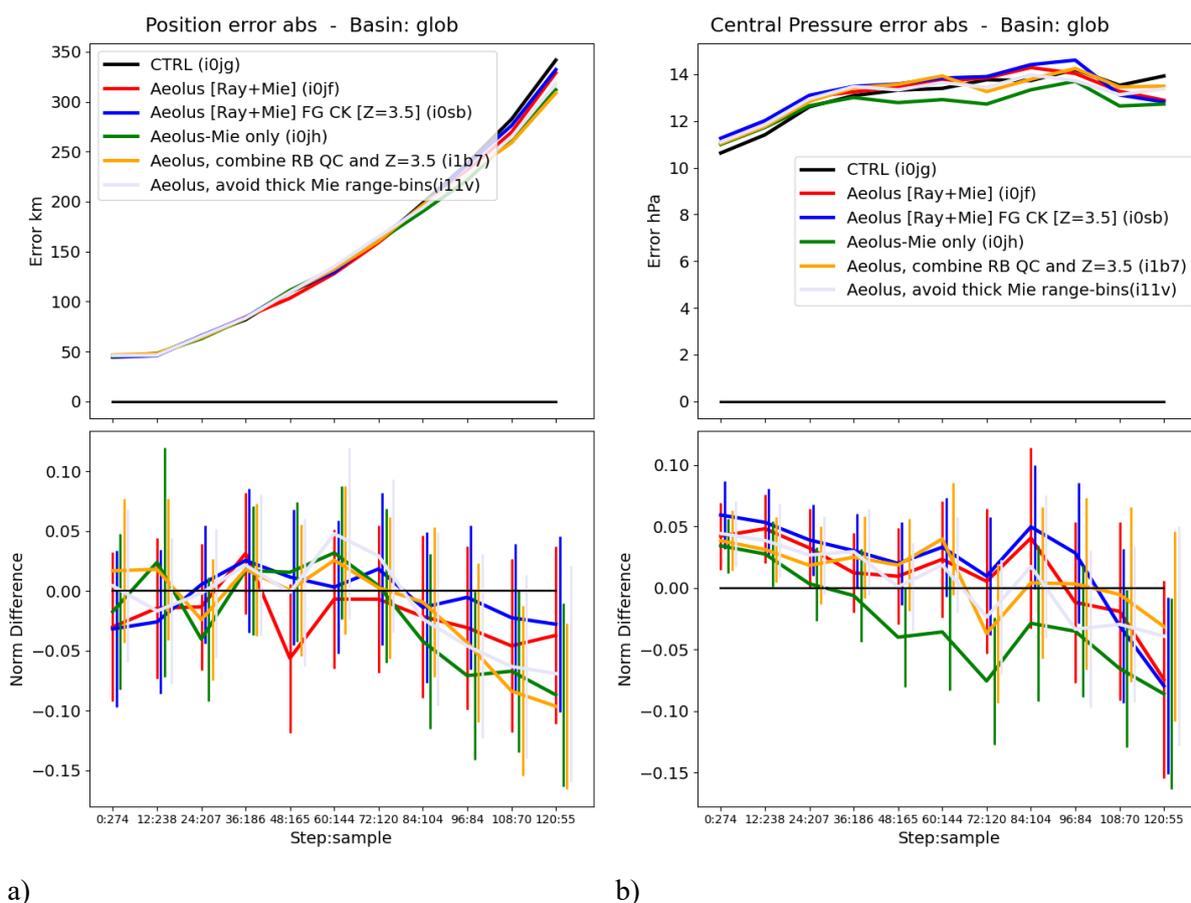


Figure 144. Global verification of tropical cyclone (TC) for the 3<sup>rd</sup> reprocessing from 4 September 2018 to 16 June 2019 (with 32-day data gap). a) Shows the statistics for TC absolute position error (km) b) shows statistics for TC central pressure error (hPa). The top plots show the errors as a function of forecast range and the bottom plots are normalised differences (negative values showing an improvement due to Aeolus) with 95% confidence interval plotted as error bars. The verification is against satellite observation derived best estimates of TC position and central pressure (which are not perfect). From 55 (at 120 hours forecast step) to 274 (at 0 hours forecast step) tropical cyclones were used in the verification. Plots produced by G. De Chiara (ECWMF).

#### 4.9 Late 2022/early 2023 NRT FM-A dataset OSEs

This section provides the verification results from a OSEs using the NRT data at baseline 15 and baseline 16 using the FM-A laser from 1 December 2022 until 30 April 2023, as discussed in sections 2.1.9, 2.4.8 and 3.3. An OSE using Rayleigh-clear plus Mie-cloudy was run over the whole period.

A separate OSE was run to test the impact of assimilating Rayleigh-cloudy winds on top of the existing observing system (including Rayleigh-clear and Mie-cloudy) for a shorter period from 3 March to 30 April 2023. It was shorter due to the Rayleigh-cloudy only reaching a decent quality after modifications were made to the AUX\_PAR\_2B on 16 February 2023 to apply a new correction-scheme to Rayleigh-cloudy winds. This correction is described in Marseille et al. 2023.

There was a data gap from 20 February until 2 March 2023 (10 days) due to the instrument entering Survival Mode. The gap has been removed from the verification results presented below.

4.9.1 Short-range forecast impact

The impact of Rayleigh-clear and Mie-cloudy on short-range forecast fit to conventional wind data is good in all areas. For radiosonde temperature the largest positive impact in the extratropics is between 300-100 hPa in the SH and 300-500 hPa in the NH (see Figure 145). It is an interesting result that the tropical radiosonde temperature impact is slightly negative around 100-200 hPa (there is also a similar feature for the late 2021/2022 OSE, see Figure 132), yet the impact on winds is positive. The impact on tropical satellite observations sensitive to temperature e.g. ATMS, GNNSRO and aircraft temperatures is positive, thus contradicting the radiosonde result.

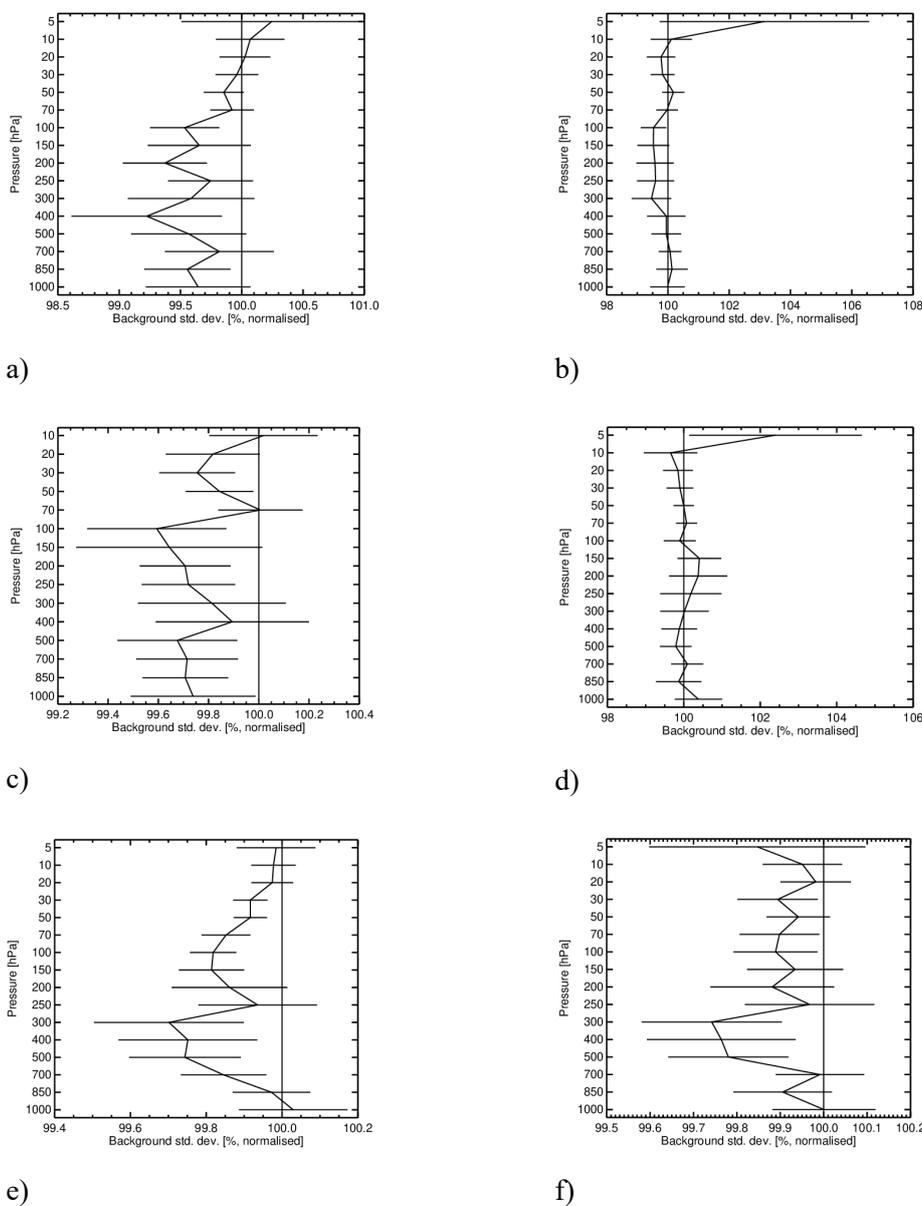
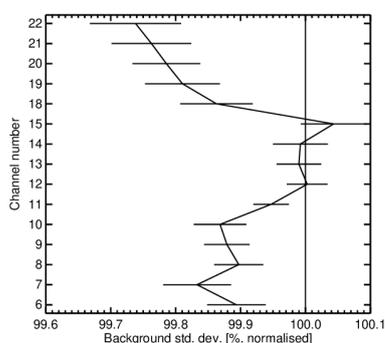


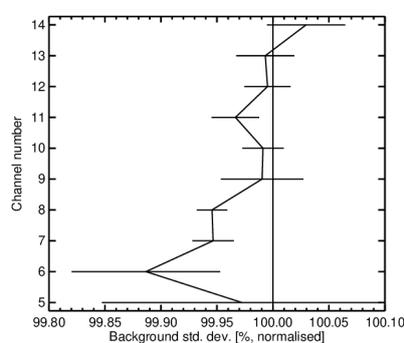
Figure 145. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are

a) conventional data vector wind in the SH extratropics, b) radiosonde temperatures in the SH extratropics, c) conventional data vector wind in the tropics, d) radiosonde temperatures in the tropics, e) conventional data vector wind in the NH extratropics and f) radiosonde temperatures in the NH extratropics. Horizontal bars show the 95% confidence range (Student’s t-test). For the period from 1 December 2022 until 30 April 2023 (with a 10-day gap).

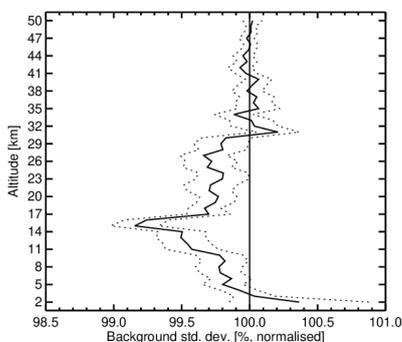
The short-range forecast fit to satellite observations (Figure 146) from Rayleigh-clear plus Mie-cloudy shows the usual improvements for those sensitive to temperature and humidity. Also the fit to AMV winds is improved (by a small amount) from surface to 200 hPa. The impact of the Hunga-Tonga Mie winds is not evident with respect to GNSSRO for this period – unlike in the earlier 2022 OSE, which is assumed to be due to the winds being increasingly sparse and noisy due to plume dispersal (see Figure 134).



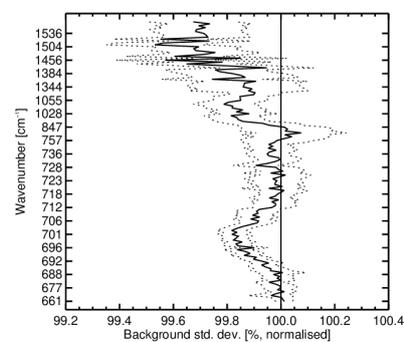
a)



b)



c)



d)

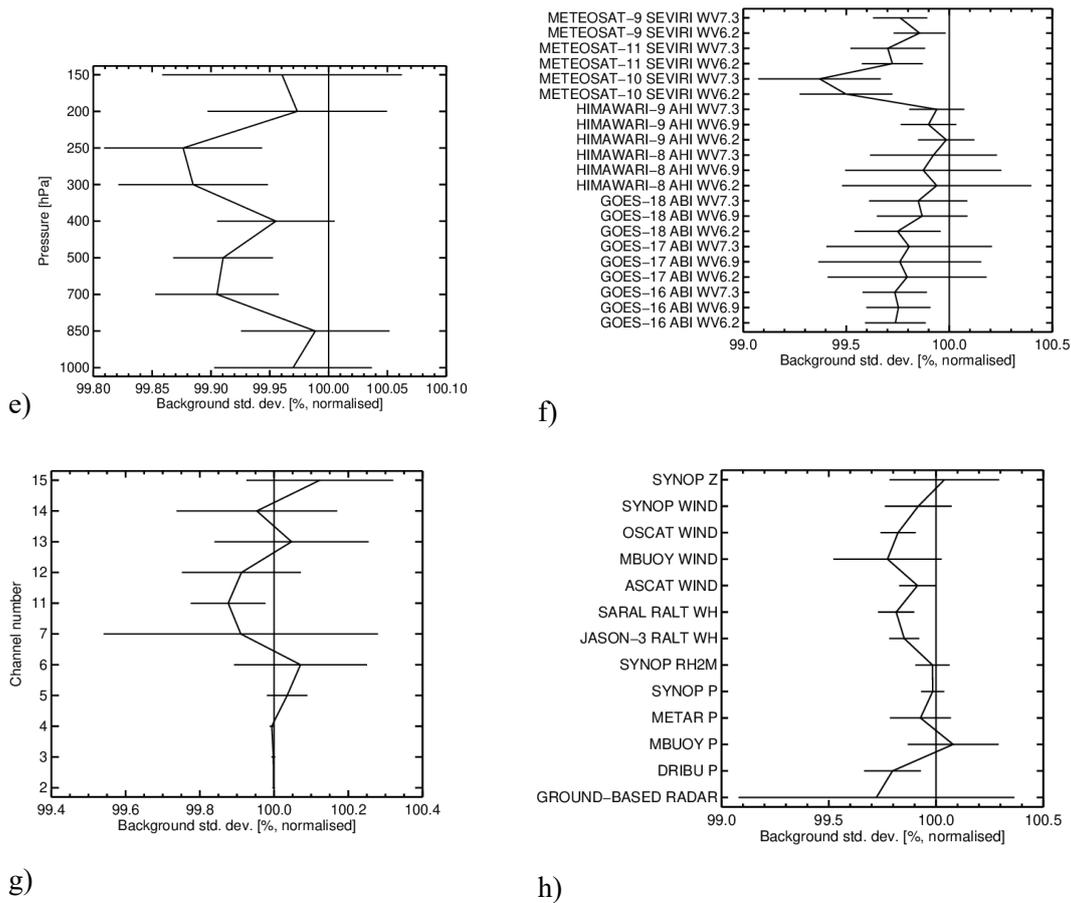
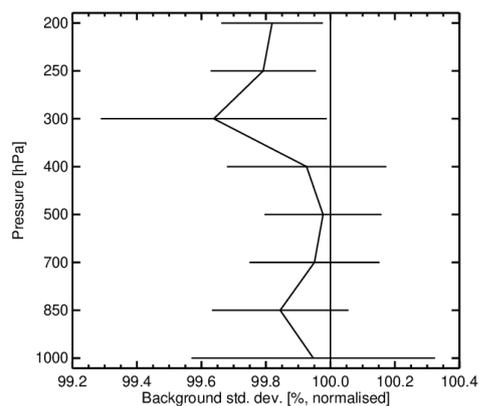
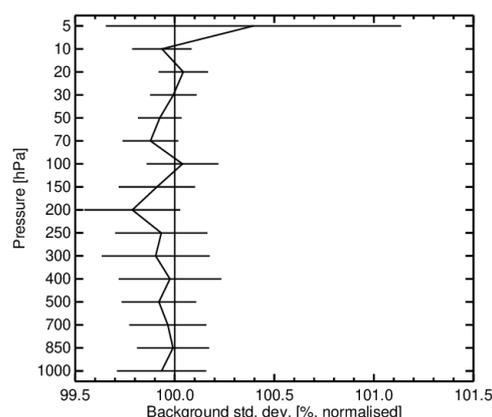


Figure 146. The change in standard deviation of O-B departures resulting from assimilating Aeolus HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Aeolus and above 100% show a degraded fit. The observation types are a) ATMS global statistics, b) AMSU-A (all-sky) global statistics, c) GNSS radio occultation in the tropics, d) CrIS global statistics, e) AMV global statistics, f) geostationary imagery global statistics, g) FY-3C, 3D, 3E MWHS2 global statistics and h) various surface observation types global. Horizontal bars show the 95% confidence range (Student's t-test). For the period from 1 December 2022 until 30 April 2023 (with a 10-day gap).

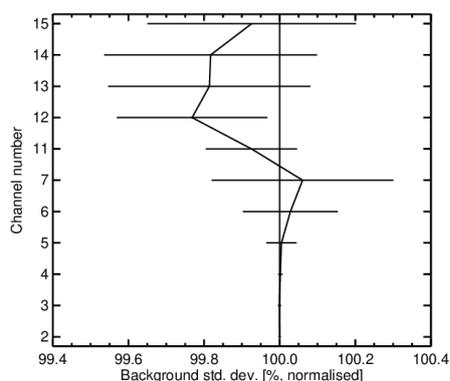
The short OSE for Rayleigh-cloudy winds showed some small positive impacts on short-range forecast fit to other observations. Some highlights are shown in Figure 147. There is small, but statistically significant improvements for temperature at 200-300 hPa and humidity sensitive passive sounders also show improvements e.g. Himawari-9 IR water vapour channels. Improvements for humidity were also found for a Rayleigh-cloudy OSE performed for the early FM-B period (see Figure 79), which adds confidence that there are real benefits to be gained from assimilating it.



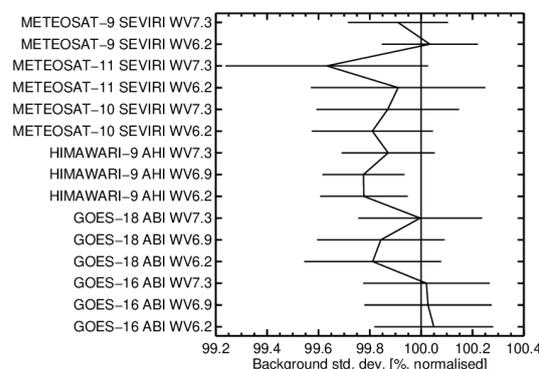
a)



b)



c)



d)

Figure 147. The change in standard deviation of O-B departures resulting from assimilating Aeolus Rayleigh-cloudy HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Rayleigh-cloudy and above 100% show a degraded fit. The observation types are a) aircraft temperature global statistics, b) radiosonde temperature global statistics, c) MWHS2 global statistics and d) geostationary imagery global statistics. Horizontal bars show the 95% confidence range (Student’s t-test). For the period from 3 March 2023 until 30 April 2023.

Given the reduced readout noise with N=5 compared to the original N=30 measurements per BRC, then it was found that the scaling of the L2Bp estimated error for use in data assimilation could be reduced; as discussed in Section 2.4.8.1. Testing of the refined scaling estimates generally produced quite neutral impact, however there was an improvement found with respect to conventional wind data in the UTLS from reducing the Rayleigh scaling factor from 1.4 to 1.25 (red line) for the period 1 December 2022 until 19 February 2023 (see Figure 148).

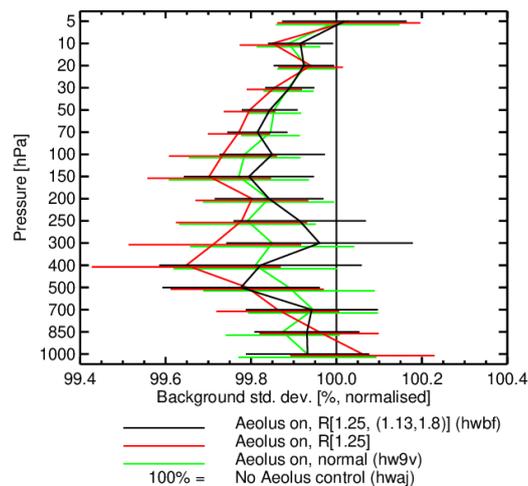


Figure 148. The change in standard deviation of O-B departures resulting from assimilating Aeolus Rayleigh-clear and Mie-cloudy HLOS wind observations, normalised so that the control is 100%. Values below 100% show an improved fit from assimilating Rayleigh-cloudy and above 100% show a degraded fit. The green line is for the nominal scaling set-up; the red for a reduction in Rayleigh scaling from 1.4 to 1.25 and the black for an additional tuning of the Mie-cloudy scaling (more weight). The observation type is global conventional vector wind. Horizontal bars show the 95% confidence range (Student's t-test). For the period from 1 December 2022 until 19 February 2023.

#### 4.9.2 Medium-range forecast impact

The verification against operational analysis for the Rayleigh-clear plus Mie-cloudy (Figure 149 and Figure 150) shows a positive impact at day 2-3, with quite a strong impact towards the poles and decent impact in the tropics. This is similar to the patterns seen for the 2<sup>nd</sup> and 3<sup>rd</sup> reprocessed datasets. The impact at one week tends to become neutral/slightly negative in the SH extratropics. This may just be a lack of statistical significance i.e. longer experiment needed.

The impact is certainly greater than the end of FM-B (2022), but smaller than 2019 FM-B period. Perhaps the significantly increased frequency of hot-pixel induced biases affecting the impact e.g. range-bin 16 was missing for most of this period (and then slightly biased when it came back). This OSE demonstrated that Aeolus still had a useful impact despite some significant improvements in the global observing system between the start of the mission (2018) and the end (2023); such as a large increase in the number of GNSS radio occultation profiles.

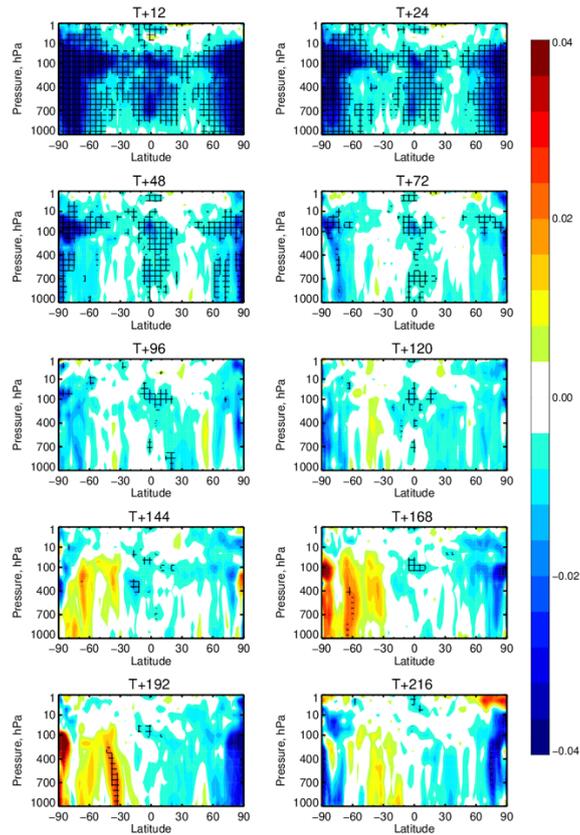


Figure 149. Zonal average normalised change in the RMS (root mean square) error of wind vector from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for period from 1 December 2022 until 30 April 2023 (with a 10-day gap). Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence. Note the scale is  $\pm 4\%$ .

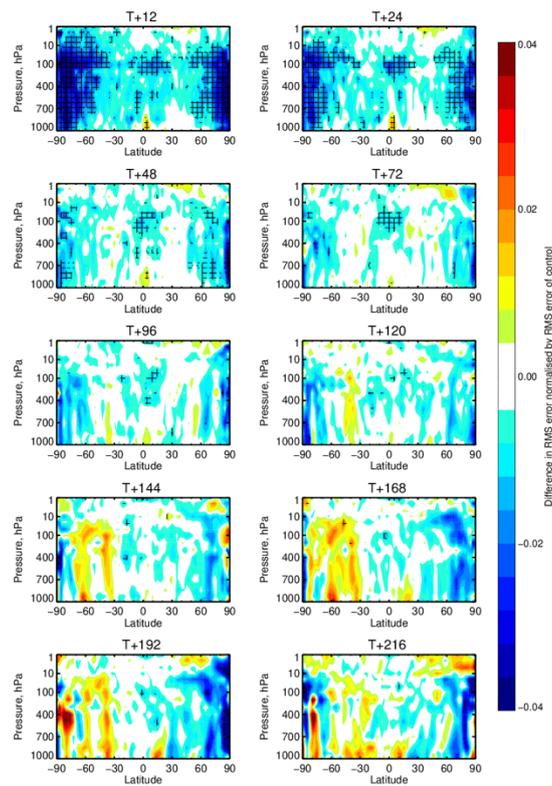


Figure 150. Zonal average normalised change in the RMS (root mean square) error of temperature from assimilating Aeolus (Rayleigh-clear and Mie-cloudy) for the period from 1 December 2022 until 30 April 2023 (with a 10-day gap). Negative values (blue) indicate a reduction in error from assimilating Aeolus and red values an increase in error. Verified against operational analyses. Cross-hatching indicating 95% confidence. Note the scale is  $\pm 5\%$ .

The Rayleigh-cloudy OSE showed neutral impact in the verification against operational analysis (not shown), despite some positive impact of the short-range forecast against observations. Perhaps this is not too surprising given the short period of the OSE and the percentage of assimilated Aeolus data being small: Rayleigh-clear = 59.0%; Mie-cloudy = 36.6% and Rayleigh-cloudy = 4.4%. Rayleigh-cloudy impact should be investigated with future reprocessed datasets with longer periods of decent data.

### 4.10 A consolidation of OSE impact from Aeolus

Metrics of NWP impact from 3-4 month periods throughout the mission were investigated to show how the impact of Aeolus has changed with time and to see if there is any dependence on the atmospheric path signal level. The OSEs used the 3<sup>rd</sup> reprocessing (early FM-A), 2<sup>nd</sup> reprocessing (first 15 months of FM-B) followed by the available NRT dataset based OSEs. The metrics used are global RMSE of 500 hPa geopotential and global 200 hPa vector wind, which are important mass and wind weather forecast metrics. Global is used to reduce seasonal effects (more impact from Aeolus in winter hemisphere) and to keep it simple. A forecast range of day 4 was chosen to reduce dependence on the choice of verifying analysis whilst maintaining a detectable magnitude of impact.

The results are shown in Figure 151. The best impact for Aeolus was found in the first 3 months of FM-B (2<sup>nd</sup> reprocessing) with 1.6% improvement of Z500 and 1% improvement for 200 hPa VW. The impact with FM-A in 2018-2019 via the 3<sup>rd</sup> reprocessing was also good. However the impact with the FM-B 2<sup>nd</sup> reprocessing soon declined and became almost neutral by late 2020 and 2021. This is thought to be mainly because of the significant decline in atmospheric path signal and hence increase in Rayleigh-clear noise and reduction in Mie-cloudy data counts assimilated. Also an increasing number of range-bin specific biases associated with imperfect DCMZ files will have degraded the impact. There was a resurgence in early 2022 using NRT dataset OSEs, and with the improved signal levels of FM-A in late 2022/early 2023 the NWP impact improved somewhat (but not to the levels of early FM-A and early FM-B).

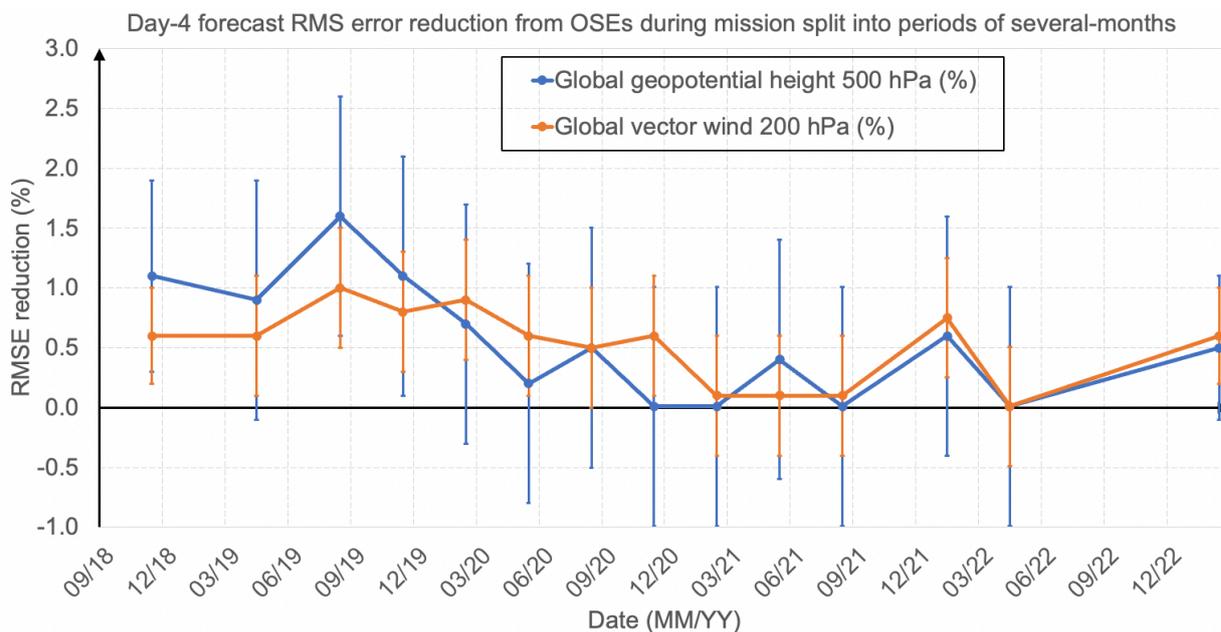


Figure 151. OSE RSME reduction from assimilating Aeolus Rayleigh-clear and Mie-cloudy winds. Using global metrics of 500 hPa geopotential and 200 hPa vector wind.

To investigate the dependence of the impact on the signal levels, the impact has been plotted as a function of the global, profile average Rayleigh-clear observation error estimate (derived from O-B statistics) used in the OSEs; see Figure 152. This shows a moderate relationship (with linear  $r^2=0.56$  for Z500) between the magnitude of impact and the Rayleigh-clear noise levels. This greater impact with smaller observation random errors is of course expected from simple data assimilation theory (see

Appendix 9.3). The dependence of impact on signal levels will be a non-linear function, however this is not trivial to derive for a realistic 4D-Var system; hence the linear fit provides a first-order approximation. An extrapolation using the crude linear fit from real Aeolus errors of typically 6 m/s to those more in line with EPS-Aeolus requirements (~2 m/s), suggests a doubling of impact is possible. However this is rather uncertain as indicated by the 1- $\sigma$  confidence intervals on the linear fit. A similar magnitude of improvement is also predicted from the FSOI extrapolations to smaller random error for the Rayleigh-clear of Figure 158; adding some confidence to this prediction. Changes in the global observing system (e.g. a lot more GNSS radio occultation data in 2023 than 2019) and changing ECMWF IFS cycles complicate this picture. These results could be made more reliable with an OSE using a consistent reprocessing and IFS cycle for the whole mission (future work).

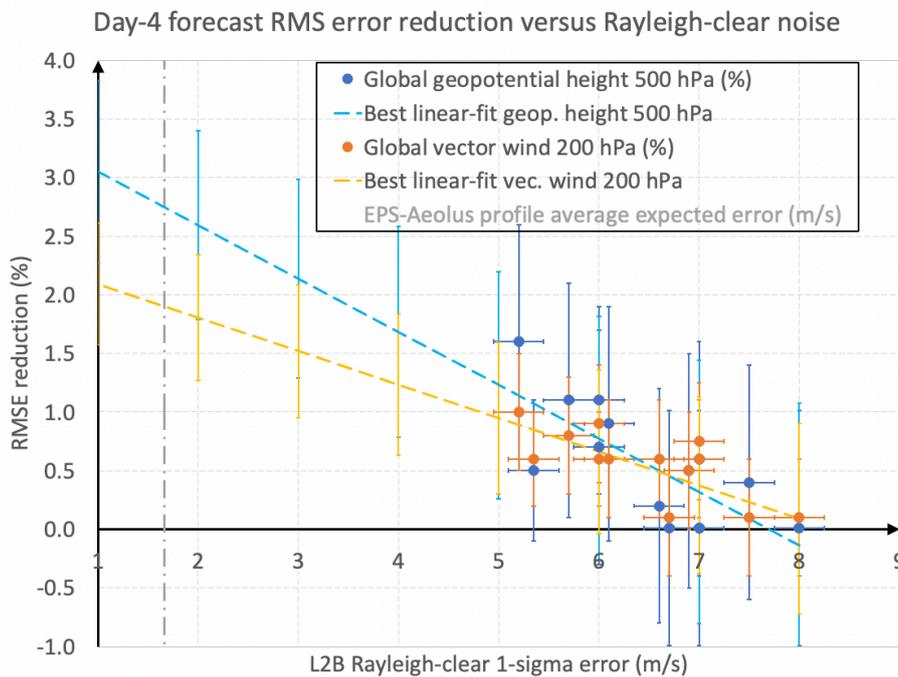


Figure 152. Plot of the OSE RMSE reduction from assimilating Aeolus as a function of the typical Rayleigh-clear random error estimate (as derived from O-B statistics). Linear fits include 1- $\sigma$  confidence intervals.

## 4.11 Aeolus impact assessment by FSOI

Section 2.5 introduces the Forecast Sensitivity to Observation Impact (FSOI) method that is applied for the assessment of Aeolus NWP impact on short-range forecasts.

### 4.11.1 Operational ECMWF data assimilation FSOI results

This section's results are determined from ECMWF's operational data assimilation FSOI (fsobs) suite, which is only relevant for Aeolus when it was operationally assimilated: from 9 January 2020 to 30 April 2023.

#### 4.11.1.1 Relative FSOI comparison of Aeolus L2B HLOS winds to other observation types

Covering a significant period of Aeolus' operational assimilation, from 1 January 2020 to 30 June 2022, Figure 153, shows the relative FSOI partitioned by instrument type and Figure 154 shows it partitioned into larger observation groups; typically accumulating many satellite instruments, using the naming conventions employed by ECMWF. *Note that there were some significant time gaps when Aeolus was not operationally assimilated due to data outage and Special Operations Requests, which leads to a smaller relative FSOI for Aeolus than during contiguous periods of assimilation.*

For data from one satellite instrument, Aeolus winds perform well, coming 12<sup>th</sup> from ~100 instruments, providing 2.3% of the total forecast improvement. In terms of space-based observations, Aeolus L2B HLOS winds provided a similar impact to the Himawari 8 AMVs, but with less impact than eight microwave and IR sounding instruments. It provides similar impact to the radiosonde BUFR data (labelled "BUFR LAND TEMP"). The overall largest impact comes from the WIGOS (WMO Integrated Global Observing System) AMDAR (Aircraft Meteorological Data Relay), which are in situ commercial aircraft observations of wind, temperature and humidity.

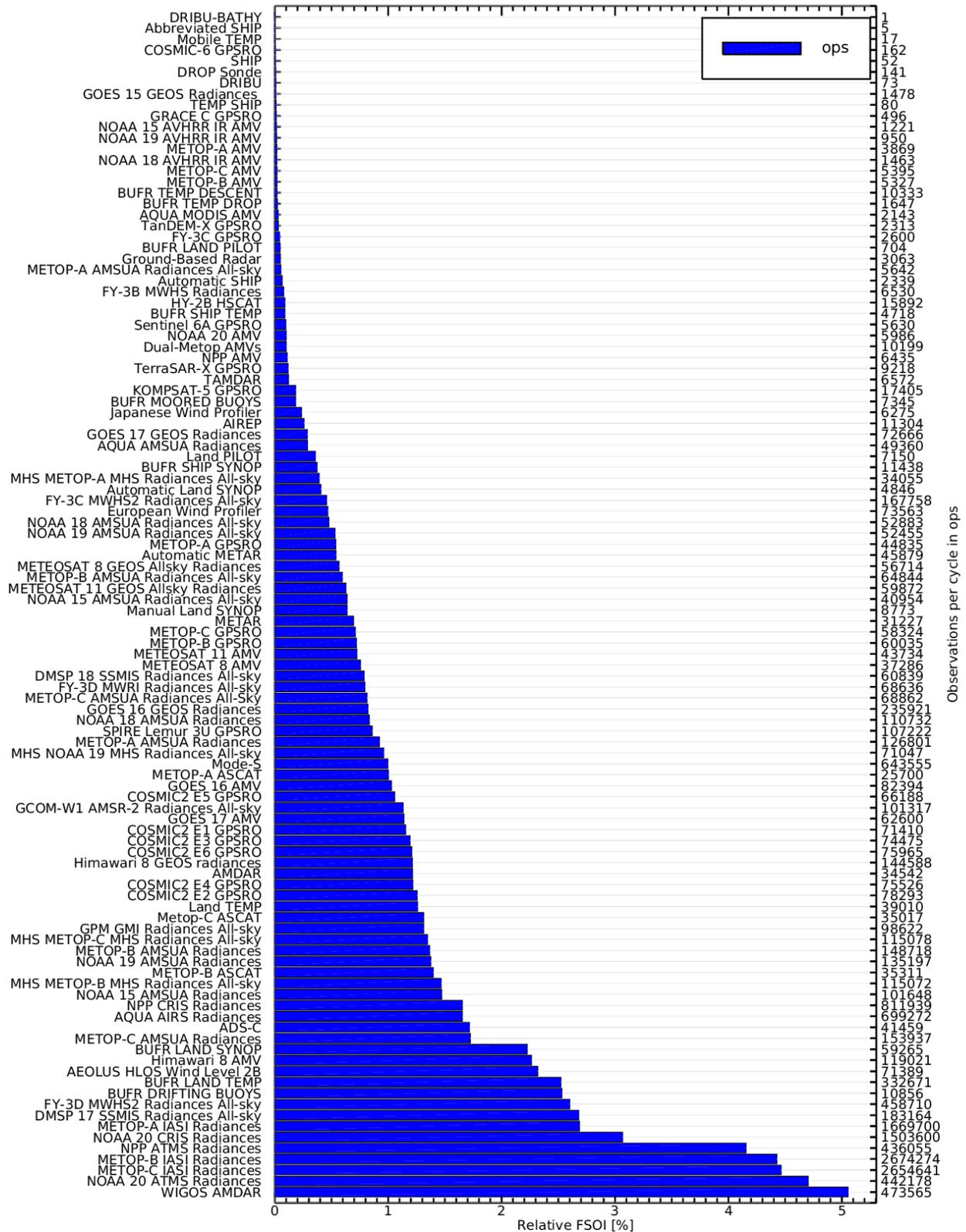


Figure 153. Relative FSOI (%) from ECMWF’s operational system, partitioned by instrument types, averaged over the period 1 January 2020 to 30 June 2022 (Aeolus went operational at ECWMF on 9 January 2020). Aeolus is listed as “AEOLUS HLOS Wind Level 2B”. The number of observations used on average per 12-hour assimilation window is shown on the right-hand side.

Figure 154 compiles the relative FSOI into larger observation groups — Aeolus is the only member of the “Wind lidar” group. Note that these observation groups are mostly made up of many satellite-based instruments e.g. Microwave WV (14 instruments); Microwave T (15); Infrared T (6); Infrared WV (12); GPSRO (15); AMV (15); scatterometer (4); although not all instruments operated throughout the whole period.

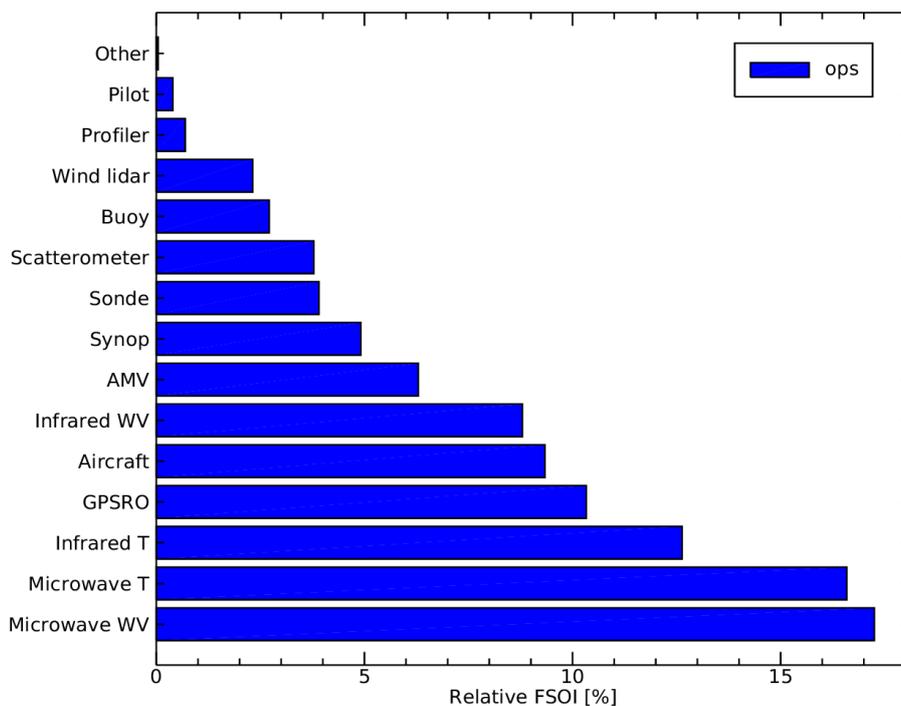


Figure 154. Relative FSOI (%) from ECMWF’s operational system, partitioned by observation groups, averaged over the period 1 January 2020 to 30 June 2022. Aeolus is the only member of the “Wind lidar” group. T stands for temperature, WV for water vapour.

The FSOI *per observation* is large amongst satellite data for Aeolus HLOS winds, see Table 4. The per observation impact is dominated by conventional (in situ) type observations. However, only scatterometer winds have more impact per observation of the satellite-based data, and Aeolus has a similar impact per observation to Japanese radar wind profilers and AMDAR (aircraft) data. These result highlights the importance of wind profile information.

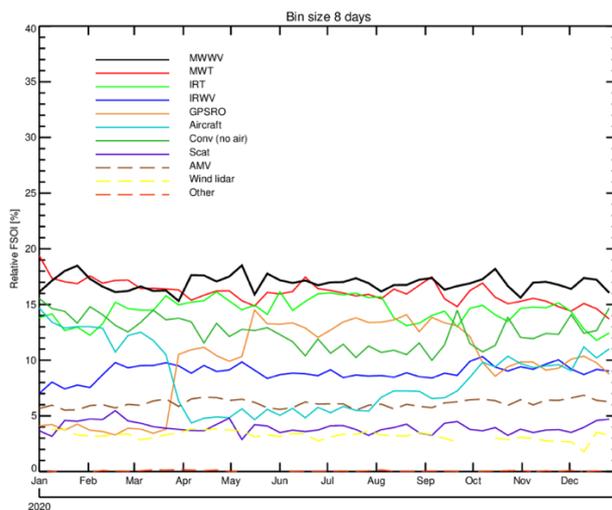
Table 4. FSOI per instrument ordered by FSOI per cycle per observation count (last column) for the period: 1 January 2020 to 30 June 2022. Only the first 30 out of 114 instruments are shown. Aeolus is highlighted in green. Satellite data is highlighted in blue.

FSOI group	Reporttype	Number per 12-hour cycle	Relative FSOI [%]	10 <sup>6</sup> *FSOI per cycle per obs count [J/kg]
Conv. (no air)	BUFR DRIFTING BUOYS	10869	2.522	4.02
Conv. (no air)	TEMP SHIP	86	0.007	1.47
Conv. (no air)	Automatic Land SYNOP	4824	0.404	1.45
Conv. (no air)	Manual Land SYNOP	8888	0.648	1.26

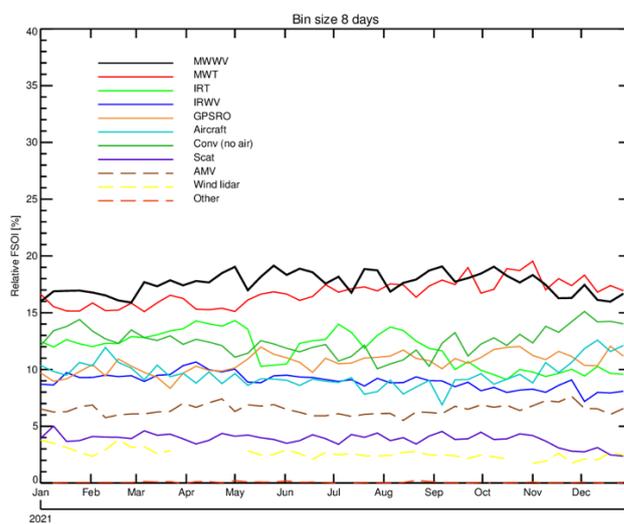
Conv. (no air)	DRIBU-BATHY	2	0.000	1.18
Conv. (no air)	DRIBU	70	0.005	1.12
Conv. (no air)	Mobile TEMP	9	0.001	1.07
Conv. (no air)	BUFR LAND PILOT	703	0.041	1.01
Conv. (no air)	Land PILOT	7193	0.353	0.85
Conv. (no air)	SHIP	52	0.002	0.78
Aircraft	ADS-C	37751	1.655	0.76
Conv. (no air)	Abbreviated SHIP	6	0.000	0.72
Scatterometer	METOP-B ASCAT	35351	1.381	0.68
Scatterometer	METOP-A ASCAT	27673	1.075	0.67
Conv. (no air)	BUFR LAND SYNOP	57985	2.201	0.66
Conv. (no air)	Japanese Wind Profiler	6103	0.231	0.66
Scatterometer	Metop-C ASCAT	35051	1.303	0.64
Aircraft	AMDAR	33747	1.217	0.63
Wind lidar	AEOLUS HLOS Wind Level 2B	70699	2.327	0.57
Conv. (no air)	BUFR SHIP SYNOP	11241	0.369	0.57
Conv. (no air)	Land TEMP	38900	1.264	0.56
Conv. (no air)	Automatic SHIP	2334	0.063	0.47
Conv. (no air)	BUFR MOORED BUOYS	7258	0.183	0.44
Conv. (no air)	DROP Sonde	150	0.004	0.42
Aircraft	AIREP	10496	0.253	0.42
Conv. (no air)	METAR	31164	0.692	0.38
AMV	METEOSAT 8 AMV	38644	0.778	0.35
AMV	Himawari 8 AMV	118709	2.264	0.33
Conv. (no air)	BUFR SHIP TEMP	4751	0.087	0.32
AMV	GOES 17 AMV	62512	1.138	0.32
Aircraft	TAMDAR	6086	0.109	0.31
AMV	METEOSAT 11 AMV	43689	0.714	0.28
AMV	NPP AMV	6445	0.105	0.28
AMV	NOAA 20 AMV	5931	0.096	0.28
GPSRO	COSMIC2 E1 GPSRO	70450	1.139	0.28

#### 4.11.1.2 Time-series of Aeolus L2B wind FSOI results

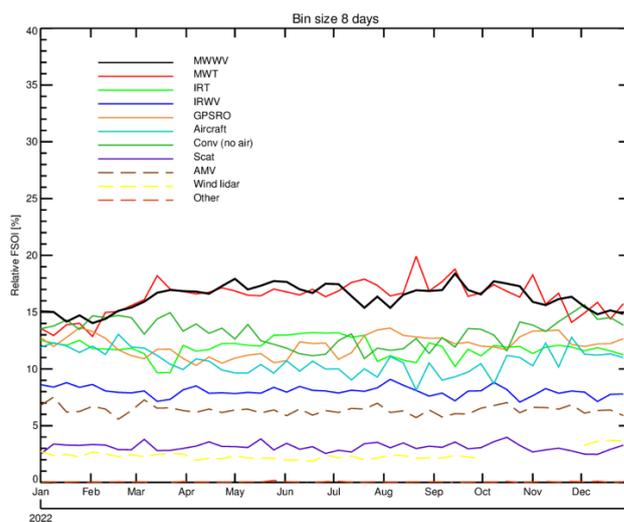
The relative FSOI from ECMWF operations for different observation types is shown over the period of January 2020 to June 2023 in Figure 155. Aeolus relative FSOI (Wind lidar, yellow) fluctuated around 3-4% level in early 2020 but gradually decreased to 2% by late 2021, then made a slight recovery in early 2022, but dipping back to 2% in summer 2022. It then increased to 3-4% with FM-A data from December 2022 to April 2023. *Note there are many periods when Aeolus was not assimilated due to various instrument Special Operations Requests (winds unavailable or unreliable); particularly in late 2020 and in 2021 as indicated by the gaps in the yellow line.*



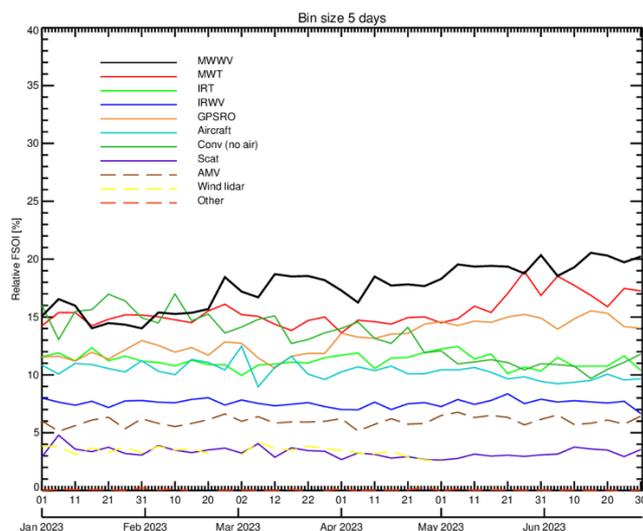
a) 2020



b) 2021



c) 2022



d) 2023

Figure 155. Relative FSOI (%) from 1 January 2020 to June 2023 with an 8 (or 5)-day bin size. Aeolus L2B winds are labelled as “Wind lidar” (yellow). MW=microwave radiances, IR=infrared radiances, WV=water vapour, T=temperature, GPSRO=GPS radio occultation, Scat=scatterometer, AMV=Atmospheric Motion Vectors, Conv=conventional. Plots courtesy of Alan Geer (ECMWF).

Figure 156 shows a time-series of *Aeolus*-only FSOI results from the ECMWF operational data assimilation from 9 January 2020 to 30 April 2023; and from July 2019 to early January 2020 the 2<sup>nd</sup> reprocessed dataset fsofs experiment was used to complete the FM-B period (as indicated by yellow background shading). It consists of global daily FSOI absolute values summed over all assimilated winds, with a 2-week rolling average to reduce noise. The negative of FSOI is taken, so positive values indicate a positive impact. The individual impacts from Mie-cloudy (blue) and Rayleigh-clear (red) are also shown. There were typically 130,000 Rayleigh-clear and 50,000 Mie-cloudy winds assimilated per day in 2020. The number of Rayleigh-clear dropped to 100,000 per day in July 2022 due to decreasing signal and QC, but with the improved SNR of FM-A since December 2022 it was 115,000. The number of Mie-cloudy winds operationally assimilated has generally increased with time even with the FM-B laser. It was around 70,000 per day with the recent FM-A data, mainly due to ground processing improvements and the improved SNR of recent FM-A. However, there were significantly more Mie-cloudy winds assimilated in the early 2019 FM-B period, reaching 100,000 per day (as discussed in section 3.5), partly thanks to large aerosol loads in that period. Periods with low data counts, due to special operations or instrument outages, are excluded in the rolling average calculation.

The Rayleigh-clear impact reduced with time with the FM-B laser because of the increasingly noisy winds as a result of the decreasing atmospheric path signal. However, since December 2022 with the return of the FM-A laser, the FSOI almost doubled to return to a mid-2020 level of impact. The Mie-cloudy impact also tended to reduce with FM-B, but more slowly —with some stabilisation in 2022. The Mie-cloudy impact increased around April 2020, due to the introduction of a new processor baseline which produced ~5,000 extra Mie-cloudy winds per day. The Mie-cloudy FSOI has increased by about 50% with the FM-A was reintroduced, reaching its largest of the mission for NRT operational data.

In early 2020 the Rayleigh-clear FSOI was larger than the Mie-cloudy, but in the rest of 2020 they were similar in magnitude, until 2022 when the Rayleigh-clear dropped to have a reduced FSOI compared to the Mie-cloudy. Even with the greatly improved Rayleigh-clear noise with FM-A in late 2022, then Mie impact has remained higher. This tends to disagree with OSE results which shows more

impact in e.g. 2020 with the Rayleigh-clear winds than the Mie-cloudy. It may point to some exaggeration of the Mie-cloudy impact with FSOI due to the high-spatial density of the observations and the relatively small, assigned observation errors; similar to issues affecting densely sampled Mode-S aircraft data over Europe.

The Aeolus FSOI appears to have some natural variability with oscillations occurring on a period of a few months. This may be influenced by changes in the rest of the operationally assimilated observing system if those observations affect the analysis in similar ways to Aeolus. For example, influence may have come from additional GNSS radio occultation data from COSMIC-2 on 25 March 2020 which also has a very strong positive impact on tropical (particularly stratospheric) winds (Ruston and Healy 2020); however, there is no obvious reduction in Aeolus relative impact shown in Figure 155 associated with this change. However, around the same time there were significant reductions in aircraft observation data counts due to the Covid-19 pandemic, particularly from late March to October 2020, which increased Aeolus relative impact a little (not shown) in the areas where aircraft were normally present. Also, seasonal weather variations are expected to lead to some FSOI fluctuations e.g. it was suggested in Žagar et al. (2021) that the phase of the QBO can lead to variations in the impact of Aeolus. Finally, as already mentioned, changes in the NRT processing baseline and data quality issues may have caused changes in FSOI.

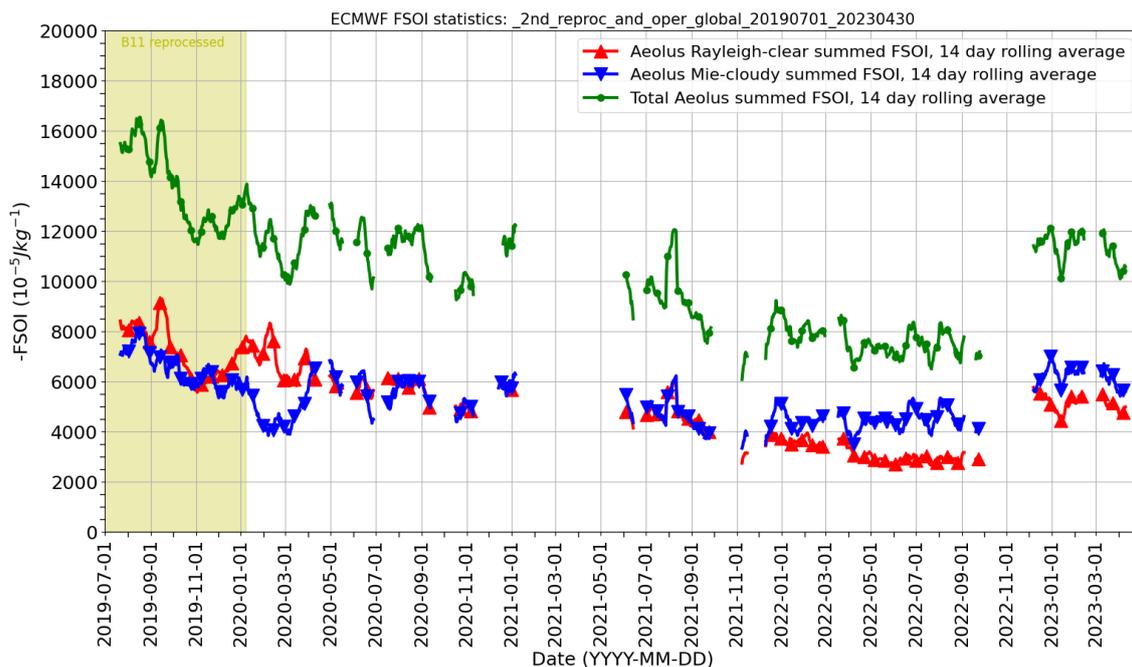
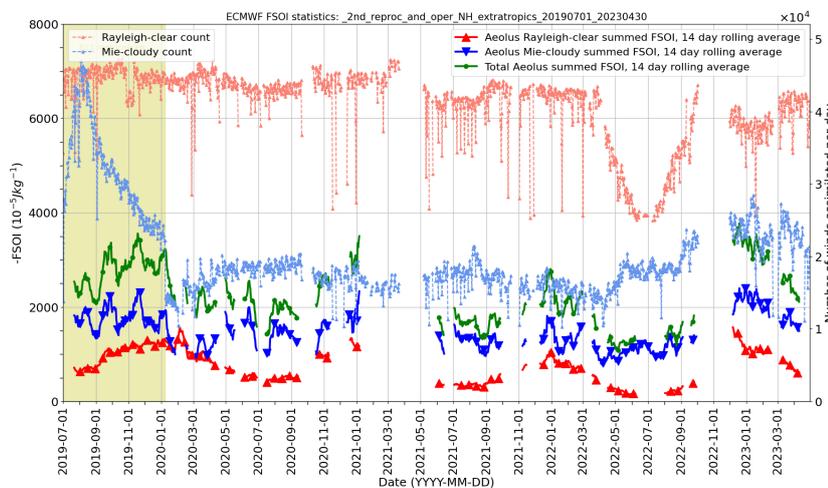


Figure 156. Time-series of global, daily summed FSOI for Aeolus L2B HLOS winds, for the whole FM-B period until 30 April 2023. A 2-week rolling average is applied to reduce noise. The y-axis is the negative of the FSOI and therefore positive values represent positive impact. Operational FSOI data is used after 9 January 2020 and 2<sup>nd</sup> reprocessing FSOI experiment is used before this.

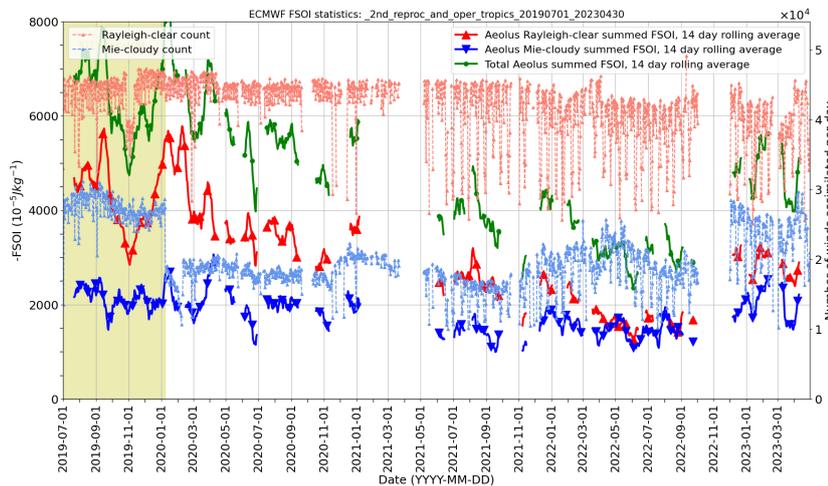
The FSOI time-series statistics can be further split into latitudinal bands as shown in Figure 157; which reveals some interesting features:

- Mie-cloudy FSOI exceeds that of Rayleigh-clear in the extratropics. This agrees with the spatial pattern of short-range impact of Mie-cloudy winds via OSEs (larger impact towards poles).

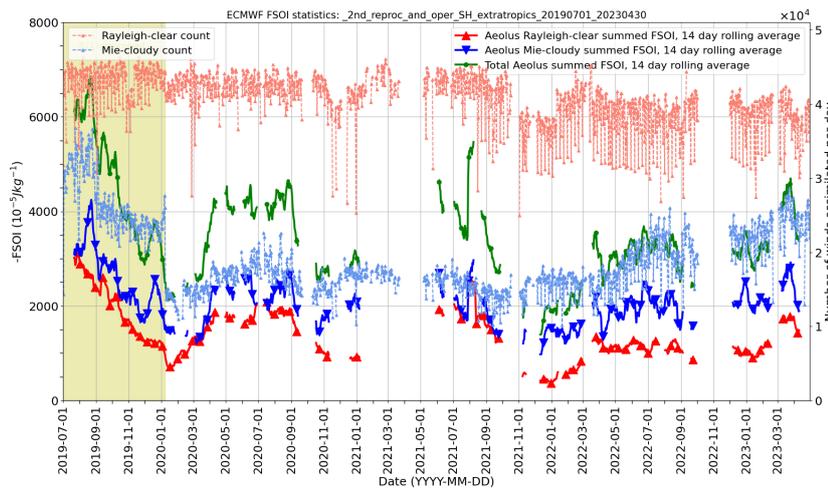
- Rayleigh-clear FSOI mostly exceeds Mie-cloudy in the tropics, apart from the last 5 months of FM-B when they became roughly equal. However, with FM-A in December 2022 the Rayleigh-clear FSOI increased again to late 2020 levels.
- There is a seasonal extratropical variation in FSOI for Aeolus, with more impact in the winter than the summer. This seasonality is stronger for Rayleigh-clear, probably because the Rayleigh channel solar background noise term is large in summer daylight.
- Aeolus has larger FSOI in the SH than NH extratropics; this is commonly seen for satellite data, due to better conventional data coverage over the NH than SH extratropics.
- The Rayleigh-clear FSOI was negligible in the NH extratropics in summer 2022 – coinciding with a large decrease in the number of assimilated data due to the noise being excessive (low useful signal and large solar background causing QC to reject many observations over N pole).



a) NH extratropics



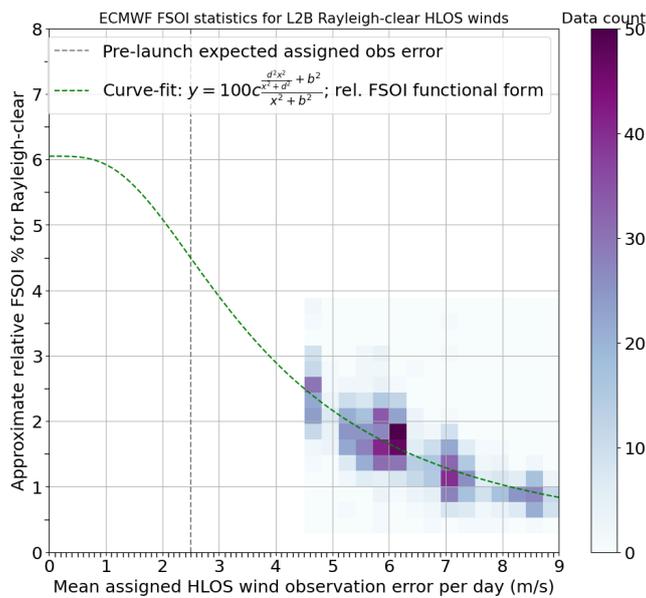
b) Tropics



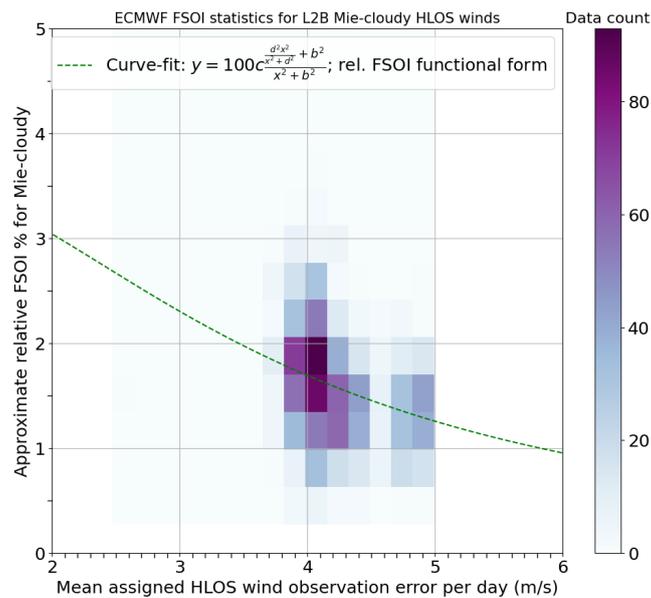
c) SH extratropics

Figure 157. As in Figure 156, but split into a) NH extratropics b) tropics and c) SH extratropics. Also showing the data counts per day on the y-axis on the right.

Given the strong variation with time during the mission of the L2B HLOS wind noise levels (particularly for Rayleigh-clear), we can assess how the FSOI varies with the magnitude of random error. This is shown in Figure 158 in terms of relative FSOI (approximated from the total FSOI per day, calibrated on a specific period) as a function of the daily mean of the assigned observation error. The assigned observation error is the 1-sigma observation error (R matrix) assigned to each observation in ECMWF’s data assimilation, which varies dynamically for Aeolus by being a function of the L2B processor estimated error. The Rayleigh-clear assigned observation error is simply a factor 1.4 times the L2Bp estimate error. It is a more complicated function for the Mie-cloudy due to the inclusion of representativeness error in the 2<sup>nd</sup> reprocessed dataset (until end of 2019) and in ECMWF operations since October 2021 (with CY47R3).



a) Rayleigh-clear



b) Mie-cloudy

Figure 158. 2D-histograms of daily Aeolus L2B winds relative FSOI (approximate via absolute FSOI) versus the mean daily assigned observation error per wind result; for a) Rayleigh-clear and b) Mie-cloudy, for the period 1 July 2019 to 30 April 2023.

There is a large range of assigned observation errors for the Rayleigh-clear due to the decreasing atmospheric path signal levels with the FM-B laser with time, and a relatively small range for the Mie-cloudy. The curve-fit model (green dashed line) is derived in the Appendix 9.8. The curve-fit model fit for the Rayleigh-clear seems to be more plausible than for the Mie-cloudy due to the larger range of assigned error.

It is possible to make a rough estimate of what the relative FSOI for Rayleigh-clear winds could be for smaller observation errors than Aeolus has achieved in practice i.e. to levels aimed at with the pre-launch mission requirements. For example, it was expected that assigned errors of 2.5 m/s were plausible pre-launch. Such a level of random error is predicted to give ~4.5% relative FSOI compared to the 1-3% achieved in practice for Rayleigh-clear. Assuming the Mie-cloudy impact remains ~2% (this is conservative, since a better coverage of Mie winds should occur with higher SNR from the low scattering ratio clouds/aerosols) then the overall Aeolus impact would be ~6.5% relative FSOI with mission requirement signal levels, compared to the 2-5% achieved. The extra impact may have made Aeolus the most important satellite instrument in relative FSOI rankings. This level of impact should be possible (or even exceeded) for an operational follow-on mission.

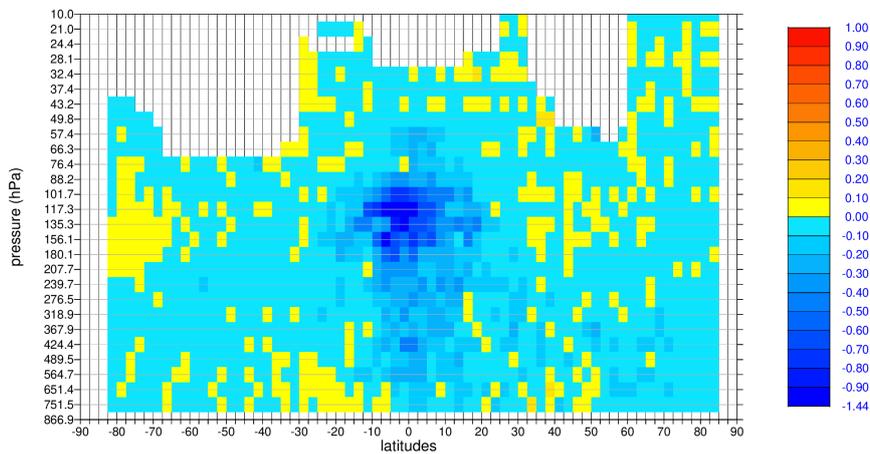
#### 4.11.1.3 Geographical distribution of Aeolus L2B HLOS wind FSOI

Figure 159 and Figure 160 show zonally averaged absolute FSOI scores *per observation* (in units of  $10^{-5}$  J/kg) for Rayleigh-clear and Mie-cloudy HLOS winds respectively (only for ascending orbits, because descending is shows similar results); for the month of January in the years 2020 to 2023.

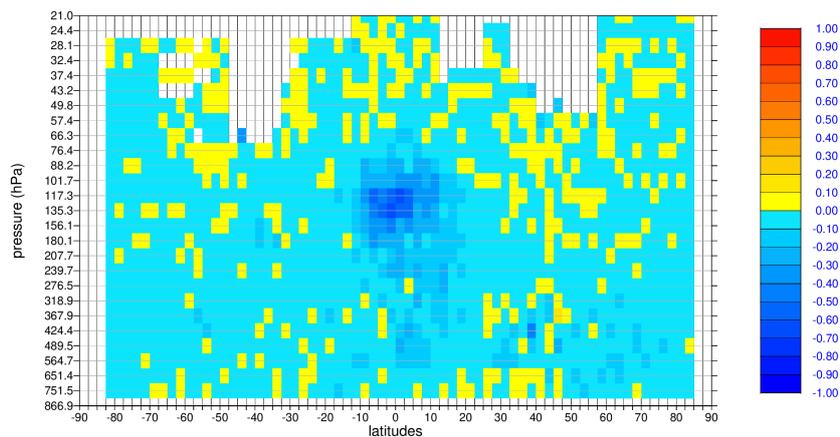
The Rayleigh-clear winds have the largest impact in the tropics (within  $\pm 20^\circ$  latitude), peaking at around 150 hPa (~14 km altitude); this peak tropical impact agrees with the short-range forecast fit to radiosonde winds and GNSSRO from assimilating Aeolus, as shown in the OSE results e.g. Figure 111c. Elsewhere the impact is a small positive effect (but with hints of more impact in 2023 for the mid-

troposphere in the extratropics). The Rayleigh-clear impact per observation in the tropics clearly decreases with time for FM-B, being a lot weaker in 2022 compared to 2020. Despite the data quality improving significantly in 2023 with FM-A the impact in the tropics is not as strong at 135 hPa as in 2020. This could be due to improvements in other observing systems e.g. a lot more radio occultation data, or perhaps due to the phase of the QBO.

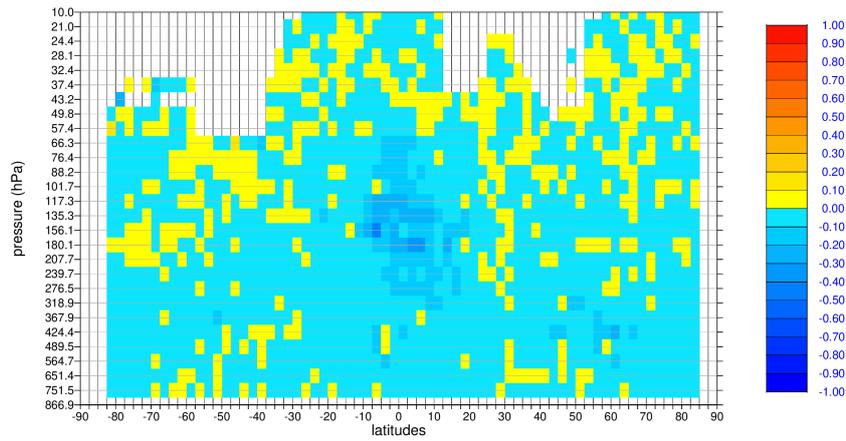
The Mie-cloudy impact looks to be more spatially noisy (due to reduced coverage compared to Rayleigh) but certainly has a positive impact overall. The coverage of Mie-cloudy winds in the vertical is improved in the early 2022 and early 2023 period due to north pole Polar Stratospheric Clouds and the Hunga-Tonga volcanic eruption plume in the tropics. The north pole PSC coverage seems to be missing in January 2021. This is probably because of the occurrence of a major sudden stratospheric warming event on 5 January 2021; see Wright et al. (2021). To form PSC requires the very cold temperatures of a strong polar vortex. There is no clear trend of reduced impact with time for the Mie-cloudy winds.



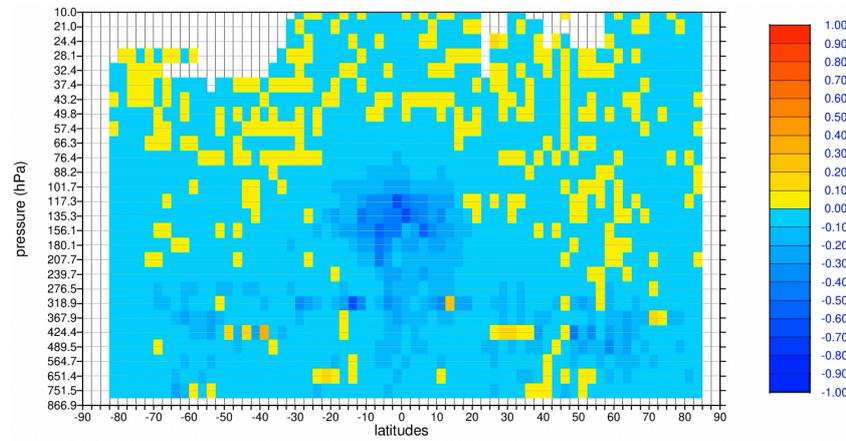
a) January 2020



b) January 2021

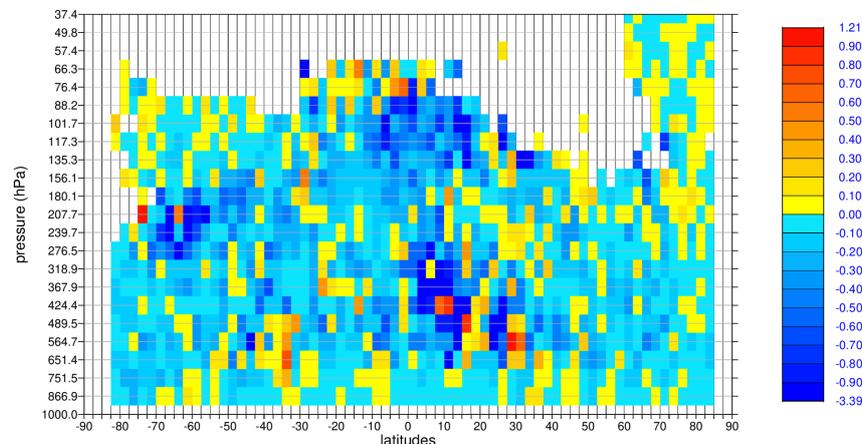


c) January 2022

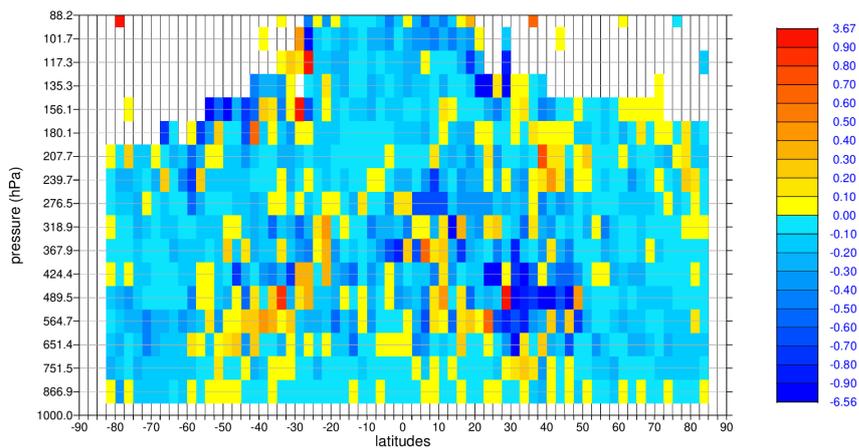


d) January 2023

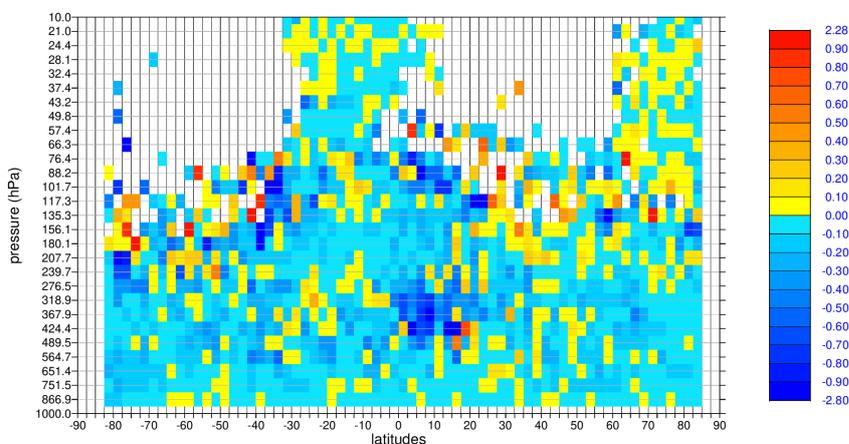
Figure 159. Zonal average plot of the Level-2B Rayleigh-clear (ascending orbits only) FSOI ( $10^{-5}$  J/kg) per observation for the periods a) January 2020; b) January 2021; c) January 2022; and d) January 2023. Blue colours are an improvement, yellow/orange a degradation.



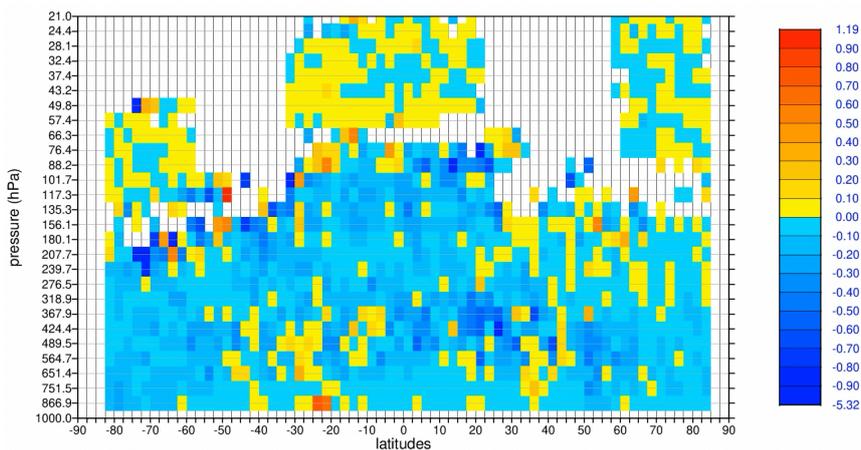
a) January 2020



b) January 2021



c) January 2022



d) January 2023

Figure 160. Zonal average plot of the Level-2B Mie-cloudy (ascending orbits only) FSOI ( $10^{-5}$  J/kg) per observation for the periods a) January 2020; b) January 2021; c) January 2022; and d) January 2023. Blue colours are an improvement. Note that the y-axis pressure axis varies due to differing vertical coverage with time.

Figure 161 shows maps of profile average (all pressure levels) FSOI *per observation* for Rayleigh-clear and Mie-cloudy winds, partitioned by orbit phase for the period of 9 January to 1 September 2020. For Rayleigh-clear winds the largest impact is in the tropics and over extratropical oceans. The Mie-cloudy impact is large in parts of the tropics and extratropical oceans. The FSOI impact per observation of the Mie is roughly twice that of Rayleigh due to the on average smaller observation random errors but given the greater data count and coverage of the Rayleigh-clear, the overall FSOI of the Mie-cloudy and Rayleigh-clear is similar.

For both the Rayleigh-clear and Mie-cloudy the impact is reduced in longitudinal bands centred over Asia and the Americas. This is due to the UTC sensing time of the Aeolus winds in those geographical regions taking place towards the start of ECMWF's 12-hour data assimilation windows. Observations that occur at the at the end of the data assimilation window have much larger positive impact than those at the start as confirmed by McNally (2019).

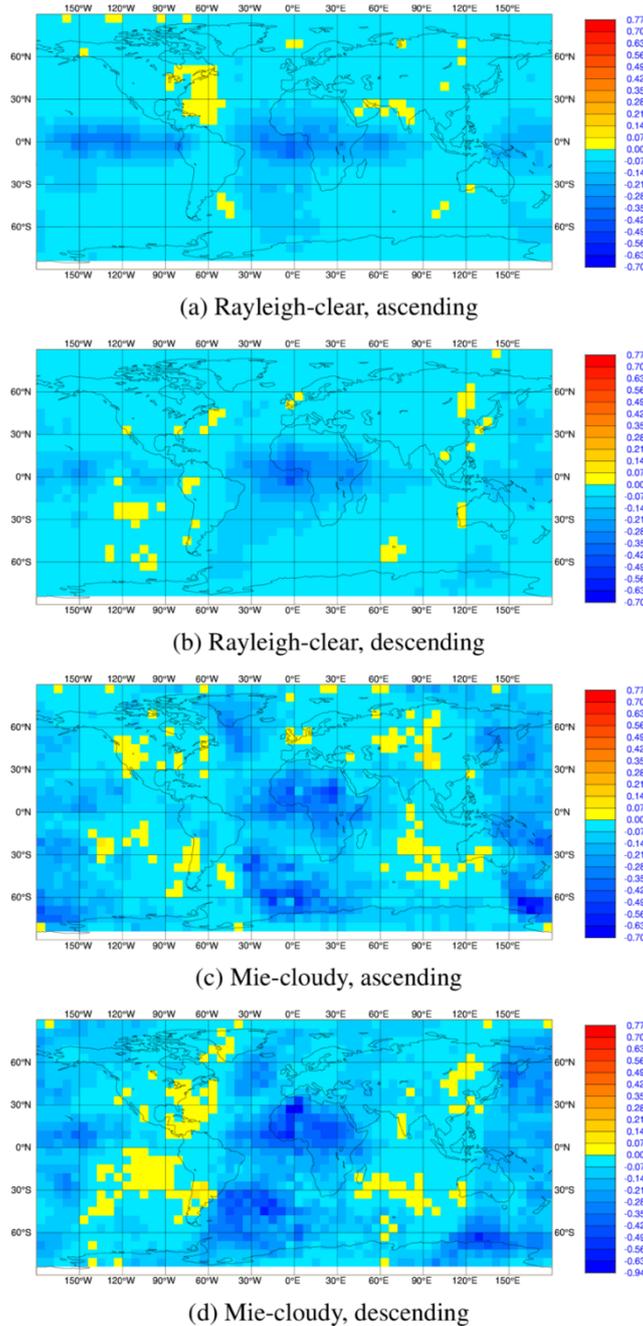


Figure 161. Maps of profile average (all pressures), mean FSOI per observation ( $10^{-5} \text{ Jkg}^{-1}$ ) for the period 9 January to 1 September 2020 for a) Rayleigh-clear in ascending orbit phase, b) Rayleigh-clear in descending orbit phase, c) Mie-cloudy in ascending orbit phase and d) Mie-cloudy in descending orbit phase. Negative values (blue) are positive impact.

The Aeolus (both Rayleigh-clear and Mie-cloudy) average operational FSOI binned by latitude and longitude for the long period from 1 January 2020 to 30 June 2022 is shown in Figure 162. The sharp transition from strong positive impact in the Atlantic and west Pacific (when Aeolus happens to sample the end of the 12-hour LWDA 4D-Var windows) to neutral impact in the Americas and Asia (when it samples the start of the windows) is again evident. Short-range forecast impact is also neutral over Europe (as is commonly seen with satellite data) because of the dominance of the strong impact from aircraft and radiosonde data.

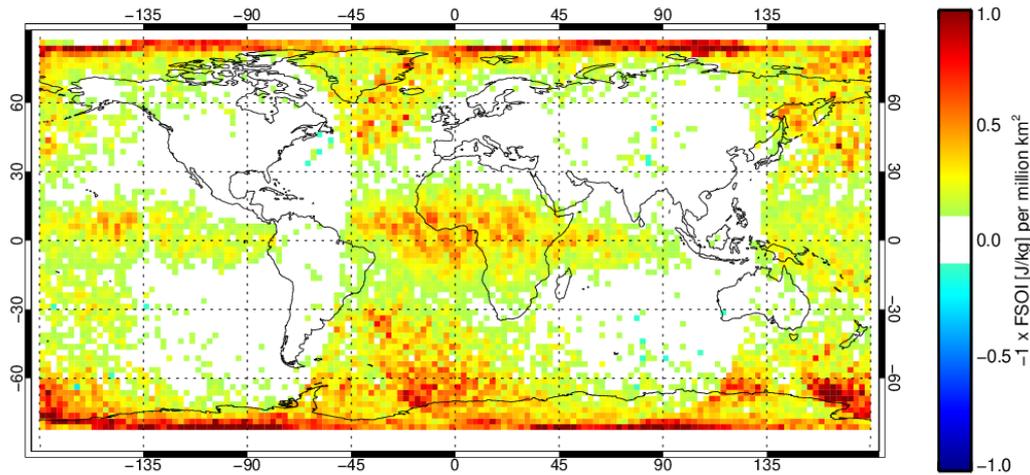


Figure 162. A map of the Aeolus (combined Mie-cloudy and Rayleigh-clear) negative FSOI per million km<sup>2</sup> for the period 1 January 2020 to 30 June 2022 from operations. Red colours (positive values) indicate positive impact, due to this being the negative of the FSOI.

#### 4.11.2 FSOI results for the first reprocessed early FM-B dataset

FSOI verification results were obtained by running an FSOI suite (fsobs) from the FM-B 1<sup>st</sup> reprocessed OSE experiment with Aeolus “on” (Rayleigh-clear plus Mie-cloudy; with improved assigned observation error; experiment ID: hil5). The Early FM-B period is particularly interesting, because Aeolus had the greatest atmospheric signal and hence the smallest L2B Rayleigh-clear random errors of the mission. Some Aeolus data was missing (due to Special Operations) in the first few days of the reprocessed early FM-B dataset, therefore for a fair comparison it was decided to start the FSOI experiment on 3 July rather than 29 June 2019.

Figure 163 shows that Aeolus provided 4.4% of the relative FSOI, which is almost double the 2.3% from operations over the period January 2020 to June 2022 (see section 4.11.1.1). Aeolus is the fourth most important instrument type in this metric (out of 79) for this period. It follows WIGOS AMDAR, NOAA 20 ATMS radiances and MetOp B IASI radiances; the latter provides a similar value to Aeolus of 4.5%. This highlights the importance of relatively good random errors and hence more data passing QC for improving the magnitude of NWP impact for a spaceborne DWL; such as was achieved in the early FM-B period (Figure 40) compared to the very noisy data of particularly 2021 and 2022.

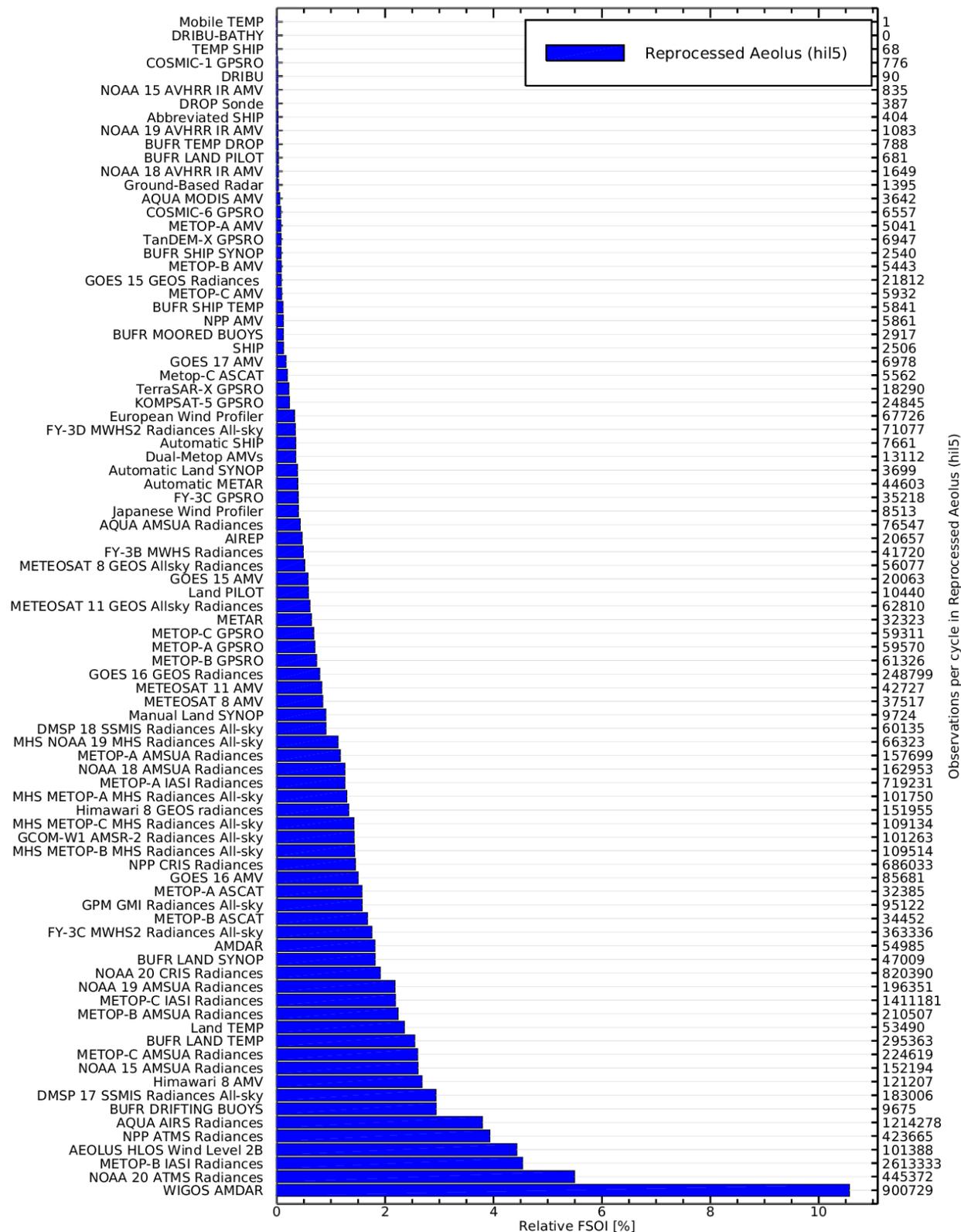


Figure 163. Relative FSOI (%) from the early FM-B period 1<sup>st</sup> reprocessed dataset experiment, partitioned by instrument types, averaged over the period 3 July 2019 to 30 December 2019. Aeolus is listed as “AEOLUS HLOS Wind Level 2B” (both Rayleigh-clear and Mie-cloudy assimilated). The number of observations used on average per 12-hour assimilation window is shown on the right-hand side.

The grouped observation-type results of Figure 164 show that Aeolus (labelled “Wind lidar”) had a relative FSOI similar to radiosondes (a bit less), but better than scatterometer and GPSRO (both groups are made up from several satellites, GNNS RO coverage has improved a lot since 2019) for this period. One of the main aims of the Aeolus mission was to provide a similar impact to the radiosonde network; so, this result suggests the achievement of that aim.

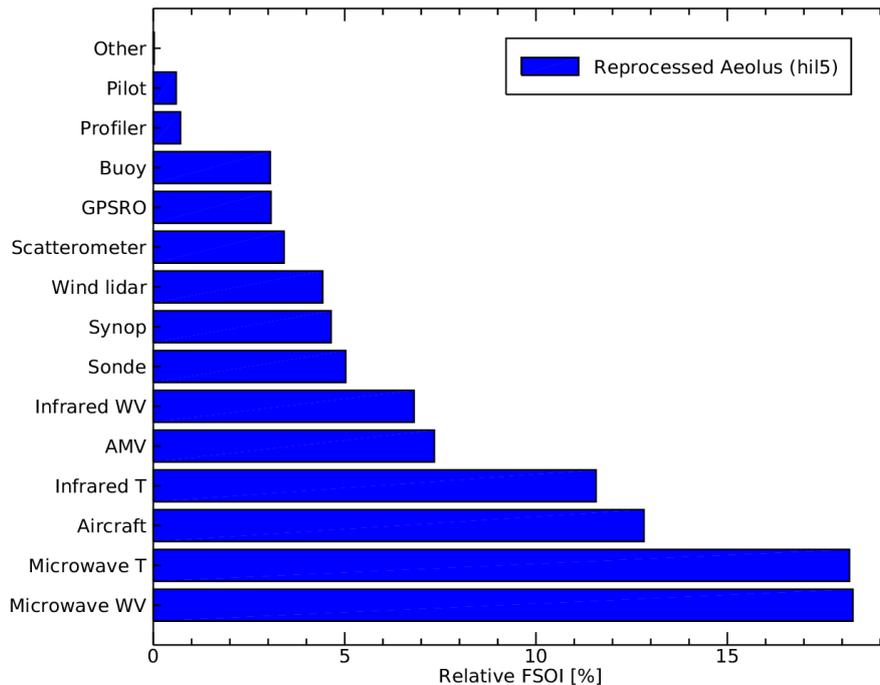


Figure 164. Relative FSOI (%) from the 1<sup>st</sup> reprocessed dataset experiment, partitioned by observation groups, averaged over the period 3 July 2019 to 30 December 2019. Aeolus is the only member of the “Wind lidar” group. T stands for temperature, WV for water vapour.

A time-series of the Aeolus absolute FSOI (split into Rayleigh-clear and Mie-cloudy) in Figure 165, shows that the impact was largest in the early part of this period. The relative FSOI was ~5% in July 2019 and decreased to 4% in December 2019, see Figure 166. The causes of the decreasing impact with time are probably due to: useful signal reduction and increasing noise for the Rayleigh-clear winds; the non-optimal “AMV range-bin” settings which occurred from 28 October to 10 November 2019; seasonal changes in tropical impact associated with level of wind shear in UTLS due to QBO phase (Žagar et al. 2021); variations in the number of Mie-cloudy winds due to high aerosol loads in the early period with northern high latitude wildfire smoke in the lower stratosphere and Polar Stratospheric Clouds in Antarctic winter (see section 3.5); variations in the rest of the observing system.

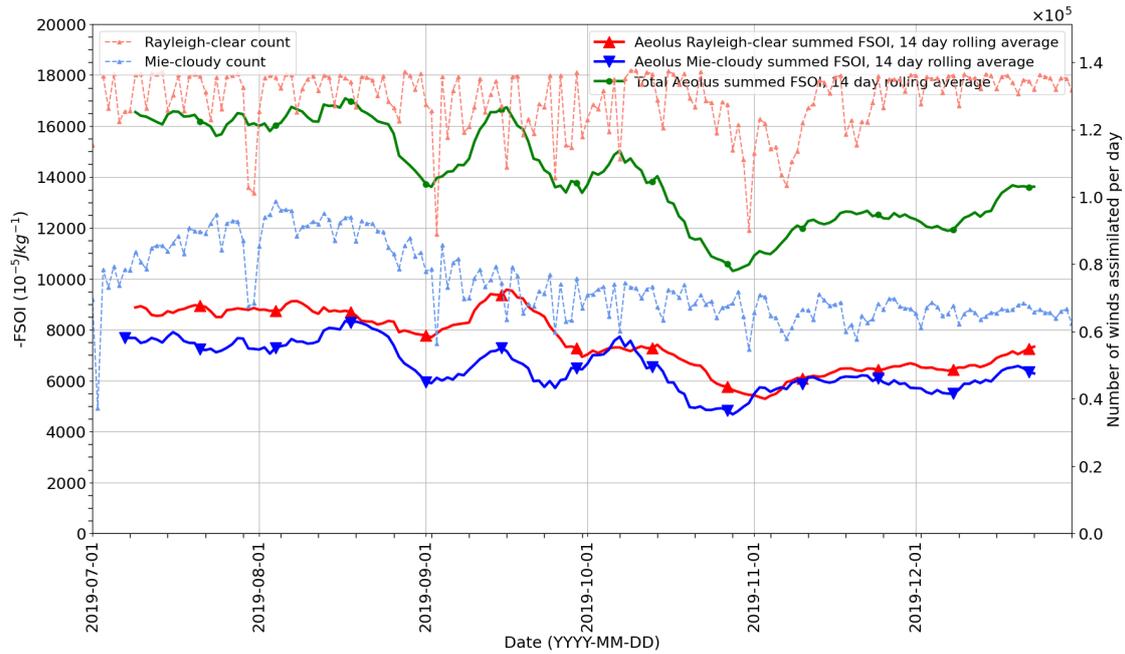


Figure 165. Time-series of global, daily summed FSOI for Aeolus L2B HLOS winds, for the period 3 July to 30 December 2019 (using the Aeolus “on” experiment from the 1<sup>st</sup> reprocessed dataset OSE). A 2-week rolling average applied to reduce noise. The y-axis is the negative of the FSOI and therefore positive values represent positive impact. The data counts are shown on the right-hand side y-axis.

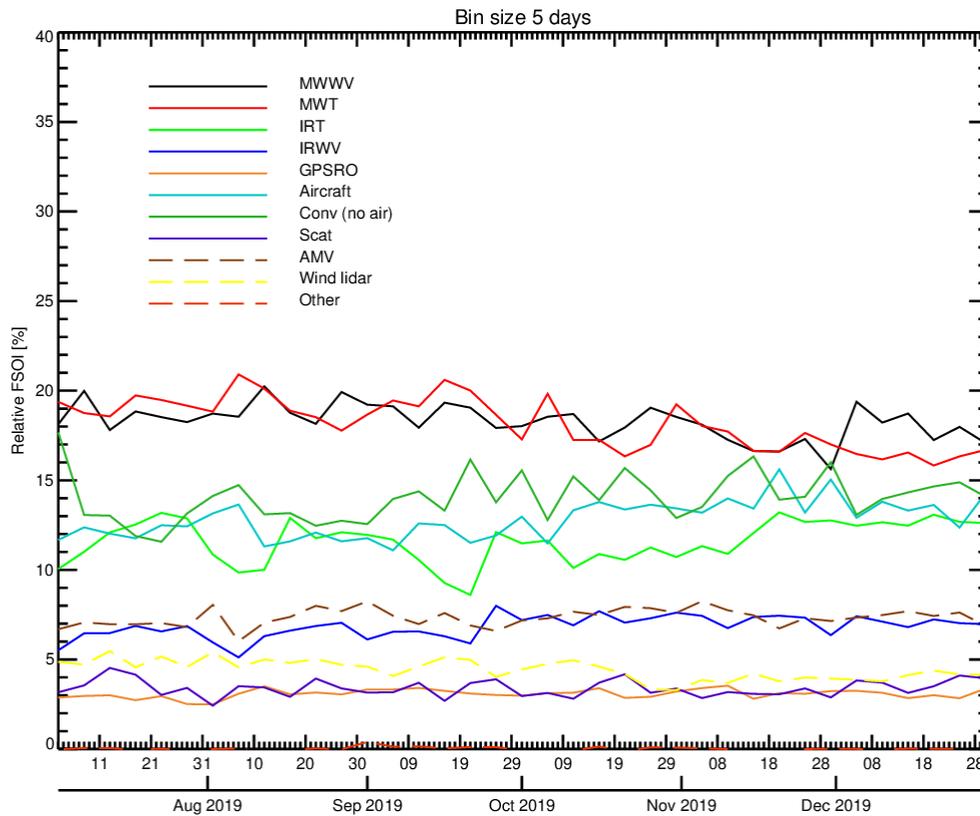


Figure 166. Relative FSOI (%) from 3 July 2019 to 30 December 2019 for the 1<sup>st</sup> reprocessed dataset, with 5-day bin size. Aeolus L2B winds are labelled as “Wind lidar” (yellow). MW=microwave radiances, IR=infrared radiances, WV=water vapour, T=temperature, GPSRO=GPS radio occultation, Scat=scatterometer, AMV=Atmospheric Motion Vectors, Conv=conventional.

#### 4.11.3 FSOI results for the second reprocessed FM-B dataset

The results of the FSOI for the second reprocessed dataset seem reasonably similar for Aeolus to those found for the first reprocessed dataset. From 3 July to 18 October 2019 Aeolus provides 5.2% relative FSOI; this is a bit larger than found for the 1<sup>st</sup> reprocessing for the same period.

***It is unclear why the NOAA 20 ATMS radiances impact has decreased significantly compared to the 1<sup>st</sup> reprocessing and operational FSOI results – the relatively low counts per cycle may point to a bug<sup>34</sup>.*** Aeolus is the second most important satellite instrument for this period/experiment with 4.2% relative FSOI, after MetOp-B IASI radiances. Aeolus provides more impact than scatterometers but less than radiosondes (“Sonde”) as shown in Figure 168.

The time-series of the impact as shown in Figure 169 shows a larger impact in 2020 (~4%) compared to the operational time-series of Figure 155 (3-4%). This may be due to the improved data quality from reprocessing (note that the M1 T bias correction only went into operations in April 2020). There is a gradual decrease in impact of L2B winds with time; from about 5% in July 2019 to 3.5% in September 2020, so qualitatively agreeing with the operational FSOI trend. A few gaps in the number of data assimilated is evident in Figure 170, particularly in 2020. The global FSOI of Rayleigh-clear exceeded the Mie-cloudy until April 2020, after which they remained rather similar; which agrees with the operational FSOI in Figure 156.

Figure 171 splits the FSOI into NH and SH extratropics and the tropics. This highlights some interesting features that agree with those found in association with Figure 157.

---

<sup>34</sup> The number of ATMS used in hls0/hlpv (2<sup>nd</sup> reprocessing) OSE looks consistent with 1<sup>st</sup> reprocessing (h15) OSE, which suggests perhaps the fsobs experiment per se (rather than OSE) had the issue.

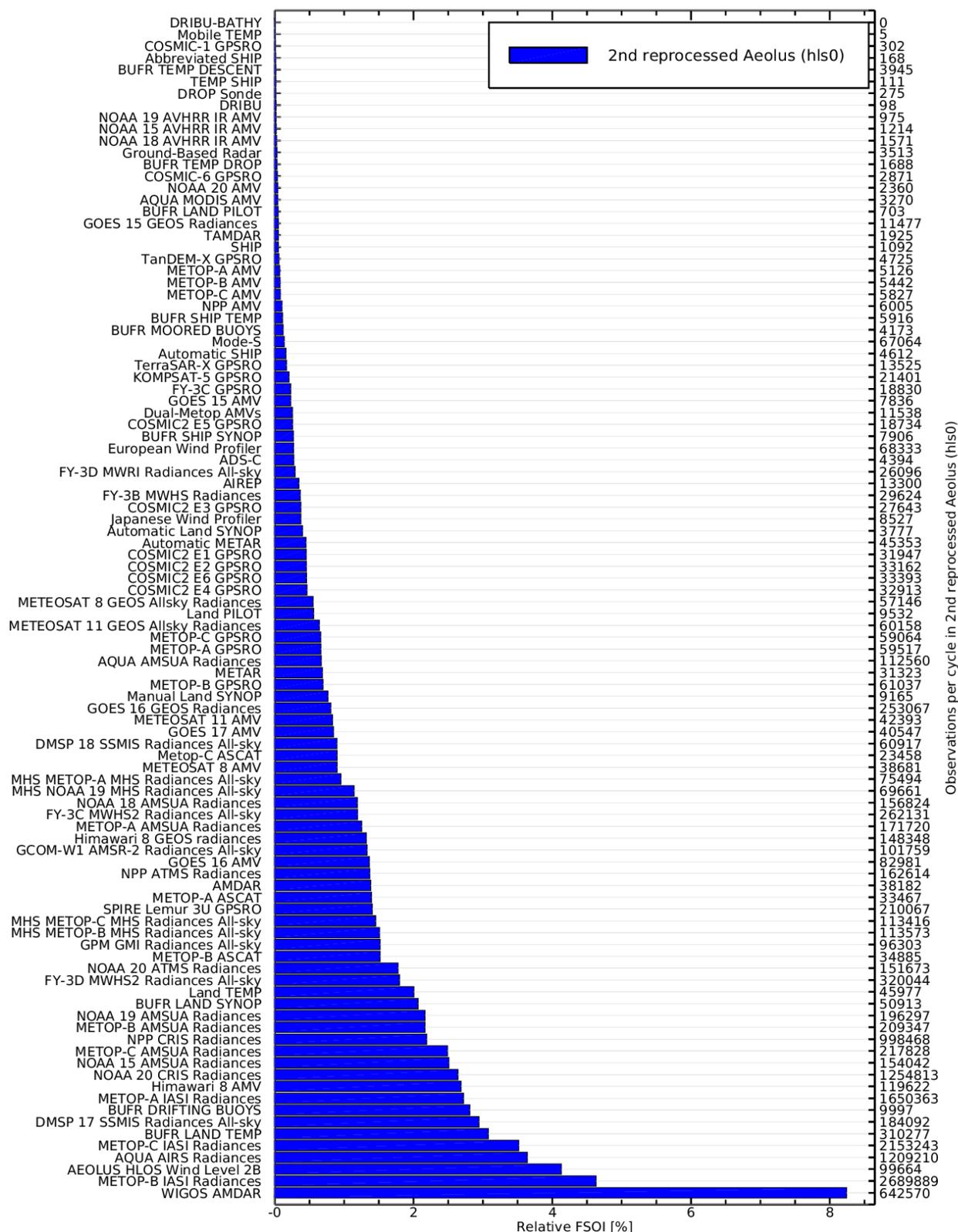


Figure 167. Relative FSOI (%) from the early FM-B period 2<sup>nd</sup> reprocessed dataset experiment, partitioned by instrument types, averaged over the period 3 July 2019 to 9 October 2020. Aeolus is listed as “AEOLUS HLOS Wind Level 2B” (both Rayleigh-clear and Mie-cloudy assimilated). The number of observations used on average per 12-hour assimilation window is shown on the right-hand side.

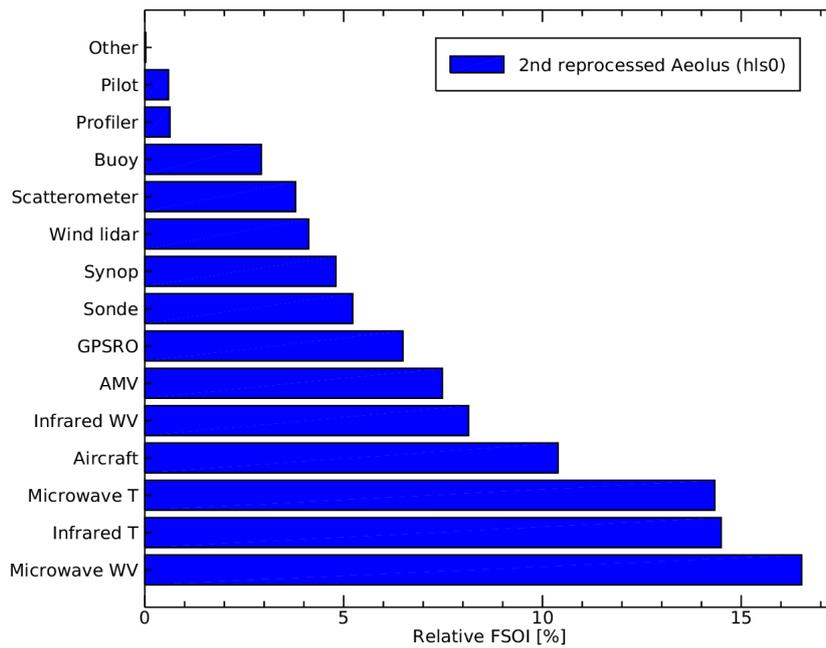


Figure 168. Relative FSOI (%) from the 2<sup>nd</sup> reprocessed dataset experiment, partitioned by observation groups, averaged over the period 3 July 2019 to 9 October 2020. Aeolus is the only member of the “Wind lidar” group. T stands for temperature, WV for water vapour.

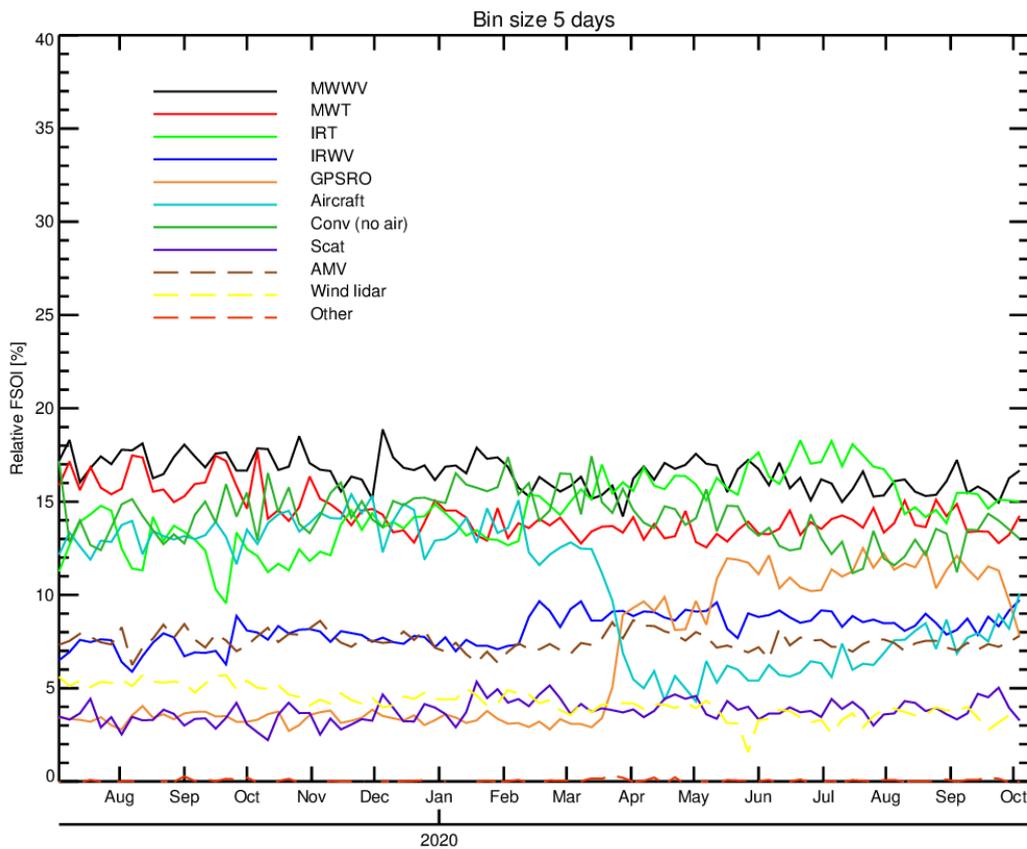


Figure 169. Relative FSOI (%) from 3 July 2019 to 9 October 2020 for the 2<sup>nd</sup> reprocessed dataset, with 5-day bin size. Aeolus L2B winds are labelled as “Wind lidar” (yellow). MW=microwave radiances, IR=infrared radiances, WV=water vapour, T=temperature, GPSRO=GPS radio occultation, Scat=scatterometer, AMV=Atmospheric Motion Vectors, Conv=conventional.

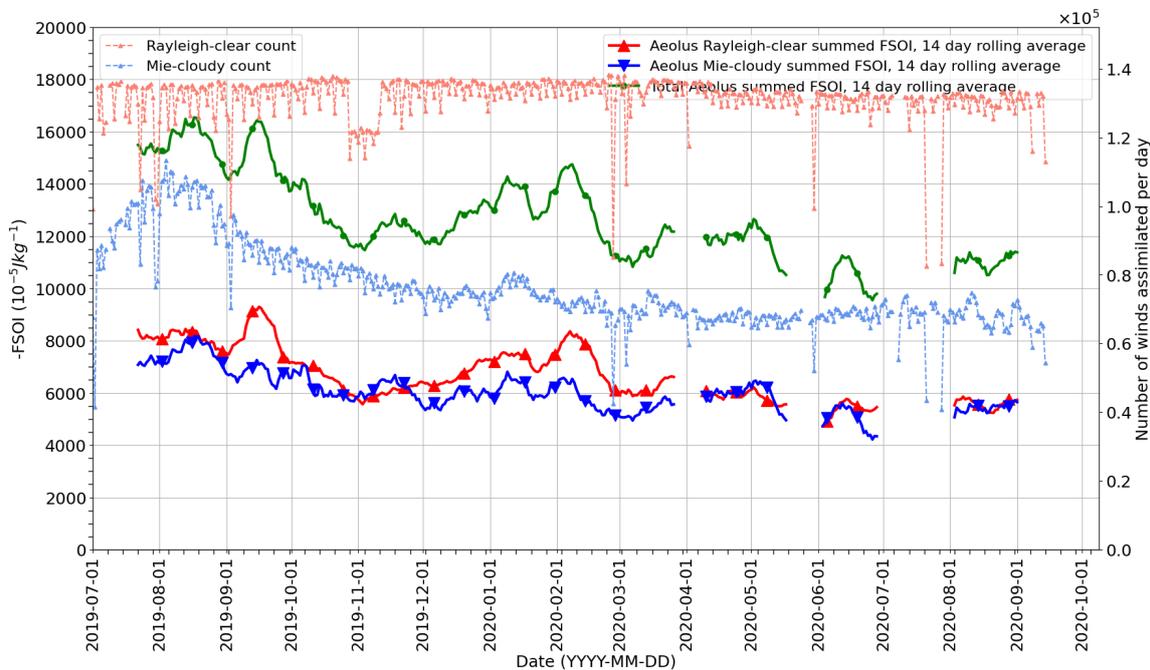
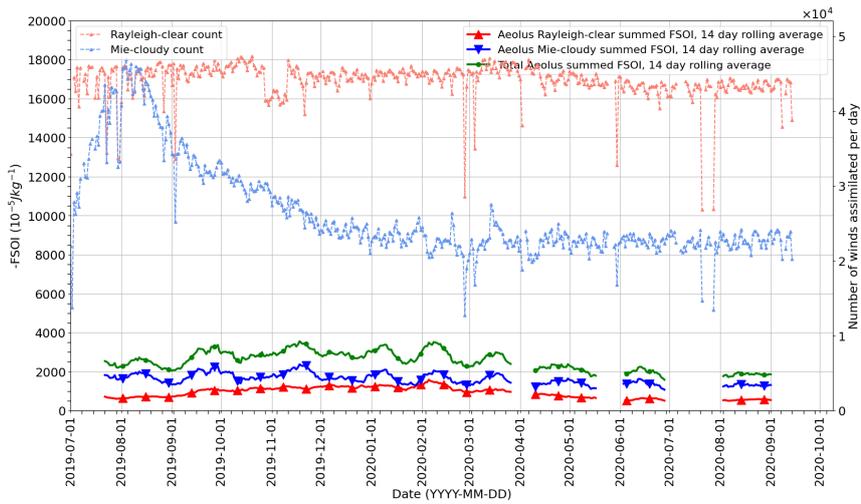
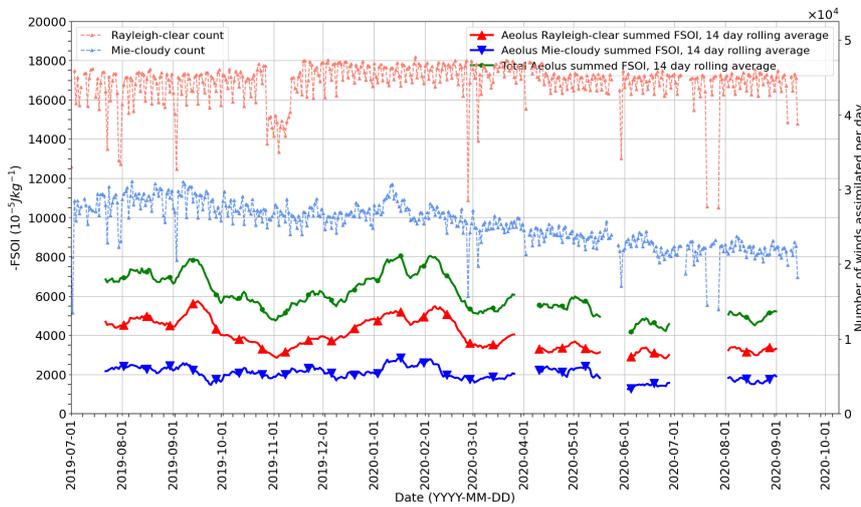


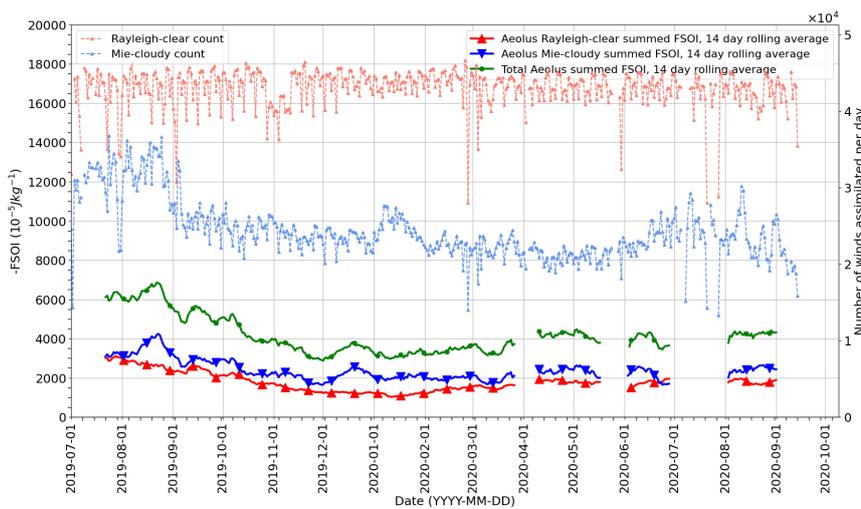
Figure 170. Time-series of global, daily summed FSOI for Aeolus L2B HLOS winds, for the period 3 July to 9 October 2020 (using the Aeolus “on” experiment from the 2<sup>nd</sup> reprocessed dataset OSE). A 2-week rolling average applied to reduce noise. The y-axis is the negative of the FSOI and therefore positive values represent positive impact. The data counts are shown on the right-hand side y-axis.



a) NH extratropics



b) Tropics



c) SH extratropics

Figure 171. As in Figure 170, but split into a) NH extratropics and b) tropics and c) SH extratropics.

## 5 Discussion

The ESA Aeolus mission requirements document (Ingmann and Straume, 2016) states a required precision of 2.5 m/s in the free troposphere and biases less than 0.7 m/s for the HLOS winds. These requirements were based on specific horizontal and vertical averaging scales and time-scales for the metrics, making it difficult to compare to error estimates for real Aeolus data. However, despite varying throughout the mission (and depending on estimates on ECMWF background random error magnitude), a precision of at best 4.0 m/s and 3.0 m/s for the L2B Rayleigh-clear and Mie-cloudy HLOS winds respectively (see section 3.7.1) is found in the free troposphere. These random errors are larger than the mission requirements, which therefore were not met.

The global average short-range forecast errors are estimated to be 2 m/s in HLOS wind space (although can reach much larger values in different areas and meteorological conditions; particularly convective outflow), so we are in a regime for the Rayleigh-clear where the instrument noise is dominant in the data assimilation, in comparison to other error components, such as representativeness error or observation operator error. The Mie-cloudy random error estimates are much closer to the mission requirements than the Rayleigh-clear in the free troposphere thanks to strong backscatter from ice and water clouds, meaning the random errors are limited by other effects. However, random errors for Mie winds from weakly backscattering and (sometimes) more attenuating aerosols still strongly depends on the emitted and received signal levels. The Mie-cloudy winds have larger representativeness error in the ECMWF model than the Rayleigh-clear due to their smaller horizontal scales and uncertainty about where the cloud is (vertically) within a range-bin. The larger noise of aerosol Mie winds have not been fully exploited so far, due to cautious quality control to remove gross errors, which also ends up removing a larger fraction of the Mie-aerosol winds. This will be improved with the 4<sup>th</sup> reprocessing.

The main reason for the mission requirements not being achieved is the lower-than-expected atmospheric path signal levels, as confirmed by DLR's and ESA's investigations for the Rayleigh channel (as communicated through DISC communications). It has been shown that the atmospheric path Rayleigh signal is around 2.7 times lower than expected for the 2018/2019 FM-A laser period, which when considering only useful signal shot-noise effects would make the Rayleigh wind random errors about 64% larger than otherwise. For early FM-B (e.g. July 2019) this factor was less, but still more than a factor of 2. Such a scaling factor would lead to 2.5 m/s times 1.64 i.e. 4.1 m/s which is in rough agreement with the precision estimate for early FM-B data. With lower signal, the solar background noise is a more dominant random error source for the Rayleigh-clear winds (see appendix 9.4), therefore making the random errors much larger for areas where Aeolus pointed into the sun-illuminated Earth. Also read-out noise and noise resulting from the Detection Chain Offset correction term becomes important with the low atmospheric path signals.

For both early FM-A (2018, 2019) and FM-B the atmospheric signal levels have dropped at large rates (as assessed by the DISC), and as is evident in the L2B Rayleigh-clear HLOS wind random error estimates in e.g. Figure 52. The cause of the lower-than-expected signal levels and ways to mitigate is not completely understood. For FM-A the atmospheric path signal drop was due to the laser itself reducing in output power – something that was better controlled for FM-A after switching it on again in November 2022. For FM-B an increasing large drop in transmission of signal after the laser transmitter was confirmed by ESA. The switch back to FM-A restored transmission levels, confirmed that some optical elements unique to FM-B's optical path are to blame for its transmission loss. This is unfortunate in that FM-B's laser energy could be raised to 182 mJ (in end-of-life testing) whereas FM-A's laser energy was limited to about 60 mJ. Lower signal levels also had negative impacts on the quality of the calibrations due to noisier ground returns, making the systematic error requirements more difficult to achieve (without NWP).

Despite the larger than expected Rayleigh-clear HLOS wind random errors and lower than expected Mie-cloudy coverage (e.g. few winds from boundary layer aerosol), **we have demonstrated that Aeolus provides a very useful positive impact in ECMWF's global NWP model**; when the signal levels are sufficiently large. It is likely that the continuous nature of the lidar curtain leads to some redundancy of information, hence larger random errors are probably not as detrimental as they could be if the profiles were all well separated in space and time (a benefit of continuous- rather than burst-mode laser operation). Also, Horányi et al. (2015b) showed that HLOS wind random error increases (via radiosondes) are not as damaging to NWP impact as perhaps expected (whereas bias is very damaging). The positive impact is probably assisted by the continuing lack of height-resolved global wind profile information in the Global Observing System.

It was demonstrated with the ECMWF data assimilation system by Horányi et al. (2015b) that using real HLOS winds (calculated from conventional wind vector observations) with artificial biases that are a large fraction of the standard deviation of observation error causes a considerable reduction in positive impact, and with sufficiently large bias an overall negative impact can result. Therefore, it is very important to correct biases in the data assimilation of Aeolus, or even better for them to be corrected by calibration procedures in the ground processing chain (preferably instrument-based calibrations).

During the mission the global average bias (accuracy) of Aeolus HLOS winds was observed to have significant drifts with time (for periods longer than a few days) due to instrumental drifts and the difficulty in obtaining high quality weekly calibrations for automated updates which would have compensated for such instrumental drifts. Global average Rayleigh-clear biases varied from being close to zero in some periods, to be several m/s HLOS in others due to the drifts. For example, for FM-B the global Rayleigh bias was drifting in the negative direction at a very high rate of around 1 m/s per month. The Mie-cloudy global average bias drifted during the FM-A period, but was, in contrast to the Rayleigh-clear, relatively stable for early FM-B. Fortunately we have found methods to correct such biases.

Very large Rayleigh-clear HLOS wind biases that vary on sub-orbital time-scales were discovered and found to be highly correlated with variations in temperature gradients across the telescope's primary (M1) mirror — particularly outer edge versus centre of the mirror. The telescope temperatures gradients are affected by the top of atmosphere radiation from the Earth, both short and long wave, and the mirror's thermal control mechanism, which aims to stabilise its temperature. Since the M1 mirror temperature gradients have a seasonal dependence, then the magnitude of this bias varies with the seasons, peaking in the Northern Hemisphere summer<sup>35</sup>. The Mie biases also have some dependence on M1 temperature gradients, but it appears to be around a factor 10 less than the Rayleigh channel and with the opposite sign.

With baseline 09, an M1 temperature dependent bias correction was applied using a new auxiliary files input: AUX\_TEL\_12 bias correction files (containing multiple linear regression parameters to be applied in the correction algorithm). This went operational on 20 April 2020 leading to much smaller biases (as shown in section 3.3.1) and it has been applied in reprocessing. However, it relies on using the ECMWF model as a reference (thus far). The AUX\_TEL\_12 generator (developed by DLR) uses multiple linear regression of the M1 thermistor temperatures versus ECMWF O-B to determine the bias correction coefficients. DLR (Fabian Weiler) showed it is possible to use ground returns as the reference (i.e. instrumental calibration), making it independent of the ECMWF model; however, the performance

---

<sup>35</sup> The biases are largest in the Northern Hemisphere summer due to the LOS pointing of Aeolus being towards the North Pole but away from the South Pole, therefore Aeolus points to brighter trop of atmosphere radiation conditions in the boreal summer.

is not as reliable as using the NWP model (something that would surely improve with greater atmospheric path signal).

For the early mission NRT data OSEs, to avoid the negative impact of bias, we initially chose to limit the CP OSE period to 40 days when the global average bias looked reasonably stable with time and could be easily corrected by a constant offset bias correction. However, in hindsight, as shown in section 3.2, biases did exist for early FM-A, varying with orbit phase angle (i.e. M1 temperature dependent) and wind speed (imperfect spectrometer response versus frequency calibration files). However, investigations comparing the 3<sup>rd</sup> reprocessing to the 2<sup>nd</sup> reprocessing have shown the sensitivity of biases to M1 T gradients are smaller with FM-A than FM-B (by up to factor 2). The CP NRT OSE was run beyond mid-October 2018 and the results confirmed our expectations (results not shown) of reduced impact. The medium range forecast impact gradually reduced as the experiment period is extended beyond 16 October 2018, which is thought to be due to the general increase in global average bias and due to the increasing number of hot pixels causing biased range-bins. The third reprocessing resolved such bias issues which led to a more consistent impact with time for the early FM-A period.

For the 2019 FM-A period OSEs using NRT data it was demonstrated that an ECMWF in-house bias correction scheme for the HLOS winds, using the ECMWF model as reference, significantly boosted the impact. The bias correction used mean O-B statistics over the previous week as a function of geolocation, in particular the orbit argument of latitude. Hence, this bias correction strategy was further refined for the early FM-B period experiments. The solution that gave the greatest impact for the Rayleigh-clear HLOS winds for the early FM-B period also included the longitude dimension to the bias correction look-up table (due to M1 temperature variations). It worked sufficiently well for Aeolus winds to go operational at ECMWF – it was applied in ECMWF operations from 9 January to May 2020. It was made redundant when the L2B processing M1 T bias correction entered the NRT processing chain. It has been demonstrated in the mid-2020 period that the improved scene-dependence of the M1 T bias correction relative to the weekly updated look-up table method provided significant improvement in the NWP impact, see section 4.4.3.1.

A surprise from the CP Aeolus OSE was that the impact in the tropics was rather small, and that the SH impact was stronger. Aeolus winds were expected to provide the largest impact in the tropics, based on dynamical arguments suggesting the importance of wind versus mass information due to the large Rossby radius of deformation in the tropics. This was also the expectation given the results using real wind observations in Horányi et al. (2015a). However, this limited tropical impact was in hindsight thought to have been caused by the uncorrected M1 temperature dependent biases. For the CP OSE in the tropics the Rayleigh winds in the descending orbit phase had a bias of around 1-2 m/s, whereas the bias in the ascending phase was small, as shown in Figure 22 a). This increased the strength of the on average easterly tropical winds, as indicated in Figure 56. The O-B bias was increased relative to tropical radiosonde zonal winds in the focus period (not shown). The third reprocessing covering early FM-A shows a positive impact in tropics which concurs with the hypothesis.

Another potential contributor to the lower CP tropical impact, was hinted at from the poorer fit to AMSU-A lower stratospheric temperature information channels in the tropics. We expected that winds would have most impact in the tropical upper-troposphere/lower-stratosphere (UTLS), an area which typically has large vertical wind shear. However, it seems likely that the 2 km thick range-bins in the tropical UTLS (above 14 km) during the CP, caused problems given that they are assimilated as point winds; perhaps creating dubious analysis increments. Ways to mitigate this include modifying the observation operator to a vertically averaging one, rejection of range-bins deemed to be too thick or

selecting thinner range-bins in NWP sensitive areas<sup>36</sup>. The negative impact verified by AMSU-A was particularly due to Rayleigh-clear winds as confirmed in Figure 57, perhaps because of the large sample of 2 km thick Rayleigh winds from the mostly clear air of the UTLS. The third reprocessing supports this hypothesis in that greater impact in the lower stratosphere occurred in the period after the RBS were modified to have better vertical resolution in the UTLS.

There are consistent geographical patterns in how Aeolus changes the analysis for different OSE periods: there are systematic changes in the tropical zonal winds (peaking at around 100-150 hPa) that persist throughout the forecasts, mostly a result of the Rayleigh-clear winds (also present when biases have improved via reprocessing). The systematic changes may be due to excessive parameterised vertical diffusion in the IFS in the lower stratosphere – meaning the model lacks vertical wind shear. The non-systematic changes in the wind field due to Aeolus are largest in areas with a high prevalence of mesoscale convective systems implying that Aeolus is correcting short-range wind forecast errors in areas where perhaps the model struggles to represent convection accurately or it is inherently unpredictable.

The Aeolus OSEs for all the periods tested so far have shown improvements in the short-range forecast fit to other observation types sensitive to wind, temperature and humidity. This is a very reliable demonstration that Aeolus is improving the analysis and short-range forecasts. The relative impact of Aeolus on short-range forecasts (versus observations) can be qualitatively compared to OSE denial experiments (Bormann et al., 2019). Aeolus' approximate 1-2% improvement in short-range forecast wind fit with respect to conventional wind data is similar in magnitude to that determined for GPSRO, infrared radiances and AMVs, which can be considered a good result for Aeolus when providing less than 1% of the observations assimilated.

The short-range forecast impact is also confirmed to be good via the FSOI results; performing well compared to other satellites from the existing observing system; the impact *per observation* is confirmed to be high for satellite data. The strongest performance was found in the 2<sup>nd</sup> reprocessed dataset early FM-B period (July-September 2019), for which Aeolus HLOS winds are the second most important instrument type, providing 5.2% relative FSOI. The operational FSOI statistics showed a continuing trend of decreasing impact; ~3-4% in early 2020 to ~2% in mid-2022; mostly due to atmospheric path signal decreasing. This has recovered to 3-4% in late 2022/early 2023 with the switch back to FM-A.

The consistent improvement in the fit to humidity sensitive observations is thought to be caused by the improvement in the advection of humidity. The reverse of the effect described for the all-sky radiance assimilation results (Geer et al., 2018) in which the assimilation of humidity and cloud sensitive observations during the 4D-Var window leads to wind increments at the start of the window, such that the humidity is advected better to improve the fit to humidity sensitive observations.

The impact of Aeolus on short-range to medium-range forecasts, as verified against analyses, shows consistent patterns for different periods, apart from 2021/2022 results (due to signal loss). That is, good magnitude positive impact has been demonstrated on the forecasts of vector wind, temperature and humidity (and hence geopotential height) for the polar regions and tropics. However, even in 2022 the impact remains positive in the tropics and at high latitudes at day 2-3, but with a smaller magnitude.

The best impact was found with the 2<sup>nd</sup> reprocessed (FM-B) dataset in the tropical upper troposphere; of order 1-2% improvement RMSE up to forecast ranges of around 9-10 days. It is the longest OSE for Aeolus. It also covers a long period of relatively high atmospheric path signal, which

---

<sup>36</sup> More favourable UTLS 1 km thick range-bin settings became the default on 26 February 2019.

explains the strong impact (see Figure 127). Running the model at higher resolution than previous OSEs also helped as did the length of the OSE (for statistical significance gains) and also improvement in the reprocessed data quality. The higher resolution model may allow wind data to be better exploited, since geostrophic adjustment theory states that winds are more important than mass observations at smaller horizontal scales. It is hoped that the removal of some remaining biases e.g. altitude-varying for Rayleigh-clear, remaining hot-pixel biases could further improve the impact if corrected in future reprocessing.

The impact from Rayleigh-clear winds was shown to be larger than Mie-cloudy winds in periods of relatively strong atmospheric signal in the tropics. The advantage for Rayleigh-clear decreased significantly when atmospheric useful signal level decreased, particularly for the FM-B period; and the FSOI suggests the Rayleigh-clear provides less impact than Mie-cloudy since late 2021. Even for periods with the largest signal levels, the Rayleigh-clear is noisier than Mie-cloudy, but the much superior atmospheric sampling for Rayleigh-clear winds compared to Mie-cloudy explains its larger impact. The Mie-cloudy impact was improved significantly by including a representativeness error term in the assigned observation error, as confirmed with early FM-B period and mid-2020 period OSEs.

The OSE results have tended to show greater impact from the Rayleigh-clear winds than the Mie-cloudy in the medium-range forecasts, whereas FSOI tends to show the Mie-cloudy having an impact on par with the Rayleigh-clear. It may be that the Mie-cloudy is able to constrain small-scale features very well due to its lower noise and limited spatial extent. Errors on smaller scales grow faster and saturate more quickly than the larger scales. The Rayleigh-clear may be constraining larger scales in the vertical and horizontal thus improving longer range forecasts. This may suggest that a constellation of DWLs could provide a lot more forecast skill by helping to correct large horizontal scales zonally rather than just along track.

The following is a list of ground processing chain updates which are thought to have improved the NWP impact of Aeolus Level-2B HLOS winds so far during the mission are:

- Increasing the number of Mie winds via reducing the L2Bp measurement grouping length-scale (improving horizontal resolution). This has been shown to lead to only a modest increase in random error but increases the number of Mie observations by a factor 2-3. See appendix 9.10 and 9.11 for details of changes in the grouping settings and N/P settings.
- The correction of the hot-pixel dark current offsets in the Level-1B processing step (since 14 June 2019), which massively reduced range-bin specific biases and avoids having to discard specific range-bins. With an increasing number of existing hot-pixels, and fluctuations of existing hot-pixels combined with low atmospheric path signals, it was necessary in September 2021 to increase the number of DUDEs per day from 4 to 9 (DUDEs are processed to DCMZ\_1B files, necessary for correcting hot-pixels).
- Better classification of Rayleigh measurement-bins into clear and cloudy, so that fewer are erroneously classified as cloudy when actually clear, thus reducing waste of clear air signal. This was done via tuning the classification thresholds; initially for L1B scattering ratio and then using the L1B Mie refined SNR for classification, which is less noisy, proved beneficial.
- Improved accounting for Mie backscatter signal on the Rayleigh channel, allowed for better quality Rayleigh-cloudy winds and hence more observations (an improvement in the way the AUX\_RBC is used).
  - In 2023 the switch-on a parameterised way of accounting for Mie contamination on the Rayleigh-cloudy winds via use of the L1B scattering ratio and using the ECMWF NWP model as a reference to derive the parameters; see Marseille et al. 2023.

- Selecting more favourable vertical sampling (range-bin settings) for NWP impact; this has been iterated on during the mission, following an ESA/DISC led advisory board. Highlights include higher-range bins in polar winter to capture the polar vortex; higher resolution in tropical UTLS; higher range-bins to capture Hunga-Tonga plume Mie winds; higher resolution in extratropics around jet stream level.
- The bias correction of L2B Rayleigh-clear winds using the telescope's M1 temperature information. This led to a substantial improvement in the quality of the L2B winds since 20 April 2020 and proved critical for more general uptake of the data for operational assimilation in NWP.
- More accurate Rayleigh HLOS wind instrument error estimates for the polar summer; via a correction in the L1B processor Rayleigh SNR estimate. A further improvement to account for readout noise with B16.
- Switching off the use of the internal path Rayleigh response in the L2B processor due to anomalous step changes in the internal Rayleigh response; first noticed to cause biases in December 2020.
- Use of NWP to help determine a more accurate Mie atmospheric path non-linearity correction and hence to improve the wind-speed dependent biases. The NWP derived non-linearity was applied to operational L2B winds on 1 July 2021; see Marseille et al. 2022
- Not strictly a ground processing chain change; but modifying the N/P settings from N=30 to N=15 on 13 December 2021 and then N=5, P=114 on 4 April 2022, which led to a significant improvement in the random error for Rayleigh-clear winds (15% and 17% respectively), mitigating the FM-B signal decrease. Most of the improvement was related to read-out noise reduction, an important noise source when the atmospheric path signal is low; however, the measurement scale Mie SNR was less noisy with longer averaging, which also helps L2Bp clear/cloudy classification.
- Reducing the number of gross error Mie-cloudy wind results which are flagged valid by use of the "residual error" Mie-core fit value for quality control. This was possible with B16. Has the potential to allow more Mie-aerosol winds in the data assimilation, since checks on L2Bp EE do not need to be as strict.

Some potential improvements to NWP impact with possible future improvements in the ground processing chain:

- Optimisation of the L2B processor settings (AUX\_PAR\_2B) more generally.
- Improved Rayleigh calibration methods as a function of temperature, pressure, scattering ratio and M1 temperature gradient; possibly with the aid of NWP assisted calibration (see Marseille 2020) and analysis of IRONICs data. Should also lead to better quality Rayleigh-clear and Rayleigh-cloudy wind (although another method has improved Rayleigh-cloudy already); possibly resolving the altitude dependent (temperature dependent) bias issues for Rayleigh-clear.
- Assessment of spatial observation error correlation.
- Improved quality control of anomalous measurement-bins due to cosmic rays, before accumulation. Potentially using the median-filter QC method of DLR (for airborne campaigns).

Potential improvements in Aeolus NWP impact via improved data assimilation methods:

- Better observation error modelling; improvements have already been achieved for the Mie-cloudy winds by accounting for a representativeness error.
- Improvements in observation operator, such as vertical/horizontal averaging of model winds to match Aeolus range-bin resolution; vertical wind; geometric height as vertical coordinate.
- Assessment as to whether spatial thinning of observations is beneficial e.g. near the Poles.
- Use of Rayleigh-cloudy winds from 4<sup>th</sup> reprocessing onwards.
- Further investigation of a stricter first-guess check QC, which has shown promising results in the third reprocessing OSE.

## 6 Conclusions

The world's first Doppler wind lidar in space, Aeolus, produced wind observations with sufficient quality to improve global Numerical Weather Prediction forecasts by a useful magnitude. This has been demonstrated with data quality assessment via comparisons of the Aeolus Level-2B (L2B) HLOS wind data with the ECMWF model equivalents and by NWP impact assessment with the OSE (Observing System Experiments) and the FSOI (Forecast Sensitivity Observation Impact) methods.

The estimated precision of Aeolus L2B HLOS winds varied considerably with geolocation, season, processing software version and range-bin settings, but typical values are 4-6 m/s for the Rayleigh-clear and 2.5-3.5 m/s for the Mie-cloudy. The observations have sufficient information content to improve the ECMWF forecasts despite the random and systematic errors being larger than pre-launch expectations.

OSEs show a positive impact for Aeolus in all periods of the mission. This impact has improved as the ground processing baseline has improved with the first and second reprocessing for the early FM-B laser period and third reprocessing for early FM-A. This was shown by the improved short-range forecast departure fits relative to the many trusted in situ and satellite observation types; decent FSOI scores in ECMWF operations and with reprocessed datasets, and by the improvement in forecasts out to the medium range verified against analyses.

The largest positive impact of Aeolus HLOS winds shown by all metrics is in the tropical upper troposphere; particularly for FM-B second reprocessing in 2019/2020. This period showed statistically significant and good magnitude positive impact on wind, temperature, geopotential and humidity forecasts in the tropical and polar troposphere and lower stratosphere by 0.5-2% in root mean square error up to 3-4 days. The positive impact extends to 9-10 days in tropics and Southern Hemisphere extratropics at 100 and 50 hPa (15-20 km). The peak in tropical impact at 100 hPa is in the east Pacific – a region typified by large background forecast zonal wind errors. The next best period of impact was the early FM-A period, with similar spatial patterns of positive impact.

Aeolus induced improvements in short-range forecasts were verified with respect to other observations sensitive to wind, temperature and humidity. They are of similar magnitude to other important satellite observing systems at ECMWF which is a positive result; considering Aeolus accounts for fewer than 0.5% of the observations assimilated. Rayleigh-clear winds provide most of the positive impact in OSEs and FSOI when atmospheric path signals are relatively strong, presumably due their much greater spatial coverage than Mie-cloudy winds. However, Mie-cloudy impact improved by more realistic assigned observation error modelling (including representativeness error). The Mie-cloudy impact overtook the Rayleigh-clear in 2022 when clear-air atmospheric path signals are low. Rayleigh-cloudy winds were shown to provide a small positive impact in the tropical winds at 500 hPa and in short-range verification against humidity sensitive observations. With improved data quality for future reprocessing expected then it may be worth considering assimilating them as default with the Rayleigh-clear and Mie-cloudy.

Consistent spatial patterns in how Aeolus modifies the analysis fields have been observed: the non-systematic changes tend to be largest in convective areas around the globe; suggesting Aeolus is mostly correcting random background forecast errors associated with convection. Very high altitude (1 hPa) non-systematic changes in the model wind field due to Aeolus seem to emanate from changes in the propagation for gravity waves from the troposphere to the stratosphere. The systematic changes are consistently largest in the tropical upper troposphere and lower stratosphere, and there also some consistent patterns towards the poles (with hints of model bias being to blame); comparisons to radiosondes suggest Aeolus is not degrading biases in statistics over the tropics.

The magnitude of Aeolus' positive impact was found to be strongly dependent on the implementation of a bias correction of the HLOS winds derived using the ECMWF model as a reference. The biases during the mission have been shown to have complicated behaviour (with geolocation and time), meaning it took a great effort by the DISC team to understand and correct them. Thankfully in Autumn 2019 an explanation for a significant contributor to the bias was found, via careful assessments of the ECMWF O-B statistics with reference to the satellite's housekeeping information. The explanation is that the Rayleigh biases are predominantly dependent on gradients in temperature across the ALADIN instruments primary mirror. The temperature gradients are thought to influence the receiver response via angular changes of the light onto the spectrometers, due to varying focus from the parabolic mirror. The variations in temperature gradients are driven by varying top of atmosphere radiation and the mirror's on-board thermal control mechanism. The sensitivity of the bias due mirror temperature gradients was greater for FM-B than for FM-A.

A bias correction method as a function of argument of latitude and longitude (look-up table) was found to work well enough using the ECMWF model as a reference. However, the more physically based bias correction using the NRT primary mirror thermistor reading as predictors was implemented in the operational ground processing chain since 20 April 2020 (and in reprocessed data), providing a more accurate bias correction. The mirror-temperature-dependent bias correction was found to improve the NWP impact of Aeolus compared to the geolocation-based method.

Given the evidence that Aeolus provides useful contribution to the global observing system it was decided to switch Aeolus on in operational data assimilation at ECMWF on 9 January 2020 and was assimilated until 30 April 2023 (due to end of mission). The operational assimilation initially used the geolocation-based look-up table bias correction method, but this could be switched off in May 2020 due to the successful operation of the M1 temperature based bias correction.

A univariate data assimilation calculation predicts that the Rayleigh-clear NWP impact should drop linearly with the relative drop in atmospheric signal (to zero impact with zero signal). This appears to be the case from OSEs throughout the mission, showing a reduced impact in the tropics when signals are low. This is borne out by the relative FSOI for the combination for Rayleigh-clear and Mie-cloudy which was ~5% in early FM-B reprocessed dataset in 2019 but is 2% in mid-2022, with the Mie-cloudy impact exceeding the Rayleigh-clear in 2022. This relative FSOI returned to 3-4% in late 2022/early 2023 with the return of FM-A. At its peak for the early FM-B period, Aeolus's relative FSOI is similar in magnitude to a MetOp IASI instrument, ranking amongst the highest FSOI per satellite instrument and it had a similar impact to the radiosonde network.

A fit to the relative FSOI of Aeolus versus assigned observation error for the Rayleigh-clear winds suggests that if the mission requirements random errors of around 2.5 m/s had been achieved, then the relative FSOI for Aeolus (Rayleigh-clear and Mie-cloudy) would have been around 6.5% making it the most important satellite instrument in the relative FSOI rankings. Global OSE impact metrics for mass and wind correlated against the varying Rayleigh-clear observation error of the mission suggest a doubling of impact if random errors closer to 2.5 m/s (profile average) rather than the more typical 6 m/s were achieved. The results are promising for the proposed Aeolus follow-on mission (EPS-Aeolus) which aims for improved precision near these levels.

## 7 Acknowledgements

This research is funded by the European Space Agency (ESA) in the framework of the activities of the Aeolus Data Innovation and Science Cluster (DISC) consortium. The presented work includes some preliminary data (not fully calibrated/validated and not yet publicly released) of the Aeolus mission that is part of the European Space Agency (ESA) Earth Explorer Programme.

Thanks to the many experts from the DISC and from ESA (particularly Oliver Reitebuch (DLR), Thomas Kanitz (ESA), Jos de Kloe (KNMI), Fabian Weiler (DLR), Thomas Flament (Météo-France), Gert-Jan Marseille (KNMI), Saleh Abdalla (ECMWF), Giovanna De Chiara (ECMWF), Dorit Huber (DoRIT), Benjamin Witschas (DLR), Uwe Marksteiner (DLR) and Olli Lux (DLR) which provided a fruitful forum for discussion during the Aeolus mission so far. Thanks to Mohamed Dahoui (ECMWF) for advice and support on observation monitoring and to Sean Healy (ECMWF) for encouragement and suggestions on what to research in 2020-2022. Thanks to Alan Geer (ECMWF) for providing the iver FSOI results for 2020-mid-2022.

## 8 References

- Abdalla, S., L. Isaksen, Peter A. E. M. Janssen, Nils Wedi, Effective spectral resolution of ECMWF atmospheric forecast models, 2013, ECMWF Newsletter, No. 137, Autumn 2013, pp. 19–22.
- Abdalla, S., de Kloe, J., Flament, T., Krisch, I., Marksteiner, U., Reitebuch, O., Rennie, M., Weiler, F., Witschas, B., 2020, Verification report of first Reprocessing campaign for FM-B covering the time period 2019-06 to 2019-12, 2020, DISC Technical Note, AED-TN-ECMWF-GEN-04, v1.0
- Abdalla, S., de Kloe, J., Flament, T., Krisch, I., Marksteiner, U., Reitebuch, O., Rennie, M., Weiler, F., Witschas, B., Verification report of the second reprocessing campaign for FM-B from 24 June 2019 till 9 October 2020, 2021, DISC Technical Note, AED-TN-ECMWF-GEN-60, v1.1
- Abdalla, S., de Kloe, J., N. Masoumzadeh, Krisch, I., Marksteiner, U., Reitebuch, O., Rennie, M., Weiler, F., Witschas, B., Verification report of the third reprocessing campaign for first FM-A period from September 2018 till June 2019, 2023, DISC Technical Note, AED-TN-ECMWF-GEN-80, v1.0
- Andersson, E., Hollingsworth, A., Kelly, G., Lönnberg, P., Pailleux, J. and Zhang, Z. (1991) Global observing system experiments on operational statistical retrievals of satellite sounding data. *Monthly Weather Review*, 119, 1851– 1864.
- Baker, WE, Atlas, RM, Cardinali, C, Clement, A, Emmitt, GD, Gentry, BM, Hardesty, RM, Källén, E, Kavaya, MJ, Langland, R, Ma, Z, Masutani, M, McCarty, W, Pierce, RB, Pu, Z, Riishøjgaard, LP, Ryan, J, Tucker, S, Weissmann, M, Yoe, JG. 2014. Lidar-measured wind profiles: the missing link in the global observing system. *Bulletin of the American Meteorological Society* 95: 4: 543–564,
- Bell, W., B. Candy, N. Atkinson, F. Hilton, N. Baker, N. Bormann, G. Kelly, M. Kazumori, W.F. Campbell and S.D. Swadley. (2008). “The Assimilation of SSMIS Radiances in Numerical Weather Prediction Models”. *IEEE Transactions on Geoscience and Remote Sensing* 46(4): 884-900. <https://doi.org/10.1109/TGRS.2008.917335>.
- Bormann, N., H. Lawrence and J. Farnan, 2019: Global observing system experiments in the ECMWF assimilation system, ECMWF technical Memorandum, 839
- Cardinali, C., 2009: Monitoring the observation impact on the short-range forecast. *Q. J. R. Meteorol. Soc.* 135: 239-250.
- Desroziers, G., Berre, L., Chapnik, B. and P. Poli, 2005: Diagnostics of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc.*, 131, 3385-3396.
- ECMWF, 2018: IFS Documentation CY45R1. Available from ECMWF: <https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/ifs-documentation>
- De Chiara, G., Magnusson, L., Rennie, M., Abdalla, S., and S. Healy, *The Contribution of Aeolus to the Prediction of Extreme Weather Events*. ESA project final report (ESA ESTEC contract No. 4000134449/20/NL/FF/tfd). June 2023.
- Engelmann, R., Ansmann, A., Ohneiser, K., Griesche, H., Radenz, M., Hofer, J., Althausen, D., Dahlke, S., Maturilli, M., Veselovskii, I., Jimenez, C., Wiesen, R., Baars, H., Bühl, J., Gebauer, H., Haarig, M., Seifert, P., Wandinger, U., and Macke, A.: Wildfire smoke, Arctic haze, and aerosol effects on mixed-phase and cirrus clouds over the North Pole region during MOSAiC: an introduction, *Atmos. Chem. Phys.*, 21, 13397–13423, <https://doi.org/10.5194/acp-21-13397-2021>, 2021.

ESA, 2008: “ADM-Aeolus Science Report”, SP-1311,'ADM-Aeolus', April 2008, ISBN 978-92-9221-404-3, ISSN 0379-6566, 130 pp, available at <https://earth.esa.int/documents/10174/1590943/AEOL002.pdf>

P. Ingmann and A.G. Straume, 2016: ADM-Aeolus Mission Requirements Document, November 216, AE-RP-ESA-SY-001

ESA, 2019: “The Living Planet Programme”, available at [https://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/The\\_Living\\_Planet\\_Programme](https://www.esa.int/Our_Activities/Observing_the_Earth/The_Living_Planet_Programme)

Feng, Z., Leung, L. R., Liu, N., Wang, J., Houze, R. A., Li, J., et al. (2021). A global high-resolution mesoscale convective system database using satellite-derived cloud tops, surface precipitation, and tracking. *Journal of Geophysical Research: Atmospheres*, 126(8), e2020JD034202. <https://doi.org/10.1029/2020JD034202>

Frost, C. and S. Thompson (2000). "Correcting for regression dilution bias: comparison of methods for a single predictor variable." *Journal of the Royal Statistical Society Series A* 163: 173–190.

Geer, A. J., K. Lonitz, P. Weston, M. Kazumori, K. Okamoto, Y. Zhu, E. H. Liu, A. Collard, W. Bell, S. Migliorini, P. Chambon, N. Fourrié, M-J Kim, C. Köpken-Watts, C. Schraff, 2018: All-sky satellite data assimilation at operational weather forecasting centres. *Q. J. R. Meteorol. Soc.*, 144, 1191-1217. <https://doi.org/10.1002/qj.3202>

Healy, S., Eyre, J., Hamrud, M., Thépaut, J-N., 2006, Assimilating GPS radio occultation measurements with two-dimensional bending angle observation operators, EUMETSAT/ECMWF Fellowship Programme Research Reports, Available from: <https://www.ecmwf.int/node/9815>

Hollingsworth, A., D.B. Shaw, P. Lönnberg, L. Illari, K. Arpe, and A.J. Simmons, 1986: Monitoring of Observation and Analysis Quality by a Data Assimilation System. *Mon. Wea. Rev.*, 114, 861–879, [https://doi.org/10.1175/1520-0493\(1986\)114](https://doi.org/10.1175/1520-0493(1986)114)

Horányi, A., Cardinali, C., Rennie, M. and Isaksen, L. (2015a) The assimilation of horizontal line-of-sight wind information into the ECMWF data assimilation and forecasting system. Part I: the assessment of wind impact. *Quarterly Journal of the Royal Meteorological Society*, 141, 1223–1232. <https://doi.org/10.1002/qj.2430>

Horányi, A., Cardinali, C., Rennie, M. and Isaksen, L. (2015b) The assimilation of horizontal line-of-sight wind information into the ECMWF data assimilation and forecasting system. Part II: The impact of degraded wind observations. *Quarterly Journal of the Royal Meteorological Society*, 141, 1233– 1243. <https://doi.org/10.1002/qj.2551>

Houchi, K., Stoffelen, A., Marseille, G. and De Kloe, J. (2010) Comparison of wind and wind shear climatologies derived from high-resolution radiosondes and the ECMWF model. *Journal of Geophysical Research: Atmospheres*, 115, D22123. <https://doi.org/10.1029/2009JD013196>.

Illingworth, A. J., Battaglia, A., Bradford, J., Forsythe, M., Joe, P., Kollias, P., Lean, K., Lori, M., Mahfouf, J.-F., Mello, S., Midthassel, R., Munro, Y., Nicol, J., Potthast, R., Rennie, M., Stein, T. H. M., Tanelli, S., Tridon, F., Walden, C. J. and Wolde, M. (2018) WIVERN: A new satellite concept to provide global in-cloud winds, precipitation and cloud properties. *Bulletin of the American Meteorological Society*, 99 (8). pp. 1669-1687. ISSN 1520-0477

Ingmann, P. and Straume, A. G.: ADM-Aeolus Mission Requirements Document, ESA, 57 pp., [https://esamultimedia.esa.int/docs/EarthObservation/ADM-Aeolus\\_MRD.pdf](https://esamultimedia.esa.int/docs/EarthObservation/ADM-Aeolus_MRD.pdf) (last access: 9 August 2023), 2016.

- Isaksen, L, Bonavita, M, Buizza, R, Fisher, M, Haseler, J, Leutbecher, M and L. Raynaud, 2010: Ensemble of data assimilations at ECMWF. ECMWF Technical Memorandum 646. Available from ECMWF <https://www.ecmwf.int/en/elibrary>
- Janiskova M. and C. Cardinali, 2016: On the Impact of the Diabatic Component in the Forecast Sensitivity Observation Impact Diagnostics, ECMWF Tech. Memo., 786,18pp
- Järvinen, H., Andersson, E. and F. Bouttier, 1999: Variational assimilation of time sequences of surface observations with serially correlated errors. *Tellus A*, 51, Issue4, 469-488.
- Korhonen, Tapio, Perttu Keinanen, Mikko Pasanen, and Aimo Sillanpaa. (2008). "Polishing and Testing of the 1.5 m SiC M1 Mirror of the ALADIN Instrument on the ADM-Aeolus Satellite of ESA". In *Optical Fabrication Testing, and Metrology III*, edited by Angela Duparré and Roland Geyl. SPIE. <https://doi.org/10.1117/12.797730>.
- Langland, RH and NL Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus A*, 56, Issue3, 189-201.
- Q. Lu, W. Bell, P. Bauer, N. Bormann, and C. Peubey. Characterizing the FY-3A microwave temperature sounder using the ECMWF model. *J. Atmos. Oceanic Technol.*, 28:1373–1389, 2011a.
- Ma, Z, Riishøjgaard, LP, Masutani, M, Woollen, JS, Emmitt, GD. 2015. Impact of different satellite wind Lidar telescope configurations on NCEP GFS forecast skill in observing system simulation experiments. *Journal of Atmospheric and Oceanic Technology* 32: 3: 478– 495,
- Malardel, S, Wedi, N, Deconinck, W, Diamantakis, M, Kuehnlein, C, Mozdzyński, G, Hamrud, M, Smolarkiewicz, P, A new grid for the IFS, 2016, ECMWF Newsletter 146, <https://www.ecmwf.int/node/17262>
- Marseille, GJ, Stoffelen, A, Barkmeijer, J. 2008. Impact assessment of prospective space-borne Doppler wind lidar observation scenarios. *Tellus A* 60: 234– 248,
- Marseille, GJ, 2020, NWP Calibration, DISC Technical Note, Ref: AED-TN-KNMI-NWP-034
- Marseille, G., de Kloe, J., Marksteiner, U., Reitebuch, O., Rennie, M., and de Haan, S.: NWP calibration applied to Aeolus Mie channel winds, *QJRMS*, 2022, <https://doi.org/10.1002/qj.4244>
- Marseille, G., de Kloe, J., Dabas, A., Flament, T., and Rennie, M.: Aeolus Rayleigh-channel winds in cloudy conditions, *QJRMS*, 2023, *accepted but not yet published*
- Martin, A., Weissmann, M., Reitebuch, O., Rennie, M., Geiß, A., and Cress, A.: Validation of Aeolus winds using radiosonde observations and numerical weather prediction model equivalents, *Atmos. Meas. Tech.*, 14, 2167–2183, <https://doi.org/10.5194/amt-14-2167-2021>, 2021.
- McNally, A.P. (2019) On the sensitivity of a 4D-Var analysis system to satellite observations located at different times within the assimilation window. *Quarterly Journal of the Royal Meteorological Society*, 145, 2806– 2816. <https://doi.org/10.1002/qj.3596>
- Podglajen, A., Hertzog, A., Plougonven, R. and N. Žagar, 2014: Assessment of the accuracy of (re)analyses in the equatorial lower stratosphere. *J.Geoph.Res.*, 119, Issue19, 11,166-11,188. <https://doi.org/10.1002/2014JD021849>
- Polichtchouk, I., Bechtold, P., Bonavita, M., Forbes, R., Healy, S., Hogan, R., Laloyaux, P., Rennie, M., Stockdale, T., Wedi, N., Diamantakis, M., Flemming, J., English, S., Isaksen, L., Vána, F., Gisinger, S. and Byrne, N. (2021). Stratospheric modelling and assimilation. <https://www.ecmwf.int/node/19875>.
- Rennie, M. (2016): Advanced monitoring of Aeolus winds, ESA Technical Note 16, AE-TN-ECMWF-GS-16, v1.3, <https://www.ecmwf.int/node/18014>

- Rennie, M., and L. Isaksen. (2020). “The NWP Impact of Aeolus Level-2B Winds at ECMWF”. ECMWF Technical Memoranda 864. <https://dx.doi.org/10.21957/alift7mhr>
- Rennie, M., Isaksen, L., Weiler, F., De Kloe, J., Kanitz, T., and Reitebuch, O.: The impact of Aeolus wind retrievals on ECMWF global weather forecasts, QJRMS, 2021, <https://doi.org/10.1002/qj.4142>
- Rennie, M., Healy, S., Abdalla, S., McLean, W., and Henry K.: Aeolus positive impact on forecasts with the second reprocessed dataset, 2022, ECMWF Newsletter, No. 173, Autumn 2022, pp. 14–20. doi: 10.21957/mr5bjs29fa
- Rodwell, M. J., and Coauthors, 2013: Characteristics of Occasional Poor Medium-Range Weather Forecasts for Europe. Bull. Amer. Meteor. Soc., 94, 1393–1405, <https://doi.org/10.1175/BAMS-D-12-00099.1>.
- Ruppert, D. (2010). Statistics and Data Analysis for Financial Engineering. Springer. p. 118. ISBN 9781441977878. Retrieved 2015-08-27
- Ruston, B, Healy, S. (2020). “Forecast Impact of FORMOSAT-7/COSMIC-2 GNSS Radio Occultation Measurements”. Atmos Sci Lett. <https://doi.org/10.1002/asl.1019>
- Stoffelen, A., 1999: A Simple Method for Calibration of a Scatterometer over the Ocean. J. Atmos. Oceanic Technol., 16, 275–282, [https://doi.org/10.1175/1520-0426\(1999\)016](https://doi.org/10.1175/1520-0426(1999)016)
- Stoffelen, A, Marseille, GJ, Bouttier, F, Vasiljevic, D, de Haan, S, Cardinali, C. 2006. ADM-Aeolus Doppler wind lidar observing system simulation experiment. Quarterly Journal of the Royal Meteorological Society 132: 1927– 1947, <https://doi.org/10.1256/qj.05.83>.
- Tan, DGH, Andersson, E, Fisher, M, Isaksen, L. 2007. Observing-system impact assessment using a data assimilation ensemble technique: application to the ADM-Aeolus wind profiling mission. Quarterly Journal of the Royal Meteorological Society 133: 381– 390, <https://doi.org/10.1002/qj.43>
- Tan, DGH, Andersson, E, De Kloe, J, Marseille, GJ, Stoffelen, A, Poli, P, Denneulin, ML, Dabas, A, Huber, D, Reitebuch, O, Flamant, P, Le Rille, O, Nett, H. 2008. The ADM-Aeolus wind retrieval algorithms. Tellus A 60: 2: 191– 205, <https://doi.org/10.1111/j.1600-0870.2007.00285.x>.
- Weiler, F., Kanitz, T., Wernham, D., Rennie, M., Huber, D., Schillinger, M., Saint-Pe, O., Bell, R., Parrinello, T., and Reitebuch, O.: Characterization of dark current signal measurements of the ACCDs used on board the Aeolus satellite, Atmos. Meas. Tech., 14, 5153–5177, <https://doi.org/10.5194/amt-14-5153-2021>, 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, Atmos. Meas. Tech., 14, 7167–7185, <https://doi.org/10.5194/amt-14-7167-2021>, 2021.
- Weissmann, M., and C. Cardinali, 2007: Impact of airborne Doppler lidar observations on ECMWF forecasts. Quart. J. Roy. Meteor. Soc., 133, 107–116.
- Wright, C. J., Hall, R. J., Banyard, T. P., Hindley, N. P., Krisch, I., Mitchell, D. M., and Seviour, W. J. M.: Dynamical and surface impacts of the January 2021 sudden stratospheric warming in novel Aeolus wind observations, MLS and ERA5, Weather Clim. Dynam., 2, 1283–1301, <https://doi.org/10.5194/wcd-2-1283-2021>, 2021.
- WMO, 2018: “Rolling Review of Requirements and Statements of Guidance”, available at <http://www.wmo.int/pages/prog/www/OSY/GOS-RRR.html>

Žagar, N., Rennie, M., and Isaksen, L., Uncertainties in Kelvin Waves in ECMWF Analyses and Forecasts: Insights from Aeolus Observing System Experiments, 2021, *Geophys. Res. Lett.*, 48, <https://doi.org/10.1029/2021GL094716>

## 9 Appendix

### 9.1 Acronyms

Acronym	Definition
4D-Var	Four Dimensional Variational data assimilation
ACCD	Accumulation charge coupled device
ALADIN	Atmospheric LAser Doppler Instrument
AMDAR	Aircraft Meteorological DAta Relay
AMSR-2	Advanced Microwave Scanning Radiometer 2
AMSU-A	Advanced Microwave Sounding Unit-A
AMV	Atmospheric Motion Vector
AOCS	Attitude and Orbital Control Systems
ASCAT	Advanced SCATterometer
ATBD	Algorithm Theoretical Baseline Document
ATMS	Advanced Technology Microwave Sounder
BRC	Basic Repeat Cycle
BUFR	Binary Universal Form for the Representation of meteorological data
CAL/VAL	Calibration and Validation
CP	Commissioning Phase
CrIS	Cross-track Infrared Sounder
CSR	Corrected Spectral Registration
DEM	Digital Elevation Model
DISC	Data Innovation Science Cluster
DLR	Deutsches Zentrum für Luft- und Raumfahrt

DS	Defence and Space
DUDE	Down Under Dark Experiment
DWL	Doppler Wind Lidar
ECMWF	European Centre for Medium-range Weather Forecasts
EDA	Ensemble of Data Assimilations
EE	Earth Explorer
EGM96	Earth Gravitational Model 96
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FDIR	Failure Detection, Isolation and Recovery
FM	Flight model
FSO	Forecast Sensitivity Observation
FY3B	Feng-Yun - 3B satellite
GOS	Global Observing System
GNSSRO	Global Navigation Satellite System Radio Occultation
GPSRO	Global Positioning System Radio Occultation
HLOS	Horizontal line-of-sight
IASI	Infrared Atmospheric Sounding Interferometer
IFS	Integrated Forecasting System (at ECMWF)
IRC	Instrument Response Calibration
IRONIC	Instrument Response Off-Nadir Calibration
ISR	Instrument Spectral Registration

KNMI	Koninklijk Nederlands Meteorologisch Instituut (The Netherlands Royal Meteorological Institute)
L1B	Level 1B
L1Bp	Level 1B processor
L2B	Level 2B
L2Bp	Level 2B processor
L2/Met PF	Level 2 Meteorological processing facility
L2C	Level 2C
LOS	Line-of-sight
LTAN	Local Time at Ascending Node
MAD	Median absolute deviation
MDA	MacDonald, Dettwiler and Associates
MetOp	Series of EUMETSAT polar orbiting satellites
MHS	Microwave Humidity Sounder
MRC	Mie Response Calibration
NH	Northern Hemisphere
NPP	National Polar-orbiting Partnership
NRT	Near real-time
NWP	Numerical Weather Prediction
OSE	Observing System Experiment
O-A	Observation minus analysis
O-B	Observation minus background
ODB	Observational DataBase (used by ECMWF)
PDGS	Payload Data Ground Segment
QBO	Quasi-Biennial Oscillation

QC	Quality control
RALT	Radar altimeter
RBC	Rayleigh-Brillouin Correction
RBS	Range bin settings
RD	Research department (at ECMWF)
RMS	Root mean square
RMSE	RMS error
RRC	Rayleigh Response Calibration
SH	Southern Hemisphere
SNR	Signal to Noise Ratio
SOR	Special Operations Request
TN	Technical Note
UTLS	upper-troposphere/lower-stratosphere
UV	Ultra-violet
VarQC	Variational Quality Control
WIGOS	WMO Integrated Global Observing System
WMO	World Meteorological Organisation
ZWC	Zero Wind Correction

## 9.2 Estimating metrics of Aeolus observation errors from O-B departure statistics

This section derives how estimates of observation error statistics can be derived from O-B statistics. A state-of-the-art data assimilation system and forecast model (such as ECMWF's) is a powerful tool for monitoring and evaluating a new observing system. The forward modelled observation equivalent (B) (also known as background) can be compared to the observation (O). O-B departure (or innovation) statistics provide valuable information on the quality of the observations.

Assuming  $O = T + \epsilon_O$  and  $B = T + \epsilon_B$ , where  $T$  is the true HLOS wind,  $\epsilon_O$  is the observation error and  $\epsilon_B$  is the background error. Then the variance and expectation of innovation statistics can, with some assumptions, be estimated as:

$$\text{Var}(O - B) = \text{Var}(\epsilon_O) + \text{Var}(\epsilon_B) - 2\text{Cov}(\epsilon_O, \epsilon_B)$$

$$E(O - B) = E(\epsilon_O) - E(\epsilon_B)$$

The observation random error estimate (standard deviation,  $\sigma$ ) is as follows (assuming the covariance of observation and background errors is negligible).

$$\sigma(\epsilon_O) = \sqrt{\text{Var}(\epsilon_O)} \approx \sqrt{\text{Var}(O - B) + \text{Var}(\epsilon_B)}$$

The observation systematic error (bias) estimate is as follows (assuming that the expectation of background error is relatively small compared to the observation bias).

$$E(\epsilon_O) = E(O - B) + E(\epsilon_B) \approx E(O - B)$$

This observation random error estimate requires knowledge of the background error variance,  $\text{Var}(\epsilon_B) = \sigma^2(\epsilon_B)$ . The ECMWF background is of course not the truth, but over large spatial scales and long-time averages it has high level of accuracy and good precision. An estimate for  $\sigma(\epsilon_B)$  was derived using "Desroziers diagnostics" (Desroziers et al. 2005) using radiosonde zonal wind component departures and Aeolus HLOS wind retrieval departures. This led to a typical range for  $\sigma(\epsilon_B)$  from 1.5 to 2.5 m/s in HLOS wind space, varying with atmospheric pressure (altitude) and geolocation. Due to the variability of  $\sigma(\epsilon_B)$ , we have used this range to define confidence intervals for the  $\sigma(\epsilon_O)$  estimates.

The maximum global average bias of radiosondes zonal and meridional wind component,  $E(O - B)$ , is approximately 0.3 m/s (measured for several periods), therefore we assume this represents the maximum possible global average background forecast bias over larger samples. Note that there are certain ephemeral model conditions which have large wind errors, e.g. larger than 15 m/s, found compared to Aeolus in the outflow from mesoscale convective systems, and winds in the tropics (e.g. Magnusson 2017).

## 9.3 Simple scalar data assimilation analysis for assessing Aeolus Rayleigh wind impact

According to basic data assimilation theory (for a scalar or univariate analysis: one observation, one model state variable) the analysis error ( $\sigma_A$ ) is a function of the observation error ( $\sigma_O$ ) and background error ( $\sigma_B$ ) as follows e.g. see:

[https://events.ecmwf.int/event/153/contributions/900/attachments/463/852/TC\\_lecture\\_1.pdf](https://events.ecmwf.int/event/153/contributions/900/attachments/463/852/TC_lecture_1.pdf)

$$\sigma_A = \sqrt{\frac{\sigma_O^2 \sigma_B^2}{\sigma_O^2 + \sigma_B^2}}$$

If shot-noise dominates then the Rayleigh-clear HLOS wind observation standard error depends on laser energy (or equivalently atmospheric path signal level) as follows (see formula in section 9.4 as the solar background term tends to zero):

$$\sigma_O = \sigma_{HLOS} = \frac{d}{\sqrt{A}}$$

Where  $A$  = signal level (or laser energy) and  $d$  = a constant

Given we have estimates of  $\sigma_O$  for a given signal level (or laser energy) from the real L2B Rayleigh-clear winds, we can determine  $c$  and plot the curve, as is shown in Figure 172 (using an estimate of error for August 2020 in the tropics).

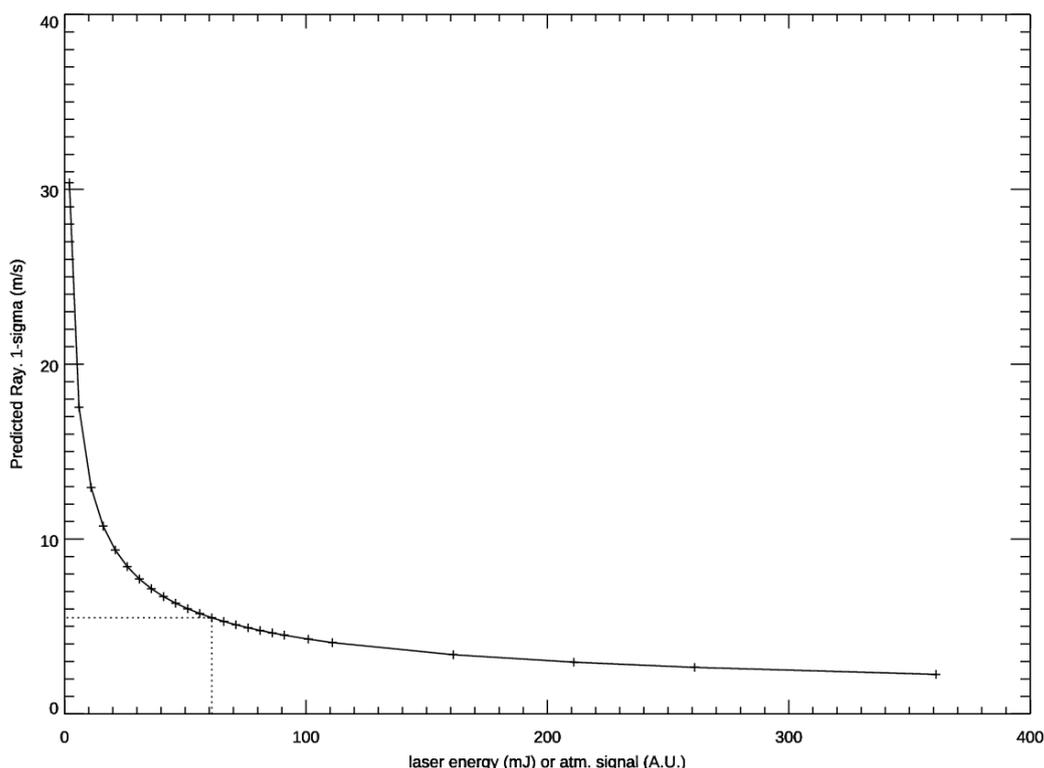


Figure 172. Expected behaviour of L2B Rayleigh-clear profile average observation error  $\sigma_O$  with laser energy (or equivalently atmospheric signal level) as calibrated for August 2020. The dotted line indicates the calibration point; 61 mJ with 5.5 m/s profile average random error (in the tropics). This calculation does not consider solar background noise (which is appropriate in the tropics given Aeolus' orbit).

Given this estimate of the observation error and the background error, we can plot the percentage relative reduction in analysis error ( $\sigma_A$ ) compared to background error ( $\sigma_B$  assumed to be 2 m/s in the tropics) as the signal level varies, as shown in Figure 173.

$$Relative\_NWP\_impact = 100\left(1 - \frac{\sigma_A}{\sigma_B}\right)$$

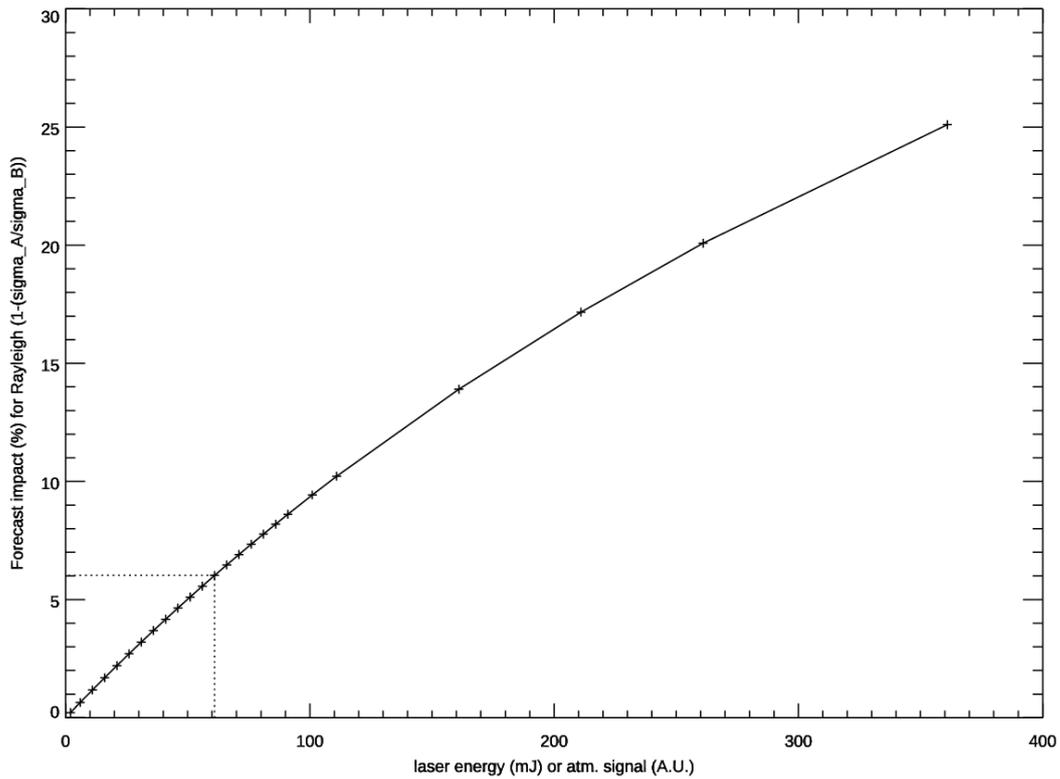


Figure 173. Relative NWP impact (relative reduction in analysis error compared to background error) assuming a  $\sigma_B$  of 2 m/s and a  $\sigma_O$  of 5.5 m/s with 61 mJ.

For a simple scalar analysis, given our typical observation and background errors in August 2020, we are in an approximately linear regime for NWP impact as a function of the atmospheric signal levels (or laser energy), as is evident from Figure 173. Therefore, the Rayleigh wind NWP impact relative changes should match the relative changes in useful signal. Only when extremely large signal levels are reached does the NWP impact tend to saturate somewhat (curve flattens). There are no changes to this conclusion (linear change in impact) if the calibration point is changed from 5.5 m/s to e.g. 4.5 m/s (appropriate for August 2019) at 61 mJ. Unfortunately, this means the potential NWP impact from Rayleigh winds has likely dropped by ~36% between August 2019 to August 2020. Of course, the complex 4D-Var data assimilation at ECMWF may not have the same behaviour; but it should be a reasonable guess.

A confirmation of the linear behaviour of NWP impact with signal level  $A$  follows from a Taylor expansion of the relative NWP impact formula for small  $\left(\frac{\sigma_B}{\sigma_O}\right)^2$  i.e. when the observation errors are much bigger than the background errors:

$$1 - \frac{\sigma_A}{\sigma_B} = 1 - \frac{1}{\sqrt{1 + \left(\frac{\sigma_B}{\sigma_O}\right)^2}} \approx \frac{1}{2} \left(\frac{\sigma_B}{\sigma_O}\right)^2 \propto A$$

#### 9.4 Rayleigh HLOS wind error propagation including solar background noise

The projection of the LOS wind onto the horizontal plane provides the formula for HLOS wind given the incidence angle  $\theta$  (90 degrees minus the elevation angle):

$$v_{HLOS} = \frac{v_{LOS}}{\sin \theta}$$

Applying error propagation leads to (assuming a negligible error in the reported  $\theta$ ):

$$\sigma_{v_{HLOS}} = \frac{\sigma_{v_{LOS}}}{\sin \theta}$$

The LOS wind depends on the Doppler frequency shift  $\Delta f$  as follows, where  $\lambda$  is the outgoing laser wavelength:

$$v_{LOS} = -\frac{\lambda}{2} \Delta f$$

For the Rayleigh channel, the Doppler frequency shift  $\Delta f$  is a function  $f(R)$  of the measured atmospheric Rayleigh response ( $R = \frac{A-B}{A+B}$ ), assuming the internal path Rayleigh response is unimportant (as demonstrated in practice in operations since December 2020):

$$\Delta f = f(R)$$

Hence, the standard error in LOS wind is:

$$\sigma_{v_{LOS}} = \frac{\lambda}{2} \left| \frac{\partial f(R)}{\partial R} \right| \sigma_R$$

Where  $\frac{\partial f(R)}{\partial R}$  is the gradient of the Rayleigh response curve.

From the L2B processor ATBD the standard error in the Rayleigh response is:

$$\sigma_R = \frac{2}{(A+B)^2} \sqrt{B^2 \sigma_A^2 + A^2 \sigma_B^2}$$

Where  $A$  and  $B$  are the useful signal counts in the direct and indirect Rayleigh spectrometer for the atmospheric path. Including the various terms, the Rayleigh HLOS wind standard error is:

$$\sigma_{v_{HLOS}} = \frac{\lambda}{\sin \theta (A+B)^2} \left| \frac{\partial f(R)}{\partial R} \right| \sqrt{B^2 \sigma_A^2 + A^2 \sigma_B^2}$$

Assuming only shot noise with a Poisson distribution is important, then the errors in the Rayleigh spectrometer useful signal counts are as follows:

$$\sigma_A^2 = r(A + (\alpha^2 + \alpha)S_A)$$

$$\sigma_B^2 = r(B + (\alpha^2 + \alpha)S_B)$$

were  $r$  = radiometric gain,  $S_{A,B}$  = solar background count (from range-bin 25) on each channel, and  $\alpha$  is the range-bin time scaling factor (time spent in range-bin  $x$  divided by time spend in range-bin 25) e.g. 2.25E-3 for a 1 km range-bin). If the indirect channel measures a constant fraction  $c$  of the photons of the direct channel (both for useful signal and solar background) i.e.

$$B = cA$$

$$S_B = cS_A$$

Then the Rayleigh HLOS wind error can be simplified to the following dependency on useful signal  $A$  and the solar background level  $S_A$ :

$$\sigma_{v_{HLOS}} = \frac{d}{A} \sqrt{A + (\alpha^2 + \alpha)S_A}$$

Where:

$$d = \frac{\lambda}{\sin \theta} \left| \frac{\partial f(R)}{\partial R} \right|$$

$d$  can be assumed to be a constant of proportionality when we have fixed incidence angle, laser wavelength and Rayleigh response calibration gradient. If solar background level is zero, then we get the well-known simple one over square root of signal dependency:

$$\sigma_{v_{HLOS}} = \frac{d}{\sqrt{A}}$$

The constant of proportionality  $d$  can be determined from a combination of measured  $A$  and  $\sigma_{v_{HLOS}}$  with zero solar background conditions from real L2B Rayleigh-clear statistics. We can then plot how the random error will vary with useful signal and solar background levels, e.g. see Figure 174. This shows a dramatic increase in noise of Rayleigh winds in the worst-case solar background noise with typical real Aeolus useful signal levels. Unfortunately, the negative effect of solar background only increases as the useful signal levels drop.

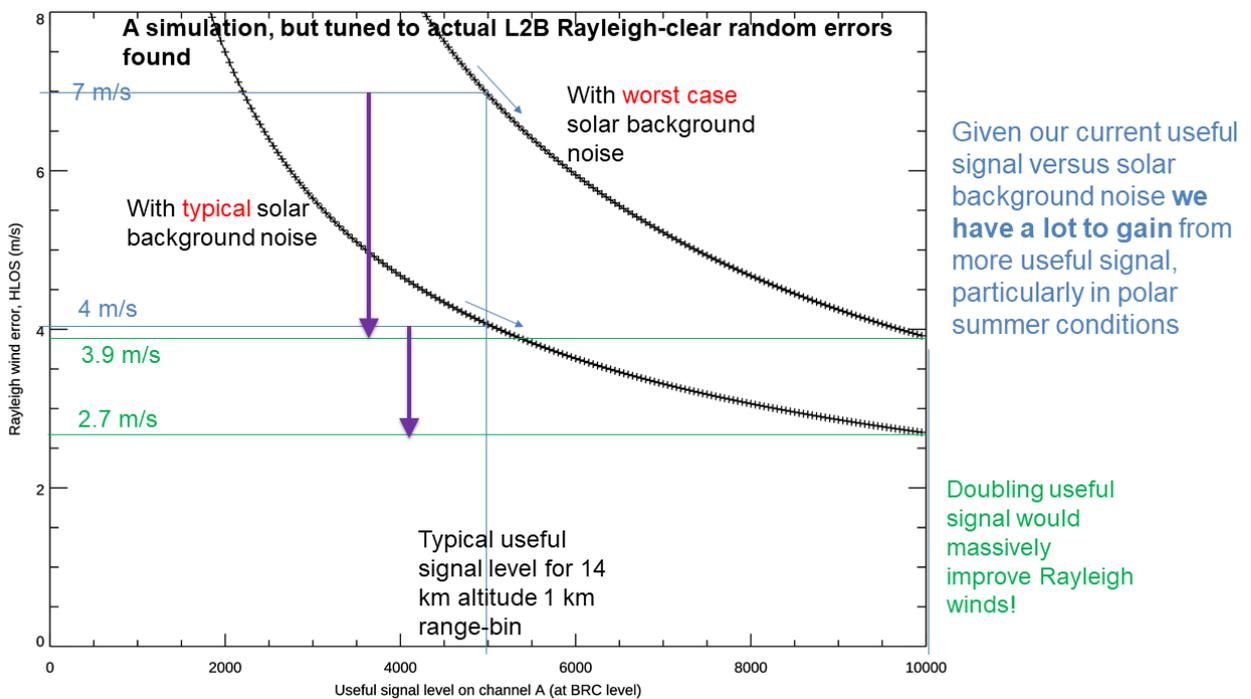


Figure 174. The expected dependence of the L2B Rayleigh-clear wind standard error  $\sigma_{v_{HLOS}}$  as a function of useful signal on channel A for a ~14 km altitude range-bin with typical and worst-case solar background levels, based on the formula derived in this section.

### 9.5 The dependence of Rayleigh HLOS wind standard error on incidence angle

Starting from the zero solar background level version of the Rayleigh HLOS wind error formula (see section 9.4), but keeping the LOS to HLOS sensitivity to incidence angle, we have the proportionality:

$$\sigma_{v_{HLOS}} \propto \frac{1}{\sin \theta \sqrt{A}}$$

Where  $\theta$  is the line-of-sight incidence angle (90 degrees minus the elevation angle) and  $A$  is the channel A useful signal count. From the lidar equation (see section 9.6) choosing the terms that vary with incidence angle only, the useful signal depends on the satellite-to-target range distance  $d_{range}$  and two-way transmission ( $T^2$ ) as follows:

$$A \propto \frac{T^2}{d_{range}^2}$$

Note that the molecular backscatter is not required for determining the sensitive with incidence angle because it is fixed for a given altitude, irrespective of the incidence angle. Neglecting the curvature of the Earth (which is assumed to be not that important for Aeolus given the low altitude orbit of 320 km relative to the Earth's radius  $\sim 6400$  km), then the satellite-to-target range is approximately:

$$d_{range} \approx \frac{d_{orbit\_alt}}{\cos \theta}$$

The two-way transmission can be approximated for the clear-sky case i.e. only molecular extinction. The molecular extinction is proportional to atmospheric density, which itself can be approximated by an exponential decaying function with altitude, with a scale height  $H$ . Therefore, the two-way transmission at the altitude above earth surface  $z_1$ :

$$\begin{aligned} T^2 &= \exp\left(-2 \int_{r=r_1}^{r=\infty} \alpha_{mol}(r) dr\right) = \exp\left(-2 \int_{r=r_1}^{r=\infty} \alpha_0 \exp\left(\frac{-r \cos \theta}{H}\right) dr\right) \\ &= \exp\left(-\frac{2\alpha_0 H}{\cos \theta} \exp(-z_1/H)\right) \end{aligned}$$

Hence the useful signal is as follows:

$$A \propto \cos^2 \theta \exp\left(-\frac{2\alpha_0 H}{\cos \theta} \exp(-z_1/H)\right)$$

Therefore, putting this dependence of  $A$  into the  $\sigma_{v_{HLOS}}$  formula we get:

$$\sigma_{v_{HLOS}} \propto \frac{1}{\sin \theta \cos \theta \exp\left(-\frac{\alpha_0 H}{\cos \theta} \exp(-z_1/H)\right)}$$

One can also account for the Earth's curvature as follows:

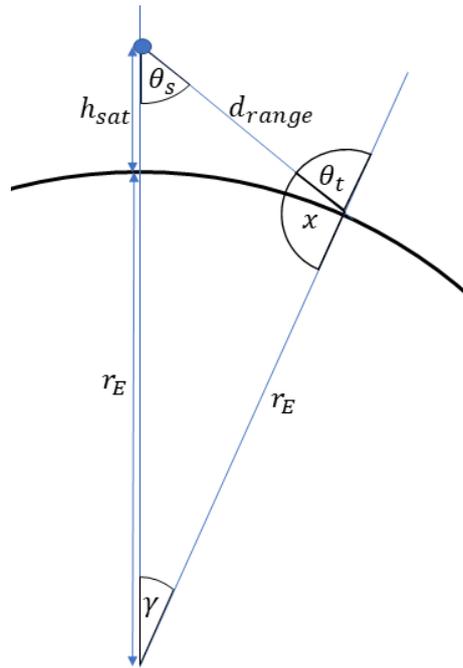


Figure 175. Geometry of incidence angles for a curved Earth.

Using the law of sines formula given the geometry in Figure 175 , the incidence angle at the target (near Earth surface)  $\theta_t$  is:

$$\theta_t = \sin^{-1}\left(\frac{(r_E + h_{sat})}{r_E} \sin \theta_s\right)$$

and the satellite to target range is:

$$d_{range} = r_E \frac{\sin \gamma}{\sin \theta_s}$$

where:

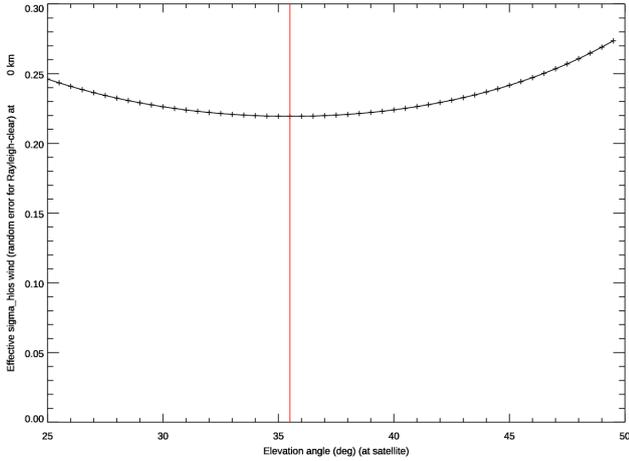
$$\gamma = \theta_t - \theta_s$$

Therefore, putting this into the formula for  $\sigma_{v_{HLOS}}$  and considering the incidence angle at the target for derivation of the HLOS wind:

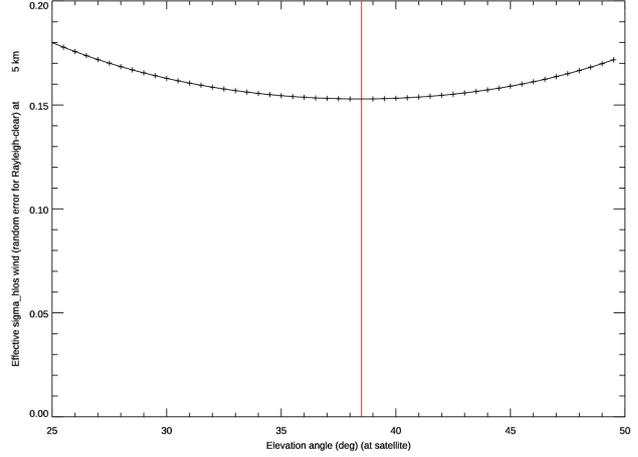
$$\sigma_{v_{HLOS}} \propto \frac{1}{\sin \theta_t \frac{\sin \theta_s}{\sin \gamma} \exp\left(-\frac{\alpha_0 H}{\cos \theta_t} \exp(-z_1/H)\right)}$$

Using a typical density scale-height  $H$  of 8 km and a typical molecular extinction coefficient near the surface of  $0.075 \text{ km}^{-1}$  (derived from the ECMWF AUX\_MET data) for a range of altitudes ( $z_1$ ) we find the optimal incidence angles in Figure 176 (red vertical lines); note this is accounting for Earth's curvature. For high altitudes e.g. 30 km the optimal incidence angle tends to condition of zero attenuation i.e. 43.5 degrees. At the surface the optimal is 35.5 degrees (note the Aeolus operational

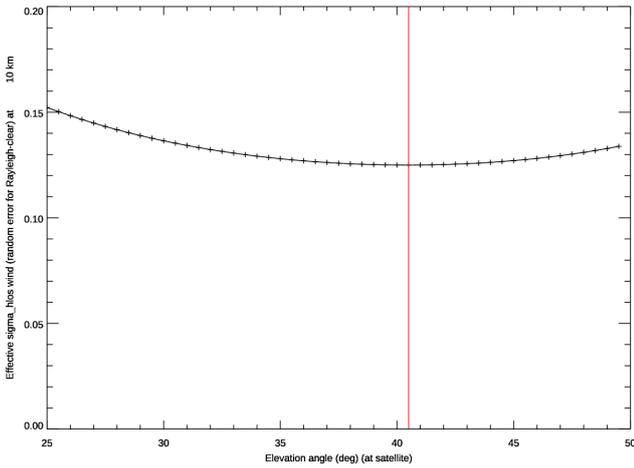
mission off-nadir pointing of 35 degrees). At 15 km, closer to where Aeolus NWP impact has been found to be a maximum; an incidence angle of 42 degrees would be optimal and would improve the Rayleigh HLOS wind random error by ~3% relative to the current 35 degree setting (assuming some other important constraints have not been missed). Note an atmospheric path signal level improvement of ~6% would be required to improve the Rayleigh HLOS wind error by ~3% with the current incidence angle setting. For reference, the relative improvement in random error is approximately 0% at 0 km, 1% at 5 km, 2% at 10 km and 4.5% at 30 km.



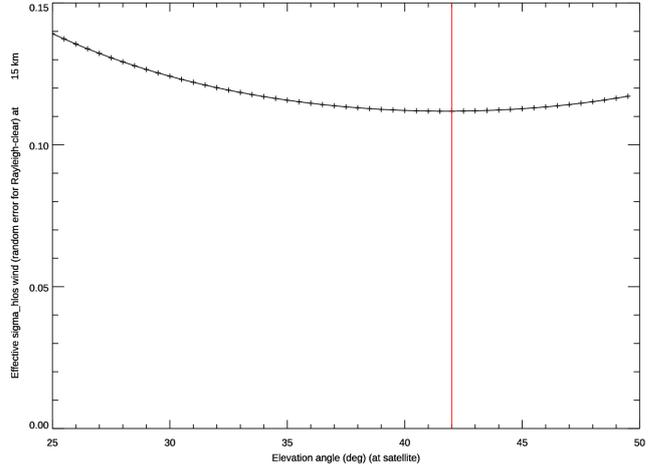
a) 0 km altitude; optimal incidence=35.5°



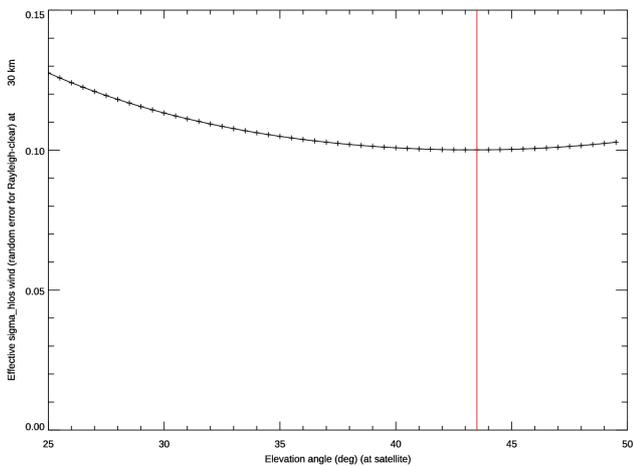
b) 5 km altitude; optimal incidence=38.5°



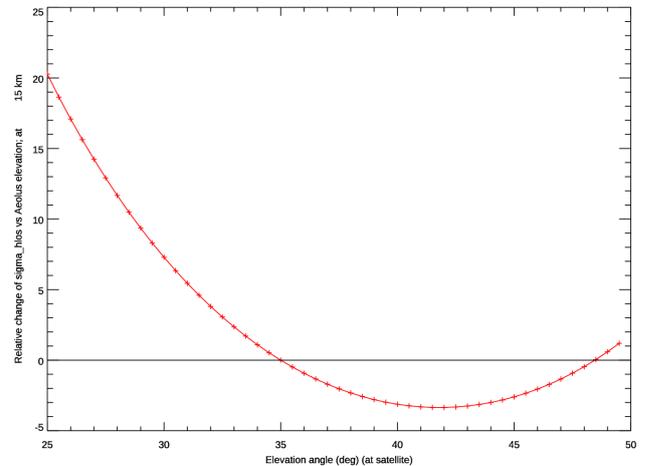
c) 10 km altitude; optimal incidence=40.5°



d) 15 km altitude; optimal incidence=42°



e) 30 km altitude; optimal incidence=43.5°



f) Percentage change in  $\sigma_{HLOS}$  with incidence compared to  $\sigma_{HLOS}(\theta = 35^\circ)$ ; at 15 km altitude.

Figure 176. Dependence of Rayleigh-clear HLOS wind standard error  $\sigma_{v_{HLOS}}$  (units irrelevant) as a function of incidence angle (at the satellite) for different altitudes a) 0 km, b) 5 km, c) 10 km, d) 15 km and e) 30km for a typical clear-air molecular density profile. Red line shows the minimum. f) shows the relative (percentage) change in the  $\sigma_{HLOS}$  for the 15 km altitude case compared to a reference of 35 degrees incidence angle i.e. that applied for Aeolus. This formulation accounts for Earth's curvature.

## 9.6 The sensitivity of Aeolus L2B O-B statistics to a model averaging versus point observation operator

As discussed in section 2.2, it should be possible to improve the Aeolus L2B HLOS wind observation operator (also known as forward model) by making it more representative of what Aeolus observes in terms of spatial averaging. Aeolus does not measure wind information from an infinitesimal geolocation point (like a delta-function weighting), but rather it is a spatial average across the input data to the L2B HLOS wind retrieval; a forward model could account for this spatial average. Assuming the background forecast has some realistic wind variability (resolution) at scales smaller than the spatial extent of Aeolus winds, such an operator would improve the O-B statistics and potentially the NWP impact of Aeolus.

Figure 177 illustrates the geometry of an Aeolus retrieved L2B HLOS wind.

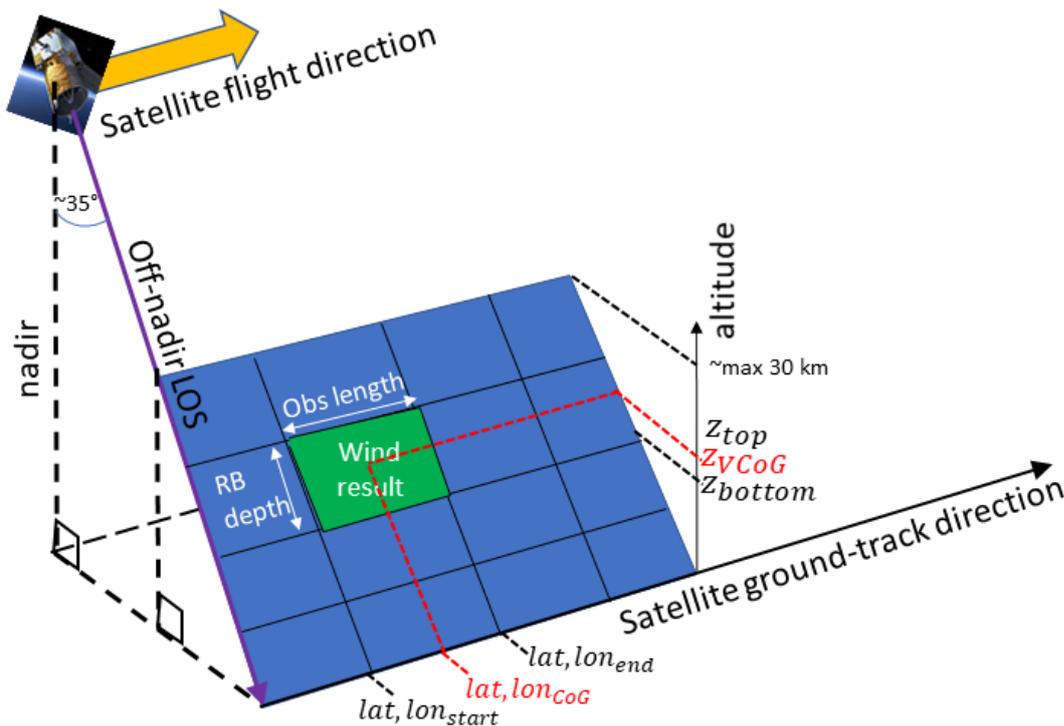


Figure 177. Illustration of how Aeolus winds are measured via an accumulation of atmospheric backscatter signal within effectively a 2D sampling rectangle. RB=range-bin; CoG=centre of gravity, vCoG=vertical centre of gravity; LOS=line-of-sight. Note that in reality there are 24 vertical range-bins available.

The maximal vertical coverage of the retrieved L2B wind is defined by the on-board commanded range-bin thickness (a time range-gating for the lidar; settings are uploaded to satellite). Range-bins typically vary in thickness from 0.25 to 2 km altitude, with the most common setting being 1 km. The

range-bin is the highest vertical resolution information available from Aeolus i.e. it is not oversampled on-board and then the ground processing produces a range-bin. The so-called “observation” of the L2B processor wind retrieval is generated from a horizontal accumulation of the spectrometer counts of “measurement-bins” (set to ~2.8 km horizontal-scale, range-bin vertical scale) according to the L2B processor measurement grouping algorithm and the classification algorithm (which classifies the measurement-bins into clear and cloudy conditions using the L1B measurement-scale information e.g. Mie SNR or Mie scattering ratio). The distance of the “observation” sampling within the plane is very small; being the width of the laser beam (Cassegrain telescope expanded and slightly divergent) near the ground (~10 m diameter). So effectively we have a tilted (off-nadir) 2D rectangle within over which the signal is accumulated.

However, within such a sample rectangle, the amplitude of the signal received will not be uniform and will vary according to the variations in the received signal amplitudes over the rectangle as described by the lidar equation (mostly driven by atmospheric optical properties):

- The total scattered power received by the lidar at a time corresponding to range  $R$  is:

$$P(\lambda, R) = P_L \frac{A_0}{R^2} \xi(\lambda, R) \beta(\lambda, R) T^2(\lambda, R) \frac{c\tau_L}{2}$$

- $\lambda$  = laser wavelength
- $R$  = range of scatterer from sensor
- $\beta$  = volume backscattering coefficient of atmosphere
- $T$  = one way transmission factor (Beer-Lambert law):  $T(\lambda, R) = e^{-\int_0^R \alpha(\lambda, R) dR}$
- $\alpha$  = atmospheric attenuation coefficient
- $P_L$  = average power in laser pulse
- $A_0$  = area of objective lens
- $c$  = speed of light
- $\tau_L$  = laser pulse duration
- $\xi$  = calibration factor (depending on spectral transmission of receiver and overlap factor)

Assuming that the instrumental parameters are fixed for the time length of an observation and that the satellite to target range is not varying much, then this can be simplified to the attenuated backscatter, which depends only on atmospheric geophysical variables, assuming parameterisations for the backscatter and extinction coefficients are available at 355 nm wavelength:

$$\beta(\lambda, R) T^2$$

Attenuated backscatter must be calculated from the top of the atmosphere to the layer of interest as an integral along the laser line-of-sight direction. When calculating many adjacent profiles, we end up with field varying both horizontally and vertically i.e. effectively leads to a 2D weighting function.

The molecular (clear air) backscatter and extinction is a function of atmospheric temperature and pressure<sup>37</sup>. The particulate backscatter (cloud) is a function of pressure, temperature, cloud ice and liquid water content (via a relatively simplistic parameterisation, more complicated ones are possible). The two sources (clear and cloud) of backscatter and extinction are added together before calculating the two-way transmission  $T^2$ .

<sup>37</sup> To be investigated if humidity (water vapour) should also be considered.

Horizontal variations of attenuated backscatter may matter when there are very strong horizontal molecular density gradients and varying cloud/aerosol backscatter along track. However, it seems from ECMWF-derived examples of attenuated backscatter that the vertical variations are more important for reducing forward model (observation operator) error<sup>38</sup>. Figure 178 shows some example vertical profiles of attenuated backscatter, forward modelled from the ECMWF background. The smoothly curved profiles result from the clear-air and the sudden increases in backscatter (and attenuation below) are due to cloud backscatter (seen at altitudes below 15 km).

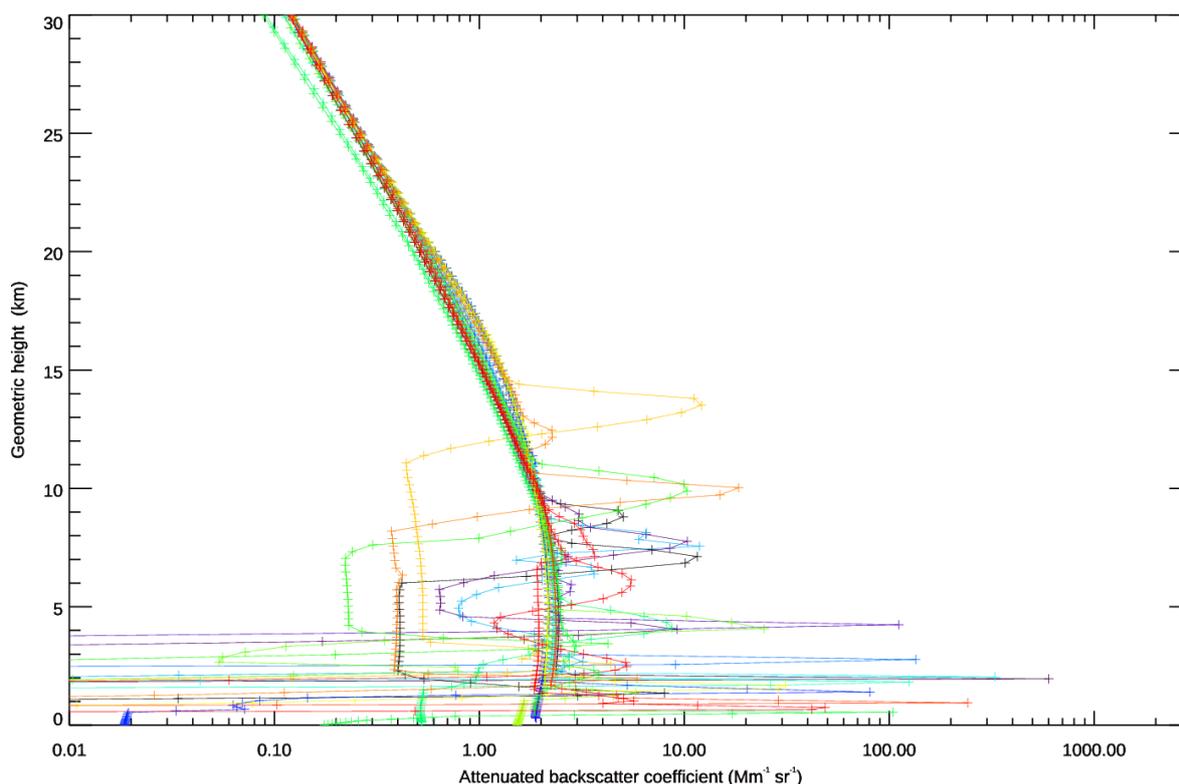


Figure 178. Examples of forward modelled attenuated backscatter profiles from ECMWF model background profiles at 355 nm wavelength. This is derived from the combination of molecular and particulate (cloud ice/liquid water content) optical properties parameterisations.

An example of the horizontal (lidar curtain) variation in the same forward modelled attenuated backscatter is evident from Figure 179. Note that ice/water clouds are seen as layers of relatively strong attenuated backscatter with zero backscatter below (due to total attenuation occurring), which are on top of a smoothly increasing function towards the surface due to the molecular backscatter, which is proportional to atmospheric density.

<sup>38</sup> Note that L2B Rayleigh-clear winds are assigned at an altitude of 0.49 (an AUX\_PAR\_2B setting) of the thickness of the range-bin to partially account for the expected pattern of attenuated backscatter in clear-air.

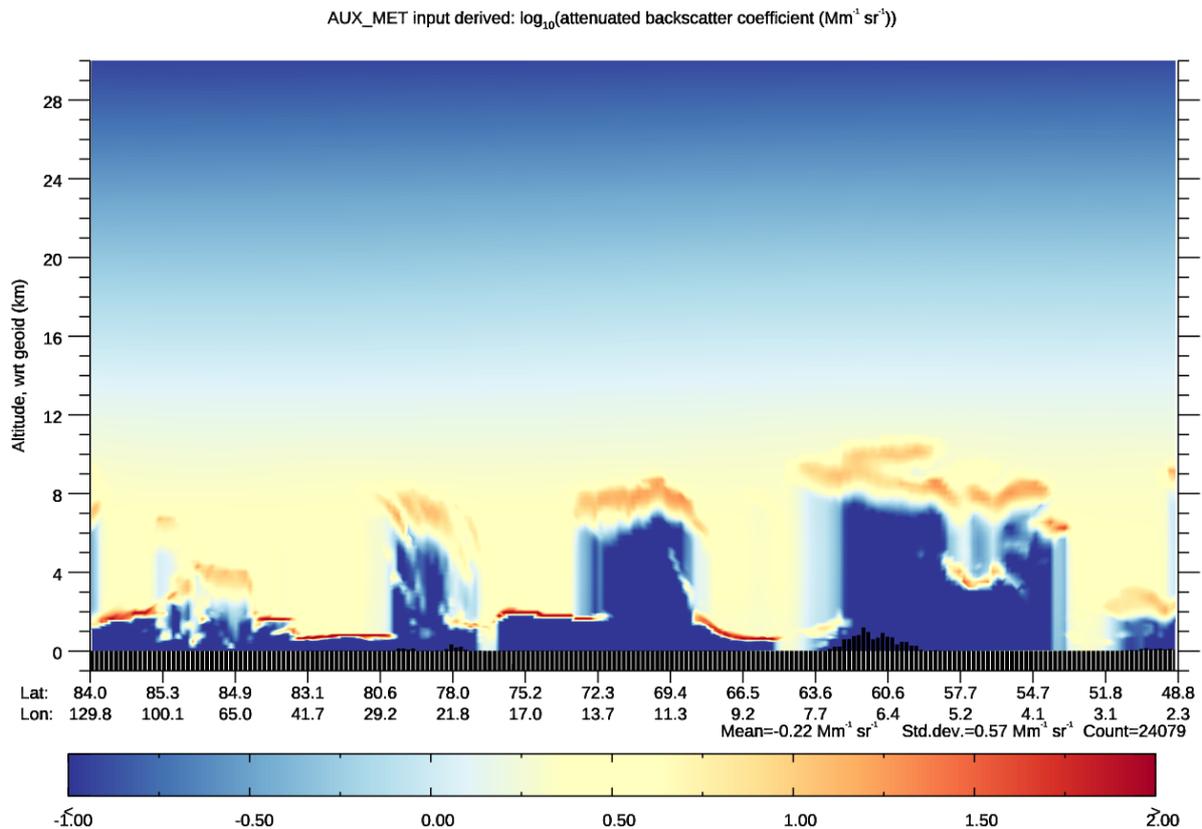
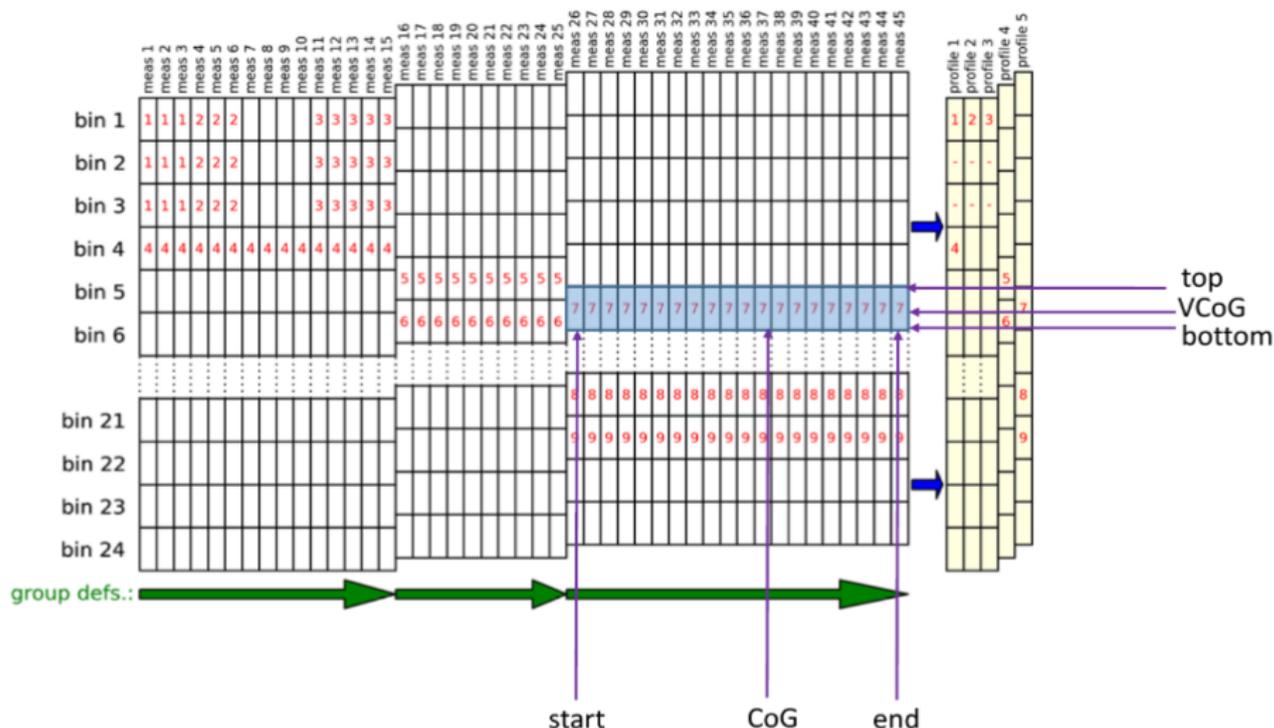


Figure 179. An example of a lidar curtain (along an Aeolus predicted descending orbit) of the ECMWF background forward modelled to attenuated backscatter (log-scale) at 355 nm wavelength. Note that  $\log_{10}$  of the attenuated backscatter is plotted. It is derived from the combination of molecular and particulate (cloud ice/liquid water content) optical properties parameterisations.

A typical L2B Rayleigh-clear HLOS wind accumulates over one BRC (Basic Repeat cycle) of measurement-bins with a duration of 12 seconds; therefore, with a satellite ground-speed of  $\sim 7.396$  km/s it has a horizontal resolution of  $\sim 88$  km. The “observation” duration is very short relative to the effective time-resolution of ECMWF’s forecast model and therefore its variation (from start to end) does not need to be accounted for in an observation operator i.e. we can consider it an instantaneous snapshot. However, the horizontal spatial extent of  $\sim 90$  km covers  $\sim 9$ -10 grid-points of current ECMWF high-resolution 4D-Var outer loop (with which the O-B departures for the analysis are calculated). The L2B Mie-cloudy winds have been typically provided (a L2B processor setting) at a relatively high horizontal resolution of  $\sim 14$  km ( $\sim 5$  measurement-bins, and just over one ECMWF grid-point spacing), and so the horizontal variation in actual attenuated backscatter due to cloud optical properties variations will not be realistically captured by the relatively coarse ECMWF background forecast. From qualitative experience, ECMWF background clouds are typically in the correct geolocation compared to Aeolus strong Mie backscatter (i.e. clouds), however the model does not always capture the details very well and lacks horizontal resolution (and there are occasional gross errors e.g. missing clouds in ECMWF).

For Rayleigh-clear winds, the vertical variation of attenuated backscatter, combined with the vertical variation in HLOS wind is assumed to be more important for the thicker range-bins e.g. 1-2 km, given the potential for many model levels to be present over one range-bin; see Figure 7. The horizontal variation in the clear air attenuated backscatter appears rather small over 90 km in Figure 179. How much the ECMWF background HLOS wind varies over different horizontal scales should be investigated in future.

The L2B BUFR provides some information to help define the 2D rectangle on which a model averaging observation operator could be applied. The following illustration for the Aeolus wind retrieval geometry is taken from the TN7.3 L2B BUFR description document (AE-TN-ECMWF-L2BP-007):



**Figure 1. Guide to the Aeolus L2B BUFR observation geometry (observation represented by the blue rectangle). The vertical axis is the LOS direction (typically 35 degrees off-nadir), the horizontal axis corresponds to the along-track direction. The small boxes correspond to vertical range-bins of individual measurements. An observation is a wind retrieval derived using a selection of such measurement-bins (the smallest granularity of data in the L1B products). The observation geometry is described by geolocations for the horizontal and vertical extent that the observation samples. The horizontal geolocations are given by the start, end and centre of gravity (CoG) of the observation and the vertical geolocations by top, bottom and vertical centre of gravity (VCoG).**

NCES			
3 40 013	<b>Atmospheric Laser Doppler Instrument (ALADIN) L2B Data</b>		
	0 01 007	Satellite identifier	
	0 02 019	Satellite instruments	
	0 01 033	Identification of originating/generating centre	
	0 01 034	Identification of originating generating sub-centre	
	0 04 001	Year	
	0 04 002	Month	
	0 04 003	Day	
	0 04 004	Hour	
	0 04 005	Minute	
	0 04 007	Seconds within a minute (microsecond accuracy)	
	0 05 068	Profile number	
	0 05 070	Observation identifier	
	0 05 069	Receiver channel	
	0 40 036	Lidar L2b classification type	
	0 08 091	Coordinates significance	2 -> Start of observation
	0 05 001	Latitude (high accuracy)	
	0 06 001	Longitude (high accuracy)	
	0 04 016	Time increment	
	0 08 091	Coordinates significance	3 -> End of observation
	0 05 001	Latitude (high accuracy)	
	0 06 001	Longitude (high accuracy)	
	0 04 016	Time increment	
	0 08 091	Coordinates significance	4 -> Horizontal Centre of gravity of observation
	0 05 001	Latitude (high accuracy)	
	0 06 001	Longitude (high accuracy)	
	0 04 016	Time increment	
	0 08 091	Coordinates significance	6 -> Top of observation
	0 07 071	Height (high resolution)	
	0 05 021	Bearing or azimuth	
	0 07 021	Elevation	
	0 40 035	Satellite range	
	0 08 091	Coordinates significance	7 -> Bottom of observation
	0 07 071	Height (high resolution)	
	0 05 021	Bearing or azimuth	
	0 07 021	Elevation	
	0 40 035	Satellite range	
	0 08 091	Coordinates significance	5 -> Vertical centre of Gravity of the observation
	0 07 071	Height (high resolution)	
	0 05 021	Bearing or azimuth	
	0 07 021	Elevation	
	0 40 035	Satellite range	
	0 40 029	Horizontal observation integration length	
	0 40 030	Horizontal line of sight wind	
	0 40 031	Error estimate of horizontal line of sight wind	
0 25 187	Confidence flag		
0 10 004	Pressure		
0 12 001	Temperature/air temperature		
0 40 037	Backscatter ratio		

Figure 180. Taken from the L2B BUFR description, showing the available geolocation information with a HLOS wind result.

For the horizontal dimension it provides the start, CoG (centre of gravity) and end latitude, longitude and time of the wind result. Also, the “Horizontal observation integration length” is available, which might be useful for monitoring purposes. For the vertical dimension, the L2B BUFR provides the top, vertical CoG and bottom altitude (geometric height) and azimuth angle (direction of pointing of the laser relative to North) for the wind result. The L2B BUFR does not provide any information about the measurement-scale signal amplitude variation across the horizontal (nor vertical) extent of the retrieval.

Before coding a more advanced observation operator in the IFS, it has first been tested using the ECMWF Advanced Monitoring software (IDL code). This uses the AUX\_MET\_12 (auxiliary meteorological) file to provide the ECMWF model background information. The AUX\_MET\_12 is required in the L2B processing to provide a priori temperature and pressure information for the Rayleigh-Brillouin Correction scheme. The AUX\_MET\_12 provides model field profiles with full vertical resolution of 137 levels via a full-resolution background forecast run (T<sub>co1279</sub>). The vertical profiles are extracted every 3 seconds (every ~20 km) along an Aeolus predicted ground-track (latitude, longitude, time). Being a vertical profile, it does not account for the ~37 degree off-nadir tilt of the Aeolus line-of-sight, which will result in a misplacement of  $\sim 30 \tan(37^\circ) = 22.6$  km for Aeolus L2B winds

at the maximum 30 km altitude, compared to the AUX\_MET\_12 vertical profile derived from the predicted off-nadir ground-track geolocation. Hence, a slightly more accurate representation should be possible in the IFS by accounting for each wind results exact latitude and longitude. The AUX\_MET profiles of u, v, T and cloud variables on geometric height are however produced using the IFS observation operator code; so, the Advanced monitoring forward model should be otherwise similar to what could be achieved directly in the IFS.

A selection of different options for the observation operator is tested with the Advanced Monitoring. They are assessed by the scaled MAD of L2B HLOS wind O-B values (m/s). Smaller values are assumed to indicate an improvement in observation operator. This is done using one days' worth of L2B HLOS winds (orbits 15822 to 15835 i.e. 14 orbits on 16-17 May 2021), processed with the operational settings for horizontal averaging (as explained earlier). This investigation should also be done using the early FM-B period e.g. July 2019, when Aeolus winds were significantly less noisy.

Different options tested:

- Nominal nearest neighbour grid-point HLOS wind observation operator.
- Mean average of all model grid-point (both horizontal and vertical) HLOS wind values found within the start and end time and within the top and bottom altitude of the L2B HLOS wind geolocation information. That is, an average in two spatial dimensions (along the orbit and within the range-bin).
- Mean average only in the vertical dimension i.e. select the model vertical profile closest to the CoG time of the L2B HLOS wind and then mean average between the top and bottom altitudes. This will help to determine if the horizontal averaging is important or not.
- Normalised-weighted average of the grid-point HLOS wind model values according to their forward modelled attenuated molecular backscatter values (considering both molecular and cloud optical properties estimates). Uses the same selection of model grid-points as in option 2.
- Mean average of u-wind component and separately v-wind component then followed by a calculation of HLOS wind (with VCoG azimuth).

The results of the short test-period are provided in the table below.

*Table 5. L2B global O-B statistics (scale MAD in lieu of standard deviation) for different observation operators. Best score highlighted in green.*

Option	Description	L2B Rayleigh-clear scaled MAD(O-B)	L2B Mie-cloudy scaled MAD(O-B)	L2B Rayleigh-cloudy scaled MAD(O-B)	Comments
1.	Nominal	6.82	3.80	6.57	
2.	Mean average over rectangle	6.78	3.75	6.52	
3.	Mean average in vertical dimension only	6.80	3.77	6.50	Perhaps need a larger

					sample for Ray-cloudy
4.	Weighted average over rectangle using attenuated backscatter	6.79	3.77	6.59	
5.	Mean average of u and v separately over rectangle, then calculate HLOS wind	6.78	3.75	6.52	Basically, the same as option 2. i.e. makes no difference

**Summary so far:** Mean average over the rectangle works the best so far. However, it is a rather small improvement — best result for Mie-cloudy (1.3% smaller  $\sigma(O-B)$ ). Should look more carefully at my attenuated backscatter calculation to check for bugs, since it was disappointing that it did not perform any better than the mean average (but still a bit better than nominal method).

An additional test was performed using FM-A data from the 3<sup>rd</sup> reprocessing (B14). This particularly orbit (13/9/2018 ~12 UTC) passed close to tropical cyclone Mangkhut. Rather thick (2 km) range-bin settings in the upper troposphere (above 12 km) were applied for this period of the Commissioning Phase. It was demonstrated that the “weighted average over rectangle using attenuated backscatter” option improved the L2B Mie-cloudy cloudy  $\sigma(O-B)$  by 5% over the whole orbit, and improvements are evident locally at the tropical cyclone outflow; as shown in Figure 181.

An example of where thick range-bins matters e.g. tropical cyclone outflow, Sept 2018

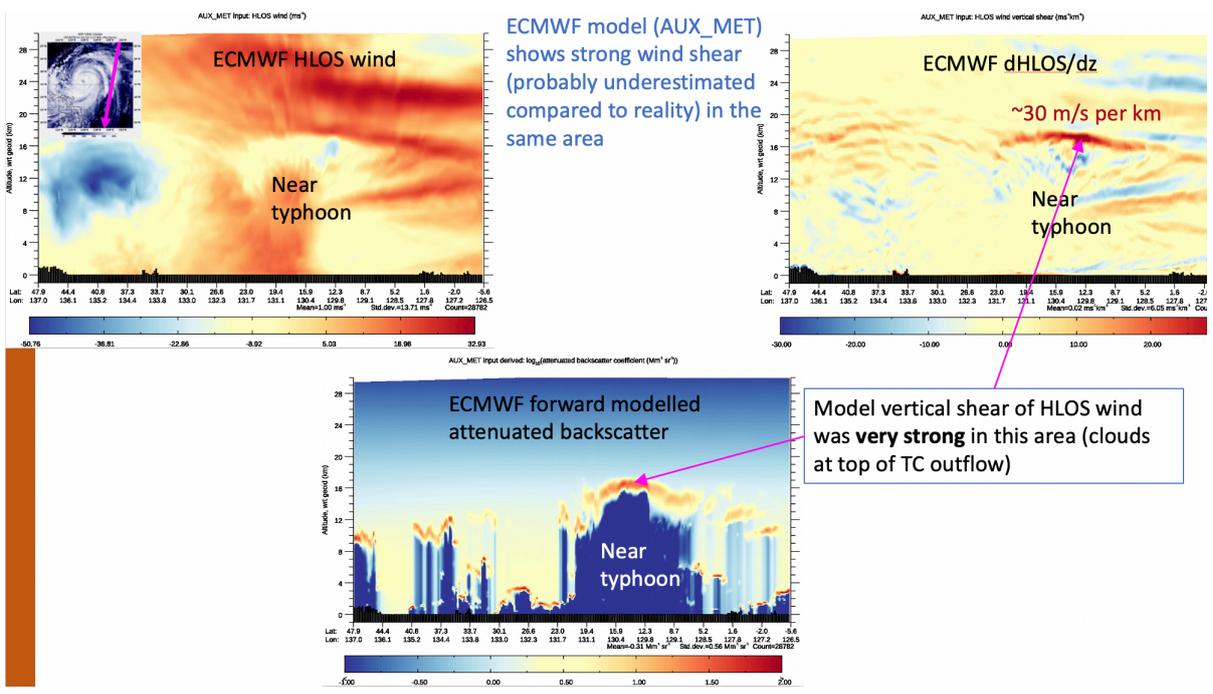
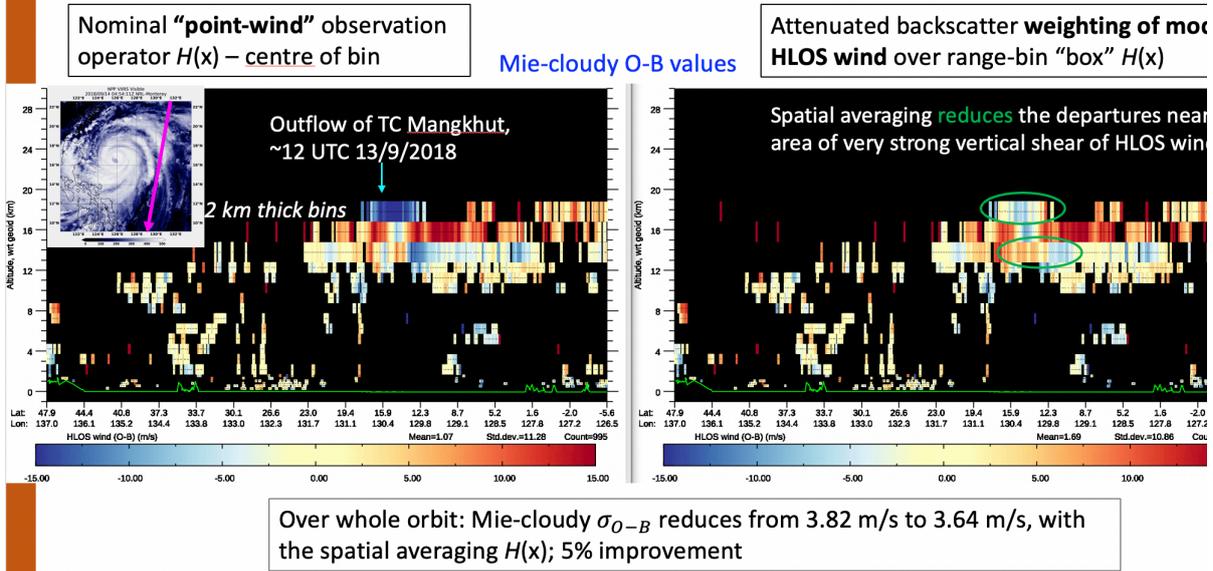


Figure 181. A test case showing that spatial averaging observation operators can reduce the O-B departures in strong vertical wind shear conditions with thick range-bin settings, suggesting they are more accurate representations of Aeolus wind data.

### 9.7 Does atmospheric refraction matter for the Aeolus wind geolocations?

A recurring question for Aeolus is whether the Earth’s atmosphere bends the lidar light significantly enough cause a noticeable error in the geolocations of the winds if assuming no bending in the Level-1

processing. To test this, a ray-tracing calculation was performed using the equations provided in the context of GNSS radio occultation calculations; see (Healy et al. 2006; equations 6 to 8).

Information on the refractive index of the Earth's atmosphere at Aeolus wavelengths of 355 nm was obtained from: <https://emtoolbox.nist.gov/Wavelength/Documentation.asp>. It is assumed that the so-called refractivity ( $N = 10^6(n - 1)$ , where  $n$  is the refractive index) decreases exponentially with altitude ( $z$ ) above the earth's surface, because to first order the refractivity is proportional to air density ( $\propto p/T$ ). At 355 nm a value of refractivity at the surface of  $N(z = 0) = 286$  was calculated for typical surface temperature and pressure values. For the exponential decrease with altitude, a scale-height of 8 km was assumed; based on typical atmospheric density profiles. The refractivity was assumed to vary with and only as a function of  $r$  (not  $\theta$ ) in polar coordinates, because the refractivity gradients in the atmosphere are much larger in the vertical than the horizontal directions.

A light ray is simulated starting from the Aeolus satellite and propagating towards the Earth's surface with the nominal off-nadir pointing angle i.e. the initial conditions are:  $r = r_{earth} + 320 \text{ km}$ ,  $r_{earth} = 6371 \text{ km}$ ,  $\theta = 0 \text{ deg}$  and  $\phi = 180 - 35 = 145 \text{ deg}$ . The ray-tracing is run twice, once with the exponential refractivity formula (as described above) and once with  $N(z = 0) = 0$  i.e. a refractive index of 1, so vacuum conditions with no refractive index gradients.

The ray-tracing calculation using the typical atmospheric refractivity gradient resulted in only a 2 m displacement of the ray path at the Earth's surface compared to a case with  $n = 1$  (vacuum). It is so small that no difference is visible in Figure 182 which shows both rays; one for vacuum (green line) and one with a typical refractivity profile (red line).

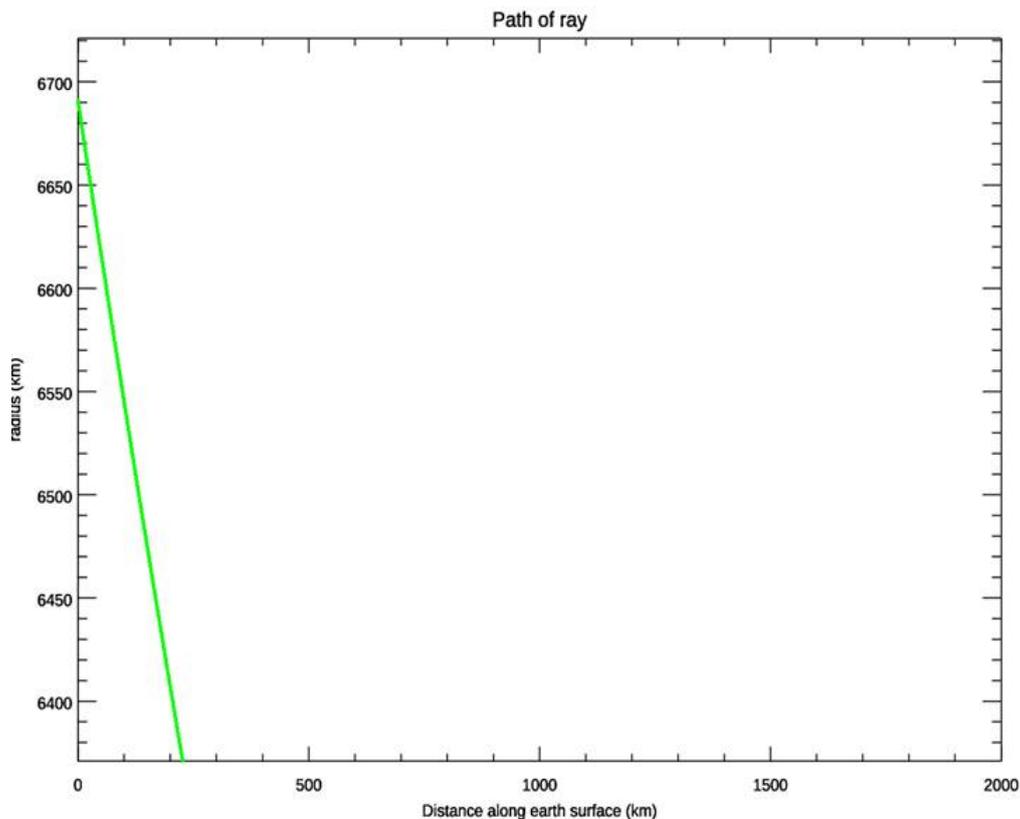


Figure 182. Ray-tracing for light coming from the Aeolus satellite pointing off-nadir by 35 degrees. The ray path in a vacuum (green line) and the typical atmospheric refractivity profile ray path (red line) are indistinguishable at this scale to the plot. They are in fact separated by 2 m at the Earth's surface.

Varying the initial  $\phi$  angle, it was found that  $\phi = 107.8 \text{ deg}$  (so  $\sim 72$  degrees in off-nadir angle) is approximately the value at which the ray in a vacuum just scrapes the earth's surface (Figure 183). The distance along earth's surface difference in that case is  $\sim 61 \text{ km}$ . This demonstrates how the vertical gradients in refractivity matter more when the ray points increasingly off-nadir i.e. closer to the geometry of radio occultation.

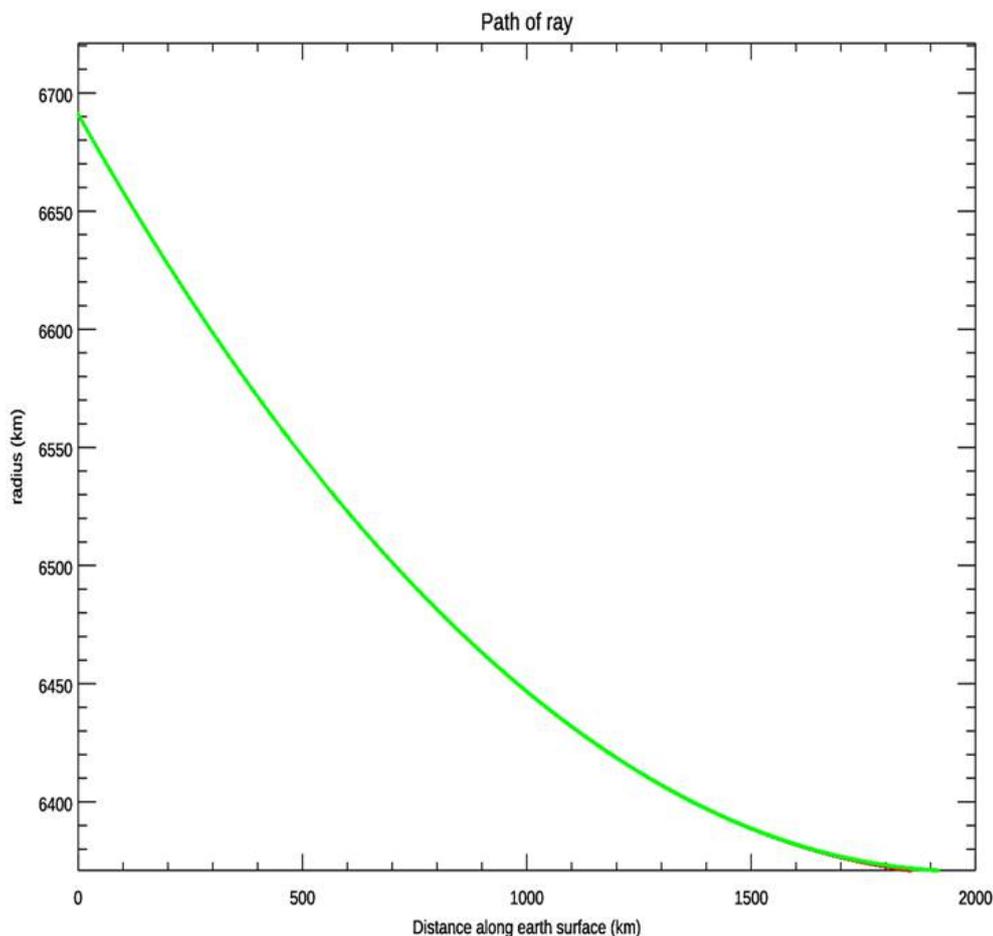
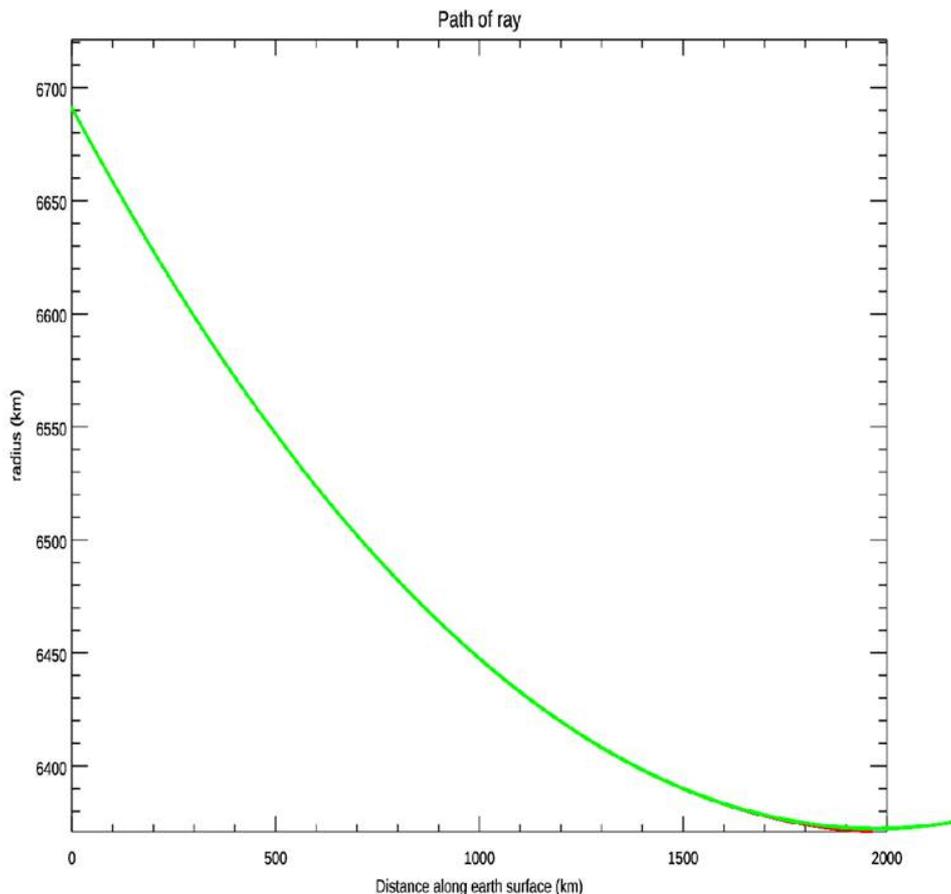


Figure 183. Ray-tracing for light coming from the Aeolus satellite pointing off-nadir of  $\sim 72$  degrees, for the which the ray path in a vacuum just scrapes the Earth's surface. The ray path in a vacuum (green line) and the ray path in a typical refractivity profile atmosphere (red line) are just about distinguished in this plot. They are in fact separated by 61 km at the Earth's surface. Note that the bending of the green ray (in vacuum) on this plot is an artefact of the x-axis being distance along the curved Earth surface.

$\phi = 107.75 \text{ deg}$  leads to the typical atmospheric refractivity profile (red line) ray hitting the surface (bent towards the earth), whereas the vacuum ray (green line) never touches the ground and propagates in a straight-line path back to space (the straight-line looks curved in the dimensions of Figure 184), so in effect an infinite ground displacement.



*Figure 184. Ray-tracing for light coming from the Aeolus satellite pointing off-nadir of  $\sim 72$  degrees, for the which the ray path in a vacuum (green line) does not reach the Earth's surface but goes back to space, whereas the ray path in a typical refractivity profile atmosphere (red line) intercepts the surface.*

With nominal Aeolus off-nadir pointing, it is necessary to increase the refractivity by a very large factor of  $1.0E5$  to generate a significant displacement of 22 km on ground; the red line being bent towards the earth is visible in Figure 185.

If the ray-tracing calculations are realistic then any detected errors in Aeolus' geolocation cannot be explained by ray bending due to atmospheric refractive index gradients; the effect is too weak.

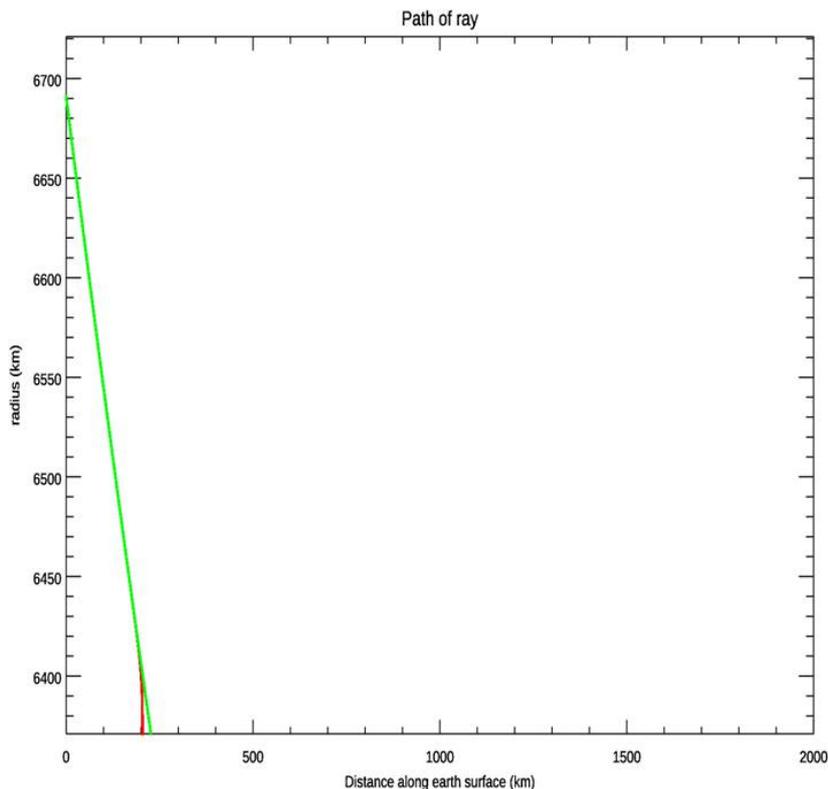


Figure 185. Ray-tracing for light coming from the Aeolus satellite pointing off-nadir by 35 degrees. The ray path in a vacuum (green line) and the unrealistically large atmospheric refractivity profile ray path (red line) are seen to separate with the refractivity causing bending toward the earth. They are in fact separated by 22 km at the Earth's surface.

## 9.8 Modelling of relative FSOI versus assigned observation error

We attempt to assess the impact of an observation type as a function of the assigned observation error ( $\sigma_o$ ) theoretically via some simplified version of the relative FSOI metric formulation. The FSOI expresses error reduction in energy norm for a 24-hour forecast versus a 36-hour forecast due to the analysis in which an observation type is assimilated. The energy norm is a quadratic error term (in model space squared).

The relative FSOI is defined as the global energy norm error reduction from a specific observation type assimilated divided by the global energy norm error reduction due to all the observations assimilated. This is depicted in Figure 186.

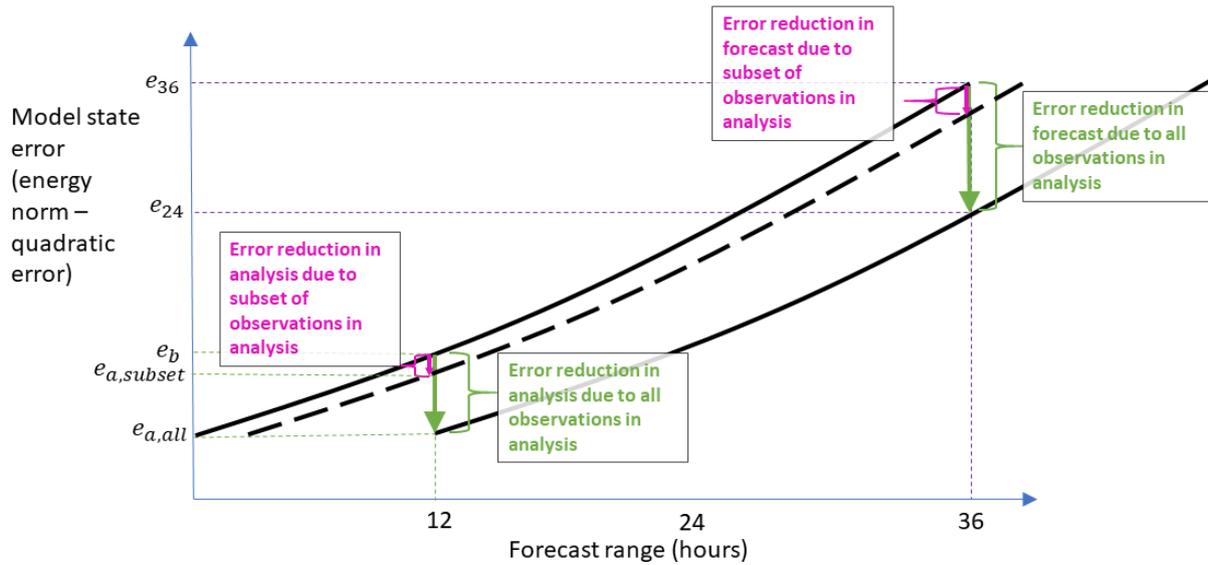


Figure 186. Schematic of how the FSOI forecast error metric (global dry energy norm) depends on the observation dataset assimilated.

Assuming the analysis error grows with time by a simple scaling factor for the 24-hour forecast (an approximation due to the typically exponential error growth with forecast range), then the relative FSOI can be written in terms of the error reduction of the analysis due to a “subset” of observations, relative to the error reduction due to “all” observations:

$$FSOI_{rel} = \frac{\Delta e_{a-b,subset}}{\Delta e_{a-b,all}} = \frac{e_{a,subset} - e_b}{e_{a,all} - e_b}$$

If the FSOI dry energy norm (quadratic error) behaves like the variance of the analysis and background error, we have:

$$FSOI_{rel} = \frac{\sigma_{a,subset}^2 - \sigma_b^2}{\sigma_{a,all}^2 - \sigma_b^2}$$

Assuming the scalar analysis formula for analysis error (observations in model space) and inserting into the above equation:

$$\sigma_a^2 = \frac{\sigma_o^2 \sigma_b^2}{\sigma_o^2 + \sigma_b^2}$$

We can simply:

$$FSOI_{rel} = \frac{\sigma_{o,all}^2 + \sigma_b^2}{\sigma_{o,subset}^2 + \sigma_b^2}$$

A refinement can be made to the formula by knowing that “all” observations is made up of the observation “subset” we are interested in, and the “rest” of the observations:

$$\sigma_{o,all}^2 = \frac{\sigma_{o,subset}^2 \sigma_{o,rest}^2}{\sigma_{o,subset}^2 + \sigma_{o,rest}^2}$$

This formula for relative FSOI allows for a curve-fit procedure to be run for plots of Aeolus relative FSOI versus Aeolus assigned observation. The function for the relative FSOI curve-fit takes the form, where y would be the relative FSOI and x is  $\sigma_o$ :

$$y = f(x) = \frac{c\left(\frac{d^2 x^2}{x^2 + d^2}\right) + b^2}{x^2 + b^2}$$

Extrapolations of such a curve-fit can give an impression of what the Aeolus relative FSOI would be if the random error improved (or degraded).

## 9.9 Variational Quality Control for Aeolus in the IFS

Aeolus L2B winds in the ECMWF operational IFS use a specific type of Variational Quality Control (VarQC) that has been in the code since at least 2010. It is useful to review the current settings and whether they are appropriate for real Aeolus L2B winds.

Before the effects of VarQC (which applies during the 4D-Var minimisation) there is a first-guess check. This permanently rejects Aeolus data from the 4D-Var minimisation if the departure:  $y - H(x) > Z \sqrt{\sigma_o^2 + \sigma_b^2}$ ;  $Z$  has been nominally set to 5 for Aeolus. For Aeolus since  $\sigma_b$  can vary from roughly  $\sim \sigma_o$  for Mie-cloudy to  $\ll \sigma_o$  for Rayleigh-clear, then the first-guess check typically rejects winds of  $> 5 \sigma_o$  to  $> 5\sqrt{2}\sigma_o = 7.07\sigma_o$ , where  $\sigma_o$  is the assigned observation error in data assimilation.

The VarQC method employed for Aeolus was inherited from that used for drifting balloon winds (PILOT) i.e. conventional wind data. It uses the assumption that the normalised departure  $((y - H(x))/\sigma_o)$  statistics take the form of a Gaussian plus a flat (gross error) distribution. This method is described e.g. here:

<https://www.ecmwf.int/en/elibrary/7759-variational-quality-control>

The flat p.d.f. assumption for gross errors is as follows, where  $d$  = a tunable parameter for determining the probability:

$$p^G = \frac{1}{2d\sigma_o}$$

The total p.d.f. assumed for the normalised departures i.e. Gaussian ( $N$ ) plus gross-error flat distribution is as follows, where  $A$  = a tunable parameter for determining the contribution of each term:

$$p^{QC} = (1 - A)N + Ap^G$$

The Gaussian distribution ( $N$ ) part for the observation cost function is:

$$J_o^N = -\ln N + const = \frac{1}{2} \left( \frac{y - H(x)}{\sigma_o} \right)^2$$

The total observation cost function when including the flat (gross error) p.d.f is:

$$J_o^{QC} = -\ln \left( \frac{\gamma + \exp(-J_o^N)}{\gamma + 1} \right)$$

The gradient of the total observation cost function with respect to the model state vector  $x$  is:

$$\nabla J_o^{QC} = \nabla J_o^N \left( 1 - \frac{\gamma}{\gamma + \exp(-J_o^N)} \right) = W \nabla J_o^N = (1 - P) \nabla J_o^N$$

Where:

$$\gamma = \frac{A\sqrt{2\pi}}{(1-A)2d\sigma_o}$$

And  $W$  is considered the “weight” of the observation in data assimilation and  $P$  is the probability of gross error.

Half-weight occurs (i.e.  $W=0.5$ ,) when:  $\frac{y-Hx}{\sigma_o} = \sqrt{-2\ln(\gamma)}$

The settings for Aeolus L2B winds in the ECMWF IFS are:  $A=0.00006$  and  $d=5$ . This leads to a p.d.f., observation cost function and weight function as shown in Figure 187. With Aeolus settings the observations are only down-weighted significantly for departures  $> 4.71\sigma_o$  and hence in most cases the VarQC will not be doing much additional QC to the first-guess check.

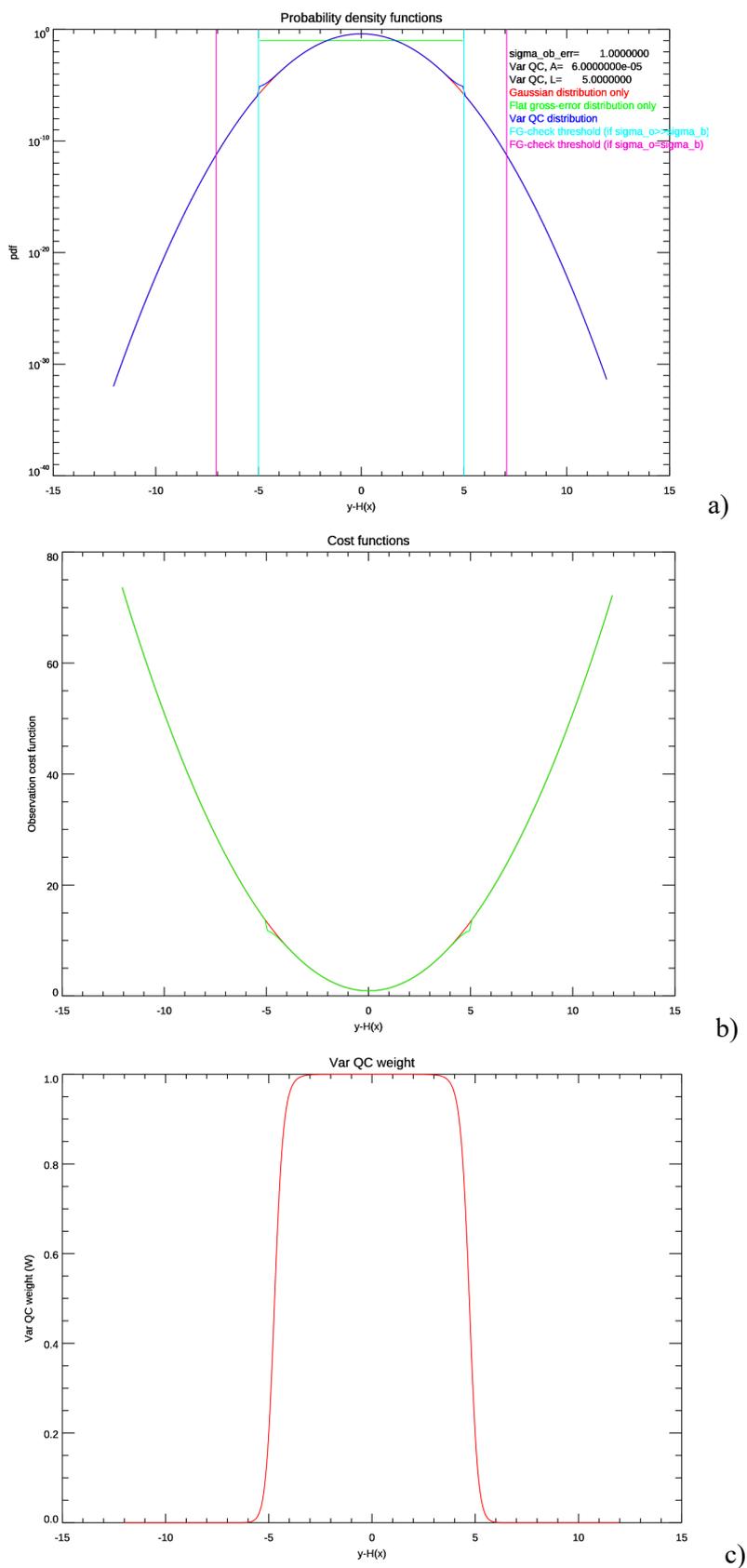
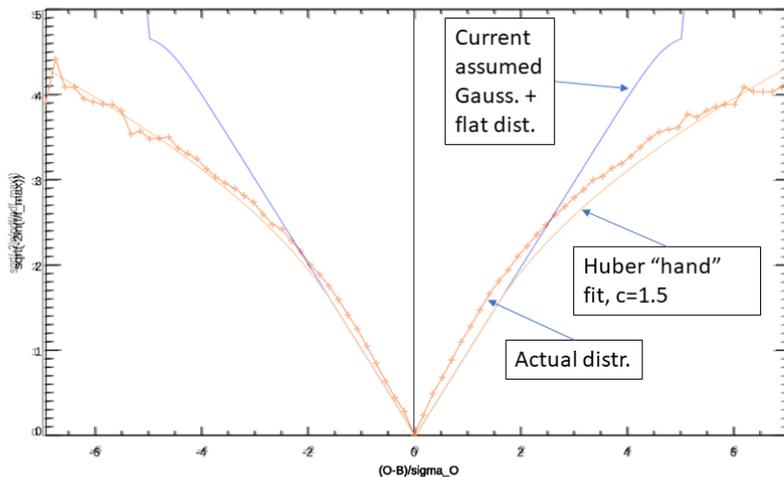


Figure 187. The a) p.d.f.; b) observation cost function; and c) VarQC weight as a function of normalised departures with Aeolus' settings of A and d (as in the operational IFS code). Also shown in

a) is the typical range of the first-guess check in cyan and pink, which will depend on the size of the assigned observation error with respect to the background error. The half-weight occurs at  $4.71\sigma_0$ .

An assessment as to whether a Gaussian plus flat pdf assumption is reasonable for real Aeolus data, is done in Figure 188.

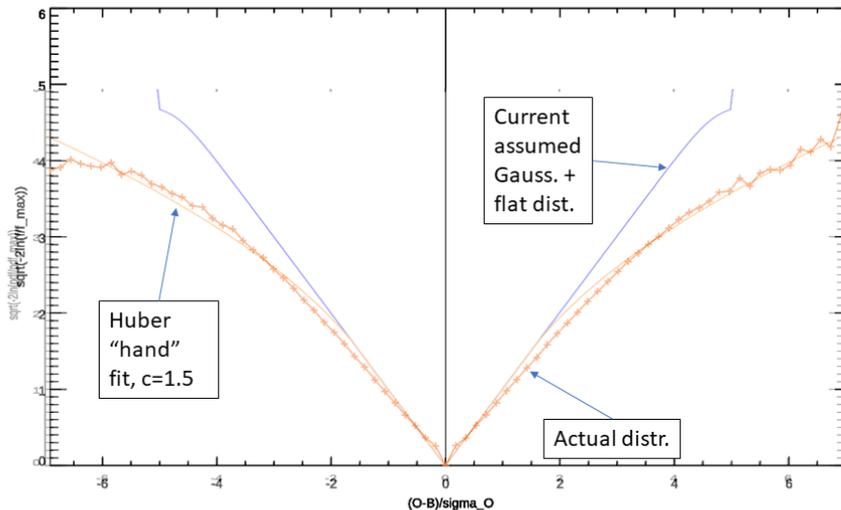
Comparing the **actual Mie-cloudy distribution** (after Gaussian transformation  $\sqrt{-2\ln(f/f_{max})}$ ) to the **transformed VarQC (Gaussian +flat) pdf** with Aeolus settings and a transformed **tuned Huber distribution**



- Huber with c=1.5 looks like a reasonable fit

a)

Comparing the **actual Rayleigh-clear distribution** (after Gaussian transformation  $\sqrt{-2\ln(f/f_{max})}$ ) to the **transformed VarQC (Gaussian +flat) pdf** with Aeolus settings and a transformed **tuned Huber distribution**



- Also, Huber with c=1.5 looks like a reasonable fit

b)

Figure 188. How the real Aeolus normalised departure statistics (orange line with crosses) compares to the Gaussian plus flat p.d.f distribution (blue) and a tuned Huber norm distribution for a) L2B Mie-cloudy and b) L2B Rayleigh-clear winds. September 2021 L2B data.

It was concluded that the current VarQC p.d.f. assumption is not a good fit to real Aeolus normalised departures, however a Huber norm with distribution with c=1.5 provides a much better fit. The Huber norm consists of a Gaussian distribution in the centre and an exponential decay distribution on the edges.

The Huber norm version of VarQC is available in the IFS code for testing and is described here: <https://www.ecmwf.int/en/elibrary/12572-use-huber-norm-observation-quality-control-ecmwf-4d-var>

The equations for the Huber norm p.d.f. are:

$$f(x) = \frac{1}{\sigma_o\sqrt{2\pi}} \exp\left(-\frac{\rho(\hat{d})}{2}\right)$$

$$\hat{d} = \frac{y - H(x)}{\sigma_o}$$

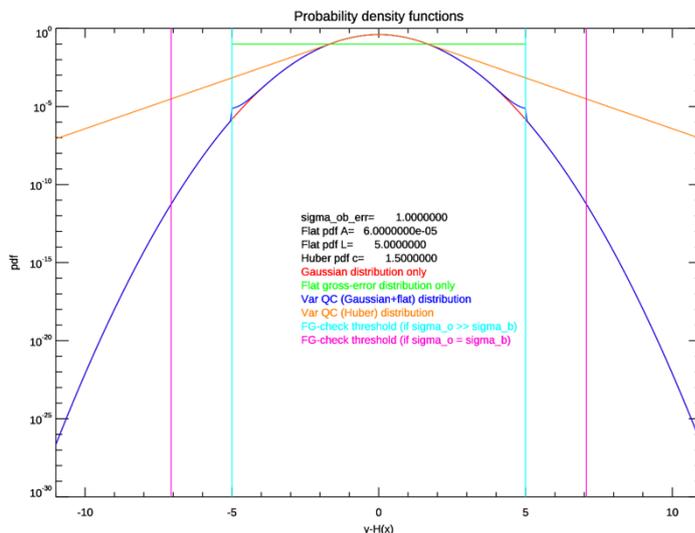
for  $|\hat{d}| \leq c$ :

$$\rho(\hat{d}) = \hat{d}^2 / \sigma_o^2$$

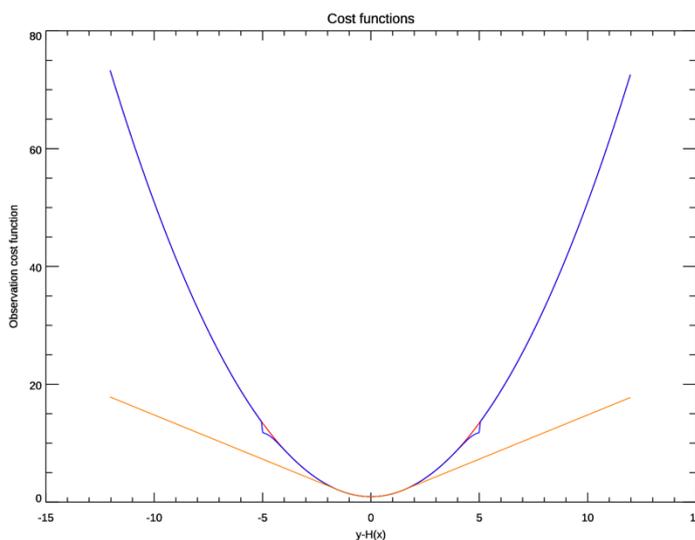
for  $|\hat{d}| > c$ :

$$\rho(\hat{d}) = (2c|\hat{d}| - c^2) / \sigma_o^2$$

A value of  $c=1.5$  was found to fit reasonably well for Aeolus, both Rayleigh-clear and Mie-cloudy. A setting of  $c=1.5$  is visualised in Figure 188. A comparison of the various pdf's is shown in Figure 189.



a)



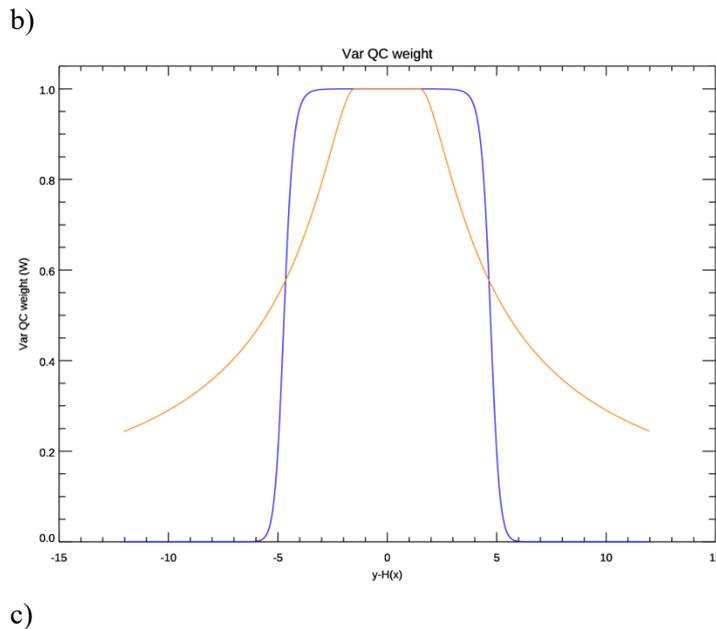


Figure 189. a) Various pdf's b) observation cost function and c) current VarQC (Gaussian plus flat p.d.f.) for Aeolus in blue compared to fitted Huber norm p.d.f. weighting in orange.

The Huber norm with  $c=1.5$  decreases the weight of the observation for  $>1.5\sigma_o$ , but then has long-tails of greater than zero weight. This was tested in an OSE for Aeolus, but it was combined with a huge first-guess check parameter ( $20\sigma$ ). This is typically done with the Huber norm i.e. in effect no first-guess check at all, to give some weight to extreme departures where the background is wrong and not the observation (as found suitable for some conventional observation types). The results however were not positive for Aeolus NWP impact, as shown in section 4.5. In hindsight, the Huber norm should be tested again but with the  $5\sigma$  first-guess check remaining; since Aeolus is known to have some genuine gross-errors (with a complicated distribution!) which should be thrown out.

## 9.10 Some notes on the horizontal scale of L2B winds and N/P settings

The horizontal scale of a L2B wind retrieval is determined by: the instrument ALADIN's laser and N/P settings; satellite orbit/LOS pointing parameters; L2Bp measurement grouping algorithm settings and the classification algorithm into clear/cloudy measurement-bins.

The ALADIN laser is operated in a pulse mode with a rate of 50.5 Hz, so there are 606 laser pulses per 12 second BRC. In terms of the detector settings: N=number of measurements per BRC (Basic Repeat Cycle), P=number of laser pulses per measurement. The detector is constrained to a maximum of  $N*P=600$  laser pulses per BRC; so there will always be at least 6 pulses not detected at the end of the BRC. Also, the detectors only accumulate (P-1) pulses per measurement. The last pulse of each measurement is not used (for reasons TBC). The missing N pulses due to the last per measurement being discarded are distributed evenly across the BRC.

The nominal detector setting to achieve  $N*P=600$  is  $N=30$ ,  $P=20$  and hence  $N*(P-1)=570$  laser pulses are accumulated per BRC (with the gap of 6 pulses not used at the end of the BRC). If  $N*P$  is reduced to  $< 600$ , then a reduced number of pulses are accumulated per BRC, leading to an increasingly large data gap at the end of the BRC. For example, if  $N=30$ ,  $P=19$ , a combination used for most of the

mission (introduced to operations on 15 February 2019 to reduce risk of laser shutdown), then  $N*(P-1)=540$ , so we have 30 fewer pulses detected per BRC than with  $N=30, P=20$ . This leads to a 30-pulse data gap at the end of the BRC, plus the 6 missing pulses that always occurs.

According to latitude/longitude/time data from an example GRND\_TRACK predicted orbit file (in July 2022), within 12 seconds (1 BRC) the Aeolus off-nadir satellite to ground intersection point moves  $\sim 88.75$  km, therefore the speed of the laser as projected on the ground is  $\sim 7.396$  km/s. Given the laser pulse rate of 50.5 Hz, then the time between pulses is  $1/50.5=0.0198$  s, which translates to a horizontal distance of 0.1464 km between pulses.

*Table 6. Estimates of the horizontal scales relevant for L2B winds for various N/P settings employed during the mission; assuming the distance per pulse is 0.1464 km.*

N/P settings	Horizontal extent (km) for 1 measurement i.e. P pulses (include the missing pulse for the gap between adjacent measurements)	Horizontal extent (km) for a classically grouped full BRC L2B wind i.e. all clear or all cloud, extent of N*P pulses	Horizontal gap in data (km) at end of BRC due to not detecting all 606 available pulses	Date N/P implemented in operations
<b>N=30, P=20;</b> <b>N*P=600; N*P-1=570</b>	2.93	87.84	0.88	Start of mission
<b>N=30, P=19;</b> <b>N*P=570; N*P-1=540</b>	2.78	83.45	5.27	15 Feb 2019
<b>N=15, P=38;</b> <b>N*P=570; N*P-1=555</b>	5.56	83.45	5.27	13 Dec 2021
<b>N=5, P=114;</b> <b>N*P=570; N*P-1=565</b>	16.69	83.45	5.27	4 April 2022

Histograms of reported L2B horizontal distance (km) for HLOS winds with L2Bp v3.70 (Baseline 14) FM-A reprocessing can be seen Figure 190. This agrees with Table 6 that for Rayleigh classic grouping, with  $\sim 88$  km for a 1 BRC accumulation with  $N=30, P=20$ , and that this reduced to  $\sim 83$  km with  $N=30, P=19$ . The advanced Mie grouping setting applied allowed accumulations up to 18 km, which consists of up to 6 measurements. The size of each measurement was a bit larger for  $N=30, P=20$  versus  $N=30, P=19$  as can be seen from the histograms.

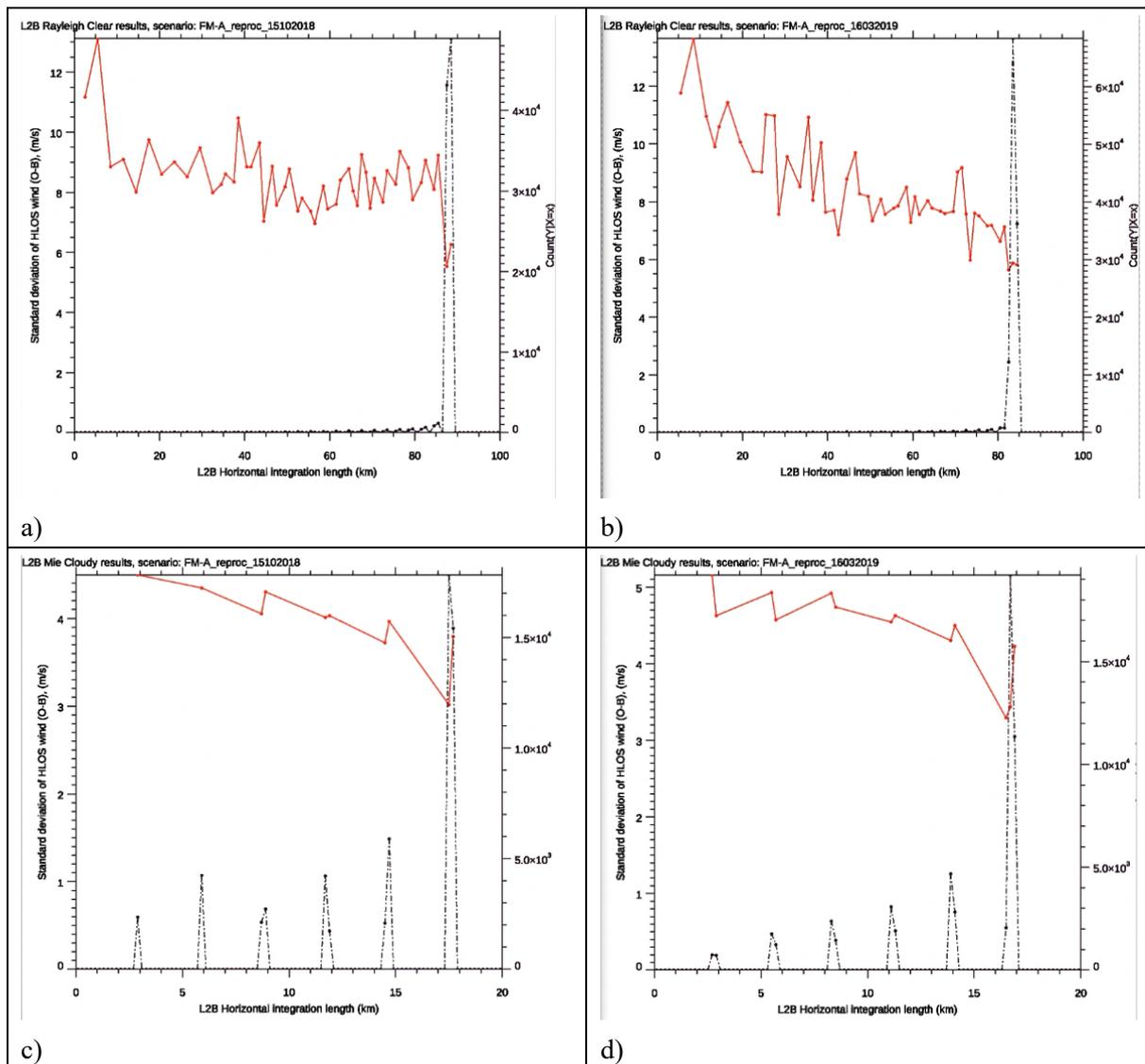


Figure 190. L2B wind reported horizontal distance value (km) histograms from L2Bp v3.70 for FM-A reprocessing. On the top row a) and b) is Rayleigh-clear winds and on the bottom row c) and d) is Mie-cloudy winds. The left column a) and c) is data from 15/10/2018 when the N=30, P=20 setting was applied, and the right column b) and d) is data from 16/03/2019 when the N=30, P=19 setting was applied. The Rayleigh-clear winds used the classic grouping option, which can accumulate measurements up to 1 BRC, and the Mie-cloudy have the advanced grouping option with a maximum accumulation size of 18 km.

### 9.11 History of changes in AUX\_PAR\_2B in NRT operational ground segment processing

Changes in AUX\_PAR\_2B settings are listed, as noted in the history of the NRT operational AUX\_PAR\_2B files. Gaps in the file version timeline are usually associated with changes by PDGS, typically in association with baseline updates and changes for blacklisting. Changes related to blacklisting, due to laser sensitivity, M1 tests and instrument outages (survival mode), are not included since there were so many, it will just confuse the story.

Table 7. List of recorded history in operational NRT AUX\_PAR\_2B files during the mission.

Date in file comments history	Date applied in operations	File version	Baseline	Change
03/10/2018	?	0002	02	Stop flagging Rayleigh-cloudy and Mie-clear invalid by default. Relax RBC_Algorithm_Params (broader range to avoid rejecting data).
11/01/2019	11/01/2019	0005	02	Modify some classification related parameters to improve Rayleigh winds e.g. SR threshold from 1.25 to 1.5(6); min, alt. for allowing SR=1 (when missing values) from 10 to 16 km; LIB Ray SNR range reduced to 2 to 35. Rayleigh-clear random error and number of observations improved. Reduces number of Rayleigh-cloudy winds.
28/02/2019	05/03/2019	0006	02	Modify Mie classification SR threshold by +0.1 (to 1.6). Change to the grouping algorithm from classic to advanced with 10 km max. accum. for Mie and 89 km max. accum. for Rayleigh (effectively 1 BRC still for Rayleigh). The Mie grouping change leads to 2.5-3.0 times more Mie winds, and a 75% increase in L2B EE DBL file size.
30/07/2019	30/07/2019	0008	02	Flag top Rayleigh range-bin, i.e. number 1, as invalid, due to DUDE above ground issue. Switch off the satellite LOS velocity correction for both Mie and Rayleigh, to improve biases.
21/10/2019	21/10/2019	0009	02	Stop flagging Rayleigh range-bin 1 as invalid, after DUDE fix.
23/10/2019	31/10/2019	0010	07	Separate grouping for Mie and Rayleigh. Rayleigh set to classic (1 BRC), Mie set to advanced with 12 km max horizontal group size. Mie classification modified to use Mie SNR with a 4.5 threshold. Mie core alg. QC thresholds tuned for real Aeolus data. Various other settings tuned (for FM-B?).

21/11/2019	29/11/2019	0011	07	For Mie winds update Peak_Height_Upper_Threshold to 1.8 based on DH's suggestion. Switch on Use_OminB_to_Flag_Hot_Pixels with a 10 m/s bias threshold.
13/12/2019	16/12/2019	0012	07	Set Rayleigh manual_hlos_bias_correction to -4.00 m/s
06/01/2020	07/01/2020	0014	07	Change manual_hlos_bias_correction to -6.00 m/s
10/02/2020	10/02/2020	0015	07	Broaden range of Mie core QC for FWHM for meas. and internal ref.
20/02/2020	24/02/2020	0016	07	Tune the Rayleigh classification settings a little; SR threshold increased to 2.0 (from 1.5(6)).
06/03/2020	02/04/2020	0017	08	Initial version with M1 bias correction off. Modify Mie core atm. FWHM range from 1.5 to 2.0.
24/03/2020	22/04/2020	0018	09	Switch on M1 bias correction.
09/07/2020	09/07/2020	0019	10	Switch off Use_OminB_to_Flag_Hot_Pixels to avoid all invalid after M1 thermal control test
16/09/2020	08/10/2020	0020	11	Ray classification changed to use Mie refined SNR with at threshold of 10. L1B_Mie_Emitted_Freq_Stdev_Threshold modified from 100 to 60 MHz.
13/11/2020	16/11/2020	0022	11	Increase range on Mie core atm FWHM thresholds by +0.1 (upper) and -0.1 (lower) to gain 2% more Mie winds
30/11/2020	30/11/2020	0027	11	Improve Rayleigh-clear bias in very low SNR by relaxing the L1B Ray SNR QC range (from 2-35 to 0-38)
14/12/2020	14/12/2020	0030	11	Modify blist to stop flagging Rayleigh winds, since bias steps removed by L2Bp v3.40.3. This was related to steps in the internal

				Rayleigh response – so we stopped using the int RR.
11/01/2021	12/01/2021	0031	11	Modify ScatRatio_Method_for_Decontamination_Clear_Winds to force SR to 1.
17/06/2021	17/06/2021	0050	12	Moonblinding Check is set to True.
01/07/2021	01/07/2021	0051	12	Fitted Nonlinear Correction is set to True.
06/07/2021	06/07/2021	0052	12	Use O-B to Flag Hot Pixels is set to True. Min Bias to Flag a Hot Pixel is set to 15.0 m/s.
01/12/2021	06/12/2021	0063	13	Mie_Core_Algorithm_Params/Start_FWHM value updated from 1.58 to 1.62  Mie_Core_Algorithm_Params_Reference_Pulse/Start_FWHM value updated from 1.33 to 1.30  Mie_Core_Algorithm_Params_Reference_Pulse/FWHM_Lower_Threshold value updated from 1.0 to 1.1
12/12/2021	12/12/2021	0064	13	Grouping_Params_Mie/Max_Horizontal_Accumulation_Length value updated from 15.0 to 18.0, to accommodate N=15, P=38 setting.
14/03/2022	14/03/2022	0066	13	Tighten QC of hot-pixel biases (from 15 to 7 m/s) due to fewer valid DUDEs available.
03/04/2022	03/04/2022	0068	14	With change to N=5, P=114 the SNR per measurement increases, therefore the upper threshold on L1B measurement-scale Rayleigh SNR needs to be increased from 38 to 58
05/05/2022	06/05/2022	0069	14	Mie core 2 algorithm QC thresholds relaxed a little to get more valid Mie winds of reasonable quality.
16/09/2022	16/09/2022	0071	15	Some better QC of outliers via adjustments to Rayl_Response_Check, L1B_Rayleigh_SNR_Check and wind climatological range check.

				Moonblinding_Check adjusted so that wind results are produced but flagged invalid (so can monitor them).
25/11/2022	25/11/2022	0074	15	New Mie NL correction for FM-A 2022 included, continue to blacklist (for FM-A switch)
05/12/2022	05/12/2022	0075	15	Tighten bias range-bin bias check to 5 m/s.
03/01/2023	04/01/2023	0076	15	Refine QC thresholds for L1B meas-scale Rayleigh response
14/02/2023	16/02/2023	0079	15	Turn on Rayleigh-cloudy correction
02/03/2023	02/03/2023	0083	15	For AMD_Matchup_Params change Max_Allowed_Time_Diff updated from 6s to 60s. To allow matching to spatially closest AUX_MET profile when errors are present in the predicted ANX time via GRND_TRACK files
17/04/2023	18/04/2023	0084	16	Change of settings appropriate for baseline 16 and B16 blacklisting
20/04/2023	20/04/2023	0086	16	End of blacklisting
28/04/2023	28/04/2023	0087	16	End-of-life test blacklisting
17/05/2023	18/05/2023	0088	16	Get it working for FM-B during EOL testing
12/06/2023	12/06/2023	0089	16	Update RR_Ref_Method, MieRespFreq_to_RR_[offset slope_unit]
20/06/2023	20/06/2023	0090	16	Update Mie NL for atm path
21/06/2023	22/06/2023	0091	16	Update Ray-cloudy correction parameters

### 9.12 Geometry of Aeolus for the derivation of the Aeolus (H)LOS wind observation operator

This section derives the equations for line-of-sight (LOS) wind and horizontal line-of-sight (HLOS) wind given the vector wind and the pointing geometry of a Doppler wind lidar in space. The geometry is illustrated in Figure 191.

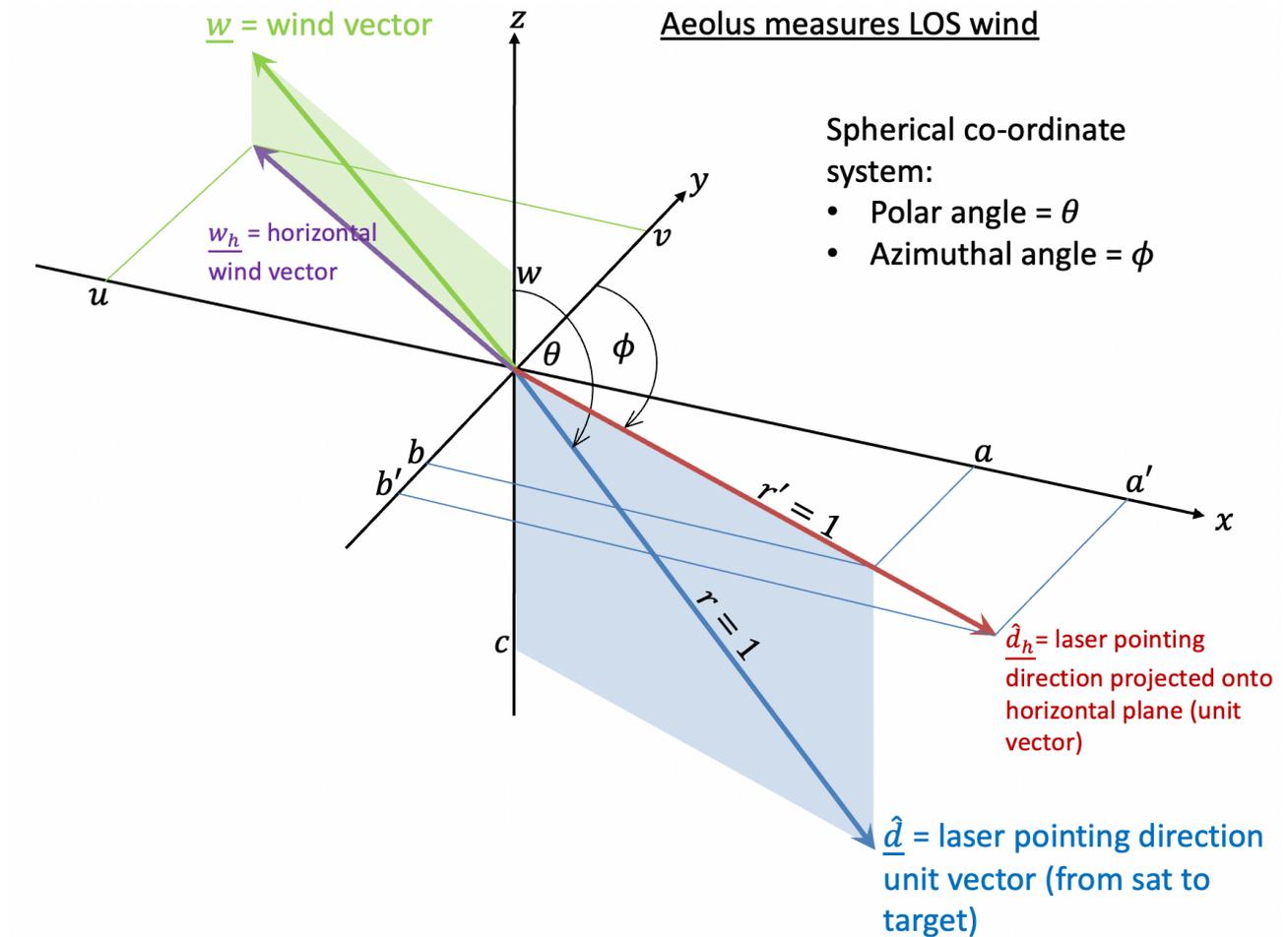


Figure 191. Geometry of Doppler wind lidar in space for use in the derivation of (H)LOS wind from vector wind.

In the spherical co-ordinate system, LOS wind is a function of wind vector field  $\underline{w}$  and unit vector  $\underline{\hat{d}}$  along the line-of-sight :

$$v_{LOS} = \underline{w} \cdot \underline{\hat{d}} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = u \sin \theta \sin \phi + v \sin \theta \cos \phi + w \cos \theta$$

Horizontal LOS (HLOS) wind is the LOS wind projected onto the horizontal plane:

$$v_{HLOS} = \frac{v_{LOS}}{\sin \theta}$$

However, Aeolus L1B/L2B data has a convention of using:

- Azimuth angle as the angle (clockwise from North) of the horizontal projection of the target to satellite pointing vector.
- (H)LOS wind is positive (negative) if the wind blowing away from (towards) the instrument.

Assuming  $w = 0$  (as is done for Aeolus); using the modified definition of the azimuth angle  $\phi' = \phi + \pi$  associated with the unit vector pointing from target to satellite ( $-\underline{\hat{d}}_h$ ); and with winds blowing away from satellite being positive we have:

$$v_{HLOS} = -(\underline{w}_h \cdot \underline{\hat{d}}_h) = -\begin{pmatrix} u \\ v \end{pmatrix} \cdot \begin{pmatrix} -a' \\ -b' \end{pmatrix} = u \sin \phi + v \cos \phi = -u \sin \phi' - v \cos \phi'$$

It is the Aeolus “special” conventions on vector direction for azimuth angle and sign of HLOS wind that leads to the minuses in the HLOS wind formula.

If the assumption of zero vertical wind component is not applied (i.e.  $w \neq 0$ ) then:

$$v_{HLOS} = -u \sin \phi' - v \cos \phi' + w \cot \theta$$

However, Aeolus L1B/L2B data does not provide the polar angle  $\theta$ , but instead the elevation angle  $\theta_e$ , which is illustrated in Figure 192, therefore:  $\theta_e = \theta - \frac{\pi}{2}$

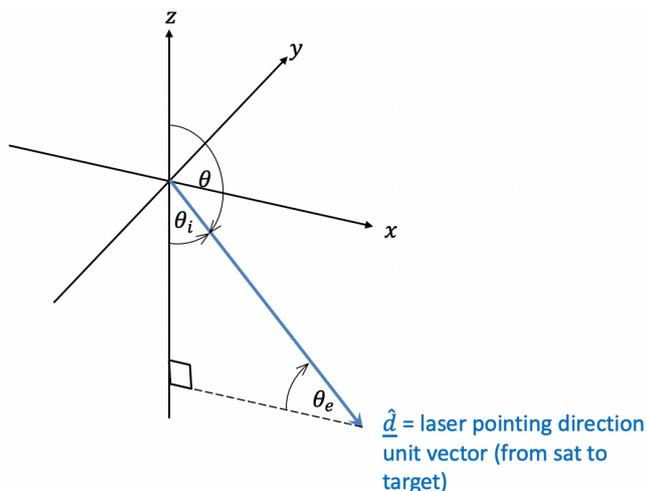


Figure 192. Relationship between the polar angle (spherical co-ordinates), the incidence angle and the elevation angle provided with Aeolus L2B data.

Inserting the elevation angle into the HLOS wind formula gives:

$$v_{HLOS} = -u \sin \phi' - v \cos \phi' - w \tan \theta_e$$

Figure 193 and Figure 194 help to illustrate the typical geometry for Aeolus.

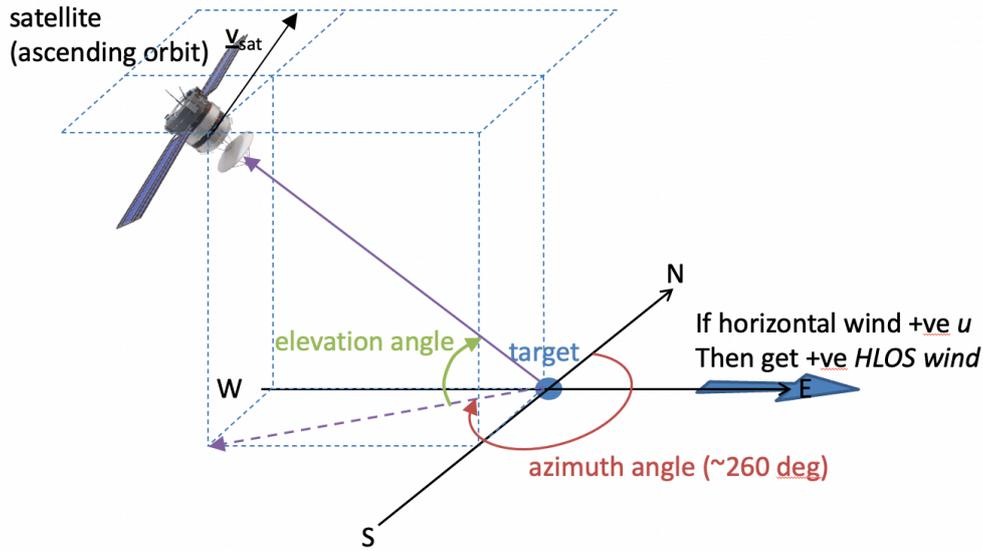


Figure 193. Geometry of Aeolus line-of-sight for ascending orbits.

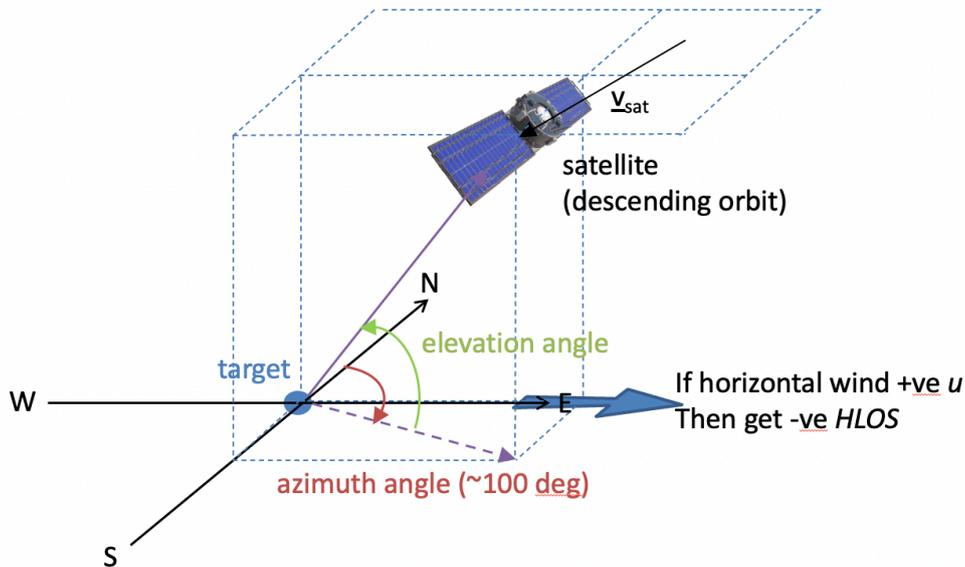


Figure 194. Geometry of Aeolus line-of-sight for descending orbits.

A propagation of uncertainty for the HLOS wind observation operator is:

$$v_{HLOS} = -u \sin \phi' - v \cos \phi' - w \tan \theta_e$$

$$\sigma_{v_{HLOS}}^2 = \left| \frac{\partial v_{HLOS}}{\partial u} \right|^2 \sigma_u^2 + \left| \frac{\partial v_{HLOS}}{\partial v} \right|^2 \sigma_v^2 + \left| \frac{\partial v_{HLOS}}{\partial w} \right|^2 \sigma_w^2 + \left| \frac{\partial v_{HLOS}}{\partial \phi'} \right|^2 \sigma_{\phi'}^2 + \left| \frac{\partial v_{HLOS}}{\partial \theta_e} \right|^2 \sigma_{\theta_e}^2$$

$$\sigma_{v_{HLOS}}^2 = \sigma_u^2 \sin^2 \phi' + \sigma_v^2 \cos^2 \phi' + \sigma_w^2 \tan^2 \theta_e + \sigma_{\phi'}^2 (u^2 \cos^2 \phi' + v^2 \sin^2 \phi' - 2uv \sin \phi' \cos \phi') + \sigma_{\theta_e}^2 w^2 \sec^4 \theta_e$$

The total differential for the HLOW wind observation operator is:

$$\begin{aligned}
 dv_{HLOS} &= \frac{\partial v_{HLOS}}{\partial u} du + \frac{\partial v_{HLOS}}{\partial v} dv + \frac{\partial v_{HLOS}}{\partial w} dw + \frac{\partial v_{HLOS}}{\partial \phi'} d\phi' + \frac{\partial v_{HLOS}}{\partial \theta_e} d\theta_e \\
 &= -du \sin \phi' - dv \cos \phi' - dw \tan \theta_e + (v \sin \phi' - u \cos \phi') d\phi' - d\theta_e w \sec^2 \theta_e
 \end{aligned}$$