ESA Contract Report

ESA Contract No. 4000131086/20/NL/FF/a

Contract Report to the European Space Agency

Final Technical Note of "Impact assessment of commercial GNSS-RO data"

Katrin Lonitz[*], Christian Marquardt[$], Neill Bowler[+]
and Sean Healy[*]
* ECMWF $ EUMETSAT + UK Met Office

November 15, 2021

Series: ECMWF ESA Contract Report Series

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

Contact: library@ecmwf.int

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

Contents

Abstract

The addition of Spire and COSMIC-2 data is a large increase in the GNSS-RO data being assimilated at ECMWF and Met Office NWP systems. Both institutions have investigated the effect of Spire and COSMIC-2 data running Observing System Experiments in their state-of-the-art NWP systems. They found that the addition of COSMIC-2 and Spire observations is beneficial, with adding Spire on top of COSMIC-2 showing further improvements for temperature, humidity and wind. In conclusion, Met Office and ECMWF would assimilate the Spire data operationally if it becomes available. Additionally, ECMWF have compared the impact of how the Spire data is processed, by comparing bending angles provided directly by Spire with Spire bending angles produced at EUMETSAT. Furthermore, this extensive GNSS-RO dataset also provides the opportunity to investigate how the impact of GNSS-RO data scales with observation number. For this reason, we have performed a set Ensembles of Data Assimilation (EDA) experiments. We have found that increasing the number of GNSS-RO data reduces the ensemble spread, especially at 100 hPa and higher and in the Tropics. In this region, we also found a linear relationship between EDA spread and variance in first guess departures of radiosonde temperature observations. For the Northern and Southern Hemisphere the decrease in EDA spread was less strong when more GNSS-RO data has been added, which might be linked inter alia to the heterogeneously distributed GNSS-RO data in the outer-tropics.

1 Introduction

This is the final report for ESA Contract No. 4000131086/20/NL/FF/an. It complements new analysis but also the work which was presented in previous technical reports. Sections 2 and 4 were part of TN3, but are complemented with some figures from TN2 and new ones. In section 3 the various experiments have been introduced, which were part of TN1. In TN1, the thinking behind the EDA experiments was introduced as well, which is now covered in section 7, where new analysis on the Ensembles of Data Assimilation experiments is summarised. Additional analysis on observing system experiments excluding Spire in the Tropics, as well as adding Spire in the absence of aircraft data are added in section 5. Furthermore, a study on re-processed Spire data by EUMETSAT and its assessment with the ECMWF data assimilation system has been added as section 6.

Background on the GNSS Radio Occultation measuring technique

The Radio Occultation technique uses radio signals passing through the atmosphere in a limb-sounding geometry to sense atmospheric properties. The GNSS-RO measurement geometry is illustrated in Figure 1. The GNSS-RO instruments in LEO orbits receive radio frequency (L-band) signals of opportunity transmitted from GNSS satellites, such as GPS and Galileo.

The history of GNSS Radio Occultation for Atmospheric Sounding started with the GPS/MET research mission in the mid 1990s (Kursinski et al. [1996], Rocken et al. [1997]). The GPS Receiver for Atmospheric Sounding (GRAS) instrument, developed by ESA starting in the late-nineties, has been flying on the MetOP-A satellite since October 2006, MetOP-B since September 2012 and MetOP-C since November 2018. GRAS is being operated by EUMETSAT. The US/Taiwanese collaboration led to the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) constellation which is an operational demonstration mission [Anthes et al., 2008]. Building upon the success of the COSMIC program, COSMIC-2 was launched in 2019 and is a network of six remote-sensing microsatellites.

GNSS-RO has a significant impact in NWP (Healy and Thépaut [2006],Cucurull et al. [2007], Aparicio and Deblonde [2008],Poli et al. [2009], Rennie [2010]) because it complements the information provided

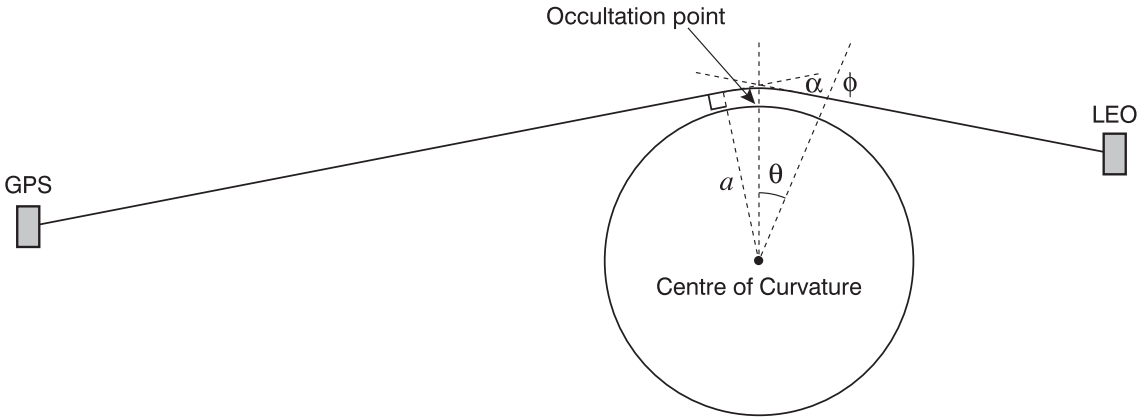


Figure 1: The geometry of the GNSS-RO measurement technique. A GNSS navigation signal is received on a low earth orbiting satellite. The ray-path is bent by refractive index gradients. The figure illustrates the bending angle, α, and impact parameter value, a. ϕ is the angle between the local radius vector and the tangent to the ray-path, and θ is the angular position within the 2D plane.

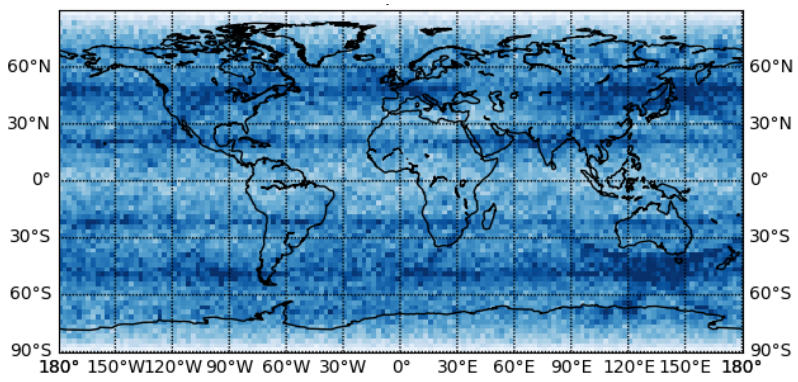
by satellite radiance measurements. This is because it can be assimilated without bias correction, and it has very good vertical resolution. The GNSS-RO technique relies on post-processing of the GNSS carrier phase measurements (level-0 data), from which it is possible to retrieve atmospheric bending angles (level-1 data) and with Abel inversion techniques refractivity profiles (level-2 data) and the related temperature, pressure and humidity profiles [e.g. Kursinski et al., 1997a].

2 Spire and COSMIC-2 Data

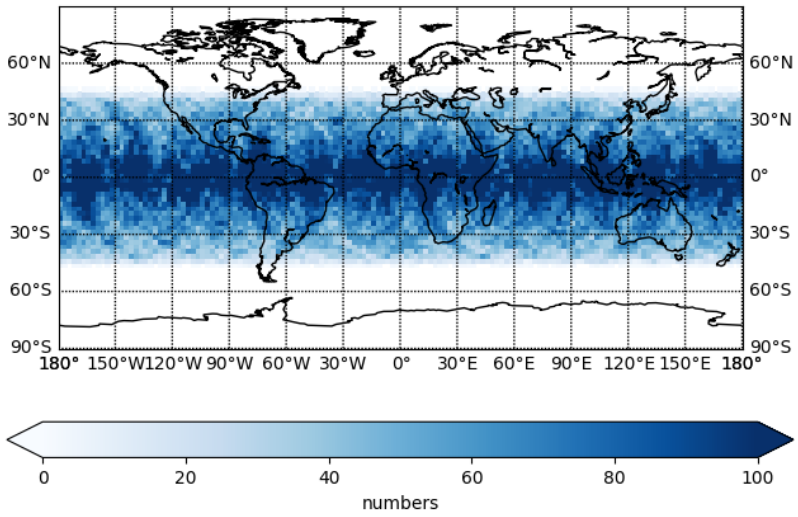
The radio occultation data from Spire is collected using their constellation of 3U cubesats, with the nominal lifetime of these satellites being 2+ years. Regularly new satellites are launched with an updated hardware [Masters et al., 2019]. For example, newer satellites have larger solar panels and two occultation antennas allowing the simultaneous collection of both rising and setting occultations. Spire develops the satellites and the GNSS receiver themselves, and also operates a network of 30+ ground stations which provide support e.g. for data downlinks.

COSMIC-2 (Constellation Observing System for Meteorology, Ionosphere & Climate – 2) is a follow up mission from COSMIC-1, which was was very successful. The cluster of six satellites were launched on June 25, 2019 and spreads along different low-inclination orbits. The main aim of this mission is to obtain temperature and humidity information as well as having a substantial contribution to space weather.

For this study we received more than 5,000 radio occultation (RO) profiles from Spire per day, globally, and for COSMIC-2 around 4,000 RO profiles per day, in a latitude band between ± 40 degrees. Fig. 2 show the geographical distribution of the data density and Fig. 3 displays how many occultations per 12 hours from Spire and COSMIC-2 have been assimilated. From those figures it can be seen that COSMIC-2 provides almost twice as much data in the Tropics than Spire.



(a) Spire



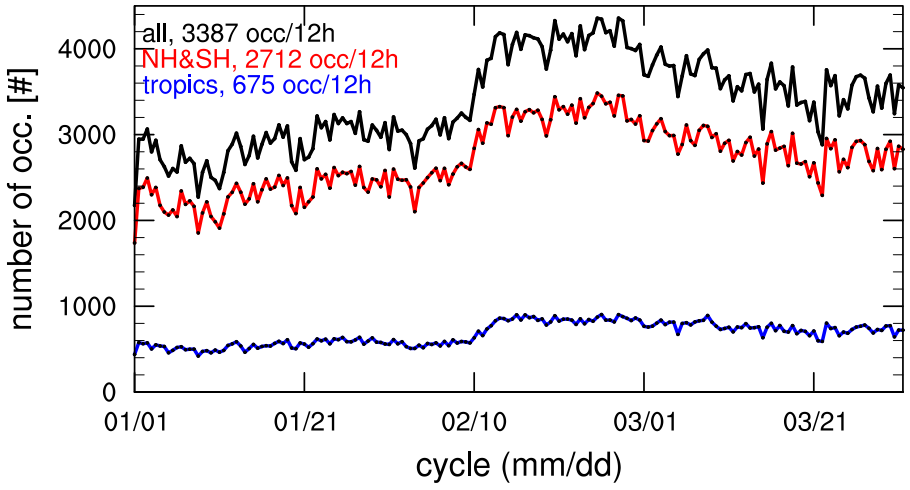
(b) COSMIC-2

Figure 2: Total number of occultations for a) Spire and b) COSMIC-2 between January and March 2020, regridded to a 2.5 x 2.5 latitude/longitude grid.

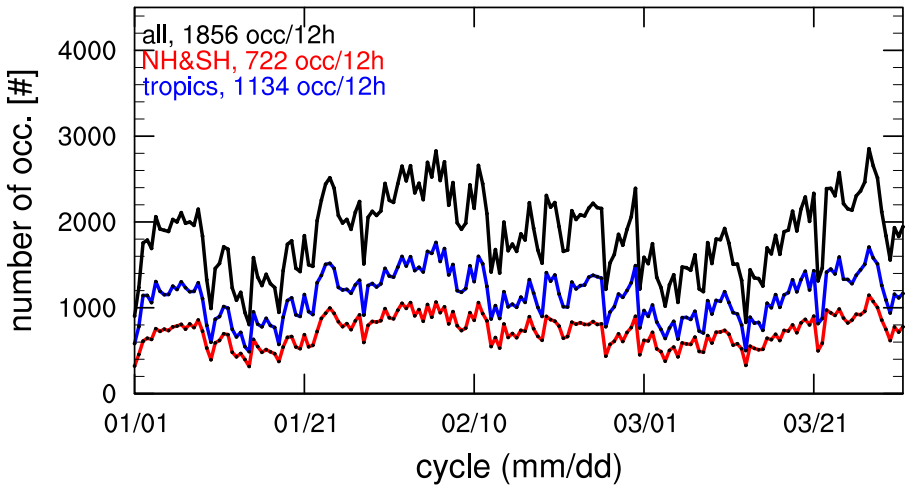
Data Acquisition

Spire have made available their GNSS-RO data from 1 January to 31 March 2020 for the purposes of this study. The Met Office and ECMWF both assimilate bending angles, and the level-1B bending angle data in BUFR format have been downloaded from Spire independently by both centres. The Met Office identified approximately 7 hours of missing data on 4 February 2020. These Spire data were available in the Spire near real time (NRT) data set which is routinely monitored at the Met Office. The missing data were quickly identified and subsequently supplied without issue. The complete data set has been analysed in this study by ECMWF and the Met Office. Additionally, the level-1b bending angle Spire data processed by EUMETSAT have been made available in May 2021.

At ECMWF and the Met Office, the Spire bending angle profiles have been processed and archived in six hour batches required for assimilation. For example, an archived file valid at time 12:00 on a given day, will contain Spire GNSS-RO profiles from 09:00 to 14:59 from that day. The Met Office and ECMWF have also downloaded COSMIC-2 data provided by UCAR prior to the data being available in NRT.



(a) Spire



(b) COSMIC-2

Figure 3: Number of occultations per 12 hour cycle for Spire and COSMIC-2 between January and March 2020 at 5 km to 6 km height. Black refers to global numbers, blue to tropical (20°S to 20°N) and red to the remaining ones. The numbers on the top left, are averages of number of occultations per 12 hour assimilation cycle.

These data have also been processed into 6 hour batches required for assimilation purposes.

EUMETSAT processed raw Spire data to level-1b bending angles in BUFR format for the purposes of this study. This has been a considerable undertaking, and EUMETSAT and Spire have been working together through various minor issues in order to achieve it. More details will follow in section 6.

3 Set of experiments

We have run various experiments at ECMWF and Met Office testing the additional assimilation of Spire and COSMIC-2 data into our four-dimensional (4D-Var) assimilation systems. The ECMWF experiments use the model cycle 47R1, whereas the Met Office uses the setup from parallel suite 43 (operational on 4 Dec 2019). The experiments have run from 1 January to 31 March 2020 using level-1b bending angle data. The chosen horizontal resolution for the ECMWF experiments is about 25 km (TCo399) and for the Met Office experiments about 40 km in the mid-latitudes (N320, 640x480 points on a regular latitude-longitude grid).

The ECMWF experiments are built on top of a **CONTROL** one, which assimilates all observations as used operationally for the experimental period. This includes the GNSS-RO data available before COSMIC-2 and Spire. The experiments are:

- **CONTROL**: This experiment assimilates all observations used operationally for the experimental period. This includes the GNSS-RO data available before COSMIC-2 and Spire.
- **Spire**: This is the **CONTROL** experiment plus Spire data.
- **COSMIC2**: This is the **CONTROL** experiment plus COSMIC-2 data.
- **Spire+COSMIC2**: This is the **CONTROL** experiment plus both Spire and COSMIC-2 data.
- **NoRO**: This is the **CONTROL** experiment but all GNSS-RO has been removed.
- **Only-GRAS**: This is the **NoRO** experiment but the three Metop GRAS instruments are assimilated.
- **Only-KOMP5**: This is the **NoRO** experiment but KOMPSAT-5 measurements made with an IGOR receiver are assimilated.
- **Only-Spire**: This is the **NoRO** experiment but Spire data is assimilated.
- **Spire-EUM**: This is the **CONTROL** experiment plus Spire data processed at EUMETSAT is assimilated.

The Met Office reproduced the **COSMIC2** and **Spire+COSMIC2** experiments in their system, to investigate the impact of Spire data when COSMIC-2 is assimilated.

ECMWF has run "10 member" Ensembles of Data Assimilation (EDA) for each experiment listed above. For clarity, recall that an EDA experiment runs a number of 4D-Var computations in parallel [e.g. Harnisch et al., 2013] . Each 4D-Var run is "perturbed", where noise has been added to the observations and some model parameters, in order to produce a set or "ensemble" of analyses and short-range forecasts. The number of perturbed 4D-Var calculations run in parallel in the EDA is usually called the number of "ensemble members". We will only run the EDA experiments for around 5 weeks because Harnisch et al. [2013] showed that the spread values averaged over the tropics and the northern and southern extratropics are stable after about 1 week.

In addition to the experiments listed above we also investigate the ability of Spire data to mitigate the loss of all aircraft observations for the experimental period. This leads to two additional experiments:

- **noair**: This is the CONTROL experiment but with all aircraft data removed.
- **Spire, noair**: This is **Spire** minus all aircraft observations.
- **Spire+COSMIC2, noair**: This is **Spire+COSMIC2** minus all aircraft observations.
- **Spire+COSMIC2, no T air**:This is **Spire+COSMIC2** minus all temperature aircraft observations.

4 Verification of ECMWF and Met Office results from Observing System Experiments

In this part of the study the following selection of experiments (listed in section 3) are analysed:

- **CONTROL**: This experiment assimilates all observations used operationally for the experimental period. This includes the GNSS-RO data available before COSMIC-2 and Spire.
- **COSMIC**: This is the **CONTROL** experiment plus COSMIC-2 data.
- **Spire**: This is the **CONTROL** experiment plus Spire data (only performed by ECMWF).
- **Spire+COSMIC**: This is the **CONTROL** experiment plus both Spire and COSMIC-2 data.
- **NoRO**: This is the **CONTROL** experiment but all GNSS-RO has been removed.

Both, the Met Office and ECMWF NWP systems can be considered state of the art but they have differences and, hence, we expect some variations in the overall impact when assimilating the GNSS-RO data. Some of the principal differences are:

- The forecasting models are formulated differently. The ECMWF model is based on spherical harmonics, whereas the Met Office model uses a finite difference scheme [Walters et al., 2017, Davies et al., 2005] on a regular latitude-longitude grid.
- The resolution of the modeling systems is different, with more than a factor of two in the horizontal resolution.
- There are large differences in the data assimilation systems. The ECMWF system uses a 12h assimilation window, whereas the Met Office system [Rawlins et al., 2007, Clayton et al., 2013] uses a 6h window.
- The forward operator used to model bending angles [Rennie, 2010]. ECMWF runs a 2D bending angle operator and Met Office a 1D version.
- The weight given to GNSS-RO observations [Bowler, 2020].

Given these substantial differences one will expect to see differences as well as common features in the impact of RO observations on the NWP systems. These similarities and differences will be highlighted in the following analysis.

4.1 Verification of medium-range forecasts against ECMWF operational analysis

Here, we compare the forecast error statistics in the experiments (**Spire**, **COSMIC-2** and **Spire+COSMIC-2**) with the corresponding statistics in **noRO** or **CONTROL**, e.g. for temperature, humidity and wind. For this reason we take the standard deviation of the forecast errors in the experiments, $\sigma_{exp}(z)$, and compare it with the standard deviation of **noRO** or **CONTROL**, $\sigma_{noRO}(z)$ or $\sigma_{CONTROL}(z)$, respectively. Hence, if the experiments have a positive impact $(\sigma_{exp}(z) - \sigma_{noRO}(z))/\sigma_{noRO}(z)$ or $(\sigma_{exp}(z) - \sigma_{CONTROL}(z))/\sigma_{CONTROL}(z)$ should be negative.

4.1.1 Met Office

Overall, the assimilation of RO observations within the Met Office system is highly beneficial. The forecast errors are greatly reduced by the addition of this data for many variables and at various lead times. The benefits are largest in the short range, since errors in the forecast model have greater influence in the medium range. Additionally assimilating Spire observations further reduces the forecast errors across almost all variables. The following section summarises key results that illustrate this overall assessment.

Temperature

In order to examine the performance of the forecasting system we plot the standard deviation of the forecast error for temperature at 100 hPa, as measured against ECMWF operational analyses. Figure 4 shows the change in the standard deviation of forecast error, calculated as a percentage change against the **no-RO** experiment. For all regions the addition of RO observations greatly reduces the forecast-error standard deviation, especially in the short-range.

In the northern-hemisphere extra-tropics the reduction in the standard deviation from assimilating RO observations reduces from approximately 16% to approximately 2% as the forecast lead time increases. The additional benefit from Spire observations reduces from approximately 4% to 1% with increasing forecast lead time. In the tropics there is a similar reduction in the standard deviation due to the assimilation of RO observations, but the additional benefit from Spire data is smaller, and reduces to approximately zero in the medium range. In the southern-hemisphere extra-tropics the effect of assimilating RO observations is much larger, reducing the standard deviation by approximately 34% in the short range and around 2% in the medium range. However, the additional benefit for assimilating Spire RO is relatively modest in this region, reducing from approximately 2% to around zero as the lead time increases. It is surprising that the impact of Spire observations should be small in the southern-hemisphere extra-tropics since this region depends heavily on satellite observations, and benefits were largest there in a previous study by Bowler [2020].

Relative humidity

The verification of relative humidity forecasts is more difficult than temperature, since this field is less well-observed and model biases can play an important role. Therefore, these results are calculated against radiosonde observations rather than against ECMWF analyses.

Figure 5 shows the change in the standard deviation of the forecast error for relative humidity at 100 hPa from assimilating RO observations. This does not show the same smooth reduction in benefit with increasing lead time that was seen for temperature. The figures are also somewhat noisy, due to the small number of observations. The northern-hemisphere extra-tropics is the only region for which the addition of Spire observations appears to reduce the standard deviation of forecast error. For the other regions the

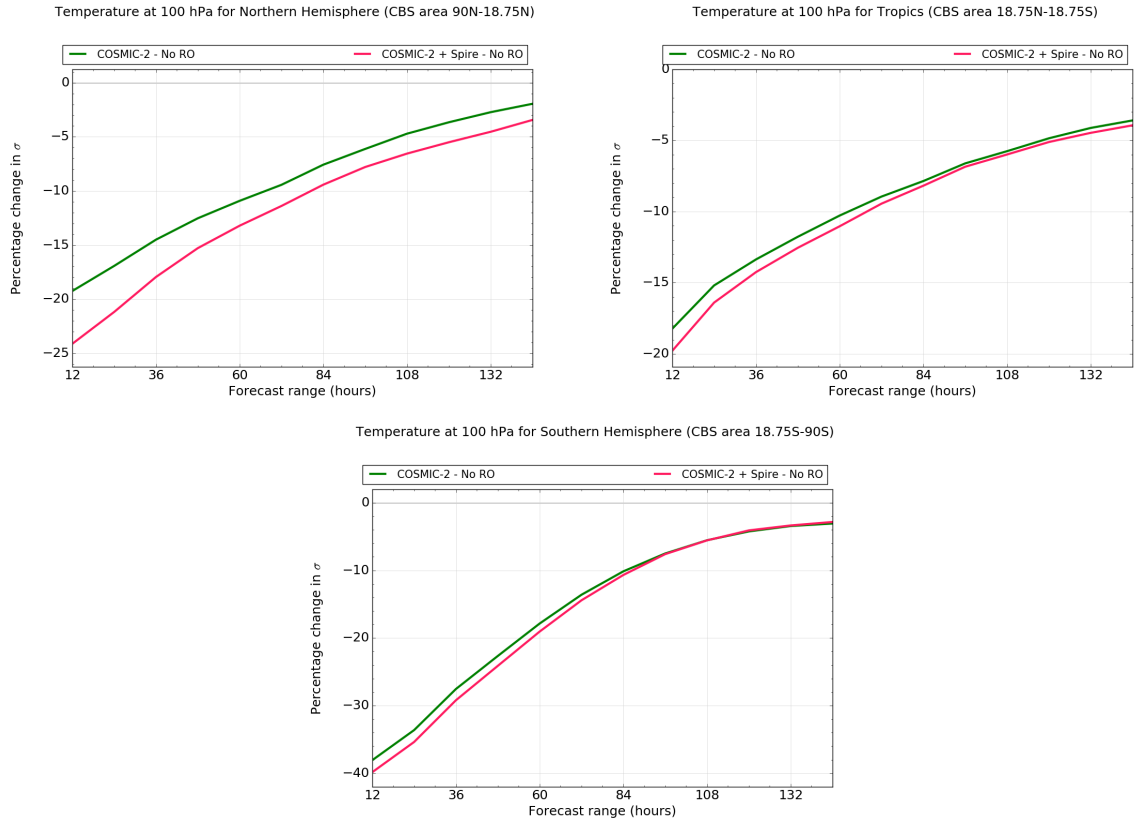


Figure 4: Percentage change in standard deviation of forecast error for temperature at 100 hPa, as measured against ECMWF operational analyses. The change is measured relative to the experiment with no RO observations. The x-axis shows the forecast lead time in hours.

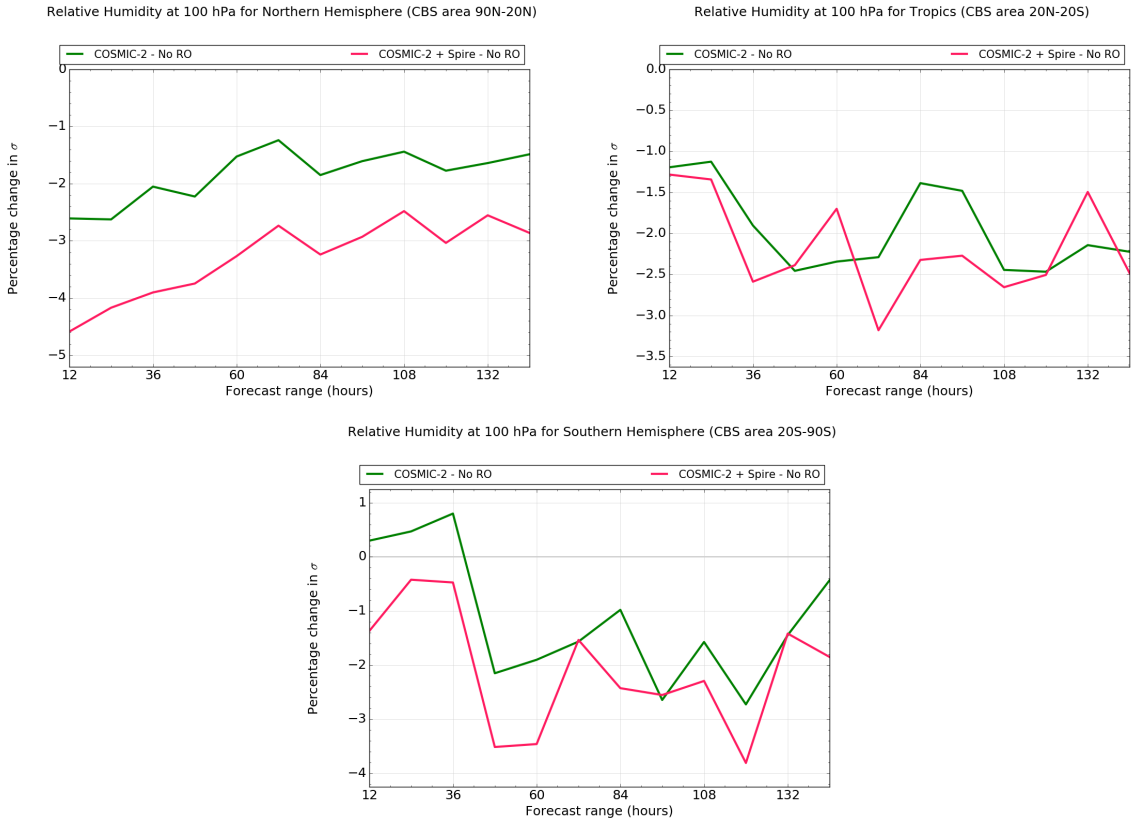


Figure 5: Percentage change in standard deviation of forecast error for relative humidity at 100 hPa, as measured against radiosonde observations. The change is measured relative to the experiment with no RO observations. The x-axis shows the forecast lead time in hours.

benefit is unclear.

4.1.2 ECMWF

Here, the medium-range forecast scores for ECMWF are shown for temperature, humidity and wind, as measured against ECMWF operational analyses consistent with the Met Office.

Temperature

In Fig. 6 we show how the normalised standard deviation for temperature changes compared to **CONTROL** (dashed line around 0) for selected pressure levels and regions (Tropics, Southern Hemisphere and Northern Hemisphere). When adding the new GNSS-RO dataset from Spire and COSMIC-2 a reduction in normalised standard deviation could be seen for all pressure levels (10 hPa, 50 hPa, 100 hPa, 200 hPa, 500 hPa, 850 hPa) and regions (not shown for 50 hPa, 200 hPa and 500 hPa) for forcast day 2 and longer. However, when removing all (existing) GNN-RO we see an increas. In other words, the temperature forecast scores have improved for **Spire**, **COSMIC2** and **Spire+COSMIC2** compared to **CONTROL** throughout the atmosphere, whereas **noRO** shows a degradation. As expected from the sensitivity of radio occultation data the biggest improvements happen for 200 hPa and higher. For example, **Spire+COSMIC2** shows a reduction of about 6% for forecast day 3 at 10 hPa in the Tropics compared to **CONTROL** (Fig. 6a). It also can be seen that the reduction in normalised standard deviation is largest

in the Southern Hemisphere for all experiments (left panel in Fig. 6) compared to **noRO**. For **Spire** the decrease at 100 hPa is about 6% in the Southern Hemisphere, 3% in the Tropics and 4% in the Northern Hemisphere compared to **noRO**. In the Tropics, the improvement of temperature forecast scores is statistically larger for **COSMIC2** and **Spire+COSMIC2** than for **Spire** at 100 hPa and higher (Fig. 6b, Fig. 6a) compared to **CONTROL**. In the Northern and Southern Hemisphere and at lower altitudes the differences are not significant (not shown).

For most height levels, the reduction in temperature forecast error is decreasing with forecast time. This is expected as the assimilation of additional GNSS-RO data only improves initial conditions, and not the forecast model. That means, the imperfect forecast model (with assumptions about physical processes, parameterisations, coarser resolution than real life, etc.) will dominate the change in forecast scores for longer forecast ranges. The decrease in reduction of forecast errors can be seen in Fig. 7 at 100 hPa for **Spire+COSMIC2** compared to **noRO**. For T+24h reductions in temperature forecast errors of more than 20% can be seen in many areas, especially in the Southern Hemisphere. At day 4 (T+96h) of the forecast range these reductions decrease to around 8-12%.

For short-range forecast scores (T+12, T+24) an increase in normalised difference in standard deviation can be seen, e.g. at 10 hPa in the Extratropics for **Spire+COSMIC2** and **Spire** compared to **CONTROL** (see Fig. 6a). This is often not a real degradation and can occur when observations are added in data-poor areas [Geer et al., 2010, section 4.2.4]. Through this additional data the analysis is perturbed relative to the forecast and it looks like that the forecast is worse, when verified against analysis than if the additional observations would not been added at all. Already Bouttier and Kelly [2001] noted this behaviour when using own analysis as a verification basis. In fact, when plotting normalised differences in standard deviations of temperature verified against own analysis a strong increase can be seen for the short-range forecast (not shown).

In other words, the short-range verification at lead times of T+12 or T+24 using analyses as a reference can be problematic. Here, errors in the analysis and forecast have similar magnitudes and can be correlated [Bormann et al., 2019]. When e.g. Spire data is added in the Extratropics correlations are reduced with a comparable smaller reduction in forecast errors leading to apparent degradations in the short-range forecast at 10 hPa. Hence, for the evaluation of short-range forecasts, fits to independent observations are usually checked. One exception seems to be in the Northern Hemisphere at 850 hPa, where the reduction of forecast errors is still big enough to show an improvement of 1% for T+12h.

Relative humidity

For relative humidity (Fig. 8) we can see similar features as for temperature (Fig. 6). with e.g. the largest reduction in normalised standard deviation of forecast errors at higher altitudes (see changes in scales) when the additional GSS-RO data is added compared to **CONTROL**. This is probably less surprising as signals in relative humidity at high altitudes (especially for 100 hPa and higher) are mostly temperature signals. In the Tropics, **COSMIC2** and **Spire+COSMIC2** show significantly larger reduction for humidity than **Spire**, especially at 100 hPa and higher (middle panel in Fig 8b). Here, also the reduction at forecast day 2 is higher for **COSMIC2** and **Spire+COSMIC2** than for **Spire** compared to **CONTROL**. We believe this is due to a better signal-to-noise ratio (SNR) for COSMIC-2 data, as a study by Shu-peng Ho from NOAA/STAR presented at IROWG-8 showed that COSMIC-2 higher SNR measurements provide better water vapour retrievals [Ho and Zhou, 2021]. Small apparent degradations for T+12h and T+24h can be found e.g. in the Tropics at 850 hPa which can be explained by the fact that an analysis is used as a reference (discussed earlier for temperature). Interestingly, in the Southern Hemisphere a larger reduction in humidity is visible for **Spire** and **Spire+COSMIC2** compared to **COSMIC2** at 850 hPa, which is mostly likely influenced by the higher number of Spire occultations (compared to

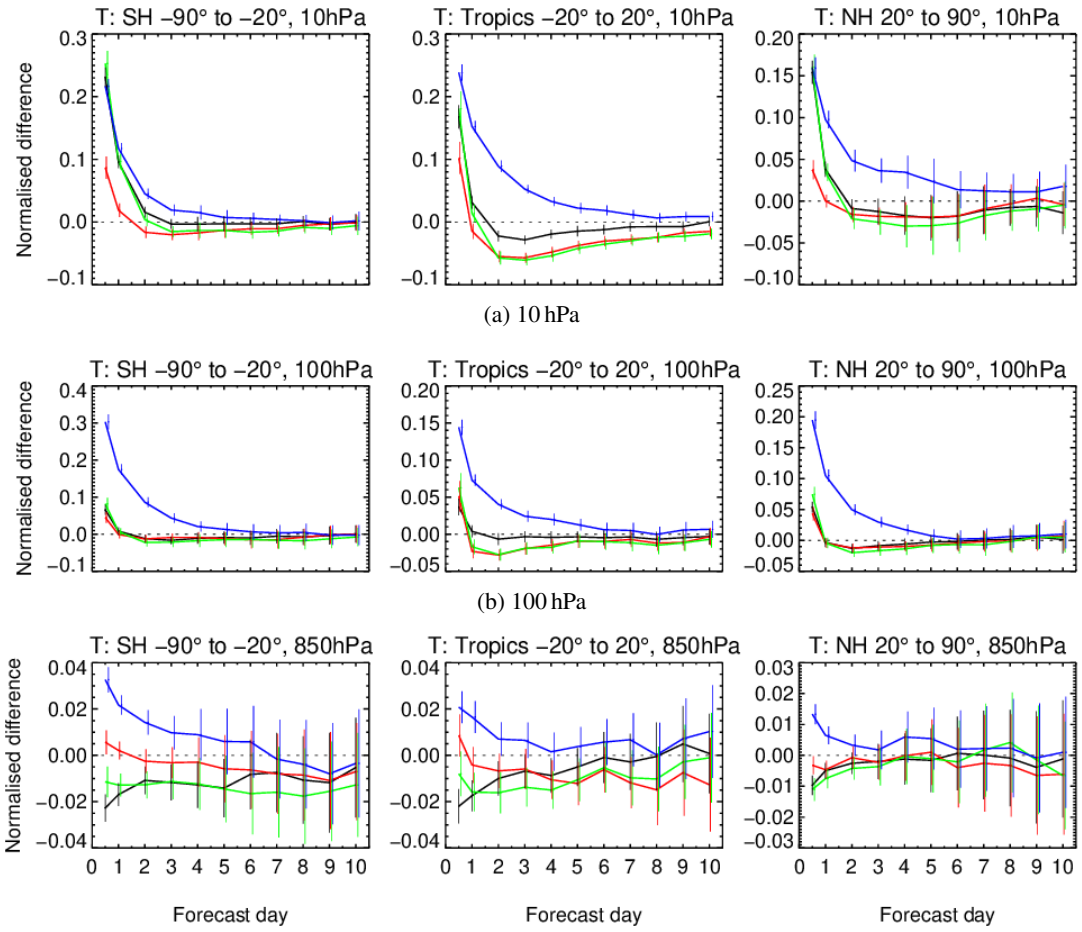
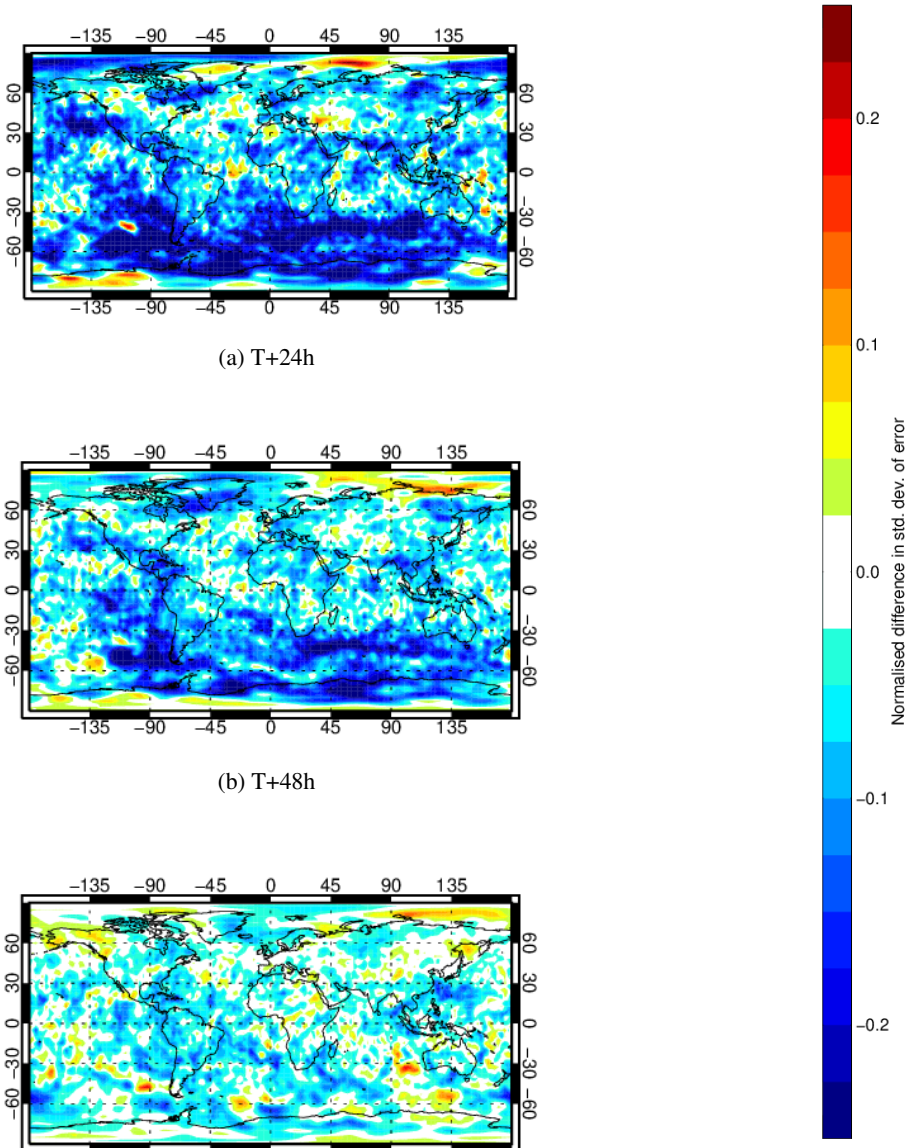


Figure 6: Normalised differences in standard deviation in temperature between **SPIRE**, **COSMIC2**, **SPIRE+COSMIC2**, **noRO** and **CONTROL**, verified against ECMWF operational analysis at different height levels (850 hPa, 100 hPa, and 10 hPa) for different forecast times and zonal regions (left: SH 90°S-20°S, middle: Tropics 20°S-20°N, right: NH 20°N-90°N). Negative values represent a decrease in standard deviation and positive values an increase in standard deviation. The confidence range is displayed by vertical bars.

COSMIC2) in the presence of less other observations compared to the Northern Hemisphere which has more other independent observations. When removing all GNSS-RO data a degradation throughout the atmosphere can be seen, which is largest in the Southern Hemisphere and Tropics.

Vector wind

Fig. 9 shows a strong reduction of up to 20% at forecast day 2 compared to **noRO** for wind at 100 hPa for the various experiments. Especially, in the Southern Hemisphere the strongest reductions can be seen. Adding Spire (Fig. 9a) improves, in particular, forecast scores over the southern oceans whereas the addition of COSMIC-2 also reduces forecast errors over Africa (Fig. 9b). Forecast scores at most lower levels do not show a significant improvement, except for **COSMIC2** and **Spire+COSMIC2** in the Tropics between forecast day 1 and forecast day 4 compared to **CONTROL** (not shown). Here, forecast scores improve by about 2% at 100 hPa and 500 hPa.



(a) T+24h

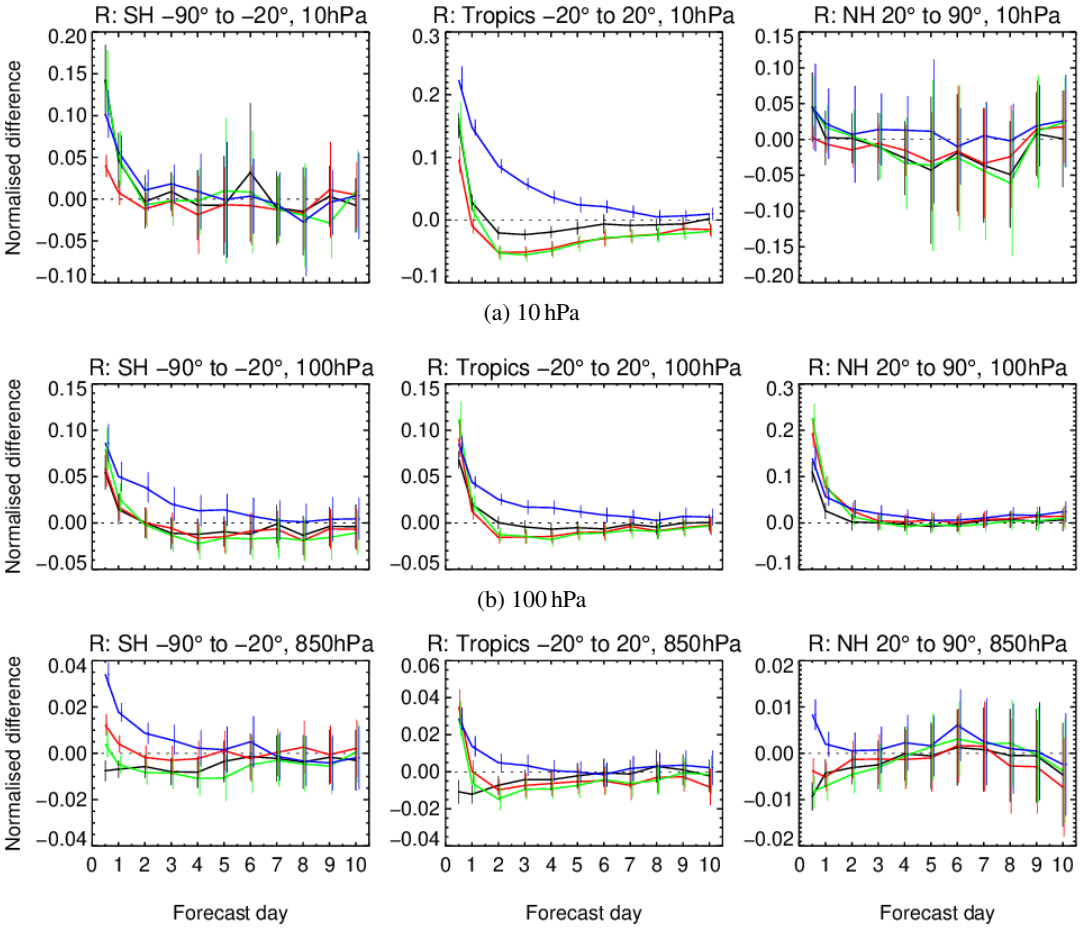
(b) T+48h

(c) T+96h

Figure 7: Normalised differences in standard deviation in temperature between **Spire+COSMIC2** and **noRO**, verified against ECMWF operational analysis at 100 hPa for different forecast times. Negative values represent a decrease in standard deviation and positive values an increase in standard deviation.

4.2 Verification of short-range forecasts against observations

To assess the short-range forecast (12-hour for ECMWF and 6-hour for Met Office) impact it is best to verify against observations (e.g. radiosondes). Here, we compute the standard deviation of the radiosonde minus short-range forecast values on a set of pressure levels (p) for **noRO**, $\sigma_{\mathrm{noRO}}(p)$, and the other experiments, $\sigma_{\mathrm{exp}}(p)$. If the quality of the short-range forecasts is improved through the addition



(a) 10 hPa

(b) 100 hPa

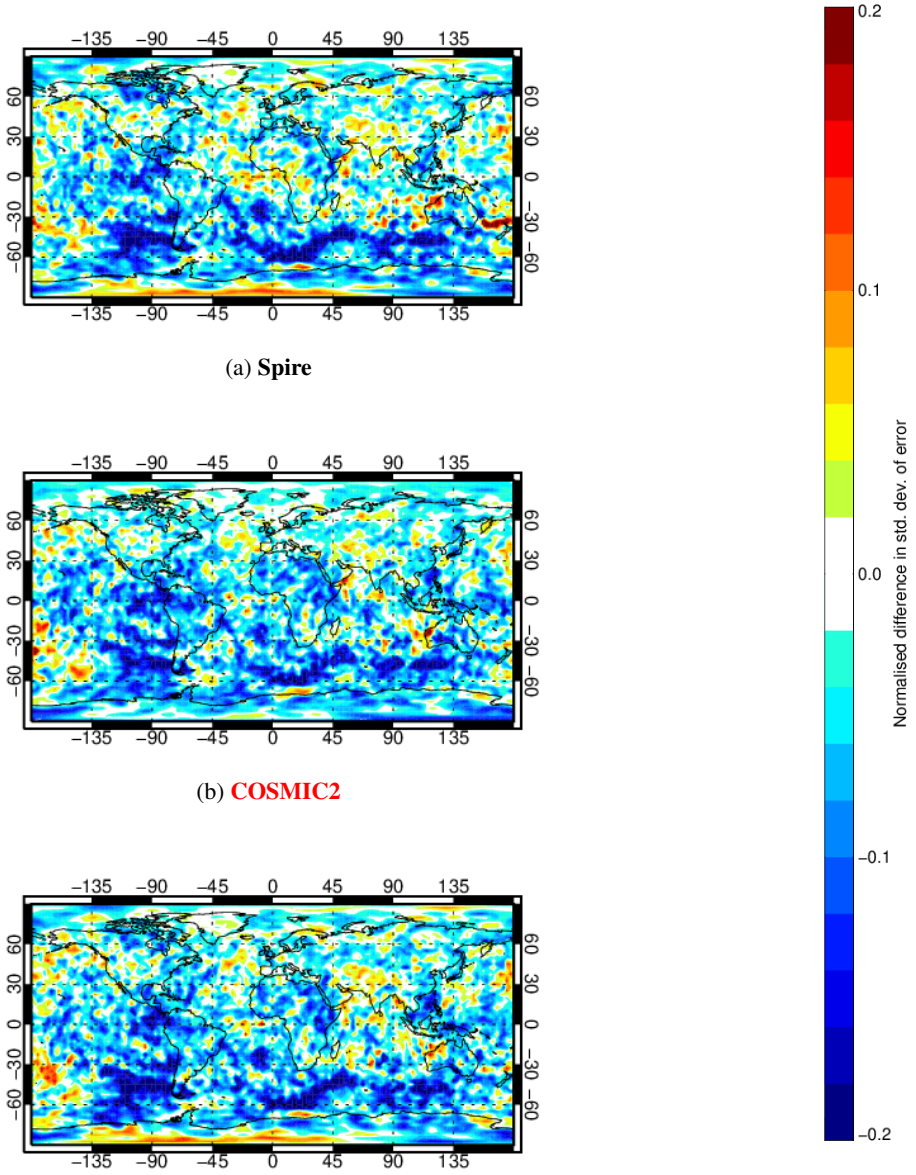
(c) 850 hPa

Figure 8: Similar to Fig.6 just for normalised differences in standard deviation in relative humidity between **Spire**, **COSMIC2**, **Spire+COSMIC2**, **noRO** and **CONTROL** for 10 hPa, 100 hPa and 500 hPa.

of the SPIRE/COSMIC-2 data, $(\sigma_{\mathrm{exp}}(p) - \sigma_{\mathrm{noRO}}(p))/\sigma_{\mathrm{noRO}}(p)$ would be negative.

Figure 10 shows $(\sigma_{\mathrm{exp}}(p) - \sigma_{\mathrm{noRO}}(p))/\sigma_{\mathrm{noRO}}(p)$ for temperatures measured against radiosondes. All the points are less than zero, indicating that short-range forecast errors are reduced by the assimilation of RO data. For the Met Office and ECMWF the standard deviations are reduced more when observations from Spire are used in addition to those from COSMIC-2 and other operational platforms. The reductions in the tropics are generally larger than the other regions, reflecting the large volume of data from COSMIC-2. Relative to this, the additional benefits from Spire observations are modest in this region. The reductions in standard deviation for the Met Office system are mostly larger than those for the ECMWF system. This could be due to a number of reasons, such as: a difference in the base skill of each system, the different weight given to RO observations, and the different resolution of the forecasting systems used in these tests. The differences between the forecasting systems can help to give confidence that the overall conclusions are likely to extend to other forecasting systems. It should be noted that the decreases in the standard deviation are very large, exceeding 10% at many points — such large changes are not seen when comparing forecasts against satellite observations.

For Fig 10 it was only possible to calculate confidence intervals for the ECMWF statistics. These are



(a) **Spire**

(b) **COSMIC2**

(c) **Spire+COSMIC2**

Figure 9: Normalised differences in standard deviation in vector wind between **Spire**, **COSMIC2** and **Spire+COSMIC2**, and noRO, verified against ECMWF operational analysis at 100 hPa for T+48h. Negative values represent a decrease in standard deviation and positive values an increase in standard deviation.

calculated using the bootstrap method (with 1000 resamples) on the statistics for each 12h assimilation cycle. For Figures 11 and 12 confidence intervals were also calculated for the Met Office statistics. These are calculated based on an estimate of the standard error of the variances used in the calculation.

Figure 11 shows $(\sigma_{\text{exp}}(p) - \sigma_{\text{noRO}}(p))/\sigma_{\text{noRO}}(p)$ for AMSU-A observations on the NOAA-19 satellite,

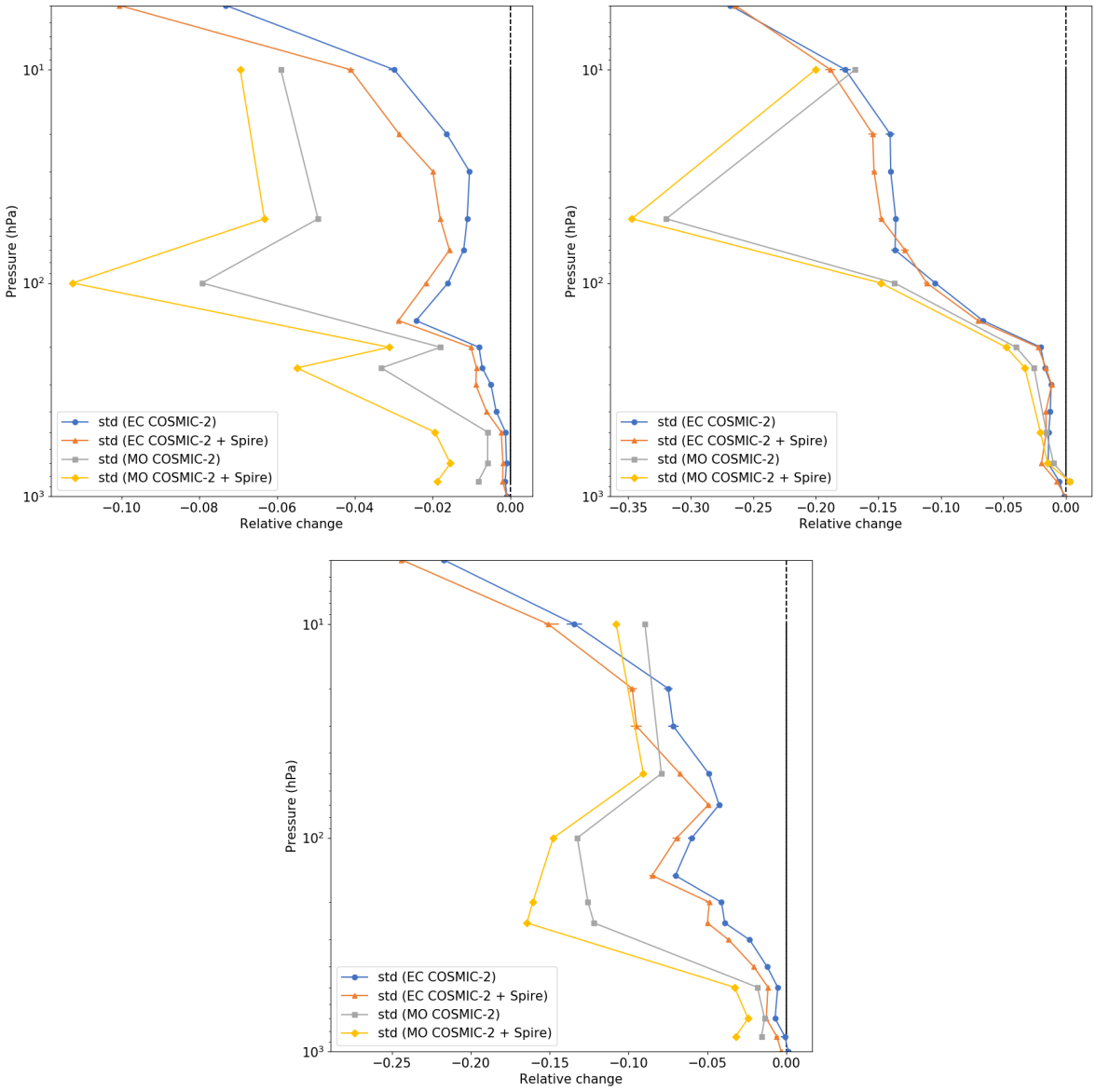


Figure 10: Normalised change in standard deviation of short-range forecasts, measured against radiosonde temperature observations. Results calculated for the northern-hemisphere extra-tropics (top left), the tropics (top right) and the southern-hemisphere extra-tropics (bottom middle). The standard deviations are normalised against the **noRO** systems. The results are shown for ECMWF **COSMIC-2** (blue circles) and **COSMIC-2 + Spire** (orange triangles), and Met Office **COSMIC-2** (grey squares) and **COSMIC-2 + Spire** (yellow diamonds). Confidence intervals are shown for the ECMWF experiments only, due to calculation limitations. They are often smaller than the size of the plotting symbol so cannot be seen.

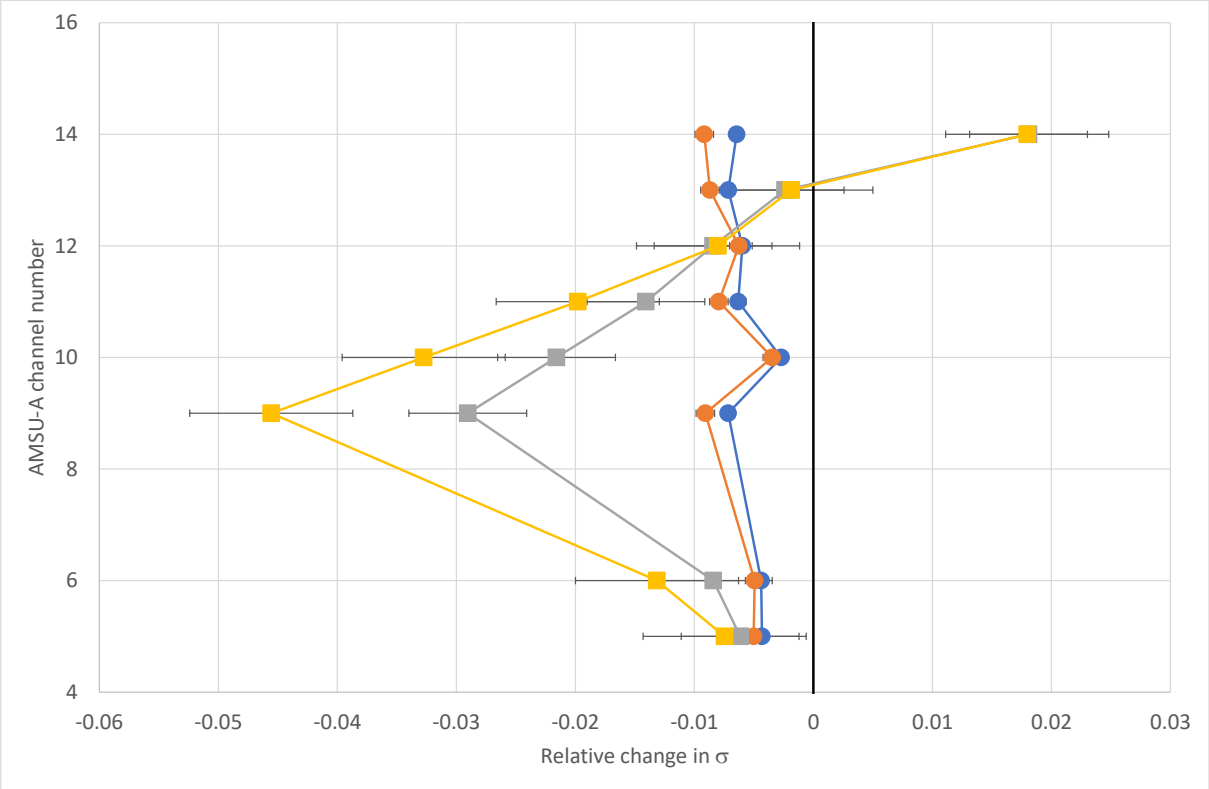


Figure 11: Normalised change in standard deviation of short-range forecasts, measured against AMSU-A observations on NOAA-19. The results are shown for ECMWF **COSMIC-2** (blue circles) and **COSMIC-2 + Spire** (orange triangles), and Met Office **COSMIC-2** (grey squares) and **COSMIC-2 + Spire** (yellow diamonds).

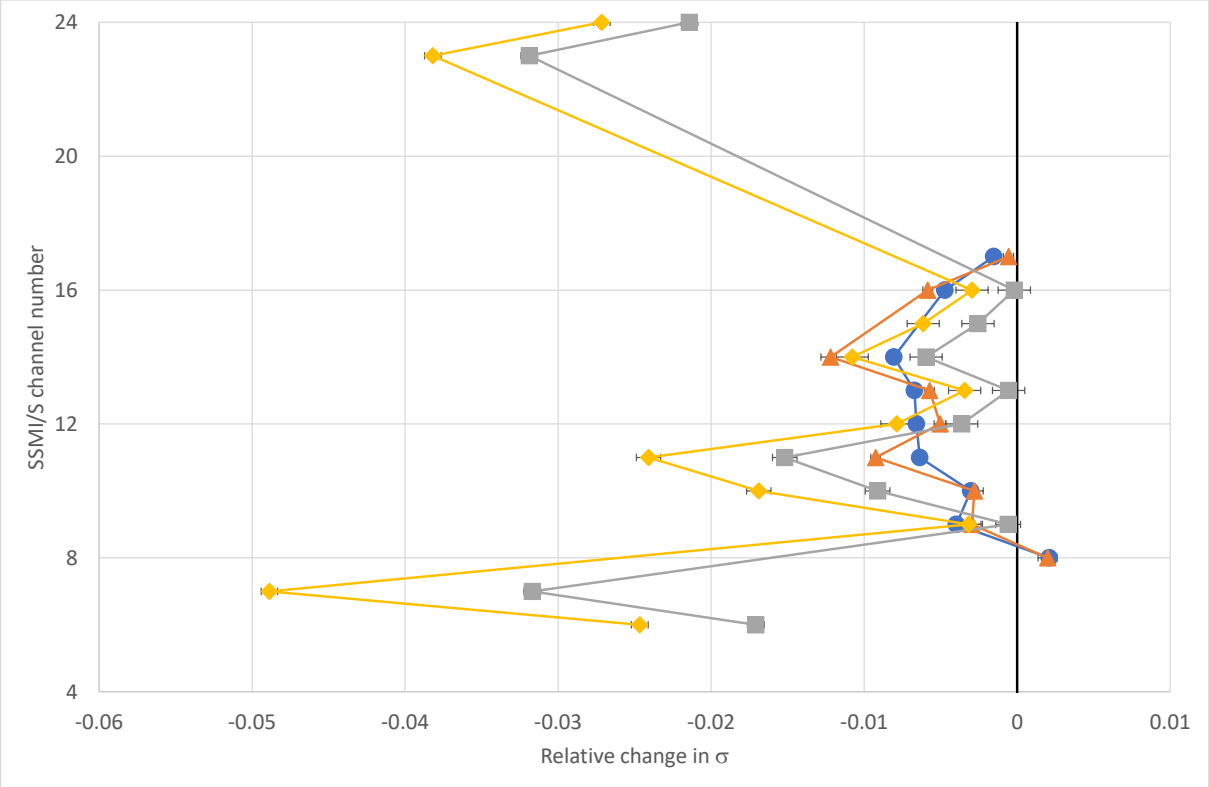


Figure 12: Normalised change in standard deviation of short-range forecasts, measured against SSMIS observations on DMSP (F-17). The results are shown for ECMWF **COSMIC2** (blue circles) and **Spire+COSMIC2** (orange triangles), and Met Office **COSMIC2** (grey squares) and **Spire+COSMIC2** (yellow diamonds).

arranged by the instrument channel for both Met Office and ECMWF. AMSU-A is a temperature sounder, where the different channels are sensitive to different ranges in height (i.e. have different weighting functions, see Fig. 49). At ECWMF AMSU-A is assimilated in clear-sky conditions in these experiments. At the Met Office channels 4 & 5 are also assimilated in cloudy conditions [Migliorini and Candy, 2019]. Apart from channels 4 (senses the surface) and 14 (senses the higher atmosphere around 2 hPa in a standard cloud free atmosphere) the addition of RO data reduces the standard deviation of the forecast errors. Compared to the radiosonde statistics the reductions in standard deviation are smaller for AMSU-A, at less than 1% for all channels in the ECMWF system. The additional benefit of adding Spire observations is also rather modest, except for certain channels with the Met Office system. For channel 9 (most sensitive to temperature around 70 hPa in cloud-free standard atmosphere) there are large benefits for adding RO data, and additionally adding Spire observations (3% and 4%, respectively). This is consistent with the regions where large impacts were seen for radiosonde observations. The Met Office system sees increased standard deviations in channel 14 for the addition of RO data. This is related to biases in the Met Office's model, and the fact that this channel is not bias corrected in order to act as an anchor to the bias correction system. ECMWF, on the other hand, applies a constrained variational bias correction for this channel [Han and Bormann, 2016].

Figure 12 shows the change in standard deviation of the short-range forecasts, measured against SSMIS observations taken from the DMSP (F-17) satellite. SSMIS is a microwave imager and sounder sensitive to humidity, water and temperature (In Fig. 50 more information can be found for some SSMIS sounding

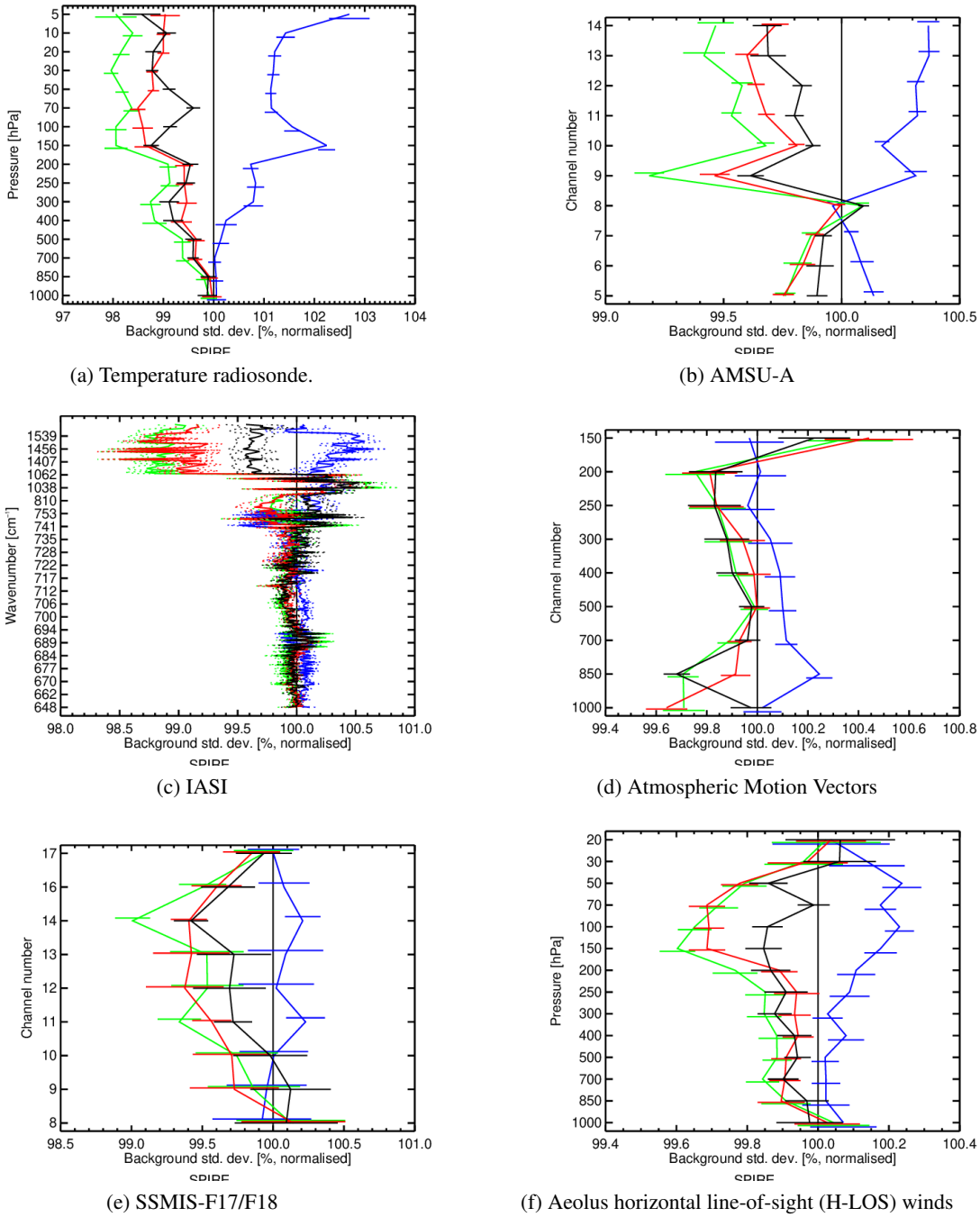
channels). At ECMWF no temperature channels (channels 23, 24) are assimilated operationally, but these are used at the Met Office. For ECMWF that means cloud, precipitation and humidity are sensed as SSMIS is assimilated in all sky conditions. Here, the benefits of assimilating RO data are mostly less than 1%. For channel 11 (sensitive to humidity around 350 hPa in a standard atmosphere), channel 14 (senses total column water vapour in clear skies) and channel 16 (sensitive to cloud liquid water and drizzle) the benefit of RO observations is slightly enhanced by the addition of Spire data, but slightly degraded for channels 12 and 13 (sensitive to larger hydrometeors like drizzle). At the Met Office SSMIS observations are only assimilated under clear-sky conditions for those experiments. Here, the benefits vary greatly for the different channels with some channels seeing greater than 3% benefit for the addition of RO observations. However, for all channels the standard deviations are smaller for the **Spire+COSMIC2** experiment than for the **COSMIC2** experiment with the Met Office system.

To illustrate in a bit more detail how the different fits to observations behave for **Spire**, **COSMIC2**, and **Spire+COSMIC2** compared to **CONTROL** for the ECMWF system Fig. 13 is shown. Here fits to observations are shown as $\sigma_{\text{exp}}(p)/\sigma_{\text{CONTROL}}(p)$, where values below 100% can be seen as an improvement and above 100% as a degradation. Also, statistically significance testing is done differently in these plots than e.g. Fig. 12. In Fig. 13 the error bars illustrate the 95% probability range of the t-distribution [Geer, 2016]. That means the confidence intervals may look different than before.

On a global scale, fits to radiosonde observations show similar improvements for **Spire** and **COSMIC2**, with **Spire+COSMIC2** having the largest impact (Fig. 13a). For the temperature sounder AMSU-A (Fig. 13b) **Spire+COSMIC2** also shows the largest reduction, especially for higher peaking channels (10 to 14) but also for lower peaking ones (channel 5). Here, **COSMIC2** shows significantly larger improvements than **Spire** with the reduction in **COSMIC2** and **Spire+COSMIC2** being up to 1.5% larger compared to **Spire**. This is in line with what could be seen in forecast scores for humidity in the Tropics with **COSMIC2** having a larger impact than **Spire** for ECMWF. Fits to other humidity sensitive observations, like SSMIS (Fig. 13e) show improvements of nearly 1% for channel 14 (measure of total column water vapour) for **Spire+COSMIC2** compared to **CONTROL**, as seen also in Fig. 12. The degradation for channel 8 (sensitive to lower atmosphere) is not significant and the differences among the experiments are also not significant. One last significant effect can be seen in Fig. 13d. Here, fits to atmospheric motion vectors is improved, especially at 850 hPa for **Spire** and **Spire+COSMIC2**. This is in line with improved forecast scores in wind. The degradation seen at 150 hPa should be interpret with caution because few AMV observations are assimilated here. Moreover, improved fits to Aeolus winds can be see in the higher troposphere/stratosphere (Fig. 13f) with the largest improvement occurring around 150 hPa.

4.3 Summary

The addition of COSMIC-2 and Spire observations is beneficial in both the Met Office and ECMWF systems, with adding Spire on top of COSMIC-2 showing further improvements for some variables and some geographical areas. The largest impact we see from the addition of COSMIC-2 and Spire data can be seen at higher altitudes, which is not surprising as the core region of radio occultation (RO) measurements is between 8 km and 30 km height. In general, improved medium-range forecast scores and improvement in short-range forecasts can be seen for temperature, humidity and wind but having different magnitudes at the Met Office and ECMWF system. For example, fits to radiosonde temperature observations in the Tropics show improvements of about 15% at 100 hPa when adding Spire and COSMIC-2 for Met Office and about 11% for ECMWF compared to **noRO**. The reasons why these differences occur are manifold and have been discussed in the previous sections. Nevertheless, they de-



(a) Temperature radiosonde.

(b) AMSU-A

(c) IASI

(d) Atmospheric Motion Vectors

(e) SSMIS-F17/F18

(f) Aeolus horizontal line-of-sight (H-LOS) winds

Figure 13: Normalised difference in standard deviation of first-guess departures between **Spire+COSMIC2**, **COSMIC2**, **Spire** (black) and **noRO** for different instruments, globally. The normalisation is done with results from **CONTROL**. Values less than 100% would indicate beneficial impacts from the experiments. The horizontal bars indicate 95% confidence range.

pend largely on how much skill the assimilation system already has even before assimilating additional radio occultation data, e.g for temperature. ECWMF and Met Office show the largest positive impact in temperature for medium-range forecast scores in the Southern Hemisphere through the addition of Spire

and COSMIC-2. For humidity, the biggest impact can be seen through the addition of COSMIC-2 in the Tropics at ECMWF. Furthermore, the positive impact for wind when adding Spire and COSMIC-2 could be seen in medium-range forecast scores for both Met Office (not shown) and ECMWF with the latter showing this improvement in the short-range, too.

In conclusion, both centres would assimilate the Spire data operationally if it becomes available .

5 Verification of additional Observing System Experiments performed at ECMWF

Additional to the ECMWF's Observing System Experiments (OSEs) analysed in section 4 we have performed supplementary ones. The first analysis of additional OSE (section 5.1) shows how much skill the addition of Spire data has by compensating the loss of aircraft data. This was motivated through the loss of aircraft data during the first part of the COVID-19 pandemic and the availability of Spire and COSMIC-2 data during this time period. The second analysis discusses if forecast scores are degraded if we have no Spire data available in the Tropics. As it has been shown in section 2, COSMIC-2 has a predominance between 40°S and 40°N, and hence the absence of Spire data in the Tropics might only has a weak impact on forecast scores. If this is the case, has been investigated in section 5.2.

5.1 Exclusion of all aircraft data

During the global COVID-19 pandemic in 2020/21 a large proportion of measurements taken on board commercial aircrafts were lost due to a significant reduction of the commercial air traffic. Ingleby et al. [2021] showed in a data denial experiment that this reduction degraded short-range forecast fits to radiosonde measurements of wind and temperature in the northern hemisphere (20°N-90°N). Having the additional Spire data, we have asked ourselves how much of the degradation due to the loss of aircraft data could be compensated through the addition of Spire data. For this reason, we have run two additional experiments. One removing all aircraft observations (temperature, wind and humidity) and another one with removing all aircraft data but adding Spire on top of **C**ONTROL.

In Fig. 15 the medium range forecast scores for temperature are displayed. As already discussed in section 4, changes in the normalised standard deviation of the forecast error at very short time scales might be only apparent and are, therefore, quantified later using observations. However, for forecast day 2 and longer the addition of Spire data on top of an experiment when no aircraft data is assimilated improves the forecast for temperature at various height levels. In the Northern Hemisphere, the temperature forecast scores are improved - even though a slight degradation compared to **C**ONTROL is still visible at 100 hPa and lower (in altitude). This might be not a surprise as most of the aircraft data is located in the Northern Hemisphere (see Fig. 14) and, hence, losing this data might be not as easily compensated by Spire data as in the Southern Hemisphere.

To evaluate how the impact of the loss of aircraft data is in the Northern Hemisphere and how well the addition of Spire data could compensate for this loss, it is best to look at fit to independent observations as done in Fig.16 in this region. For fits to temperature from radiosondes and AMSU-A (Fig.16a, Fig.16b) only in the stratosphere the loss of aircraft data can be compensated (even partially over-compensated) by the presence of Spire. In the lower stratosphere and troposphere, there is still a clear degradation even though it is reduced by up to 2.5% at 250 hPa for radiosonde temperature observations. For wind observations (AMVs in Fig. 16d and Aeolus in Fig. 16f) the compensation is even less strong, but still significant

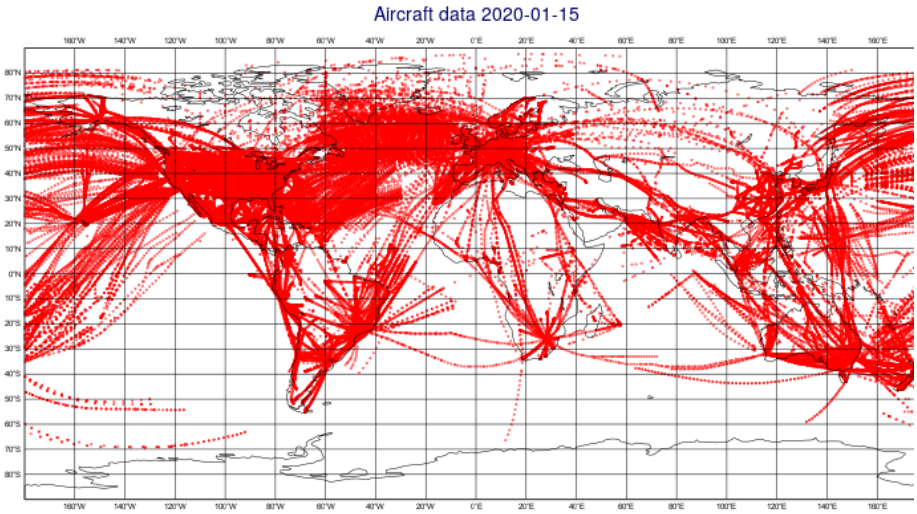
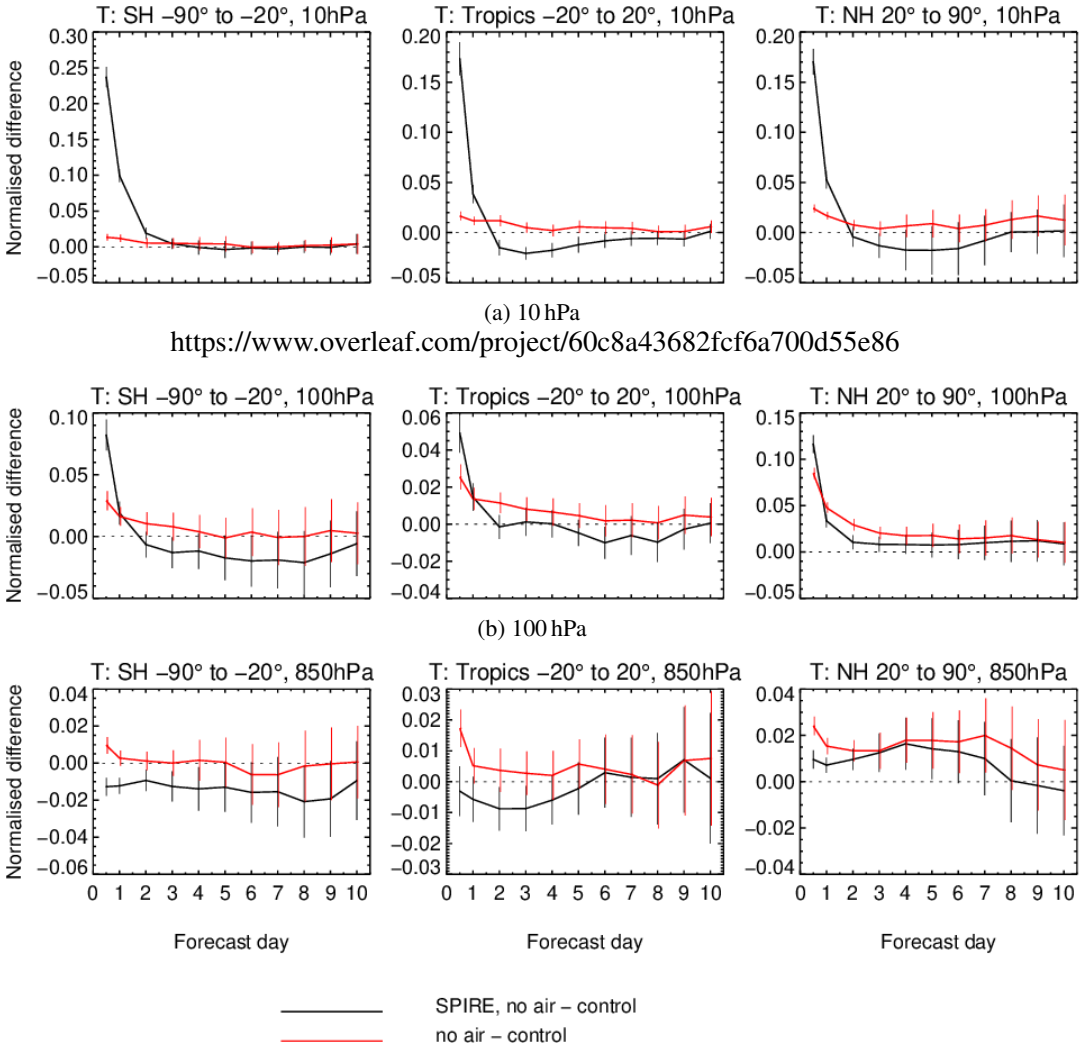


Figure 14: Location of aircraft wind reports for a 24 hour period (21Z/21Z) on 15 January 2020.

for 250 hPa to 50 hPa for Aeolus. Similar applies to the humidity channels of IASI (wavenumber 1038 and up) and microwave imager and sounder SSMIS (Fig.16c, Fig.16e, respectively).

To investigate further which of the aircraft observations (e.g. temperature, humidity, wind) can be compensated by additional GNSS-RO data, we have run additional experiments excluding only temperature aircraft observations but adding Spire and COSMIC-2. Fig. 17 shows selected fits to temperature, humidity and wind radiosonde observations in the Northern and Southern Hemisphere. For the Northern Hemisphere, the loss of all aircraft data cannot be compensated with the addition of Spire and COSMIC-2 data in the troposphere and lower stratosphere. However, in higher levels an over-compensation is clearly seen for all observations. If only temperature aircraft data would be lost the compensation effect is even stronger. In the Southern Hemisphere where less aircraft data is present, the addition of GNSS-RO at least compensates for the loss of aircraft data but often also improves fits to radiosonde observations. Only for wind in the troposphere (up to 200 hPa) the loss of all aircraft data could not be entirely compensated through the addition of Spire and COSMIC-2 data. This is in line with findings by Ingleby et al. [2021] stating that wind makes up the majority of the impact of aircraft data.

In summary, it is clear that Spire and COSMIC-2 data compensate some of the degraded forecast skill when aircraft data is lost in almost all regions and for most background fits. However, Spire and COSMIC-2 struggle to recover the very large degradation seen for the Northern Hemisphere in the troposphere and lower stratosphere. For the Southern Hemisphere and Tropics there are several cases where Spire and COSMIC-2 are better than **CONTROL** and **no air**, highlighting the value of GNSS-RO data and also the smaller amount of aircraft data in these regions.



(a) 10 hPa

https://www.overleaf.com/project/60c8a43682fcf6a700d55e86

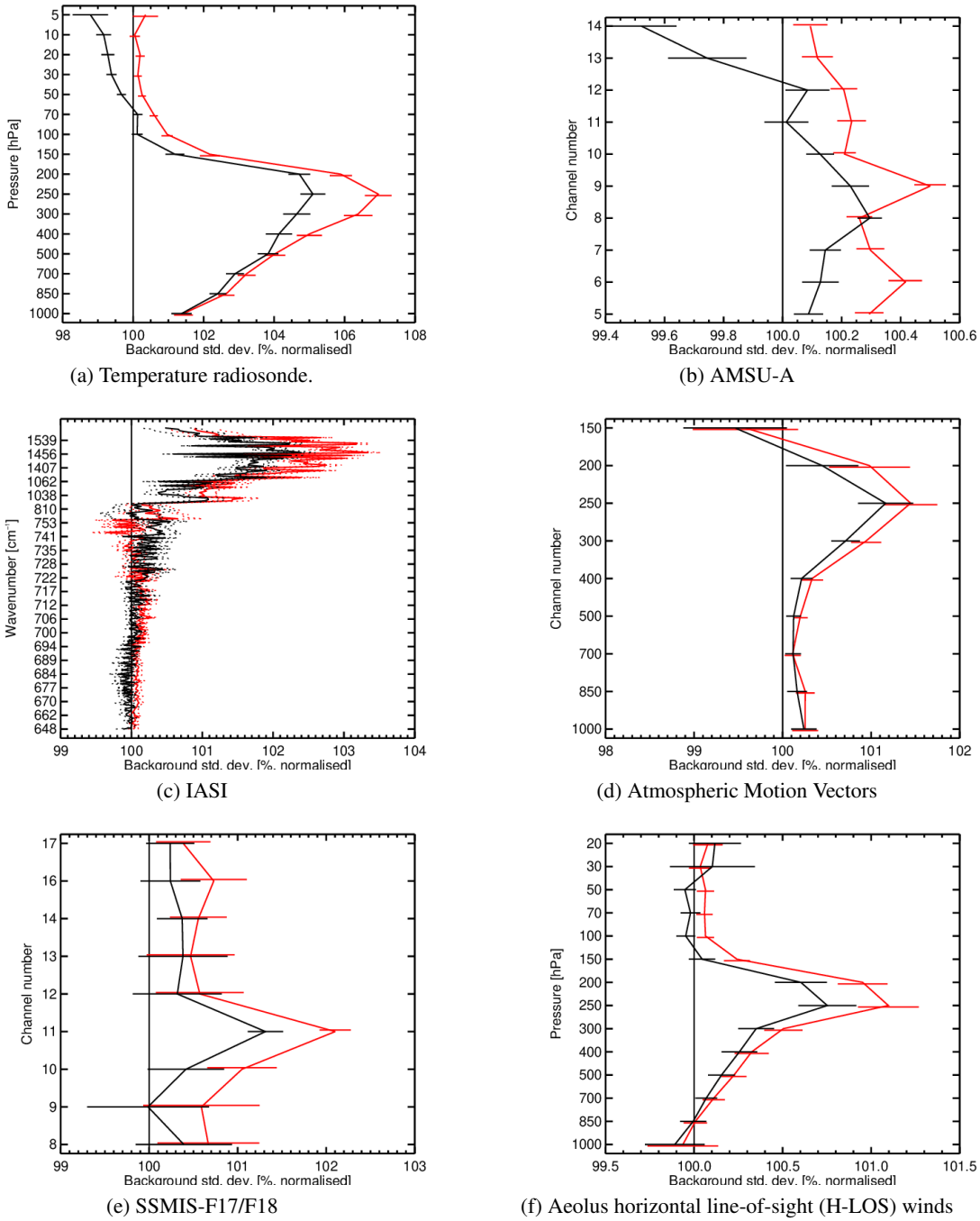
(b) 100 hPa

(c) 850 hPa

Figure 15: Normalised differences in standard deviation in temperature between **Spire no air** and **no air**, and CONTROL, verified against ECMWF operational analysis at different height levels (850 hPa, 100 hPa, and 10 hPa) for different forecast times and zonal regions (left: SH 90°S-20°S, middle: Tropics 20°S-20°N, right: NH 20°N-90°N). Negative values represent a decrease in standard deviation and positive values an increase in standard deviation. The confidence range is displayed by vertical bars.

5.2 Spire data excluded in Tropics

As shown before, the impact of the additional Spire data on improving the short-range and mid-range forecast scores is apparent but does vary with latitude. For example, in the Tropics the addition of COSMIC-2 shows a larger impact than Spire. By excluding Spire data in the Tropics we like to test in how far the forecast quality is reduced. For this reason we have run two experiments assimilating Spire and COSMIC-2 data on top of **C**ONTROL, however excluding Spire data between 30°S and 30°N, or 20°S and 20°N. Fig. 18 shows the medium-range forecast scores for temperature when using the



(a) Temperature radiosonde.

(b) AMSU-A

(c) IASI

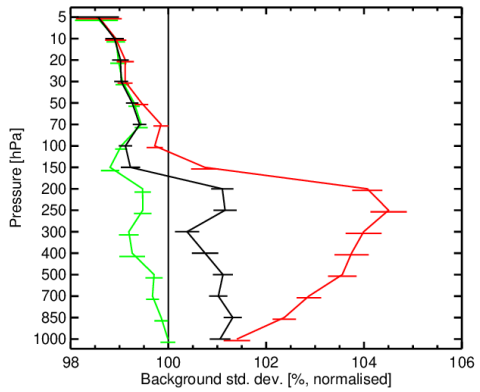
(d) Atmospheric Motion Vectors

(e) SSMIS-F17/F18

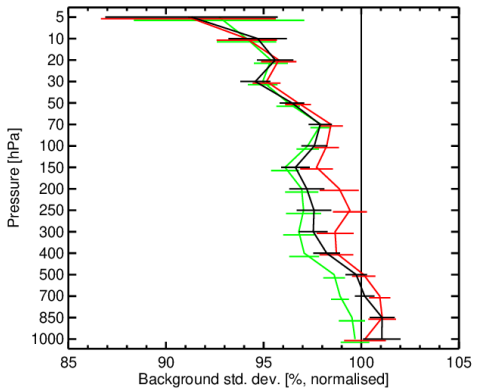
(f) Aeolus horizontal line-of-sight (H-LOS) winds

Figure 16: Normalised difference in standard deviation of first-guess departures between **Spire no air** and **no air**, for different instruments in the Northern Hemisphere. The normalisation is done with results from CONTROL. Values less than 100% would indicate beneficial impacts from the experiments. The horizontal bars indicate 95% confidence range.

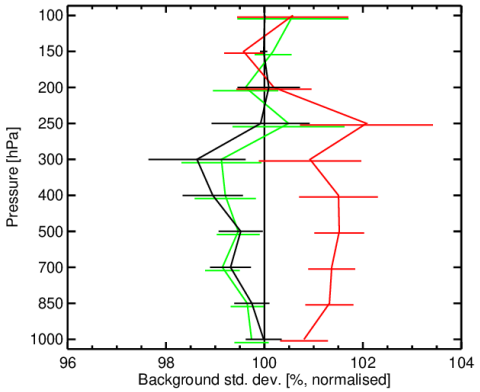
operational ECMWF analysis as a reference. The differences between the experiments are visible for forecast days three and shorter and at lower height levels (e.g. 850 hPa). However, the differences are not significant except for day 1 at 850 hPa in the Southern Hemisphere where the exclusion of Spire data



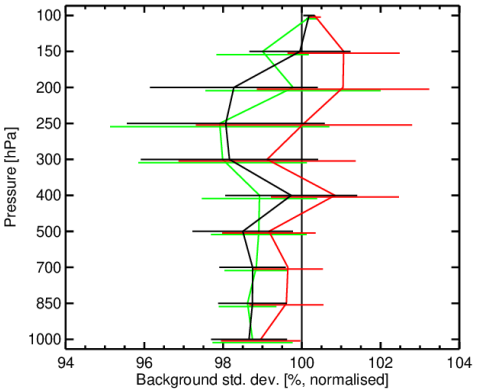
(a) Temperature radiosonde in Northern hemisphere.



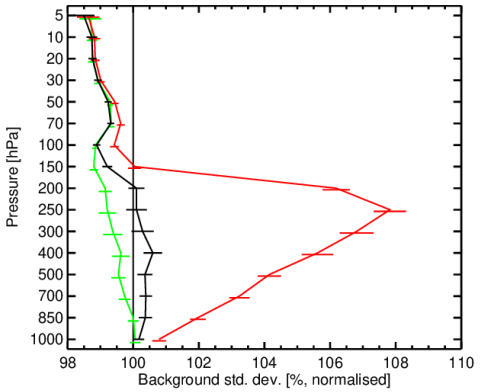
(b) Temperature radiosonde in Southern hemisphere.



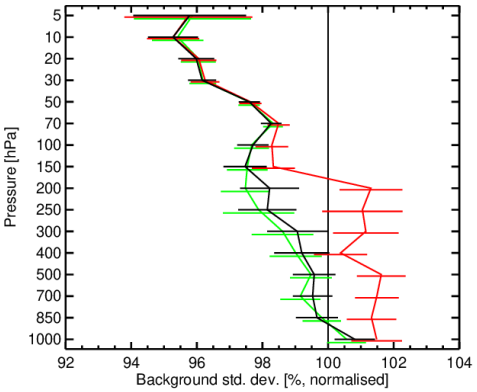
(c) Humidity radiosonde in Northern hemisphere.



(d) Humidity radiosonde in Southern hemisphere.



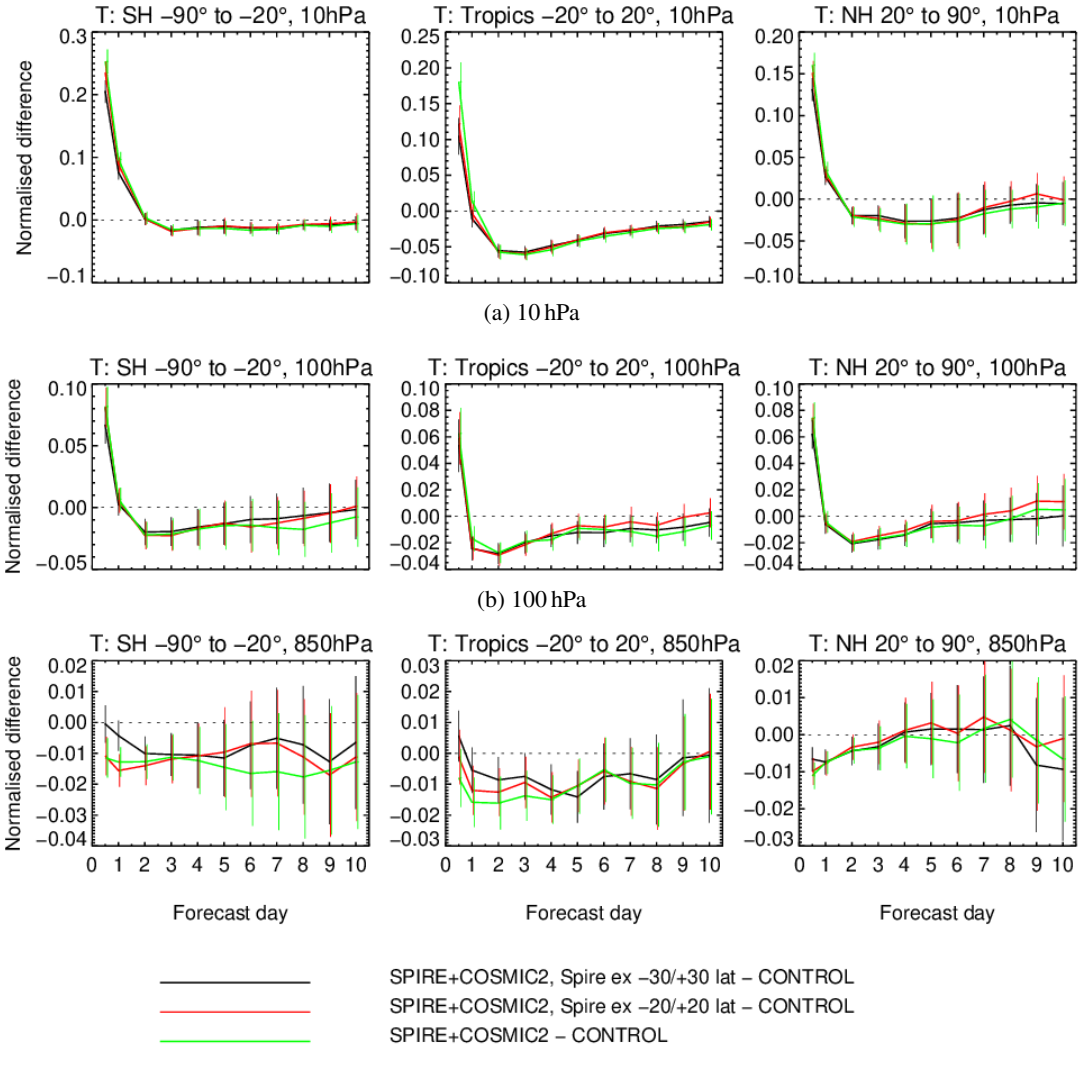
(e) Wind from Radiosonde in Northern Hemisphere



(f) Wind from Radiosonde in Southern Hemisphere

Figure 17: Normalised difference in standard deviation of first-guess departures between **Spire+COSMIC2 no T air**, **Spire+COSMIC2 no air** and **Spire+COSMIC2**, for different instruments in the Northern Hemisphere (a,c,e) and Southern Hemisphere (b,d,f). The normalisation is done with results from CONTROL. Values less than 100% would indicate beneficial impacts from the experiments. The horizontal bars indicate 95% confidence range.

between 30°S and 30°N shows less improvement than the other experiments compared to CONTROL.



(a) 10 hPa

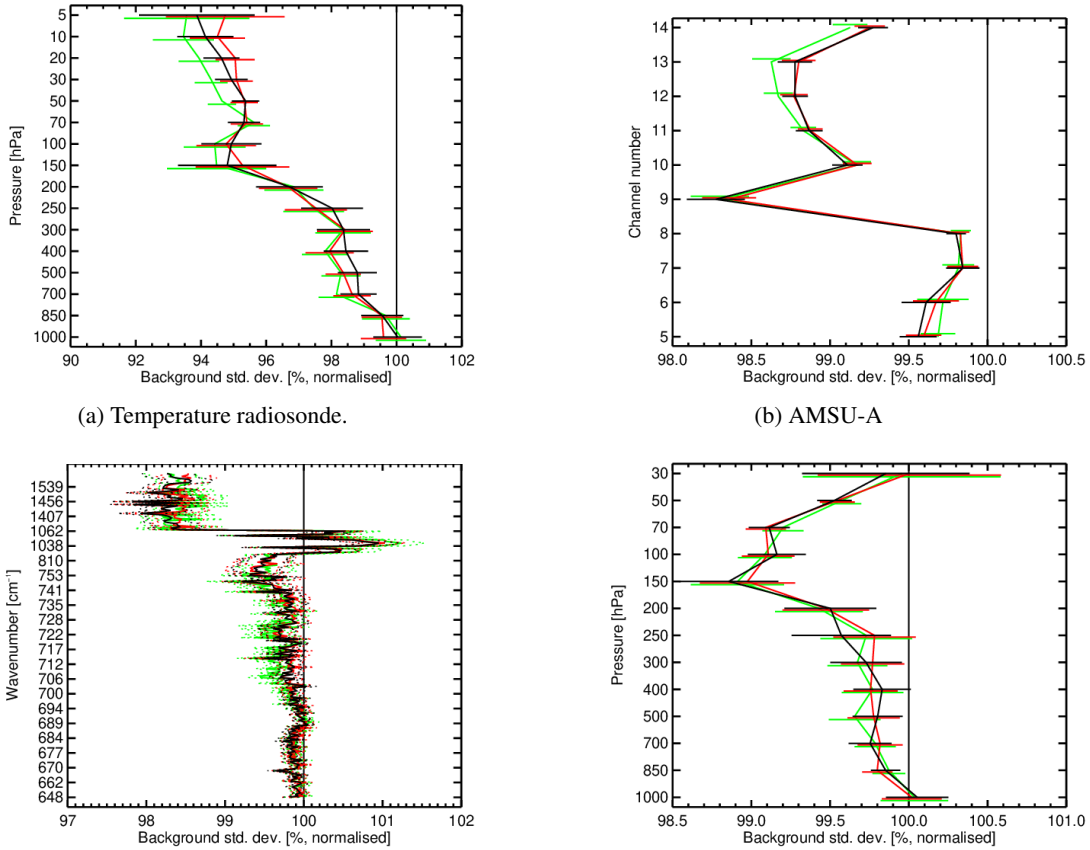
(b) 100 hPa

(c) 850 hPa

Figure 18: Normalised differences in standard deviation in temperature between **Spire+COSMIC-2, Spire excluded between 30°S and 30°N**, **Spire+COSMIC-2, Spire excluded between 20°S and 20°N** and **Spire+COSMIC-2**, and CONTROL, verified against ECMWF operational analysis at different height levels (850 hPa, 100 hPa, and 10 hPa) for different forecast times and zonal regions (left: SH 90°S-20°S, middle: Tropics 20°S-20°N, right: NH 20°N-90°N). Negative values represent a decrease in standard deviation and positive values an increase in standard deviation. The confidence range is displayed by vertical bars.

Nevertheless, when interested in shorter forecast timescales it makes sense to look into fits to independent observations, as done in Fig. 19 in the Tropics. For the higher peaking AMUS-A temperature channels 12 to 14 (Fig. 19b) the improvement for the experiments excluding Spire in the Tropics are less pronounced as when all Spire would be assimilated. A similar signal can also be seen for radiosonde temperature measurements (Fig. 19a) in about 50 hPa to 30 hPa. This slight degradation is not significant, but this is likely linked to the experimentation period of "only" three months [Geer, 2016]. Fits to wind observations (Fig. 19d) or IASI (Fig. 19c) do not show differences for experiments with Spire being assimilated

in the Tropics or not.



(a) Temperature radiosonde.

(b) AMSU-A

(c) IASI

(d) Aeolus horizontal line-of-sight (H-LOS) winds

Figure 19: Normalised difference in standard deviation of first-guess departures between **Spire+COSMIC-2, Spire excluded between 30°S and 30°N**, **Spire+COSMIC-2, Spire excluded between 20°S and 20°N** and **Spire+COSMIC-2** for different instruments in the Tropics. The normalisation is done with results from CONTROL. Values less than 100% would indicate beneficial impacts from the experiments. The horizontal bars indicate 95% confidence range.

In summary, the exclusion of Spire data in the Tropics degrades slightly fits to temperature measurements in the tropical stratosphere. However, fits to humidity sensitive observations (not shown) or wind are not affected. Also in the medium range the degradation is mostly insignificant, which might be caused by the mentioned effect of running those experiments for "only" three months. However, we also argue that the strong prevalence of COSMIC-2 data in the Tropics might compensate (at least partially) for the loss of Spire data there.

6 Assimilation Impact of Spire data processed by EUMETSAT

In addition to the bending angle profiles Spire delivered to ECMWF and the MetOffice directly, the company also provided EUMETSAT with lower level data from the same period. This data included raw observations for both occultation and zenith antenna measurements as well as excess carrier phase

data. The latter is referred to level 1a data and represents the result of combining raw measurements with precise estimates of the satellite's position. Excess carrier phase data is basis for deriving bending angle profiles, usually through applying one of several "wave optics" retrieval algorithms. However, specific algorithms and details of their implementation usually differ between data providers. EUMETSAT's activities within this study aimed at obtaining a better understanding of the lower level data characteristics of Spire data and at making them directly comparable to those of existing RO data from public missions such as the GRAS instruments flown on Europe's Metop satellites. For this, Spire level 1a data was processed using EUMETSAT's RO reference processor applying the same algorithms, filtering options, and quality control procedures as for GRAS and other RO missions such as Sentinel-6.

This section of the Final Technical Note describes the processing implemented at EUMETSAT, discusses some low level data characteristics in comparison with GRAS data, and compares statistics against ECMWF short range forecast data with a particular focus on tropospheric statistics in section 6.1. The bending angle data set produced by EUMETSAT was assimilated by ECWMF, with results being discussed in section 6.2.

Because the selection of profiles by the data provider occurred at different points of time and covered different product levels, the occultation data sets provided to the two NWP centres on one hand and to EUMETSAT on the other are not completely identical. However, the difference in daily numbers of occultations are small, with their day-to-day variations being very close to each other. We therefore believe that the differences between the two data sets do not have a significant impact on the results described in this section.

6.1 Processing of Spire data at EUMETSAT

6.1.1 Wave optics processing

The EUMETSAT wave optics processing of Spire data is based on the excess phase data provided by Spire and generates bending angle profiles based on a fast version of the Phase Transform Jensen et al. [2003]. The implementation of this algorithm at EUMETSAT is a modified version of the fast algorithm described in Gorbunov and Lauritsen [2004] which is also implemented in the ROM SAF's ROPP software package [Culverwell et al., 2015]; in particular, the approach to signal filtering and smoothing, the estimation of bending angles from the transformed measured signals and the cutoff of retrieved bending angle profiles at the bottom differ in the EUMETSAT implementation. In contrast to other wave optics implementations, there is no transition between a wave optics retrieval in the troposphere and a more traditional geometrical optics retrieval in the stratosphere. Instead, the fast Phase Transform (FPT) is applied throughout the entire vertical coverage of bending angle profiles, avoiding inconsistent error characteristics between the wave and geometrical optics part of the profiles.

The algorithms applied to the Spire data are identical to those being implemented in a future (v5.0) version of the operational GRAS processing at EUMETSAT and the current Non-Time Critical processing of Sentinel6/Michael-Freilich RO data. They were also applied in a recent (v2) reprocessing of CHAMP, GRACE and COSMIC data carried out by EUMETSAT.

Bending angle profiles produced by EUMETSAT are available in a high resolution on a vertical grid with a vertical spacing of 25 m. For the application of this data in NWP, these profiles are smoothed and downsampled to a vertical grid of 247 levels between the ground and 60 km (impact) altitude, where the vertical spacing varies from slightly more than 100 m in the lower troposphere to about 300 m in the mid- and upper stratosphere and lower mesosphere. Similar to the level spacing, the amount of smoothing is

altitude dependent and constructed to be consistent with the smoothing of carrier phase data at stratospheric heights in geometrical optics retrieval schemes. As the amount of smoothing increases with altitude, the reduced vertical level spacing reflects the reduced information content in the bending angle data higher up where the increased smoothing reduces the vertical resolution. The vertical level spacing is, by-and-large, consistent with the smoothing applied. This differs from the original bending angle data produced by Spire; while bending angles are also smoothed with height-dependent filter widths (Masters 2021, pers. comm.), the original bending angle data is provided on a regular vertical 200m grid.

6.1.2 Occultation numbers, data distribution and duty cycles

The data set provided by Spire consists of 635.595 occultations from 19 different satellites; of those, 635.134 occultations were successfully processed towards bending angle profiles by EUMETSAT. In the remaining 461 cases, the EUMETSAT processing failed due to several processor shortcomings. While these issues will be fixed in a future version of the EUMETSAT processor, we believe that the small number of occultations dropped during the processing has no impact on the conclusions of the study.

Some of the older satellites (in particular FM046, FM075, FM080 and FM088) together provided fewer than 15.000 occultations, or less than 2.5%, of the overall dataset; the majority of data thus stems from the remaining 15 satellites. Table 3 in Appendix D lists the amount of data provided by all satellites, together with statistics on the number of rising and setting occultations. In figures showing per-satellite results in the following sections, we do not present individual results for these four satellites to limit the number of subplots. Unless noted in the accompanying text, the characteristics of these four satellites also do not differ significantly from this those being observed for the remaining Spire satellites.

With the exception of two satellites (FM046 and FM113), all Spire spacecrafts provided both rising and setting occultations in similar numbers, although the percentage of rising ones is typically slightly smaller (47%) than that of setting occultations (53%), with small variations between satellites. With respect to the various GNSS constellations, the data set contains measurements from the GPS, GLONASS and Galileo systems, with GPS providing the majority of the data (42%). Galileo and GLONASS both provide both a smaller number of occultations (27% and 26%, respectively). The number of occultations from QZSS (5%) is significantly smaller due to lower number of QZSS satellites. The percentage of occultations from the three major GNSS constellation varies somewhat between the individual Spire satellites with maximum deviations from the averages cited above being below 10%.

Fig. 20 presents the number of daily Spire occultations during the study period stratified by GNSS system; the figure suggests that the relative numbers of Spire occultations with respect to the various constellations remain stable over time. In contrast to other figures in this report, Fig. 20 shows the number of occultations within 24 hours rather than a 12 hour data assimilation window.

As the Spire data stems mostly from polar orbiting satellites, the meridional distribution of occultations should look similar to that of any polar satellite performing RO measurements. This is indeed the case: Fig. 21 shows the distribution of occultation numbers over 5 degree latitude bins for the main Spire satellites. The lower right subplot contains the latitudinal distribution of GRAS occultations from the three Metop satellites. Some Spire satellites exhibit meridionally asymmetric distribution patterns: For example, FM079 exhibits significantly fewer occultations in the Southern Hemisphere Tropics, while FM080 provided fewer occultations in the Southern Hemisphere mid-latitudes. On the other hand, newer satellites from FM103 onwards provide slightly more occultations on the Southern Hemisphere in general. Also note that the satellites FM084, FM085 and FM088 (not shown) are flying on a low inclination orbit, causing their occultation coverage being limited between (about) 60°S and 60°N.

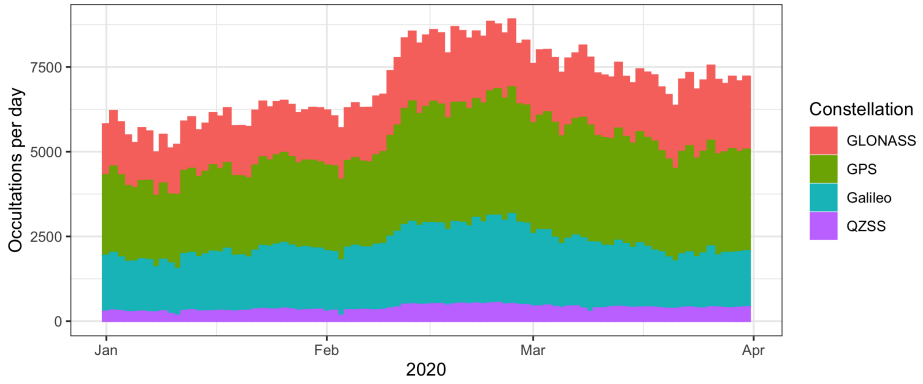


Figure 20: Number of daily Spire occultations; colours indicate the GNSS constellation from which occultations were obtained.

The combined meridional distribution of all Spire satellites during the study period is remarkably similar to the one from the GRAS instruments on the Metop satellites: Fig. 22 shows the density distribution of the two data sets on top of each other. The only systematic difference appears to be that there is slightly more Metop data in outer (20 - 30 degrees) Tropics of both hemispheres.

In a previous study on the assessing Spire data quality [Marquardt et al., 2020], we noted that the vast majority of Spire occultations being available at that time were obtained from sun-synchronous polar orbits with local observation times clustering around 9:30 and 21:30. With respect to local time, these observations were thus distributed very similar to the ones from EUMETSAT's Metop instruments. Therefore, one of the conclusions of the previous study was that the local time distribution of the Spire data should become more complementary to existing polar orbit RO missions.

Newer Spire satellites do improve on this. Fig. 23 shows the local time distribution of occultations from the main Spire satellites. From FM103 onward, local times are centred around 3:00 and 15:00 local time, and are thus in between the local times covered by the Metop instruments. For the older satellites, the local time sampling is partially similar to the Metop data (for FM079, and FM0900 - FM102) while FM084, FM084 and FM088 are flying on drifting low-inclination orbits, thus covering different local times during the study period. Overall, therefore, the newer satellites improve the distribution of Spire data with respect to the data already available from other operational RO data in polar orbit.

In the previous analysis of Spire data, we found that observation periods of Spire satellites are typically 50 minutes long, with a break of approximately the same length in between. Fig. 24 shows a timeline for Spire and Metop satellites, along with a zoom-in into three days in January 2021. In this plot, each occultation is represented by a single dot; periods of continuous operations show up a continuous coloured line segments. Older satellites show typical duty cycles in the order of one hour or slightly less, with similar periods without occultation measurements in between. Newer Spire satellites, especially from FM099 onward, exhibit fewer and longer active periods. For the satellites FM105, FM106 and FM108 in particular, interruptions have become even shorter. Longer duty cycles, in principle, allow for longer orbit arcs and hence improved upper level bending angle data quality, at least when individual observation periods are processed individually (which is the case in the Spire processing; Masters 2021, pers. comm.) An even better way would be to process data from combined consecutive arcs as proposed in the previous Spire Data Quality Assessment Study.

For reference, Fig. 24 also shows the timeline for the Metop satellites which provide a 100% duty cycle.

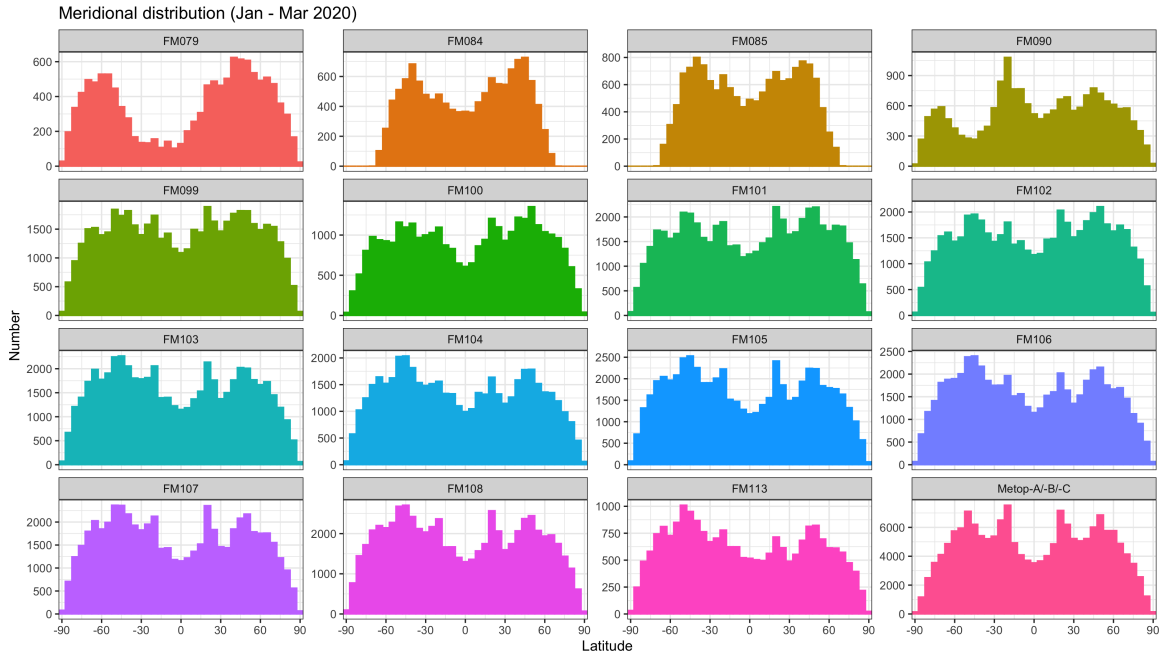


Figure 21: Meridional distribution of occultations from the main Spire satellites during January to March 2020 in 5 degree latitude bins. The lower right subfigure shows the meridional distribution of the combined Metop-A, -B and -C occultations for comparison. Note that the vertical axis is different for each satellite, reflecting the varying numbers of contributions.

6.1.3 RO error characteristics in general

Before discussing the error characteristics of Spire RO data, we'd like to remind the reader about general characteristics of RO uncertainties, in particular their dependency on height. The following is based on Kursinski et al. [1997b], taking into account more recent results from Sokolovskiy et al. [2010].

In the upper stratosphere and mesosphere (above about 25 or 40 km), the main contributions to random neutral atmospheric bending angle uncertainty are (i) the amount of instrument thermal noise entering through the carrier phase measurements, and (ii) ionospheric residual noise that could not be removed in the processing. The thermal noise of a receiver (usually characterised through its Signal-to-Noise Ratio, or SNR), if not further filtered, is white; even after some kind of smoothing applied as part of the processing, it will cause high frequency (short vertical wavelength) random fluctuations in retrieved bending angle profiles, dominating their random uncertainty. In contrast, ionospheric residual noise will typically exhibit longer vertical scales, in effect separating the impact of the two noise sources into different vertical scales.

Both error sources cause a more or less constant noise floor for bending angle data in the upper atmosphere. The size of the latter is being determined by the noise characteristics of the receiver, with occasional additional contributions from residual ionospheric activity. Because bending angle profiles fall off exponentially with height, the relative errors of bending angles increase exponentially with altitude above 35 of 40 km.

In the mid- and lower stratosphere as well as the upper troposphere (typically between 8 and 35 or 40 km), measurement accuracy is dominated by the presence of horizontal gradients in the atmosphere; the receiver's noise performance becomes less and less important as a contributor to the error budget

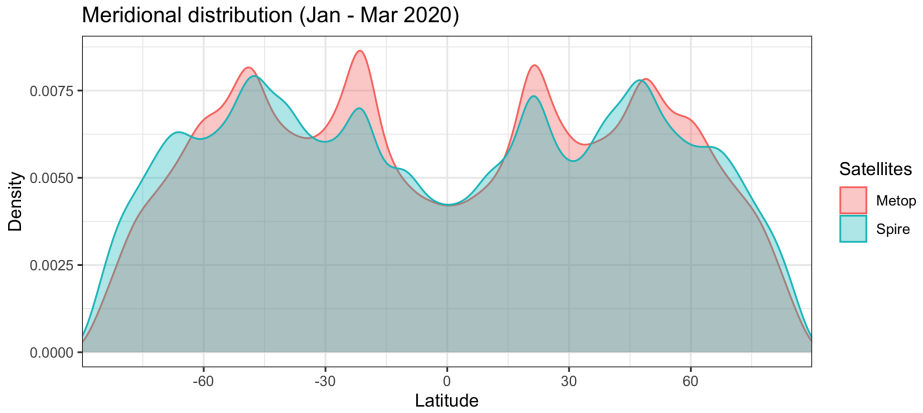


Figure 22: Meridional distribution of the occultation number density for the combined Spire and Metop satellites during January to March 2020. The density distribution was calculated using a (one-sided) bandwidth of 2.5 degree using a Gaussian kernel.

of RO observations. In this altitude region, SNR is reduced due to the defocusing of the blended rays, but receivers are still capable of tracking the GNSS signals successfully. Thus, bending angles from different RO receivers usually provide similar error characteristics, with any differences between them usually been due to the choice of smoothing bandwidths applied in the respective processing. In this altitude range, relative bending angle errors are approximately constant with height, and receiver noise only contributes weakly to the random uncertainty of bending angles.

In the lower and mid-troposphere, SNR is further reduced due to signal defocusing, but the complicated signal propagation through the atmosphere caused by the presence of water vapour adds rapid amplitude and phase variations to the measured data. In this altitude range, two additional effects become important contributors to the bending angle error: First, the ability of the receivers to track the multitude of signals present at the same time through an "open loop" approach in the presence of very low SNRs; and second, the ability of the bending angle retrieval algorithms to disentangle these signals into a single-valued bending angle profile. Note that such wave optics algorithms still assume spherical symmetry of the atmosphere, i.e. the absence of horizontal gradients. Therefore, deviations from this conditions continue to be major contributor to the bending angle error budget.

Due to increased levels of water vapour in the mid and lower troposphere, the complexity in the signal propagation increases from below the tropopause towards the ground, and so does the uncertainty of bending angle retrievals.

While carrier phase noise (and thus, SNR) does not directly contribute to the error budget of bending angle retrievals in the lower troposphere, it may play a subtle role in affecting lower tropospheric biases through the processing. Sokolovskiy et al. [2010] demonstrated that biases at the bottom of bending angle profiles depend on the cut-off of carrier phase data in the Earth's shadow. If signals are not tracked sufficiently low into the Earth's shadow, bending angles may be biased negatively; if too much much noise is let into the retrieval, positive biases can be produced. For this reason, wave optics retrievals perform some kind of signal cut-off before calculating a bending angle profile. In the EUMETSAT processing as well as in other systems (e.g., the one by Sokolovskiy et al.), the lower tail is analysed to estimate the noise level of the signal-free part of the measurements. Data is then only used if SNRs are above this noise level. If a receiver produces a higher signal-free noise level, signal cut-offs may occur systematically higher, causing larger negative biases. Conversely, if thresholds are lowered to

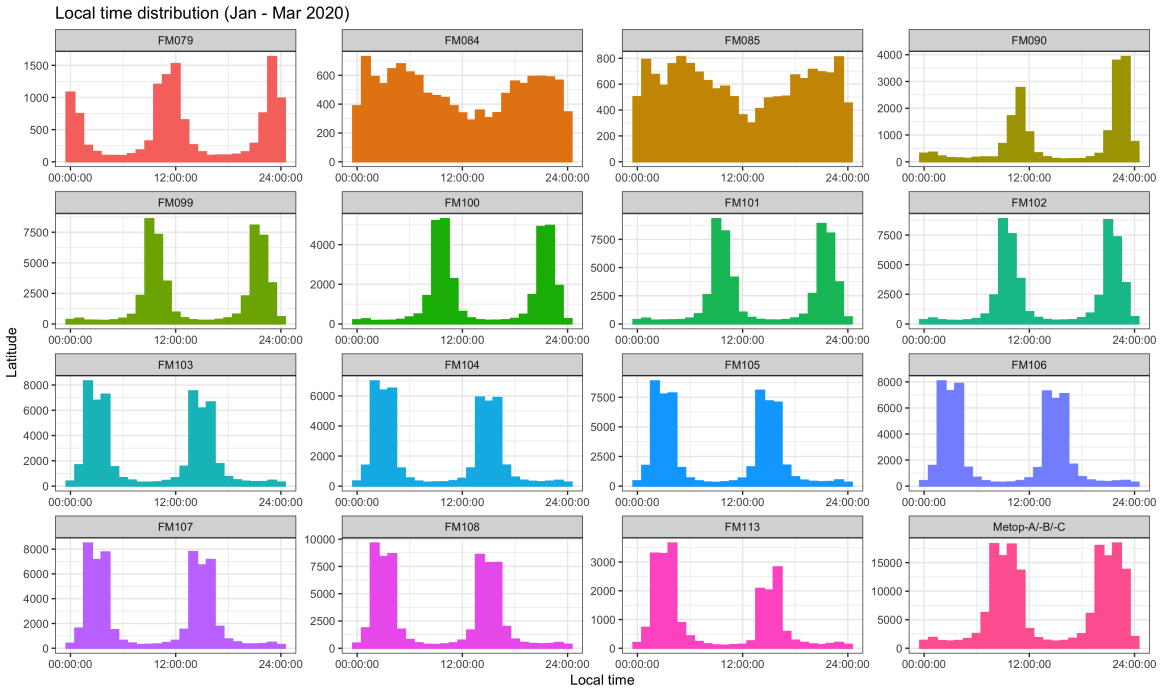


Figure 23: Local time distribution of occultations from the core Spire and the combined Metop satellites in hourly bins from January to March 2020.

accommodate higher signal-free noise levels, the additional noise in the data processed might produce positively biased bending angle data in the lower troposphere. The threshold for the signal cur-off is obviously a tuneable parameter, suggesting that bending angle biases immediately above the Earth's surface are mostly due to processing choices independent from the actual measurement.

6.1.4 Signal-to-Noise Ratios, carrier phase and bending angle noise

As the radio occultation antennae of the Spire satellites are significantly smaller than those for, for example, the GRAS instrument, and also because the Spire data processing requires the use of single differencing to account for receiver clock errors, Spire measurements are expected to exhibit smaller Signal-to-Noise Ratios (SNRs) than occultation data from the Metop satellites. Figs. 25 to 27 demonstrate that this is indeed true. All figures show histograms of the mean 60-80 km altitude SNR distribution of the Spire and Metop data. For Spire, the data is stratified by satellite; as discussed before, we left out the four satellites (FM046, FM075, FM080 and FM088) which together provide fewer than 15.000 occultations to the overall dataset to keep the number of subplots reasonable. SNRs are shown for both L1 (Fig. 25) and L2 measurements (L5 for Galileo data; Fig. 26). For comparison, Fig. 27 shows the same statistics for the GRAS instruments on Metop satellites. In order to make the plots comparable, the horizontal axes are identical for the Spire and Metop data, but differ for the L1 and L2(5) figures. The different colours indicate data from observations using different GNSS satellite types.

The figures show a number of interesting features for both data sets. To start with, Spire SNRs are - as expected - smaller than those of Metop, for both L1 and L2 measurements: While the bulk of Metop measurements exhibit SNRs between 500 and 1000 V/V for L1 (100 and 400 for L2) measurements, Spire data exhibit typical SNRs between 50 and 500 V/V for L1 (and 50 and 300 for L2 or L5) signals,

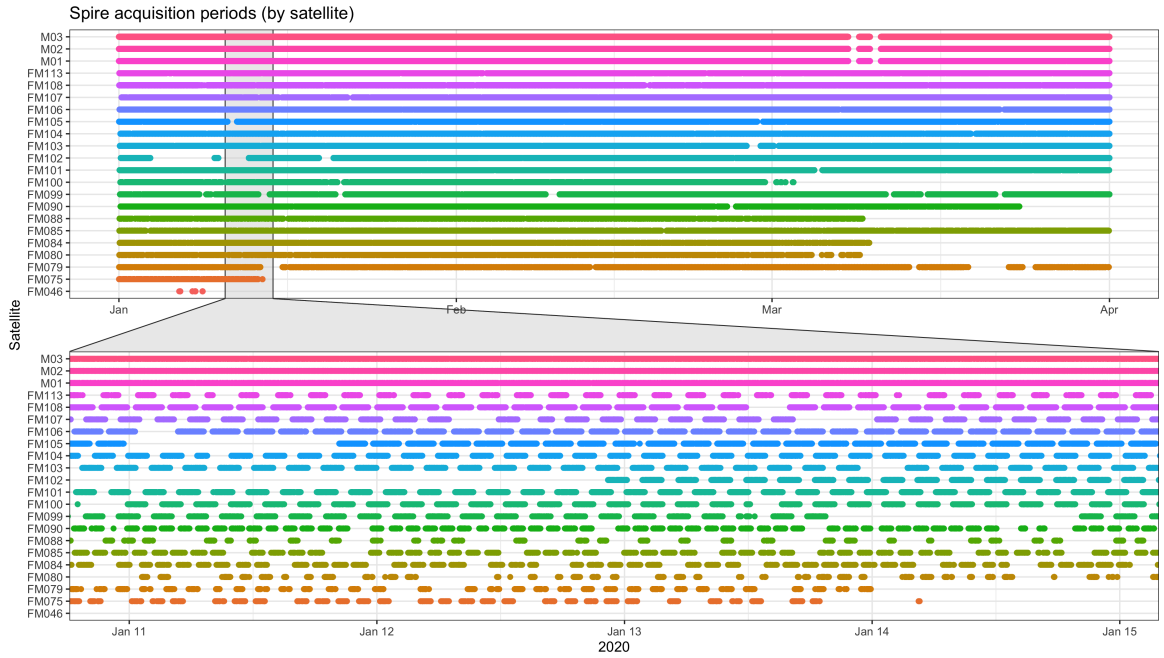


Figure 24: Timeline of the available occultation soundings for Spire satellites during January, February and March 2020; the vertical axis shows the Spire satellite IDs. Data from the three Metop satellites is shown on top of the figure. The lower part of the figure shows a zoom into the timeline between 11 and 14 January.

respectively. Note that Spire satellites track only non-encrypted signals; this explains why the Spire GPS measurements are limited to Block IIF and IIR-M GPS satellites transmitting the public L2C signals of the modernised GPS. Because the tracking of these newer unencrypted signals doesn't require code-less tracking algorithms, they should offer benefits in terms of SNR for receivers being able to use them. Nevertheless, Metop SNRs, despite being based on the encrypted legacy L2 signals, still provide higher SNRs in the secondary frequency compared to the Spire instruments.

A specific characteristic of Metop data is that for a small number of occultations, L2 SNRs can be exceptionally high (e.g., 800 to 1000 V/V). Here, the GRAS instruments (which track encrypted legacy GPS signals in the L2 band only) benefit from the flex-power mode of recent II-RM and II-F satellites that have their L2 transmit powers increased above their nominal values over certain parts of the Earth. Spire satellites, only tracking the modern L2/C GPS signals, do not benefit from flex power variations which are limited to the GPS legacy signals. Thus, their distribution is more homogeneous and, overall, lower. In addition, variations in the transmit power do not occur for other GNSS constellations.

For Spire satellites, the SNR distributions are not homogeneous. The peak of the L1 NR distributions is around or above 250 V/V for most newer satellites (from FM099 onward), but some exhibits their peak below (e.g., FM105); older satellites can exhibit significantly lower SNRs, e.g. FM085 and also FM088 (not shown) with a peak around 100 V/V only. For L2, the newer FM103 has a comparatively poor SNR performance.

A parameter directly related to the accuracy of bending angle data is the inherent noise present in carrier phase measurements. In theory, SNR is inversely related to the phase noise; high values of SNR, therefore, are associated with lower noise and thus better bending angle performance at high altitude. The overall lower Spire SNRs at both measurement frequencies thus translate into slightly poorer bend-

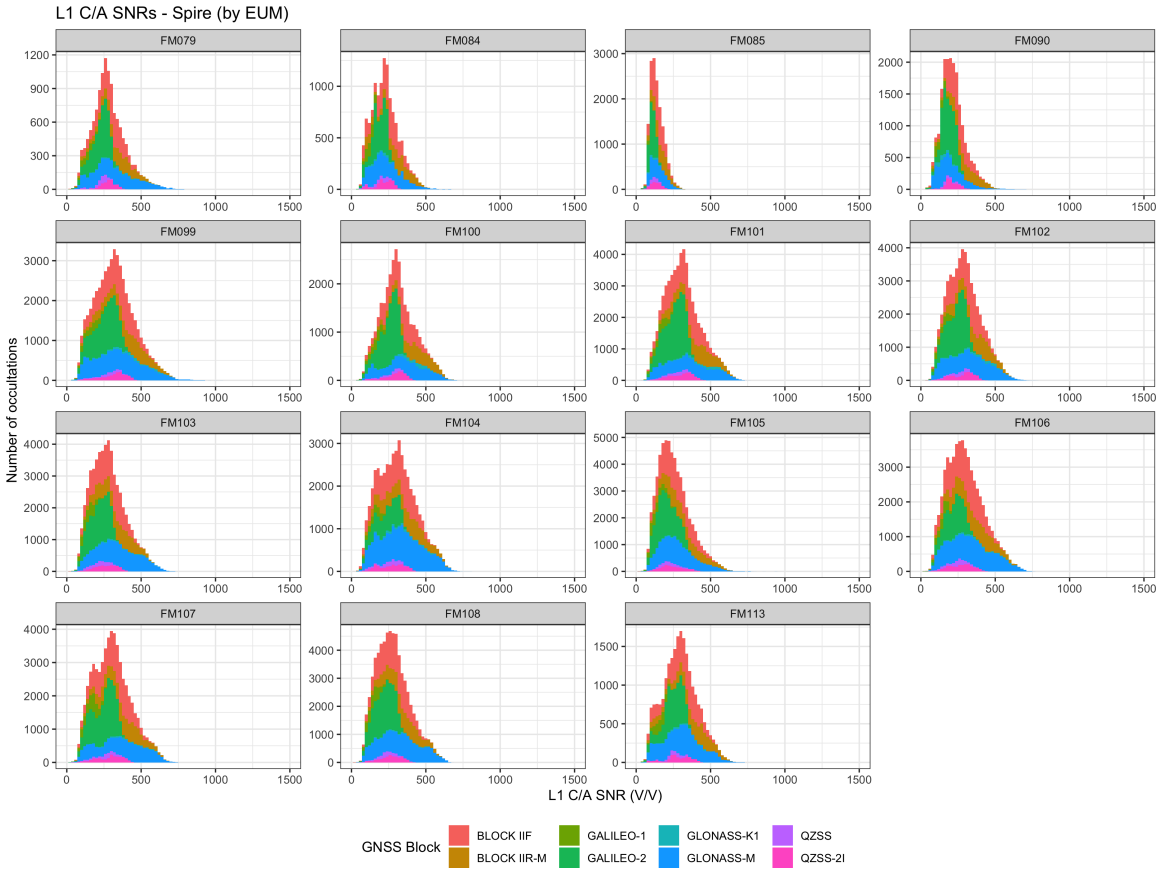


Figure 25: Distribution of the upper altitude (60 - 80 km) mean SNRs (in V/V) for L1 C/A carrier phase measurements for the main Spire satellites. Colours indicate the satellite type ("block") of the transmitting GNSS satellite; see the legend at the bottom of the figure.

ing angle performance compared to higher SNR missions. It is possible to independently estimate the (excess) carrier phase noise present in raw RO measurements. The EUMETSAT processing includes a carrier noise estimate based on a smoothing cubic spline fit to excess phase measurements for altitudes between 60 and 80 km height. An optimal smoothing parameter is estimated for each individual profile and frequency by minimising the Generalised Cross Validation (GCV) score of the fit; the average carrier phase noise over the entire altitude range can then be estimated from the residuals of the fit [Wahba, 1975, Craven and Wahba, 1979]. The limitation of the statistics to the height range between 60 and 80 km ensures that variations in the carrier phase levels are mostly related to the instrument noise (which is characterised by the receiver's SNR), but do not reflect carrier phase variations due to, e.g., neutral atmospheric effects on the signal propagation.

Results for the two frequencies involved in RO measurements of Spire data, stratified by satellite, are shown in Figs. 28 and 29 for the Spire satellites, as well as in Fig. 30 for Metop. As for the SNR distributions, Spire satellites providing only few occultations were left out of the figures for simplicity.

By and large, the independently estimated carrier noise levels allow similar conclusions as the SNR plots. For the Spire data, carrier phase measurements exhibit significantly larger noise levels, in line with their lower SNRs. Notable are again the differences in the noise distributions that occur between different satellites. The satellites FM085 and FM088 (not shown) stand out due to their very large L1

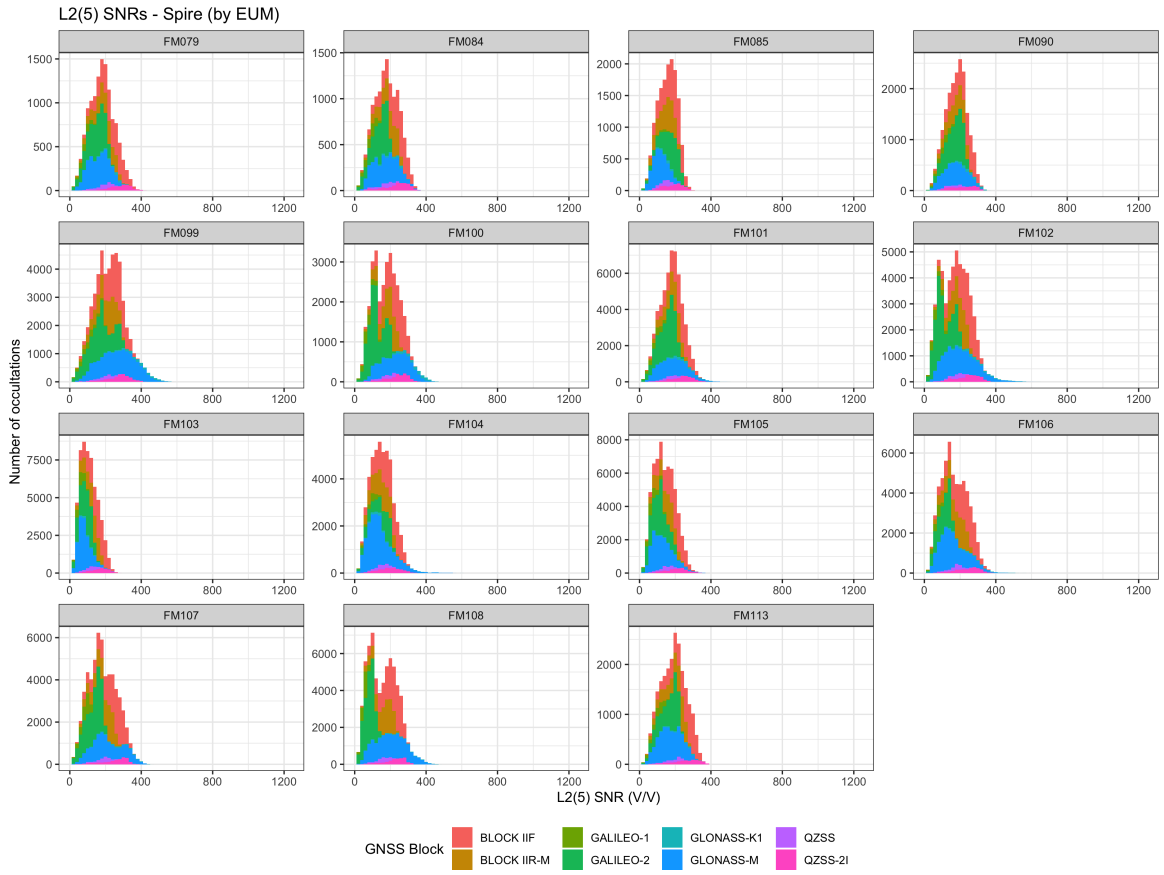


Figure 26: As Fig. 25, but for L2 (L5 in case of Galileo) carrier phase measurements.

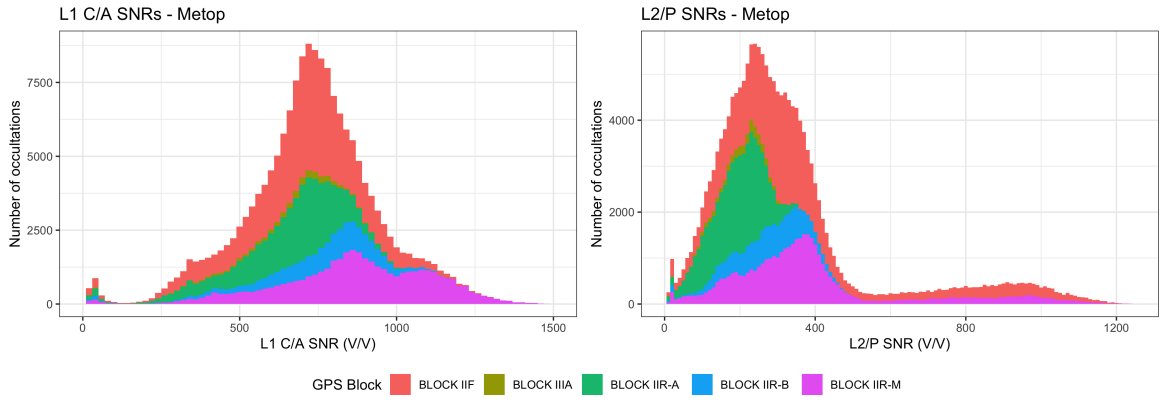


Figure 27: Distribution of the upper altitude (60 - 80 km) mean SNRs (in V/V) for L1 C/A (left) and L2/P (right) carrier phase measurements for Metop occultations. Colours indicate the satellite type ("block") of the transmitting GPS satellite.

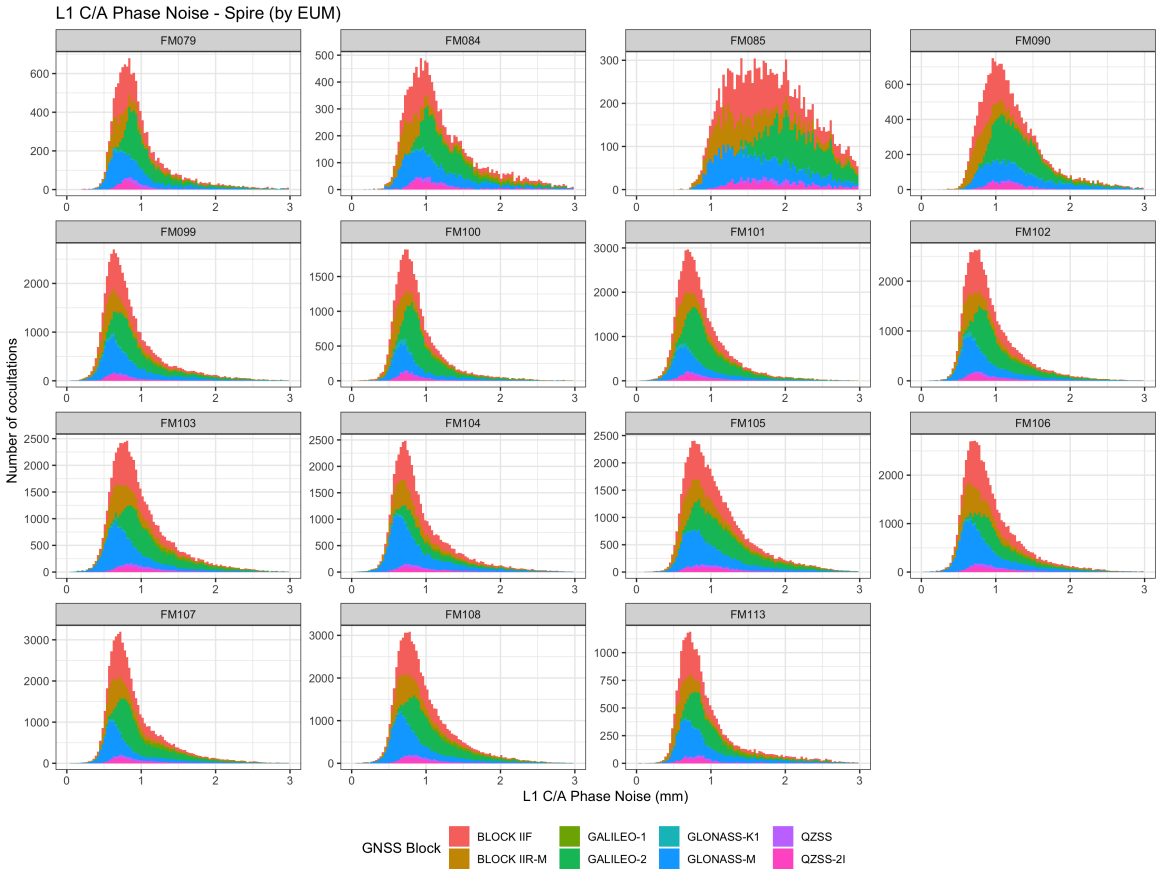


Figure 28: Distribution of the upper altitude (60 - 80 km) phase noise (in mm) for L1 C/A carrier phase measurements for the main Spire satellites. Colours indicate the satellite type ("block") of the transmitting GNSS satellite; see the legend at the bottom of the figure.

carrier phase noises; all remaining Spire satellites exhibiting peaks in the phase noise distributions below 1 mm. In comparison, Metop phase noise peak well below 0.5 mm. For L2 (or in case of Galileo, L5) data, FM103 is one of the newer satellites that stands out with large carrier phase noise values with a peak in the distribution around 2 mm; the remaining satellites exhibit distribution peaks between 1 and 2 mm. In contrast, Metop L2 carrier phase noises exhibit a peak around 0.75 mm. Note that newer Spire satellites (from FM100 onward) appear to exhibit longer tails in their L2 carrier phase noise distributions compared to older versions of the Spire receiver. The reason for these inter-satellite differences is not clear to us; Spire usually argues that data from different satellites has rather similar error characteristics. At least for SNR and phase noise, our findings are contradictory to this statement (but see the analysis of bending angle noise below). Overall, it seems advisable to regularly monitor the noise performance of existing and upcoming Spire satellites.

Carrier phase noise (or uncertainty) translates into bending angle uncertainty in altitude regions where phase noise contributes significantly to the bending angle budget, i.e. in the upper stratosphere and mesosphere. In order to estimate the random noise of the bending angle retrievals, EUMETSAT fits a low (3rd) order polynomial to the neutral bending angle profile between 60 and 80 km height and uses a statistically robust estimate of the standard deviation as proxy for the high frequency / small vertical scale noise in the altitude range (see Wilcox [2017] or Maronna et al. [2019] for an introduction into robust statistics). The use of a robust estimator aims at rejecting large peak-like deviations from the smooth

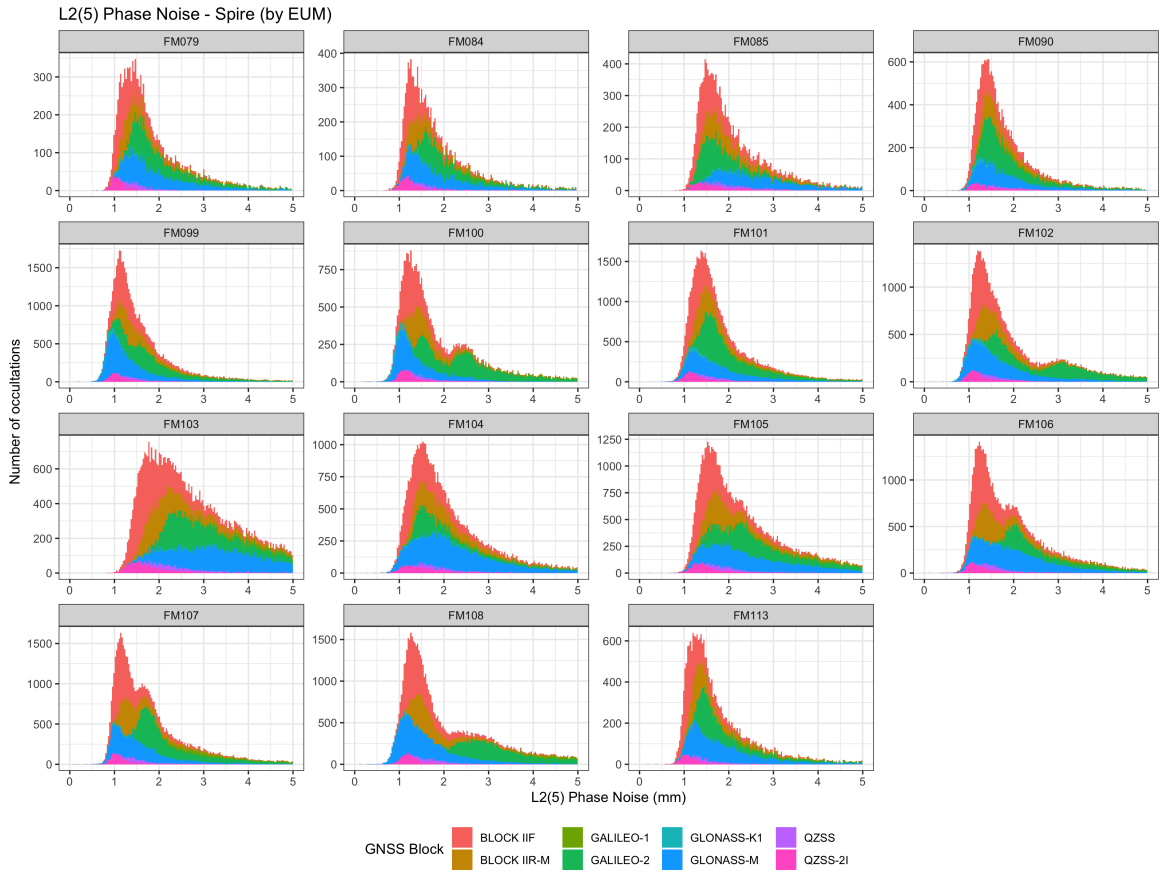


Figure 29: As Fig. 28, but for for L2 (L5 in case of Galileo) carrier phase measurements.

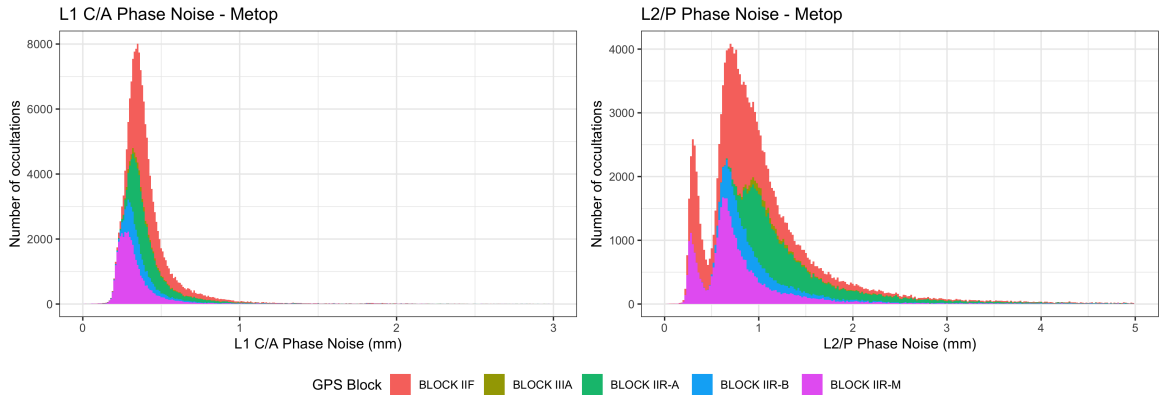


Figure 30: Distribution of upper altitude (60 - 80 km) phase noise (in mm) for L1 C/A (left) and L2/P (right) carrier phase measurements for Metop occultations. The colour indicates the type ("block") of the transmitting GPS satellite.

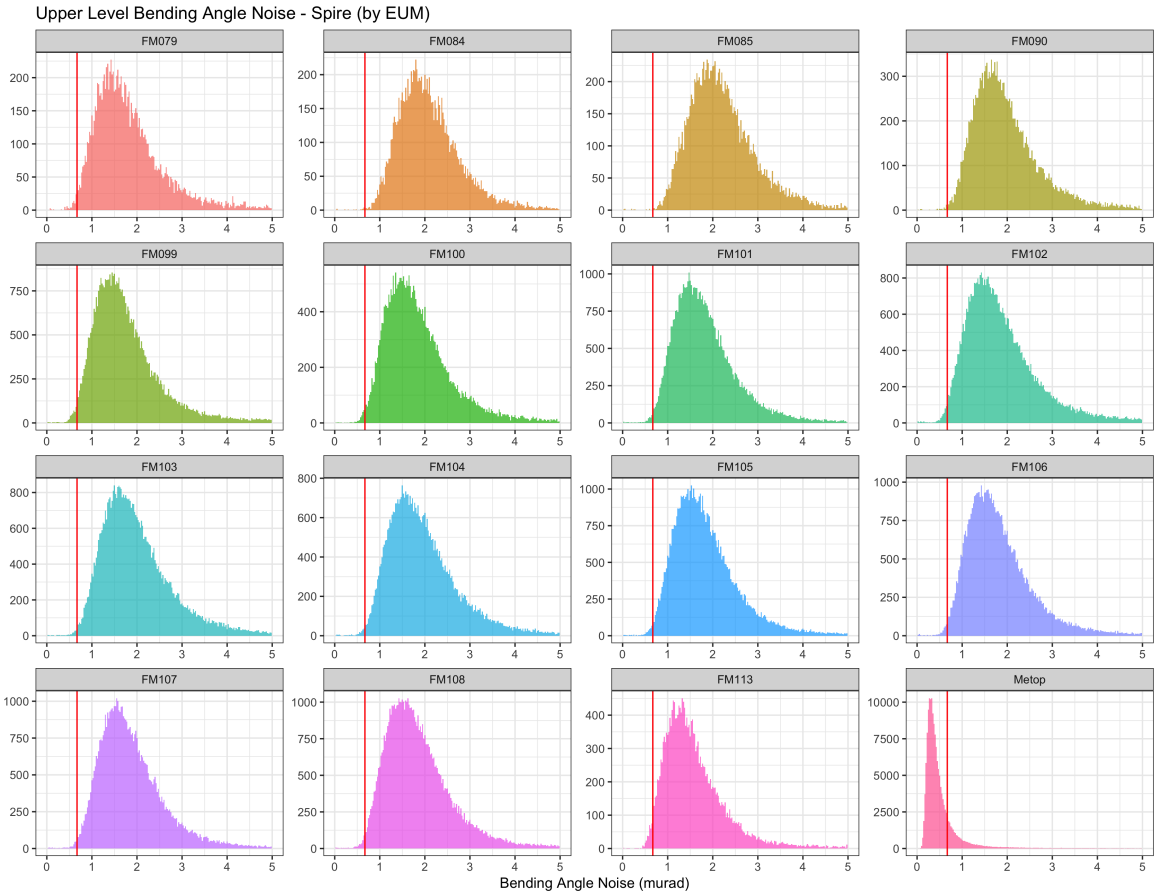


Figure 31: Bending angle noise (in μrad) between 60 and 80 km altitude for the main Spire satellites. The plot in the lower right corner shows the distribution of bending angle noise values for Metop occultations, while the red line indicates the formal one-sigma requirement for neutral atmosphere bending angles for the Metop instruments.

polynomial fit caused by ionospheric residuals. Thus, the noise numbers obtained are an estimate of the instrument's impact on the bending angle performance at high altitudes. They form a lower bound to the actual bending angle error as ionospheric residual perturbations will further increase the bending angle uncertainty in general.

Fig. 31 summarises the results of the bending angle noise estimation. Apart from the distribution of the noise estimates for the main Spire satellites, the figure also shows the bending angle noise for Metop data estimated in the same way in the lower right subplot. The formal 1-sigma requirement for Metop bending angle performance is indicated by a vertical red line.

For the majority of the Spire satellites, the bending angle noise distribution peaks around 1.5 μrad, with the bulk of the data exhibiting noise values between 1 and 3 μrad. For some satellites (e.g., FM085 and FM088 (not shown)), the noise distribution is shifted to peak values of up to 2 μrad. Nevertheless, the shape and width of the bending angle noise distributions for the various Spire satellites shows a greater homogeneity than the SNR and carrier phase noise distributions. This suggests that the filtering algorithms implemented in the EUMETSAT processing are able to mitigate a large part of the high frequency carrier phase noise present in the Spire data.

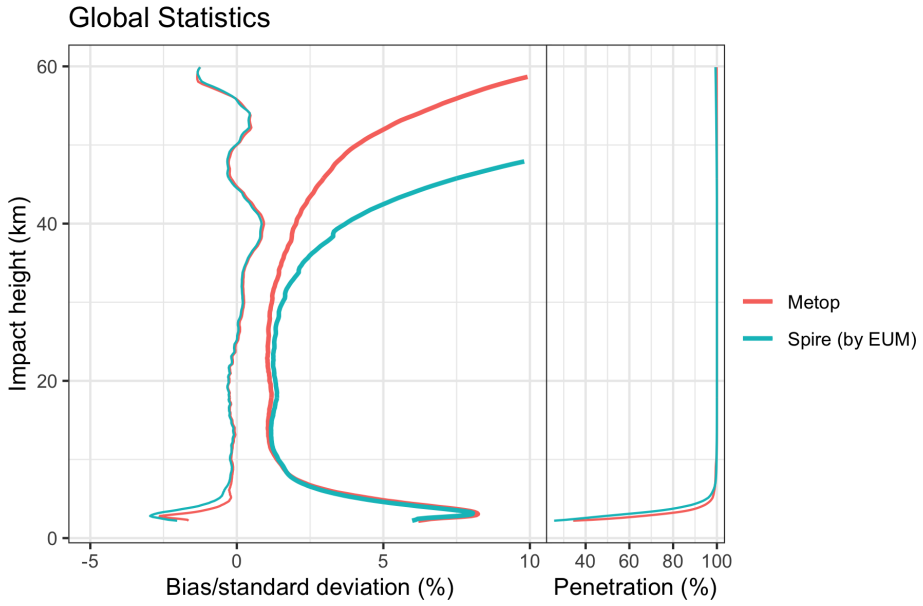


Figure 32: Global observation minus background statistics for Spire (blue) and Metop (red) bending angle data against ECMWF operational short range forecasts between January and March 2020. Note that for Spire, the statistics includes data from all four GNSS constellations supported by the Spire receivers, and also data from all Spire satellites. Metop only provides GPS measurements.

We finally note that in comparison to Metop, bending angle noise levels are significantly higher in the Spire data. This is an expected result, given the smaller antennas (and thus SNRs) of the Spire instruments as well as the need for zero-differencing in the Spire data processing. With bending angle noise peaking at around 0.5 μrad and the bulk of the Metop data exhibiting noise levels below 1 μrad, Spire bending angle uncertainties appear to be larger by a factor of about 3. The formal Metop accuracy requirements for bending angle data are not satisfied by the Spire data.

6.1.5 Global statistics against ECMWF

Fig. 32 shows the global statistics of both Spire and Metop data against co-located ECMWF short-range forecast profiles; presented are robust bias and standard deviations against the NWP data. For both data sets, the vertical bias structure is consistent; deviations from zero can be attributed to systematic errors in the ECMWF data, at least in the upper troposphere, stratosphere and mesosphere. The only visible difference in the bias structure occurs in the lower troposphere: while both data sets are slightly negatively biased, the Spire biases are slightly larger than the Metop ones. Also, Metop bending angle profiles appear to penetrate around 1 km lower into the lowest few km of the troposphere.

In the stratosphere and mesosphere, Spire deviations from the ECMWF data are significantly larger than those for Metop, in line with the higher carrier phase noise found in the low level data. These larger deviations are present throughout the entire stratosphere, with significant deviations beginning at an altitude around 30 km and increasing with altitude.

A known issue in Metop data is an asymmetry in the statistics between rising and setting occultations in the lower tropical troposphere. Recent updates to the EUMETSAT processing mostly resolved the issue partially, but some asymmetry remains and can be seen in the global statistics; see Fig. 33 (left).

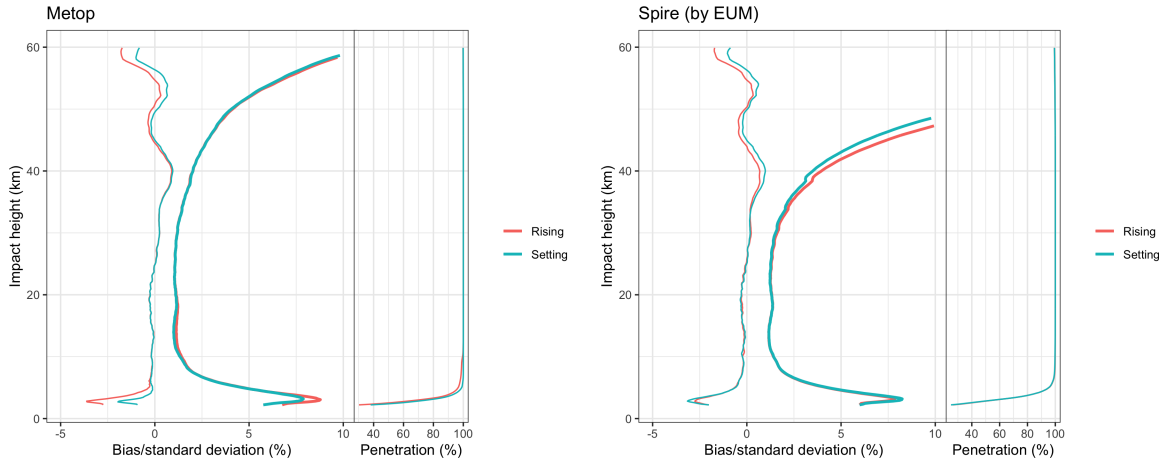


Figure 33: As Fig. 32, but showing the Metop (left) and Spire (right) statistics stratified by rising (red) and setting (blue) occultations.

Metop data also exhibits a small bias between rising and setting occultations above 45 km. In contrast, the Spire data does not suffer from systematic or random asymmetries in the statistics against ECMWF data in the troposphere. However, the standard deviation of rising occultations above 35 km is slightly increased compared to the setting ones. Above 40 km, rising and setting biases of Spire data differ by similar amounts as for Metop.

The left plot in Fig. 34 shows the global statistics for Spire data, this time stratified by the GNSS constellation. The vertical bias structure is consistent for the various GNSS systems, with minor variations in both the lower troposphere as well as the upper stratosphere and mesosphere. Above 35 km, GLONASS occultations exhibit slightly larger standard deviations compared to the other three GNSS systems. In this altitude region, the atmosphere is relatively homogeneous; variations in the sampling usually do not affect statistics against NWP data strongly. The increased standard deviations for GLONASS occultations thus are likely caused by the data characteristics of the GLONASS measurements. A common explanation is that GLONASS satellite clocks perform poorer compared to the clocks on other GNSS satellite systems, causing a small but measurable degradation of GLONASS measurements [Schreiner et al., 2020].

The situation is more complicated in the lower stratosphere and troposphere. Here, Spire data based on QZSS measurements appears to outperform occultations from the other GNSS constellations. On the other hand, error characteristics in these altitude regions are more sensitive to sampling issues: In the tropical lower stratosphere, gravity waves are omnipresent in the vicinity of tropical convection. RO measurements resolve these small-scale vertical fluctuations, while NWP models usually do not, or at least under-represent gravity wave activity. As a consequence, standard deviations between RO and NWP data in the lower and mid-stratosphere are usually larger in the tropics. In the troposphere, the high abundance of tropical water vapour causes larger errors in lower tropospheric RO data at low latitudes. As occultations derived from QZSS constellations occur primarily in the extra-tropics, they exhibit lower standard deviations in these altitude regions. The apparently better performance of QZSS occultations in global validation statistics is thus likely caused by their different meridional distribution compared to occultations obtained from other GNSS constellations.

The large number of occultations in the Spire data set available in this study allows to analyse the consistency of bending angle data from different Spire satellites. The right plot in Fig. 34 shows the global

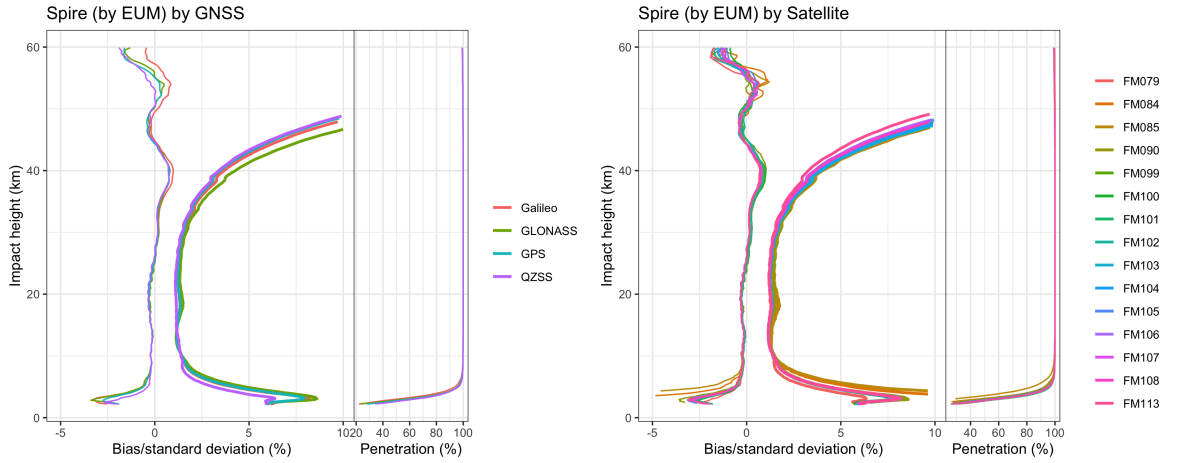


Figure 34: As Fig. 32, but showing the Spire statistics stratified by GNSS constellation (left) and by Spire satellite (right).

statistics for the main Spire satellites. In the stratosphere and mesosphere, the vertical bias structures of the Spire satellites are consistent with each other, although the individual bias curves spread slightly above 35 km altitude; the spread is of similar size as the spread between the biases obtained from different GNSS systems, or between rising and setting occultations. We also find variations in the standard deviations for different satellites in the upper stratosphere and mesosphere, again similar in magnitude to the spread for different GNSS constellations.

In the troposphere, Fig. 34 shows larger differences for some of the Spire satellites: FM085 and FM084 exhibit significantly increased (negative) biases and standard deviations below 8 km. On the other hand, FM079 shows the smallest bias and standard deviation of all satellites in the same altitude range. As before, however, tropospheric variations in the statistics against NWP data may be related to their sampling patterns. For example, both FM084 and FM085 (as well as FM088; not shown) are flying on low-inclination orbits, resulting in a larger number of tropical occultations and hence larger tropospheric uncertainties. For FM079, we noted above that this satellite provides very few occultations in Southern Hemisphere low latitudes, in effect reducing the number of tropical occultations compared to other Spire satellites.

6.1.6 Tropospheric performance

The results from the last section suggest that for analysing the performance of tropospheric radio occultation data, the sampling distribution has be taken into account - especially when comparing different instruments, GNSS constellations or individual satellites. To compensate for different sampling patterns, we calculate observation minus background statistics in separate latitude bands. The comparison between the Spire and Metop data in low (between 30°S and 30°N), medium (between 30° and 60° latitude in both hemispheres) and high (poleward of 60°) latitudes is shown in Fig. 35. Statistics for rising and setting occultations are shown separately.

Note that the figure uses impact height and not geometrical altitude as vertical height scale. The difference between these is negligible in the stratosphere and above, but significant close to the Earth's surface. As a rule of thumb, the sea surface is typically located around 2 km impact height. Fig. 35 thus shows the vertical range between the sea surface and the upper troposphere.

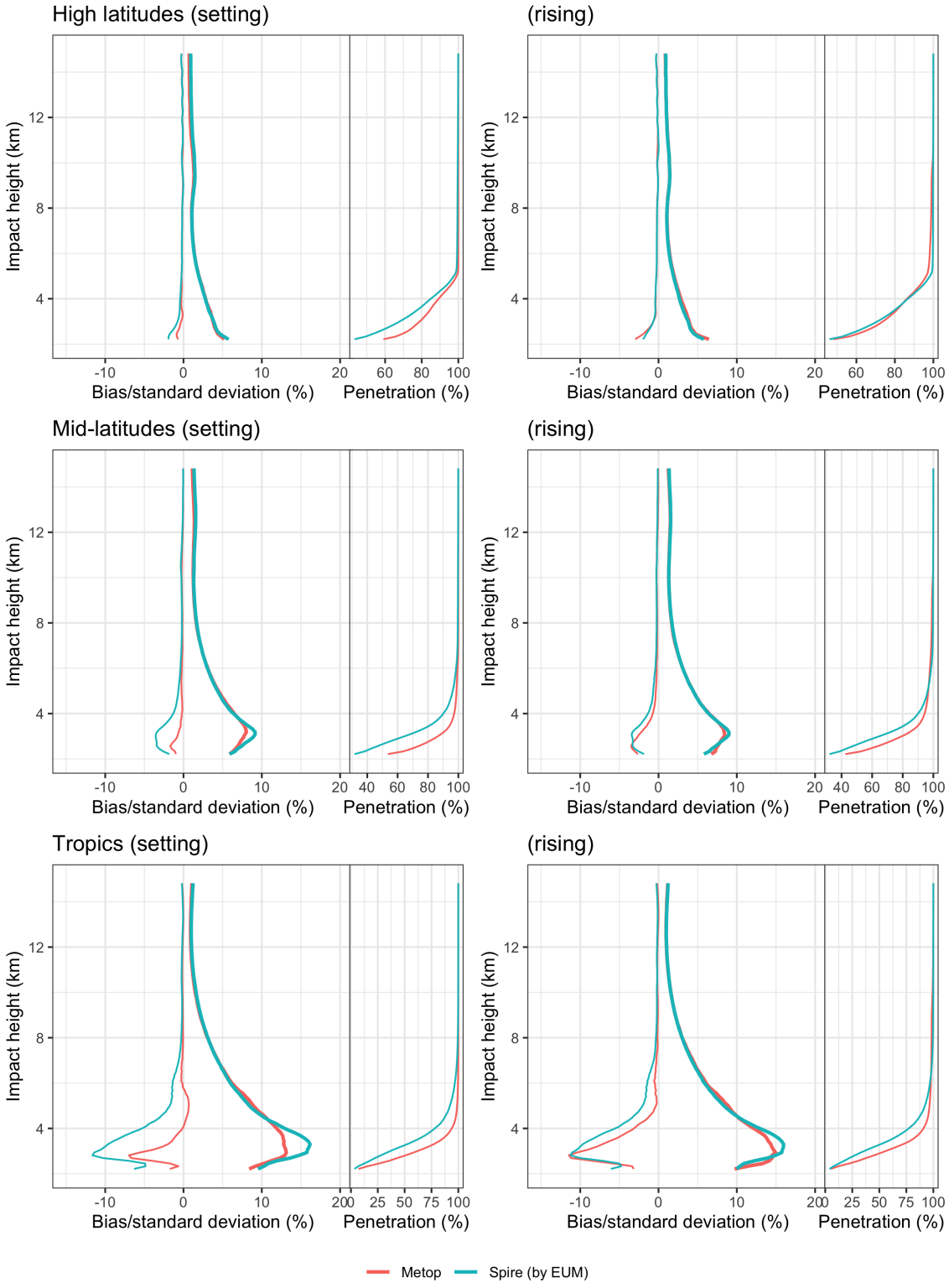


Figure 35: Tropospheric observation minus background statistics for Spire (blue) and Metop (red) bending angle data against ECMWF operational short range forecasts between January and March 2020. The figure shows the statistics in high (top), middle (centre) and low latitudes (bottom), stratified by rising and setting occultations. For Spire, the statistics includes data from all four GNSS constellations supported by the Spire receivers and from all Spire satellites. Metop only provides GPS measurements.

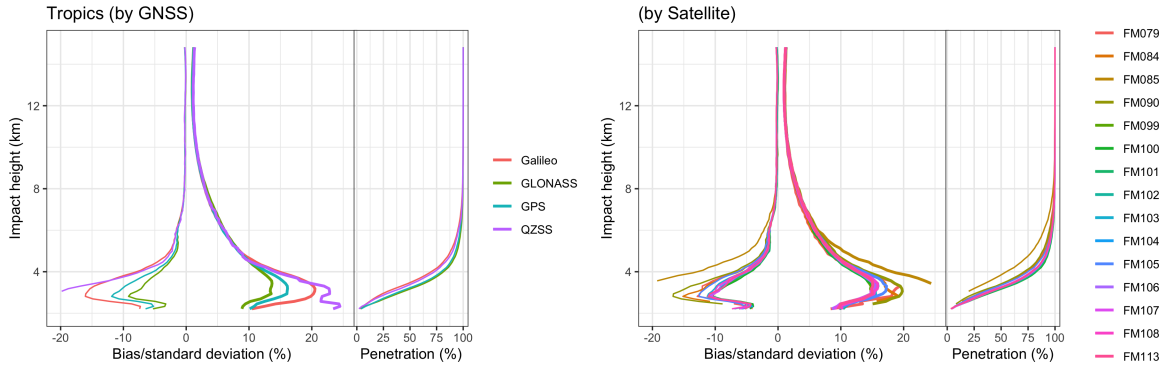


Figure 36: As Fig. 35, but for the tropics and Spire data only. The statistics is stratified by GNSS constellation (left) and individual Spire satellite (right). Rising and setting occultations were combined; also note that the horizontal axis has been extended in range compared to Fig. 35 to allow for larger deviations.

Overall, the statistics are similar to results from other comparisons and validations of RO data, and also to our previous analysis of the Spire data quality. Both systematic and random deviations in the troposphere are smallest in high latitudes and increase towards the tropics where water vapour abundance is high. Biases in the lowest 2 to 4 km above the surface are typically negative and largest in the tropics. Note that in the previous analysis of Spire data quality, we found that original Spire data are mostly positively biased in the lower troposphere, in contrast to other RO missions. In the EUMETSAT processing, the Spire biases are consistent with those of other instruments.

Compared to Metop, the EUMETSAT processed Spire data exhibits larger negative biases in mid and low latitudes, especially for setting occultations. Lower tropospheric Spire biases reach up to -10 % close to the surface in the tropics, and begin to be present at somewhat higher altitudes than for Metop (for example, significant negative biases in the tropical occultations are present below 6 km impact altitude for Spire, but only below 4 km for Metop). Standard deviations are comparable between Spire and Metop data, apart from the lowest 2 km above the Earth's surface in the tropics where Spire data exhibits slightly larger standard deviations. The Spire data does not reach down to the surface as close as the Metop data does; Spire bending angles appear to miss the lowest 0.5 - 1 km that are still covered by Metop data. An exception are rising occultations in high latitudes, where both datasets are comparable in terms of their penetration into the lowest few km of the troposphere.

Differences in Spire's observation minus background statistics in the Tropics for different GNSS constellations and the various Spire satellites are shown in Fig. 36. While the statistics is consistent between GNSS constellations and also Spire satellites in the free troposphere (i.e., for impact heights above 4 - 6 km), significant differences in both biases and standard deviations remain in the lowest 2 km of the troposphere for both different GNSS constellations and also for some Spire satellites. They thus cannot be solely explained by varying sampling patterns.

Concerning GNSS constellations, both (negative) biases and standard deviations are largest for QZSS; compared to GPS and GLONASS occultations, both are are nearly twice as large in the lowest 2 km of the tropical troposphere. In case of Galileo data, biases and standard deviations are larger by about 50% compared to GPS and GLONASS. The Galileo findings are qualitatively similar to results in the first Spire data quality assessment, where we found that Galileo data processed by Spire exhibit larger (although positive) biases and standard deviations in the lower tropical troposphere. At the time, there were not enough occultations available in the tropics for conducting a similar analysis for QZSS.

For individual Spire satellites, inconsistencies occur in the lowest two km of the troposphere as well. Here, both biases and and standard deviations vary by as much as 5 % between Spire satellites. One satellite (FM085) remains standing out with significant larger biases and standard deviations against the ECMWF data. We found that the same satellite (together with FM088 for which no statistics is shown) also stood out with rather low SNRs and high carrier phase noise values. Fig. 36 (right) further suggests that data from this satellites penetrates less deep into the lower troposphere compared to the other Spire satellites.

Overall, the results in this section indicate that Spire data characteristics do vary between both GNSS constellations and between different satellites, with the main effects seen in the boundary layer of the atmosphere.

6.1.7 Summary

EUMETSAT processed Spire excess carrier phase data using the same algorithms, filter settings and (mostly) identical quality control limits as for the upcoming operational processing of GRAS data from its Metop instruments. The low level data analysis confirmed (as expected) that the SNR of Spire data is significantly lower than that of Metop. Lower SNR values translate into higher carrier phase noise, which in turn results in increased bending angle noise levels in the stratosphere and mesosphere. Therefore, Spire bending angle data exhibits significantly increased standard deviations against ECMWF short range forecasts in the upper stratosphere and mesosphere. In the mid-/upper troposphere and lower to mid-stratosphere, Spire bending angle performance approaches that to Metop data as the role of instrument noise for the bending angle uncertainty decreases with altitude.

In the troposphere, EUMETSAT-processed Spire data exhibits negative biases, similar to other major RO missions. This contrasts with positive biases of similar size found in original Spire data in a previous assessment of Spire data quality. While lower tropospheric biases of Spire data are slightly larger than for Metop, standard deviations of Spire and Metop data are mostly consistent in the troposphere, at least for the free troposphere above 2 km above the Earth's surface.

In the free troposphere and lower to mid-stratosphere, Spire data for different GNSS constellations as well as from different Spire satellites is more or less consistent with each other. However, while the data characteristics of different constellations and Spire satellites are similar in these altitude regions, they are not identical. In particular, we found larger deviations above 35 km altitude as well as in the lowest 2 km of the troposphere.

Above 35 km altitude, inter-satellite and and inter-constellation variations are of the same order of magnitude; bias variations are comparable with the difference in the statistics between rising and setting occultations for both Metop and Spire bending angle data. Given the size of RO standard deviations at these altitudes, we do not consider these differences problematic, at least in an operational NWP context.

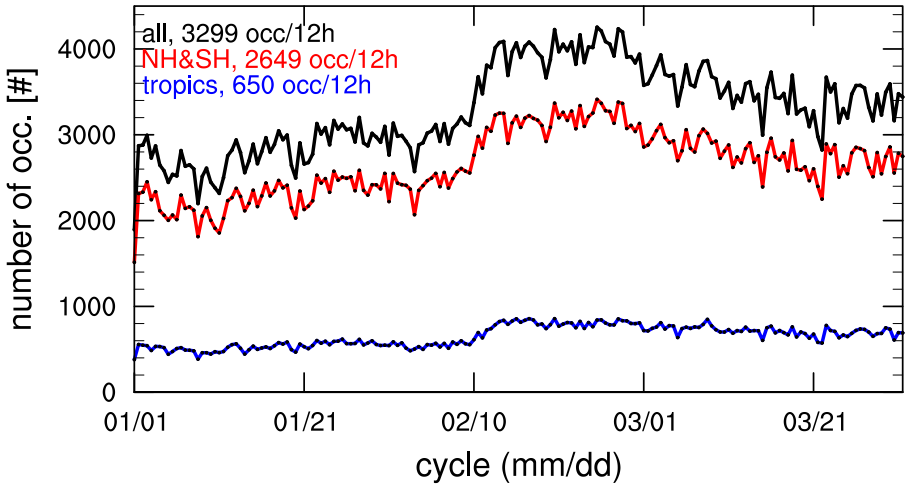
In the lower troposphere, both systematic and random deviations vary by as much as 5 % between individual Spire satellites close to the surface. As in the previous assessment of Spire data quality, we also found that Galileo data performs somewhat poorer in the lower troposphere compared to GPS and GLONASS constellations, as does QZSS data.

The deviations between Spire satellites in the lower troposphere are within the spread known from the comparison of data from other RO missions. Thus, there is no reason to not use them in NWP applications. However, the finding highlights that data from individual Spire satellites is not automatically exchangeable. Thus, close monitoring of the data characteristics for individual Spire satellites seems

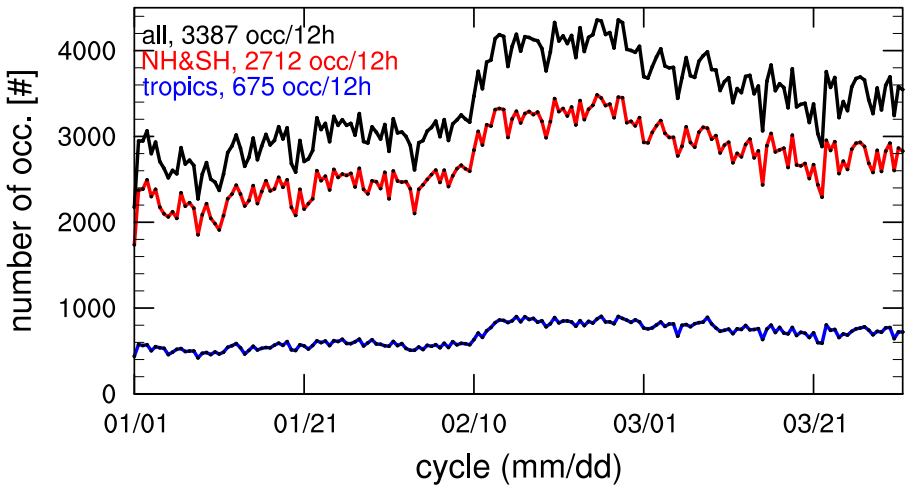
advisable, as is routinely done for other RO missions as well, e.g. by the ROM SAF.

6.2 Results from assimilating Spire data processed at EUMETSAT

The assimilation results using the Spire data processed at EUMETSAT have run also for the full three month experimentation period (1 January to 31 March 2020). Fig. 37 shows the comparison of how many Spire occultations have been assimilated at ECMWF using the EUMETSAT processing and the original one by Spire. It is clearly shown, that the numbers are very similar.



(a) Spire EUM



(b) Spire

Figure 37: Number of occultations per 12 hour cycle for Spire for January to March 2020 at 5 km to 6 km height.

6.2.1 *Verification of medium-range forecasts against ECMWF operational analysis*

First, we compare the forecast error statistics in the experiments (**Spire**, **Spire-EUM** and **noRO**) with the corresponding statistics in **CONTROL** for temperature. As it will be done also later in the remaining sections, we take the standard deviation of the forecast errors in the experiments, $\sigma_{exp}(z)$, and compare it with the standard deviation of **CONTROL**, $\sigma_{CONTROL}(z)$. Hence, if the experiments have a positive impact $(\sigma_{exp}(z) - \sigma_{CONTROL}(z))/\sigma_{CONTROL}(z)$ would be negative.

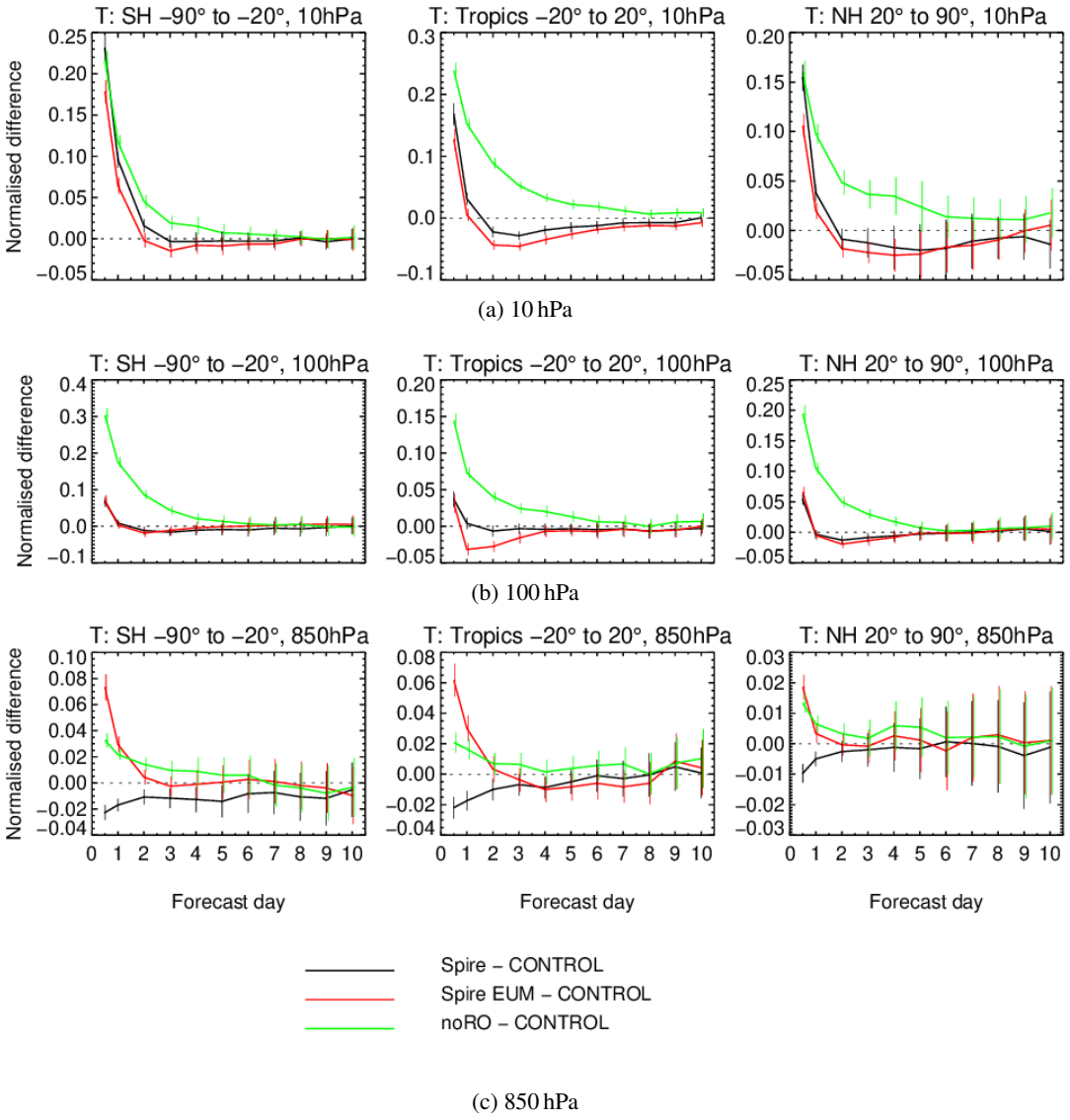


Figure 38: Normalised differences in standard deviation in temperature between **SPIRE**, **Spire EUM** and **noRO**, and CONTROL, verified against ECMWF operational analysis at different height levels (850 hPa, 100 hPa, and 10 hPa) for different forecast times and zonal regions (left: SH 90°S-20°S, middle: Tropics 20°S-20°N, right: NH 20°N-90°N). Negative values represent a decrease in standard deviation and positive values an increase in standard deviation. The confidence range is displayed by vertical bars.

In Fig. 38 we show how the normalised standard deviation for temperature changes compared to **CONTROL** (dashed line around 0) for selected pressure levels and regions (Tropics, Southern Hemisphere

and Northern Hemisphere) using the ECMWF operational analysis as a reference. For all height levels and regions, the change in forecast errors is very similar for both Spire datasets at least from day 2 onwards. The biggest impact can be seen in the tropical stratosphere (100 hPa and 10 hPa) where Spire EUM does show a slightly larger decrease in temperature forecast error than the original Spire dataset. Interestingly, looking at the change in forecast errors for vector wind reveals that **Spire-EUM** shows a strong decrease (2-4%) in tropical forecast errors at 100 hPa for day 3 compared to **CONTROL**, whereas **Spire** does not (not shown). For other regions the forecast errors for vector wind do not differ much between **Spire-EUM**, **Spire**, **CONTROL** and **noRO**. Changes in the short-range forecasts (up to day 1/2) can be apparent using analysis based reference, as discussed earlier (section 4). Here, using observations as comparison is best.

6.2.2 Verification of short-range forecasts against observations

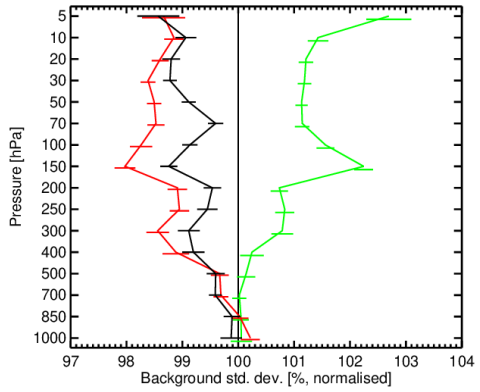
To illustrate in a bit more detail how the different fits to observations behave for **Spire**, **Spire-EUM**, and **noRO** compared to **CONTROL** for the ECMWF system, Fig. 39 is shown. Here fits to observations are shown as $\sigma_{\text{exp}}(p)/\sigma_{\text{CONTROL}}(p)$, where a decrease can be seen as an improvement and an increase as a degradation. For temperature sensitive observations (radiosonde temperature and AMSU-A; Fig. 39a and Fig. 39b, respectively), both Spire datasets clearly show improved fits compared to **CONTROL** and **noRO**, especially in the middle troposphere and higher. Here, **Spire-EUM** does show larger improvements than **Spire** by up to 1% at around 100 hPa for radiosonde temperature observations compared to **CONTROL**. For wind fits (Aeolus, Fig. 39e) the addition of Spire data does improve in the short-range but differences between **Spire-EUM** and **Spire** are small. Just between 150 hPa and 70 hPa, the fits for Aeolus using **Spire-EUM** are better than for **Spire**. For water vapour sensitive observations (Water vapour channels) of IASI; Fig. 39c the EUMETSAT processed Spire data gives larger reductions in normalised fits of standard deviation in first guess departures. Channel 14 of SSMIS (Fig. 39d), sensitive to total column water vapour, however shows a different picture with improvements of 0.5% by **Spire** and a small degradation of 0.2% by **Spire-EUM**. Nevertheless the humidity and cloud sensitive channels 9 to 11, which sense the atmosphere preferably between 700 hPa and 300 hPa, respectively show an improvement for **Spire-EUM**.

6.2.3 Summary

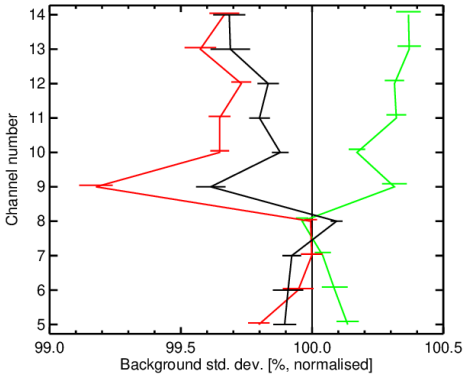
The EUMETSAT processed Spire data shows improvements in short-range and medium-range forecasts which are similar but sometimes even larger as the original Spire dataset. Especially for temperature in the mid troposphere and higher the improvements are much stronger for **Spire-EUM** than for **Spire**.Also for wind between 70 hPa and 150 hPa and some humidity or hydrometeor sensitive observations improvements are stronger for **Spire-EUM**.

7 ECMWF Ensembles of Data Assimilation (EDA)

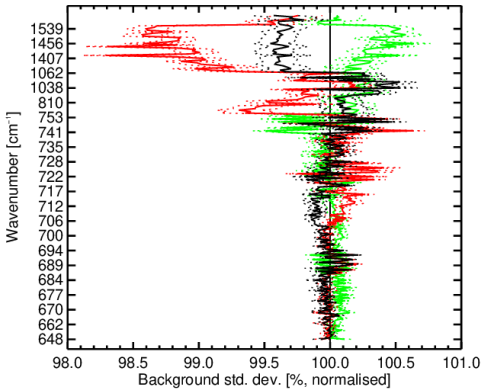
To estimate the statistical uncertainty in the short-range forecast we performed Ensembles of Data Assimilation (EDA, Bonavita et al. [2012]) experiments. The EDA configuration we used consists of ten independent cycling assimilation systems, where observations (including soil moisture), sea surface temperature, and model physics are perturbed. If those perturbations are correctly specified, the EDA will provide good estimates of short-range forecast error uncertainties. Harnisch et al. [2013] used the EDA



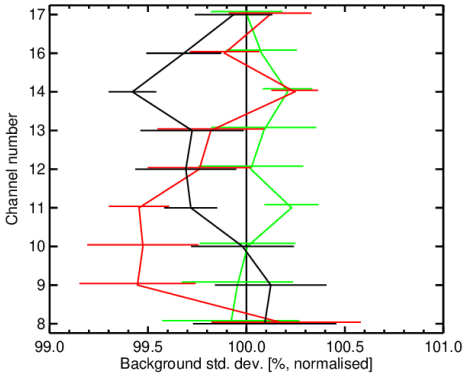
(a) Temperature radiosonde.



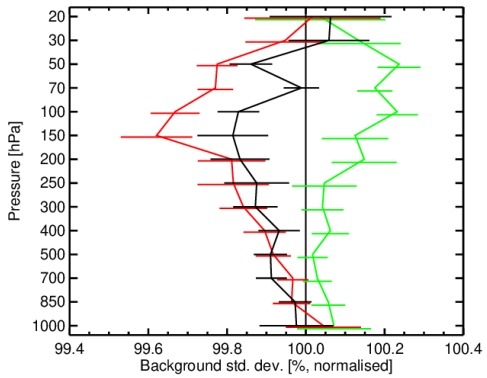
(b) AMSU-A



(c) IASI



(d) SSMIS-F17/F18



(e) Aeolus horizontal line-of-sight (H-LOS) winds

Figure 39: Normalised difference in standard deviation of first-guess departures between **noRO**, **Spire-EUM** and **Spire** for different instruments, globally. The normalisation is done with results from CONTROL. Values less than 100% would indicate beneficial impacts from the experiments. The horizontal bars indicate 95% confidence range.

spread values (standard deviation amongst ensemble members) to estimate how the impact of GNSS-RO data scales with observation number using simulated GNSS-RO data. More recent work by Bowler [2020] assimilating Spire data shows some consistency with the Harnisch et al. [2013] scaling plots, but

Bowler [2020] uses a combined RMS value averaged over meteorological variables and forecast range, so the quantities are not directly comparable and therefore further investigation is justified. Here, we aim to investigate how the relationship between EDA spread and the forecast error statistics changes as the number of real GNSS-RO measurements assimilated increases.

7.1 Background

Data assimilation is the process in operational NWP where information from a short-range NWP forecast is combined with the latest observations in a statistically optimal way, in order to produce an "analysis" [e.g. Lorenc, 1986]. The analysis is then used as the initial conditions for a new forecast. The ECMWF NWP system is based around a four-dimensional variational (4D-Var) approach [e.g. Rabier et al., 2000]. Data assimilation requires estimates of the statistical uncertainty in both the short-range forecast state and the observations, and the analysis should fit both sources of information to within their expected errors. In recent years, the estimation of the short-range forecast uncertainty has improved significantly with the introduction of ensemble techniques, and the ensemble approach adopted for this purpose at ECMWF is the EDA.

The EDA system has been used operationally at ECMWF since June 2011 (Isaksen et al. [2010], Bonavita et al. [2011]), as a way of estimating flow dependent background error covariance information for the operational ECMWF NWP system. A schematic is shown in Figure 40. The EDA implementation used for this study is close to the original operational implementation in 2011, and it consists of a 10-member ensemble of data assimilation cycles, plus one control experiment. These are run twice daily at 00 UTC and 12 UTC with an outer loop horizontal resolution of $T_{Co}399$ ($\sim$ 50 km horizontal resolution), and the 4D-Var assimilation window is 12 hours. The EDA experiments in this study are based on model cycle IFS 47R1 with 137 vertical levels and three inner loops at a resolution of $T_L95/T_L159/T_L255$ for the control run.

The EDA system uses perturbed model tendencies, perturbed observations – including sea surface temperatures – and perturbed model physics to represent the model error [Palmer et al., 2009]. The motivation for running an EDA system is that if the applied perturbations are consistent with the statistics of the actual error sources in the NWP system, then the ensemble information can be used to estimate the corresponding analysis and short-range forecast error *statistics*. The theoretical justification of the EDA technique is given in Žagar et al. [2005] and Isaksen et al. [2010], and it is reproduced in the Appendix C for completeness.

7.2 Methodology

We have performed various EDA experiments adding Spire and COSMIC-2 data over about a month, excluding the first 9 days because of spin up. That means the results include data for the period of 10 January to 10 February 2020. The following eight EDA experiments have been set up, similarly to their corresponding OSEs: **CONTROL**, **COSMIC2**, **Spire**, **Spire+COSMIC2**, **Only-Spire**, **Only-KOMP5**, **Only-GRAS** and **NoRO**. Here we show spread values derived on pressure levels, truncated at spectral wavenumber 255 and gridded on to a reduced gaussian grid N128. Harnisch et al. [2013] showed that it is best to capture the signal of the observation impact on rather short lead times. At longer lead times the model uncertainties have a growing impact on the signal. Hence, we choose to calculate the EDA spread values for T+12h.

In the context of assessing new observation types, we estimate how the new added observations reduce

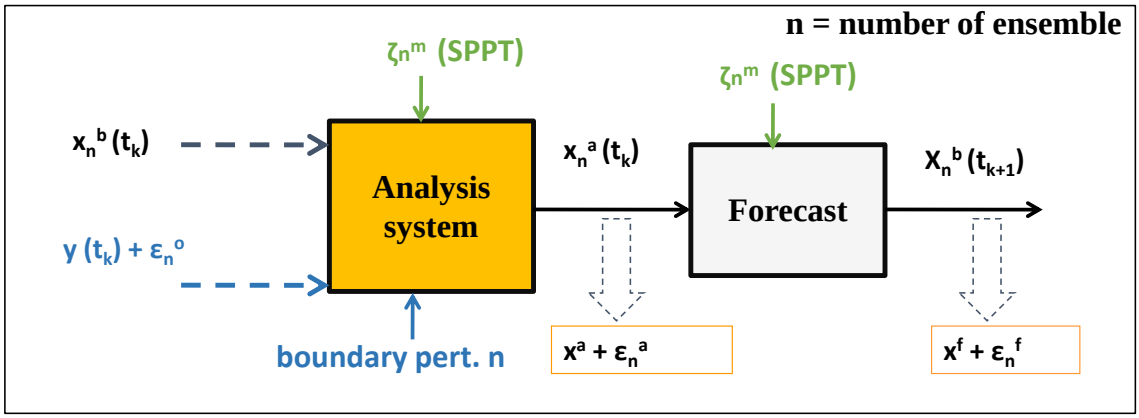


Figure 40: A schematic of the EDA method. Ten 4D-Var assimilation systems are run in parallel. Each 4D-Var has a vector of random noise ($\boldsymbol{\varepsilon}^{\mathbf{o}}_{\mathbf{n}}$) added to the observations, $\mathbf{y}(\mathbf{t_k})$, the boundary conditions like sea surface temperature, and the model physics ($\boldsymbol{\zeta}^{\mathbf{m}}_{\mathbf{n}}$, SPPT is the stochastically perturbed parameterized tendency, used to generate model errors.) An ensemble of 4D-Var analyses are produced, $\mathbf{x}^{\mathbf{a}}_{\mathbf{n}}$, each with a distinct error vector, $\boldsymbol{\varepsilon}^{\mathbf{a}}_{\mathbf{n}}$. These then result in an ensemble of short-range forecasts ($\mathbf{x}^{\mathbf{f}}_{\mathbf{n}}$), each with an error vector $\boldsymbol{\varepsilon}^{\mathbf{f}}_{\mathbf{n}}$, when propagated forward in time with the forecast model.

the EDA spread [Tan et al., 2007]. The EDA spread s_d for each assimilation cycle d is defined as the ensemble standard deviation σ_d about the ensemble mean, and it provides an estimate of the statistical uncertainty for each assimilation cycle d

$$s_d = \sigma_d = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}\left[x_n - \left(\frac{1}{N}\sum_{n=1}^{N}x_n\right)\right]^2} = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x_n-\bar{x})^2} \tag{1}$$

where x_n denotes the single ensemble member and $\bar{x}$ the ensemble mean.

Similarly, the averaged EDA spread s_{EDA} over D assimilation cycles is

$$s_{EDA} = \sqrt{\mathbb{E}\left[\sigma_d^2\right]} = \sqrt{\frac{1}{D}\sum_{d=1}^{D}\left(\frac{1}{N-1}\sum_{n=1}^{N}(x_n-\bar{x})^2\right)} \tag{2}$$

with the expectation operator $\mathbb{E}$.

To summarise, we use the EDA spread to estimate the *statistical uncertainty* in the short-range forecasts. This uncertainty is related to the probability density function (PDF) of the forecast errors. A specific forecast error on a given day should be interpreted as a draw (or realisation) from these PDFs. Clearly, the EDA output is dependent on the assumed observation error covariance matrices used to perturb the system, and it provides a realistic estimate of the uncertainty of the atmospheric state *if, and only if,* these assumed error covariances are realistic.

We also compare the EDA spread, s_{EDA}, on a given pressure level, p (omitted here for clarity of notation), with the corresponding radiosonde departure statistics on that level for the set of experiments, i, outlined above. The standard deviation of the radiosonde minus forecast departures, σ_{o-b}, includes terms related to the forecast error statistics, σ_b, and the standard deviation of radiosonde measurement errors, σ_o, so the variances can be written as,

$$\sigma^2_{o-b} \simeq \sigma^2_b + \sigma^2_o \tag{3}$$

assuming the forecast errors and radiosonde observation errors are not correlated.

The radiosonde observation error statistics, σ_o, should not depend on the number of GNSS-RO data assimilated, and the EDA is providing only an estimate of σ_b. This suggests modeling the variation of the σ^2_{o-b} values from the set experiments with different GNSS-RO numbers with

$$\sigma^2_{o-b} = m \times s^2_{EDA} + c \tag{4}$$

where m is scaling factor and c should be related to variance of the radiosonde measurement errors, and it can be compared with other estimates [e.g., Desroziers et al., 2005].

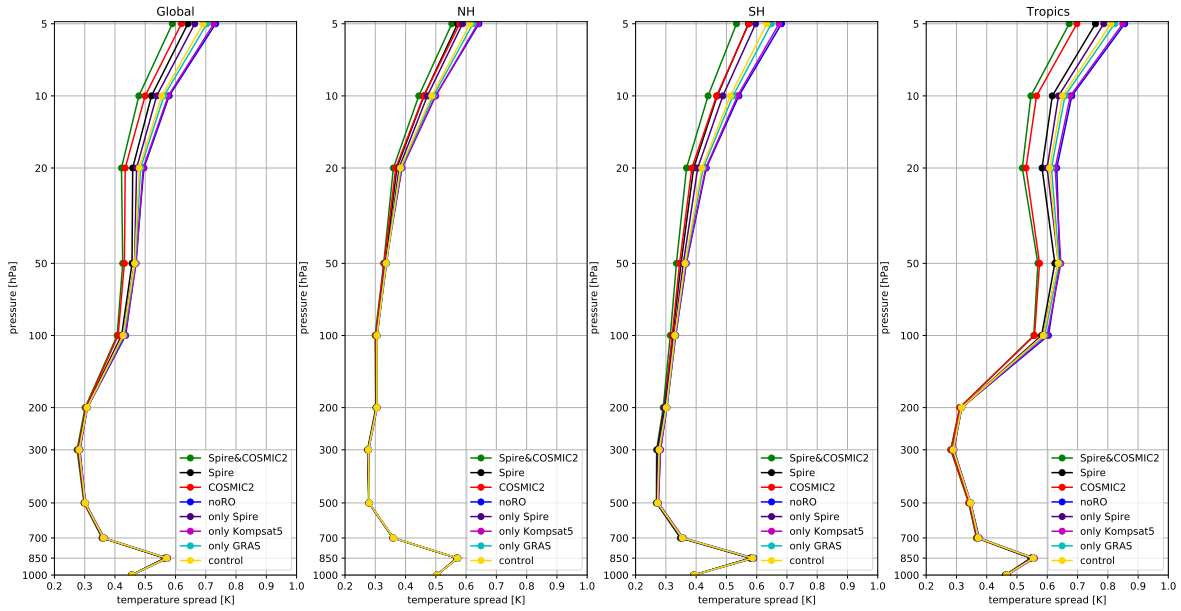
7.3 Results

The vertical distribution of the EDA spread values for temperature values are shown in Fig. 41, with a decrease in EDA spread values with the addition of more GNSS-RO data. This is in particular true for higher levels of 300 hPa (about 9 km) and higher, where the reduction in spread becomes larger with height. For example, at 10 hPa (about 31.5 km) the spread in the tropical temperature values is reduced by about 14% for **Spire+COSMIC2** compared to **CONTROL**.

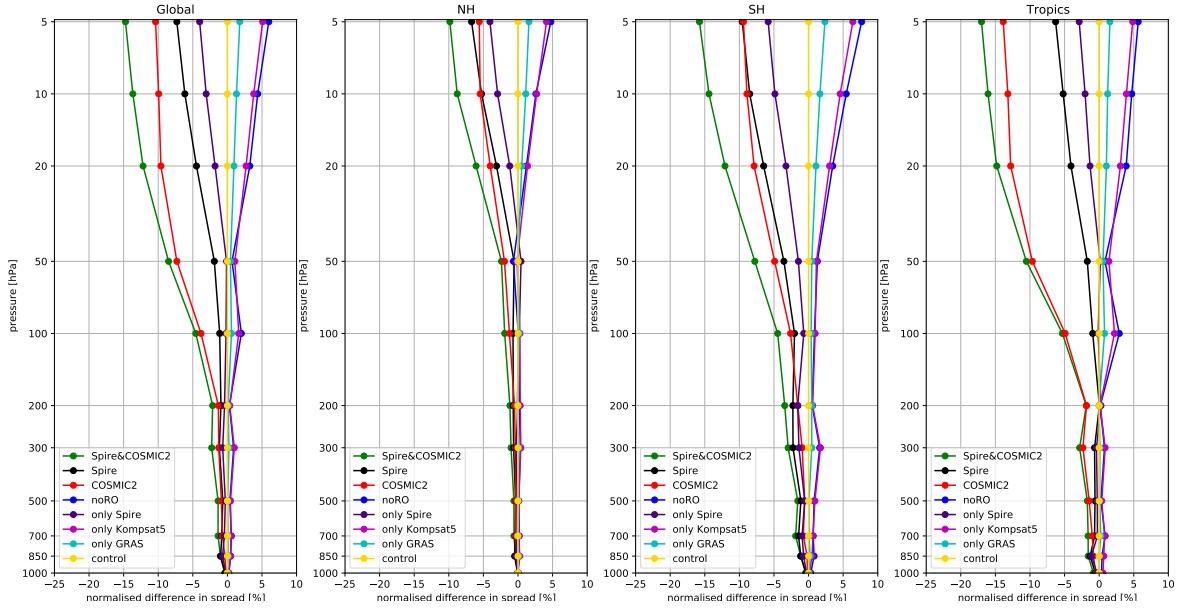
How the spread values for temperature at 10 hPa and 100 hPa change in more detail can be seen in Fig. 42 and Fig. 43, respectively. The largest differences occur in the Tropics, where most the COSMIC-2 data is added (see section 2). At 10 hPa, locally more than a 30% reduction in spread can be seen with the addition of Spire and COSMIC-2 data. This reduction is less strong in lower height levels, as shown before. Already at 100 hPa (see Fig. 43) the reduction is more centred along the Intertropical Convergence Zone (ITCZ) $\pm 20^\circ$ but reduced to around 20% and less locally. Here, the larger spread values could be linked to the variability in height of the tropical tropopause layer at this height level in the austral summer.

To shed a bit more light on this topic, we investigated how exactly the EDA spread values for temperature reduce as a function of global GNSS-RO occultation numbers, shown in Fig. 45. In general, the EDA spread reduces with an increase of GNSS-RO profiles assimilated within the 12-hour assimilation window. The biggest decrease is in the Tropics at 100 hPa (Fig. 45c), whereas the smallest decrease is in the Northern Hemisphere at lower levels (e.g. Fig. 45f). This might be not surprising as it could be deduced from Fig. 41. However, comparing this decrease in EDA spread in temperature at 100 hPa with the equivalent Fig. 9b by Harnisch et al. [2013] (see Fig. 44) for simulated GNSS-RO profiles indicates some similarities and differences. The reduction in the Tropics is about 7% in their study (using a 24h forecast time) and 8% in this study (using a 12-hour forecast time) for global daily GNSS-RO profiles of around 6000 (3000 per 12-hour assimilation cycle). However, in the Northern Hemisphere, Harnisch et al. [2013] have a reduction of 10% and we have 1%. Also, it can be seen that the addition of COSMIC-2 mostly reduces the spread values compared to the other instruments, especially in the Tropics at 100 hPa. The numbers in both studies differ in the extra-tropics but show similarities in the Tropics. The main reason for these differences are probably the heterogeneous distribution of real GNSS-RO data in this study and using a newer model setup (cycle 47R1) compared to Harnisch et al. [2013]. Also, differences in the assigned observation errors, experimentation time and number of ensembles could also play a role.

The next interesting question is how this decrease in EDA spread scales with the statistics of forecast errors from the corresponding OSEs. To evaluate forecast errors, we use a forecast lead time of 12 hours and verify against ECMWF operational analysis. Fig. 46 shows the vertical distribution of EDA values for temperature in the Tropics and the corresponding standard deviation of forecast errors for **noRO** and



(a) Total EDA spread values



(b) Normalised EDA spread values

Figure 41: Vertical profiles of EDA spread values for temperature. a) Total spread values and b) normalised by **CONTROL** experiment.

Spire+COSMIC2. In general, the EDA spread in temperature behaves rather similar to the standard deviation in forecast errors with larger spread values in the stratosphere (Fig. 46a). Relatively, the spread values and forecast errors decrease for **Spire+COSMIC2** compared to **noRO** (Fig. 46b), with a larger decrease for the EDA experiments. However, we have to keep in mind that the forecast errors shown here are from fields truncated at T42 (2.5 deg), meaning they are not directly comparable to the EDA spread values derived from temperature fields truncated at T255. Some smaller scales will be missing for the

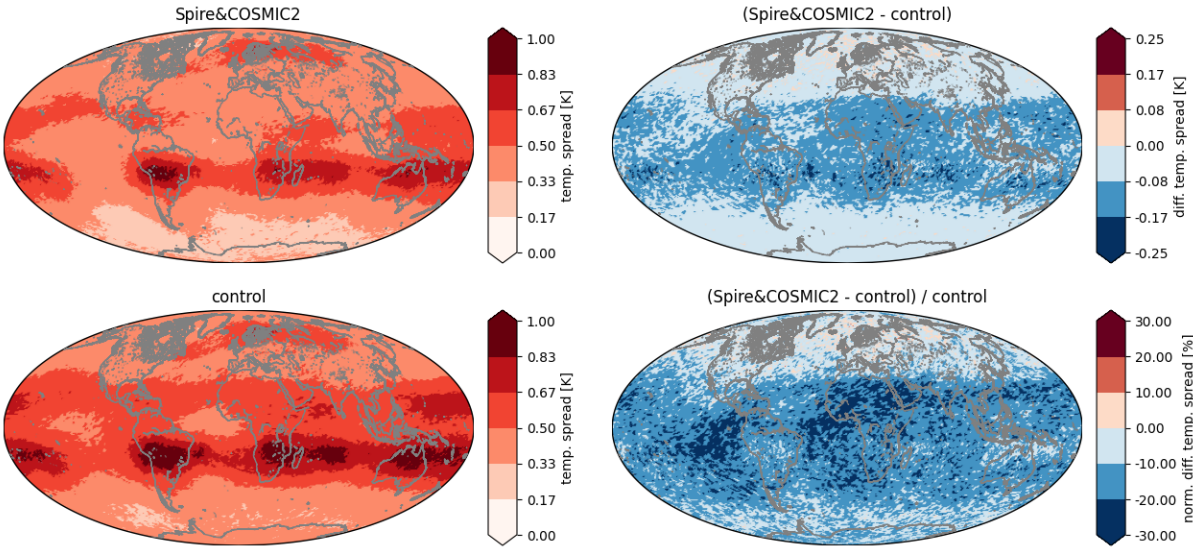


Figure 42: Geographical distribution of temperature spread values at 10 hPa. Top left: for **Spire+COSMIC2**, bottom left: for **CONTROL**, top right: difference between **Spire+COSMIC2** and **CONTROL**, and bottom right: normalised difference between **Spire+COSMIC2** and **CONTROL** (normalised by **noRO**).

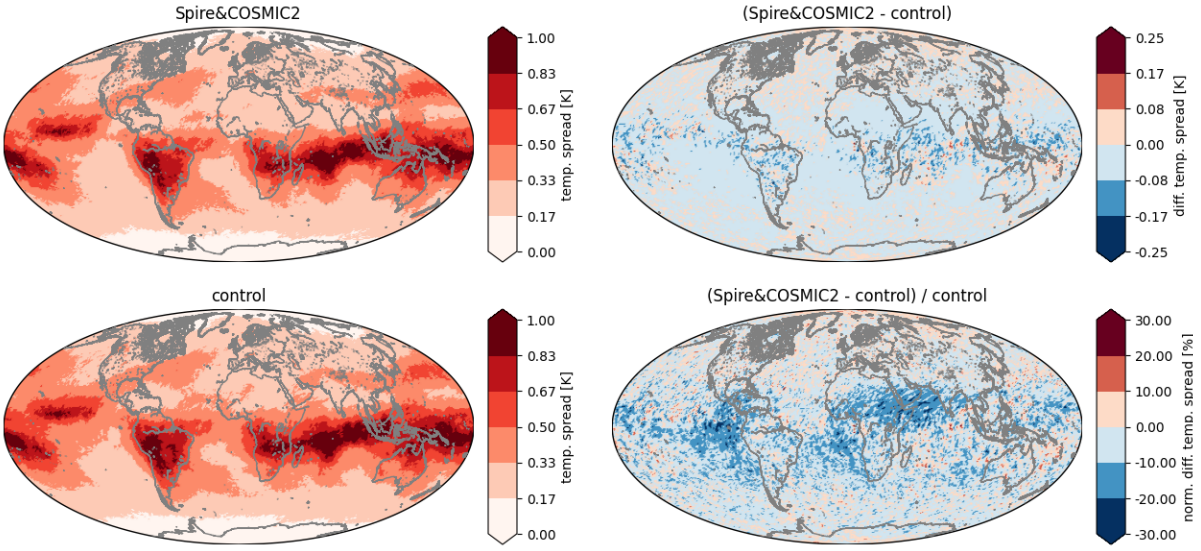


Figure 43: Same as Fig. 42 but at 100 hPa.

forecast errors which are seen in the EDA spread values.

Another limitation in the use of very short forecast lead times e.g., T+12 h, is the use of analyses for the evaluation of forecast errors. Here, the errors in the analysis and forecast can have comparable magnitudes and might be as well highly correlated [Bormann et al., 2019]. This can be compensated by using observations as a reference, instead. For example, we can compare the EDA spread on a given height level with the corresponding radiosonde departure statistics on that level for the set of experiments outlined above. The variance of the radiosonde minus forecast departures is the sum of the true forecast error and radiosonde measurement error variances, as shown in equation 3. This suggests the EDA

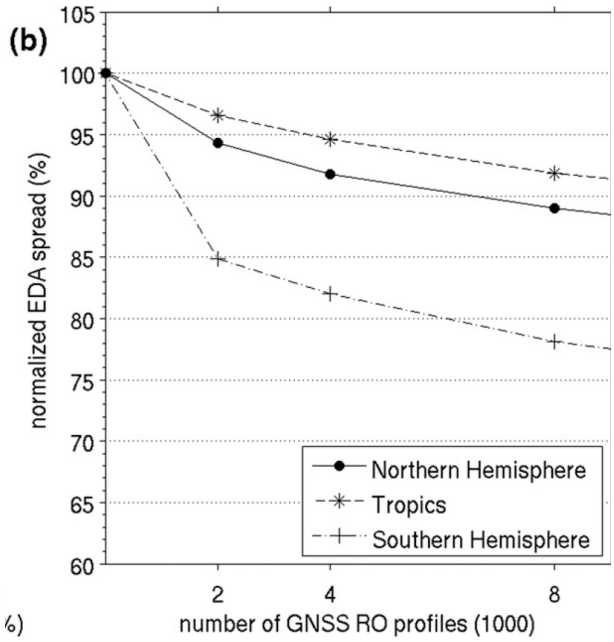


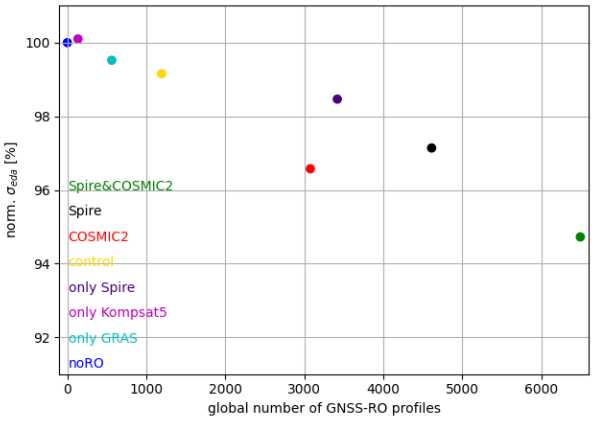
Figure 44: Figure 9b from Harnisch et al. [2013]: normalized EDA spread (%) at 24-h forecast time as a function of the assimilated number of simulated GNSS RO profiles. Results are for temperature at 100 hPa for the period 8 Jul–15 Aug 2008.

spread, as a measure of forecast error statistics, can be related to first guess departures with the use of estimates of the radiosonde measurement errors (e.g., Desroziers et al. [2005]).

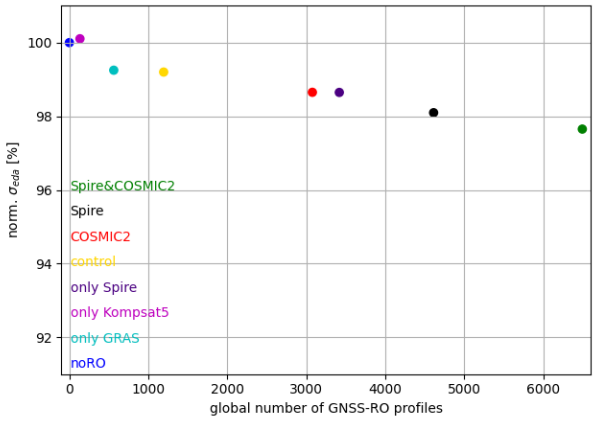
Examples of how the variance in FG departure scales with the EDA variance when GNSS-RO data is added for radiosonde temperature measurements at 100 hPa and 500 hPa are shown in Fig. 47. In the Tropics at 100 hPa (Fig.47c) a linear relationship between the two measures can be seen which means that a fixed decrease in the EDA spread^2 gives a fixed (but different magnitude) decrease in the variance of the FG departures for the OSEs. However, this linear relationship cannot be seen in all geographical areas and altitudes as for example at 100 hPa in the Northern Hemisphere (Fig. 47e). The reason why we see such behaviour can be manifold. For example, the distribution of GNSS-RO data, the location and number of radiosonde observations and maybe the scales included in the shown truncated EDA values could be a reason for this behaviour. Also, using estimated observation errors according to Desroziers et al. [2005], includes the assumption that the background errors in the OSEs are the same – which is not the case.

In the Tropics, it can been seen that R^2, the coefficient of determination, is quite high with around 0.90 to 0.99, meaning the best fit can actually be seen here between s^2_{EDA} and σ^2_{o-b}, especially at 100 hPa (see Fig. 47). Furthermore, the calculated slope is quite similar with about 5 in the Tropics for both shown height levels. In the Southern and Northern Hemisphere R^2 is significantly lower than in the Tropics. We believe this might be related to the global distribution of GNSS-RO data, as COSMIC-2 data is located between 40°S and 40°N. That means in turn the calculated spread values for the outer-tropics might be different to an ideal scenario where all the GNSS-RO data would be distributed more homogeneously. Also, the distribution of the radiosondes (as shown in Fig. 48) are not evenly distributed globally.

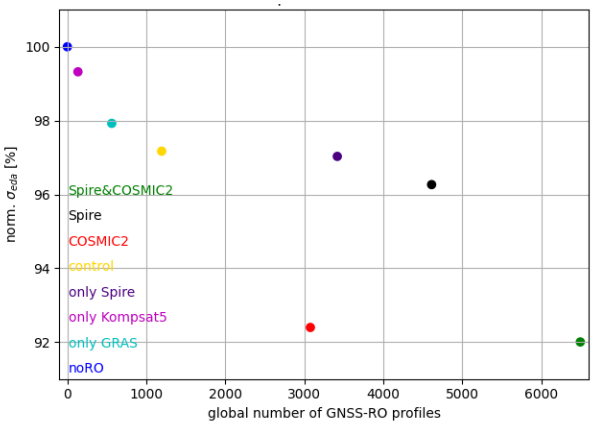
Furthermore, the shown first guess departures are calculated at full resolution in the trajectory, so for our model setup having a resolution of TCo399 this corresponds to T399 truncation for the forecast error part. For our EDA experiments, we retrieve the temperature fields of all ensemble members with a T255



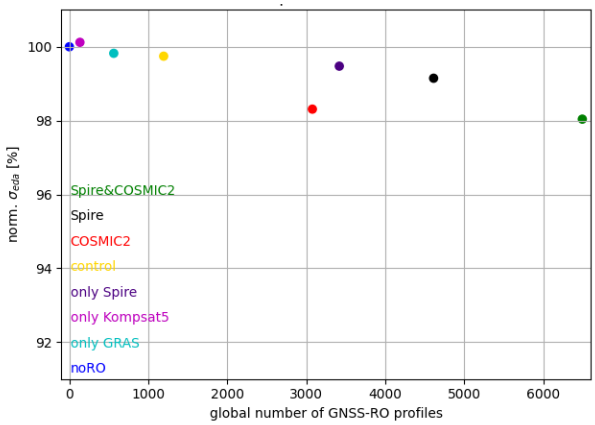
(a) Southern Hemisphere, 100 hPa



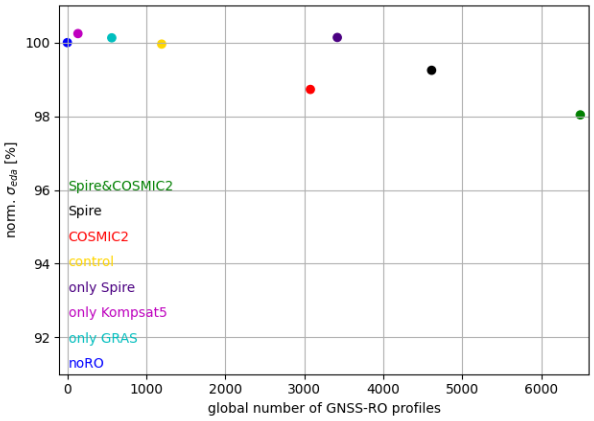
(b) Southern Hemisphere, 500 hPa



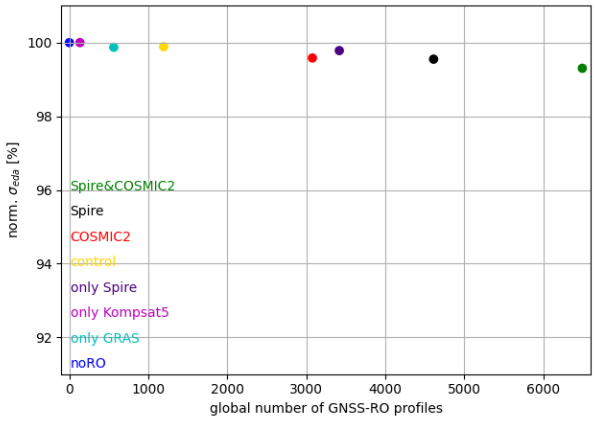
(c) Tropics, 100 hPa



(d) Tropics, 500 hPa

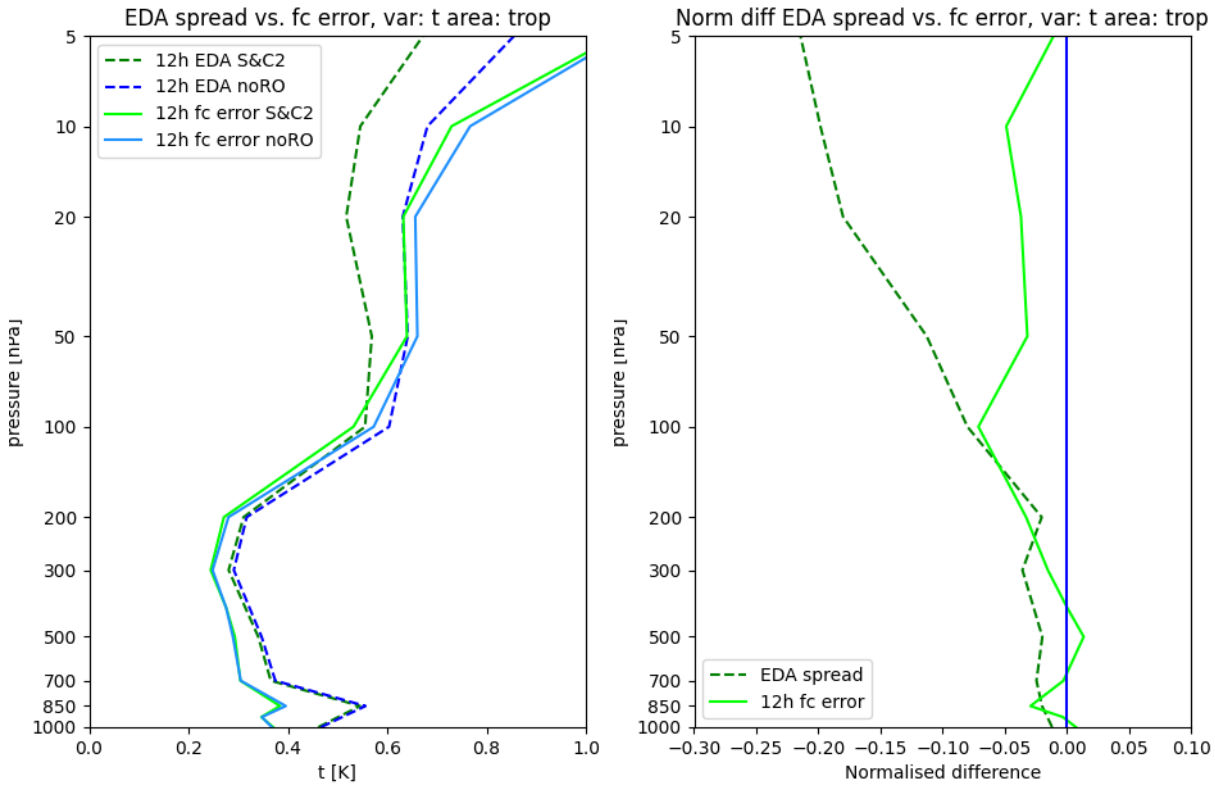


(e) Northern Hemisphere, 100 hPa



(f) Northern Hemisphere, 500 hPa

Figure 45: Normalised EDA spread for temperature as a function of average number of assimilated GNSS-RO occultations per 12-hour assimilation window at 100 hPa and 500 hPa in the Tropics and, Southern and Northern Hemisphere. Different coloured dots refer to different experiments.



(a) Absolute values

(b) Normalised values

Figure 46: Vertical profiles of EDA spread and standard deviation of the forecast error of tropical temperature values for **Spire+COSMIC2** and **noRO**. Operational analyses are used as a reference for calculation of the forecast error. a) absolute values and b) normalised difference between **Spire+COSMIC2** and **noRO** (normalised by **noRO**).

truncation. That means, we cannot directly use the EDA spread as a measure of forecast errors used in the OSEs. But first we have to obtain the EDA spread which is actually used by the background error covariance matrix:

$$s_{used}(lev) = SES_{T159}(lev) \times \frac{\bar{s}_{T255}(lev)}{\bar{s}_{T159}(lev)} \times rednmc \times \sqrt{2}, \tag{5}$$

with SES being the post-processed EDA spread with e.g. adding climatology, $\bar{s}$ being the global averaged spread, and *lev*, being a specific height/pressure level. Effectively, equation 5 is the EDA spread, we have calculated from T255 truncated fields times $rednmc \times \sqrt{2}$, but may have slightly different structure. Here, $\sqrt{2}$ comes from how the used background error statistic is calculated from differences of forecasts (= twice the variance) and $rednmc = 0.95$ is a global tuning parameter. The second step is to acknowledge that calculating the first guess departures at outer loop resolution, we have full resolution for the observations. That means if we use the calculated first guess departure and assume the observation errors known, we would not get the used background error but something larger because the forecast is at T399 and the EDA spread^2 is at T255 truncation. However, one can compensate for this difference in global sense by using the ratio of T399 and T255 spread. For our purpose that would mean to multiply the ensemble spread (using T255 truncated forecasts) by $rednmc \times \sqrt{2} = 0.95 \times \sqrt{2} = 1.34$ times the ratio

of EDA spread at T399 and T255 to compare with the used background error that comes out from the first guess departure calculation.

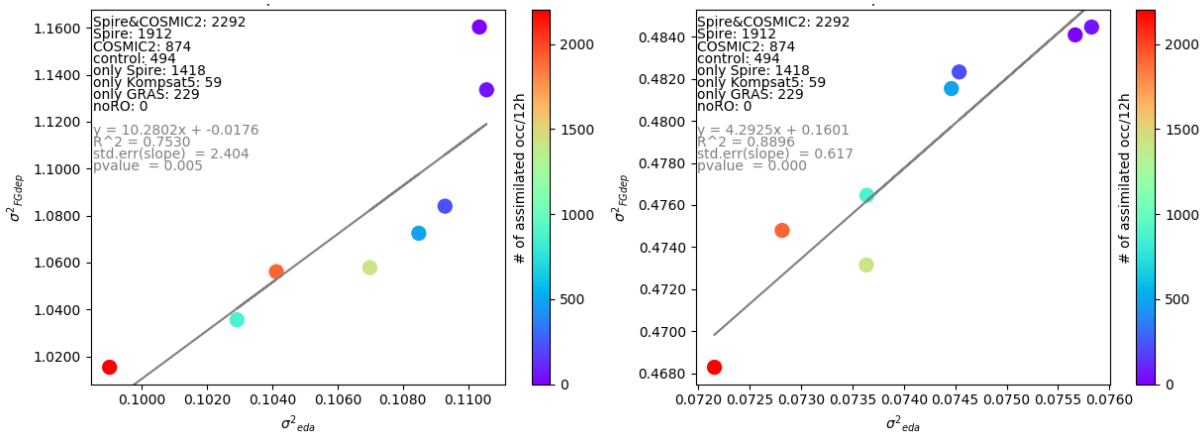
Another factor probably lies in the source of spread: model error perturbations versus observation perturbations. For example, it could be that the observations errors are quite sensitive and, hence, could effect the spread. More research is needed to fully understand the details.

In Fig. 47 the calculated fits according to formula 4 are shown. Here, y equals σ^2_{o-b} and x equals s^2_{EDA}. For the different regions and different height levels the slope varies between 3 and 10. Taking account for the different resolutions as just discussed for s^2_{EDA} and σ^2_{o-b}, the slope computed in Fig. 47 is actually $\widetilde{m} = (\sqrt{2} \times 0.95)^2 \frac{s^2_{T399}}{s^2_{T255}} \times m = 1.805 \frac{s^2_{T399}}{s^2_{T255}} \times m$, meaning formula 4 revises to:

$$\sigma^2_{o-b} = m \times 1.805 \frac{s^2_{T399}}{s^2_{T255}} \times s^2_{EDA} + c. \tag{6}$$

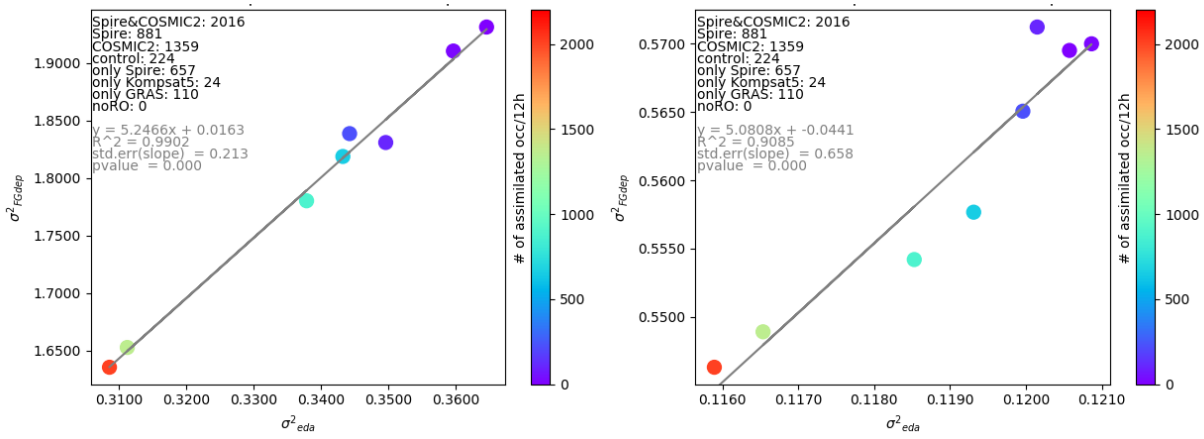
7.4 Summary

The reductions in EDA spread when adding Spire and COSMIC-2 data are the largest in the stratosphere and are qualitatively consistent with reductions in forecast error statistics. Also, results show that this is partially true when EDA statistics are evaluated against radiosonde observations. The challenges when studying ensemble spread values and comparing them with forecast error statistics or observations are numerous. For example, one must be fully aware that for the EDA experiments the variability of the perturbations does not grow sufficiently through the forecast (under-dispersive) in some regions and height levels. This means the EDA can underestimate the impact of the addition of GNSS-RO data in these areas. Furthermore, the evaluation of forecast error statistics depends on the choice of analysis as a reference, which has limitations as discussed. Also, the model resolution of the experiments does matter for which scales can be captured at the various height levels. Nevertheless, in the Tropics where most of the GNSS-RO data is located a nice linear relationship between EDA spread^2 and variance in first guess departures could seen at higher altitudes. Here, s^2_{EDA} and σ^2_{o-b} can be used to see the effect from adding GNSS-RO data which shows a reduction of in their values. The order of reduction in the Tropics is similar to the one found by Harnisch et al. [2013] for simulated data. Also, it could be seen that s^2_{EDA} can be used to estimate background errors.



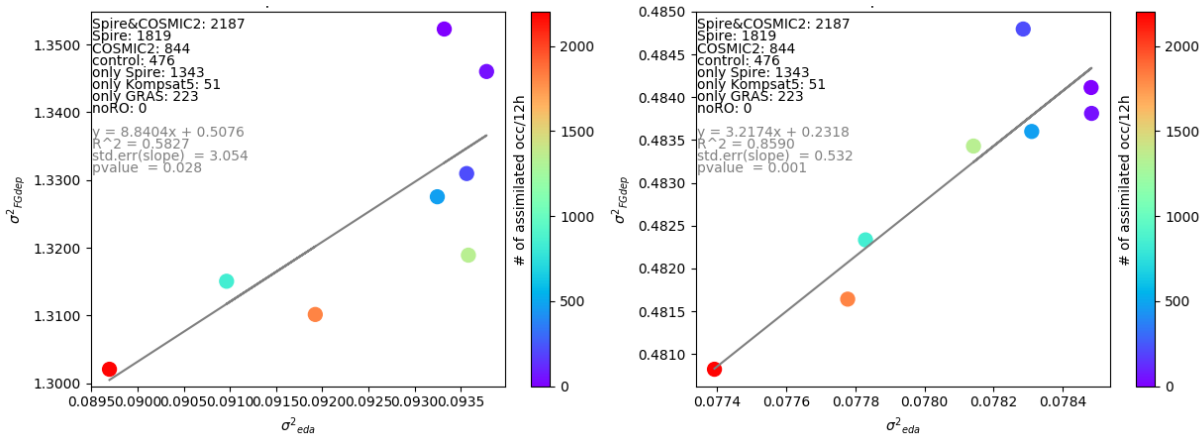
(a) Southern Hemisphere, 100 hPa

(b) Southern Hemisphere, 500 hPa



(c) Tropics, 100 hPa

(d) Tropics, 500 hPa



(e) Northern Hemisphere, 100 hPa

(f) Northern Hemisphere, 500 hPa

Figure 47: Change in variance of FG departure from OSEs versus EDA variance for radiosonde temperature measurements at 100 hPa and 500 hPa in the Tropics and, Southern and Northern Hemisphere. Different coloured dots refer to number of GNSS-RO occultations assimilated on average per 12 hour-cycle in certain region.

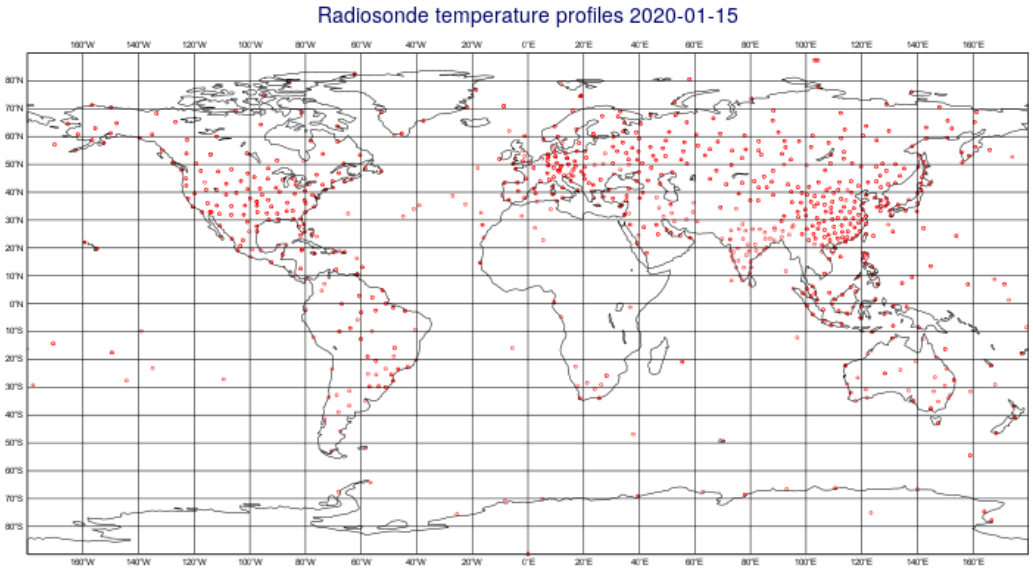


Figure 48: Location of radiosonde data for a 24 hour period (21Z/21Z) on 15 January 2020.

8 Summary

GNSS-RO measurements have a significant impact in numerical weather prediction (NWP). Unlike other instrumentation (e.g. mircowave imagers and sounders), GNSS-RO measurements do not require bias correction to the NWP model. In addition, GNSS-RO complements other measurement techniques by providing high quality vertical profiles of the atmosphere, in particular temperature information in the upper-troposphere and stratosphere.

The addition of Spire and COSMIC-2 data is a large increase in the GNSS-RO data being assimilated at ECMWF and Met Office NWP systems. With this significant increase of radio occultation data we have been provided with an opportunity to investigate the effect of this dataset on short- and medium-range weather forecast scores. Furthermore, this extensive GNSS-RO dataset also provides the opportunity to investigate how the impact of GNSS-RO data scales with observation number, and compare with the simulation results of Harnisch et al. [2013], which were part of an earlier ESA study, and have been used to justify the case for more GNSS-RO.

In a joint study Met Office and ECMWF have investigated the effect of Spire and COSMIC-2 data running Observing System Experiments (OSEs) in their state-of-the-art NWP systems. Both institutions have found that the addition of COSMIC-2 and Spire observations is beneficial, with adding Spire on top of COSMIC-2 showing further improvements for temperature, humidity and wind and some geographical areas. For example, for radiosonde temperature measurements the addition of Spire and COSMIC-2 data improves short-range forecast scores up to 2% at 150 hPa compared to a control run. It was also seen that the largest impact happens at higher altitudes, which is not surprising as the core region of radio occultation (RO) measurements is between 8 km and 30 km height. In general, improved medium-range forecast scores and improvement in short-range forecasts can be seen for temperature, humidity and wind but having different magnitudes at the Met Office and ECMWF system. For example, fits to radiosonde temperature observations in the Tropics show improvements of about 15% at 100 hPa when adding Spire and COSMIC-2 for Met Office and about 11% for ECMWF. Both centres show the largest positive impact in temperature for medium-range forecast scores in the Southern Hemisphere through the addition of Spire and COSMIC-2. For humidity, the biggest impact can be seen through the addition of COSMIC-2 in the Tropics at ECMWF. Furthermore, the positive impact for wind when adding Spire and COSMIC-2 could be seen in medium-range forecast scores for both centres. In conclusion, Met Office and ECMWF would assimilate the Spire data operationally if it becomes available.

Additionally, ECMWF have compared the impact of processing centre on the Spire impact, by comparing bending angles provided directly by Spire with Spire bending angles produced at EUMETSAT, updating results from a previous ESA study assessing the Spire RO data quality. The reprocessed data was assimilated into ECWMF, showing similar and sometimes even better improvements in short- and medium-range forecasts than the original Spire dataset.

We also performed a set Ensembles of Data Assimilation (EDA) experiments corresponding to the OSEs. The EDA configuration we used in this study consists of ten independent cycling assimilation systems, where observations (incl. soil moisture), sea surface temperature, and model physics are perturbed. If those perturbations are correctly specified, the EDA will provide good estimates of analysis and short-range forecast error uncertainties. Harnisch et al. [2013] used the EDA spread values (standard deviation amongst ensemble members) to estimate how the impact of GNSS-RO data scales with observation number using simulated GNSS-RO data. Here, we investigate how the relationship between EDA spread and the forecast error statistics changes as the number of real GNSS-RO measurements assimilated increases. We have found that increasing the number of GNSS-RO data reduces the ensemble spread, especially at 100 hPa and higher (in altitude) and in the Tropics where most of the GNSS-RO data is located. At

100 hPa in the Tropics the reduction of EDA spread is in the same order as to the one found by Harnisch et al. [2013] for simulated data. Here, the reduction is about 7% in their study and 8% in this study for global GNSS-RO profiles of around 6000 during a day. Also, we find a linear relationship between EDA spread and variance in first guess departures of radiosonde temperature observations in the Tropics. This proves that the observation impact for GNSS-RO data based on the EDA spread generally agrees with the impact found in OSEs, at least for the higher altitudes in the Tropics. For the Northern and Southern Hemisphere the decrease in EDA spread is less strong when more GNSS-RO data has been added, which might be linked inter alia to the heterogeneously distributed GNSS-RO data in the outer-tropics. Still, the EDA spread values provide a reliable tool to assess the observation impact of GNSS-RO data in the higher troposphere and stratosphere.

A International Civil Aviation Organization (ICAO) Standard Atmosphere

Altitude [km]	Pressure [hPa]
0	1,013
1	899
2	795
5	540
8	356
10	264
15	120
20	55
30	11

Table 1: Selected altitudes with the approximate pressure level for ICAO Standard Atmosphere.

Pressure [hPa]	Altitude [km]
10	31.5
50	20.5
100	16.0
200	11.8
500	5.6
700	3.0
850	1.5
1000	0.1

Table 2: Selected pressure levels with the approximate altitudes for ICAO Standard Atmosphere.

B Sensitivities of satellite observations

In Fig. 49 and Fig.50 the Jacobians of AMSU-A and SSMIS are shown, which are representative for a standard atmosphere under clear sky conditions. The Jacobian (or weighting function) gives you an idea at which height level the satellite instrument is most sensitive. Note that for humidity-sensitive observations there will be considerable variation in the height of the Jacobian peaks depending on the humidity in the atmosphere, with peaks higher in the atmosphere in moist atmospheres, and lower in dry atmospheres.

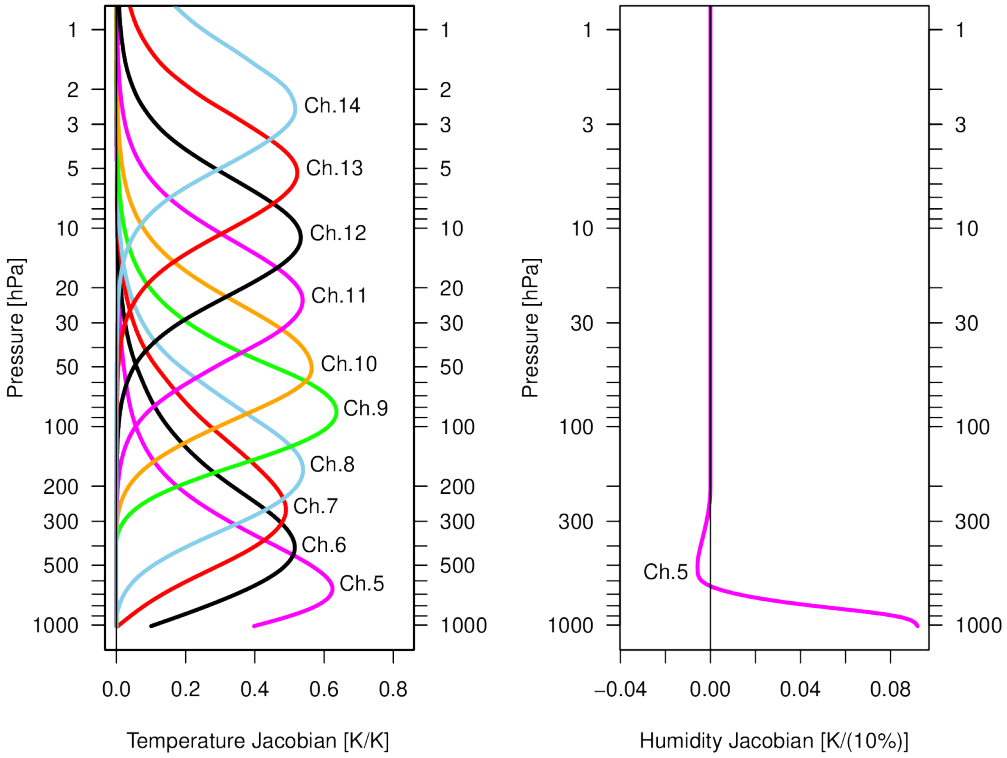
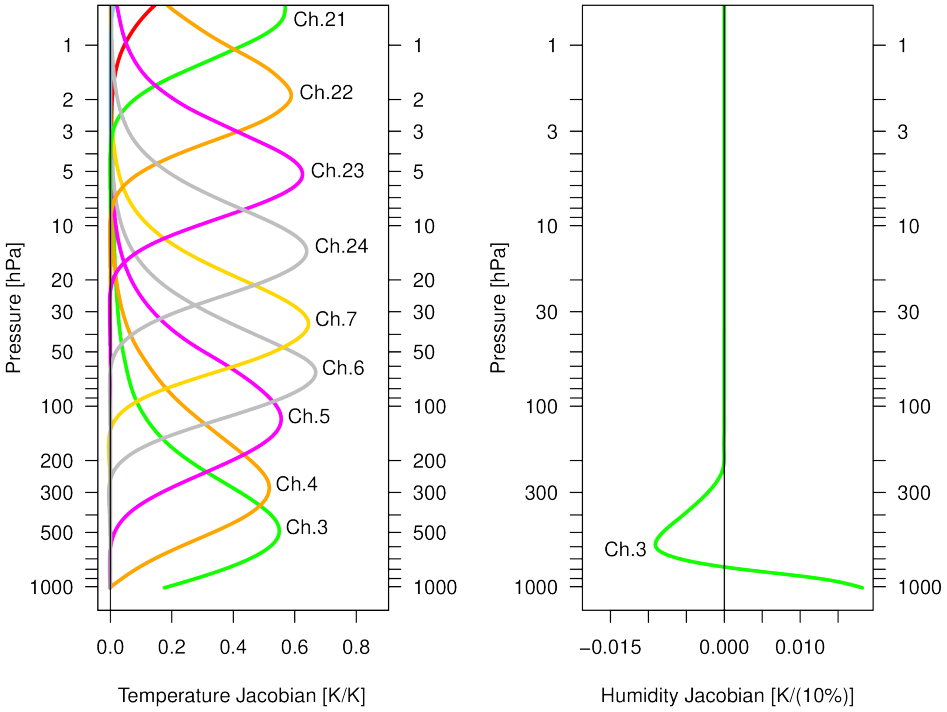
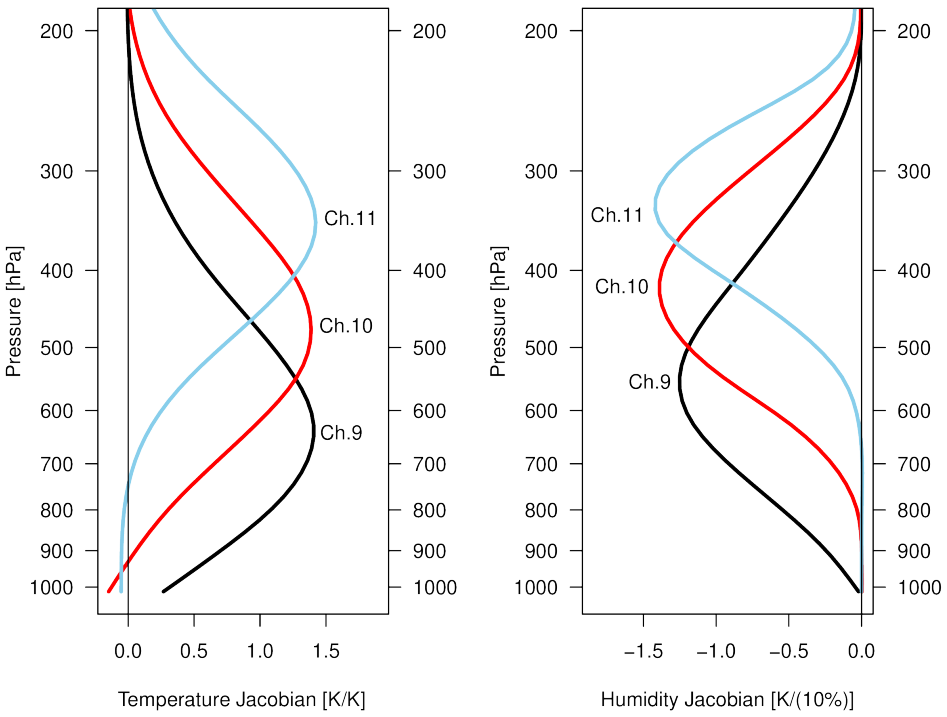


Figure 49: Jacobians for selected channels of AMSU-A.



(a) "Temperature channels"



(b) "Humidity channels"

Figure 50: Jacobians for selected channels of SSMIS.

C Theoretical Justification of EDA technique

This appendix outlines the theoretical justification of the EDA technique given in Žagar et al. [2005] and Isaksen et al. [2010]. It is also reproduced in Appendix A of Harnisch et al. [2013].

Consider the general linear analysis-forecast system

$$\mathbf{x}_a^k = \mathbf{x}_b^k + \mathbf{K}_k \left(\mathbf{y}^k - \mathbf{H}_k \mathbf{x}_b^k \right) \tag{7}$$

$$\mathbf{x}_b^{k+1} = \mathbf{M}_k \mathbf{x}_a^k \tag{8}$$

where k denotes the analysis time, $\mathbf{y}^k$ the observation vector, $\mathbf{x}_b^k$ the model background state and $\mathbf{x}_a^k$ the model analysis state. $\mathbf{M}_k$ and $\mathbf{K}_k$ are matrices describing the model forecast and a general gain matrix, not specifically the Kalman gain.

Assuming that observation, background and model errors are independent of each other, the analysis and background error covariance matrices, $\mathbf{P}_k^a$ and $\mathbf{P}_k^b$, respectively, can be shown to evolve according to

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^b (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \tag{9}$$

$$\mathbf{P}_{k+1}^b = \mathbf{M}_k \mathbf{P}_k^a \mathbf{M}_k^T + \mathbf{Q}_k \tag{10}$$

where $\mathbf{R}_k$ is the covariance matrix for the observation error and $\mathbf{Q}_k$ for the model error.

For a perturbed analysis-forecast system using perturbed observations and a stochastic forcing term that represents model errors, the general equations (Eqs. (7) and (8)), can be written as

$$\tilde{\mathbf{x}}_a^k = \tilde{\mathbf{x}}_b^k + \mathbf{K}_k \left(\mathbf{y}^k + \eta^k - \mathbf{H}_k \tilde{\mathbf{x}}_b^k \right) \tag{11}$$

$$\tilde{\mathbf{x}}_b^{k+1} = \mathbf{M}_k \tilde{\mathbf{x}}_a^k + \zeta^k \tag{12}$$

where $\tilde{\mathbf{x}}_a^k$ and $\tilde{\mathbf{x}}_b^{k+1}$ are the perturbed analysis and background states, respectively.

The differences between the perturbed and unperturbed states leads to the equations for the perturbations

$$\varepsilon_a^k = \tilde{\mathbf{x}}_a^k - \mathbf{x}_a^k = \varepsilon_b^k + \mathbf{K}_k \left(\eta^k - \mathbf{H}_k \varepsilon_b^k \right) \tag{13}$$

$$\varepsilon_b^{k+1} = \tilde{\mathbf{x}}_b^{k+1} - \mathbf{x}_b^{k+1} = \mathbf{M}_k \varepsilon_a^k + \zeta^k \tag{14}$$

which implies that the perturbations evolve with the same update equations and matrices as in the general unperturbed case (Eqs. (7) and (8)).

The covariance matrices for the perturbations of the analysis and background state are given by

$$\overline{\varepsilon_a^k \left(\varepsilon_a^k\right)^T} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \overline{\varepsilon_b^k \left(\varepsilon_b^k\right)^T} (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \tag{15}$$

$$\overline{\varepsilon_b^{k+1} \left(\varepsilon_b^{k+1}\right)^T} = \mathbf{M}_k \overline{\varepsilon_a^k \left(\varepsilon_a^k\right)^T} \mathbf{M}_k^T + \mathbf{Q}_k \tag{16}$$

where the applied perturbations η^k and ζ^k are required to have the error covariance matrices $\mathbf{R}_k$ and $\mathbf{Q}_k$, respectively. From the comparison with Eqs. (9) and (10) it can be seen that if $\overline{\varepsilon_b^k \left(\varepsilon_b^k\right)^T} = \mathbf{P}_k^b$ at some time k, $\overline{\varepsilon_a^k \left(\varepsilon_a^k\right)^T} = \mathbf{P}_k^a$ and $\overline{\varepsilon_b^{k+1} \left(\varepsilon_b^{k+1}\right)^T} = \mathbf{P}_{k+1}^b$. This implies that the error statistics of the analysis-forecast system can be generated from an ensemble of perturbed assimilation cycles. However, it is important to emphasise here that the resulting error statistics will only be realistic estimates of the true error statistics, if all input error covariances are realistic.

D Additional Statistics from EUMETSAT's Spire Processing

Table 3 shows the number of occultations for each Spire satellite during the study period, stratified by rising and setting occultations as well as by the GNSS system from which observations were taken. Percentages refer to each satellite. Numbers are also given for the Metop satellites.

Table 3: Occultation numbers and percentages for the various satellites, stratified by occultation type (i.e., rising and setting) and GNSS system. The last line in each group ("Total") shows the total number of occultations for the Spire and Metop satellites, respectively. Percentages are rounded to a single decimal; empty cells denote missing (i.e., zero) occultation numbers.

		Type				GNSS System							
Satellite	Number	Rising		Setting		GPS		Galileo		GLONASS		QZSS	
ID / Name	#	#	%	#	%	#	%	#	%	#	%	#	%
Spire													
FM046	20			20	100.0	20	100.0						
FM075	1785	845	47.3	940	52.7	693	38.8	561	31.4	473	26.5	58	3.2
FM079	12581	7564	60.1	5017	39.9	5039	40.1	3530	28.1	3262	25.9	750	6.0
FM080	5490	2467	44.9	3023	55.1	2208	40.2	1693	30.8	1252	22.8	337	6.1
FM084	12467	6406	51.4	6061	48.6	5093	40.9	3346	26.8	3133	25.1	895	7.2
FM085	15107	7312	48.4	7795	51.6	7343	48.6	3507	23.2	3154	20.9	1103	7.3
FM088	7123	3434	48.2	3689	51.8	3525	49.5	1813	25.5	1324	18.6	461	6.5
FM090	19270	6531	33.9	12739	66.1	7822	40.6	6154	31.9	4291	22.3	1003	5.2
FM099	50706	24613	48.5	26093	51.5	20551	40.5	13846	27.3	13858	27.3	2451	4.8
FM100	32424	16455	50.7	15969	49.3	13122	40.5	10699	33.0	6823	21.0	1780	5.5
FM101	56619	28253	49.9	28366	50.1	23164	40.9	18732	33.1	11611	20.5	3112	5.5
FM102	53273	26428	49.6	26845	50.4	21390	40.2	15408	28.9	13541	25.4	2934	5.5
FM103	56038	27985	49.9	28053	50.1	23257	41.5	15236	27.2	14493	25.9	3052	5.4
FM104	48026	22316	46.5	25710	53.5	20323	42.3	7591	15.8	17353	36.1	2759	5.7
FM105	60577	30050	49.6	30527	50.4	25085	41.4	17473	28.8	14859	24.5	3160	5.2
FM106	57063	25625	44.9	31438	55.1	26077	45.7	10591	18.6	17151	30.1	3244	5.7
FM107	58175	29167	50.1	29008	49.9	24539	42.2	18109	31.1	12481	21.5	3046	5.2
FM108	65963	33269	50.4	32694	49.6	26846	40.7	18674	28.3	17111	25.9	3332	5.1
FM113	22427	2502	11.2	19925	88.8	9297	41.5	5729	25.5	6271	28.0	1130	5.0
Total	635134	301222	47.4	333912	52.6	265394	41.8	172692	27.2	162441	25.6	34607	5.4
Metop													
Metop-A	57740	27949	48.4	29791	51.6	57740	100.0						
Metop-B	57545	27643	48.0	29902	52.0	57545	100.0						
Metop-C	57065	27116	47.5	29949	52.5	57065	100.0						
Total	172350	82708	48.0	89642	52.0	172350	100.0						

Acknowledgements

We acknowledge Spire for making their data available quickly at the start of this project. The authors also thank Elias Hòlm and Simon Lang for discussion on aspects of the EDA, and Niels Bormann for the figures showing the weighting functions. Also, we thank Bruce Ingleby for providing maps of aircraft and radiosonde data.

References

R. A. Anthes, P. A. Bernhardt, Y. Chen, L. Cucurull, K. F. Dymond, D. Ector, S. B. Healy, S.-P. Ho, D. C. Kuo, Y.-H. Hunt, H. Liu, K. Manning, C. McCormick, T. K. Meehan, W. J. Randel, C. Rocken, W. S. Schreiner, S. V. Sokolovskiy, S. Syndergaard, D. C. Thompson, K. E. Trenberth, T.-K. Wee, N. L. Yen, and Z. Zeng. The COSMIC/FORMOSAT-3 Mission: Early Results. *Bull. Am. Meteorol. Soc.*, 89:313–333, 2008.

J. M. Aparicio and G. Deblonde. Impact of the assimilation of CHAMP refractivity profiles in Environment Canada global forecasts. *Mon. Wea. Rev.*, 136:257–275, 2008.

M. Bonavita, L. Raynaud, and L. Isaksen. Estimating background-error variances with the ECMWF Ensemble of Data Assimilations system: some effects of ensemble size and day-to-day variability. *Q. J. R. Meteorol. Soc.*, 137:423–434, 2011.

M. Bonavita, L., and E. Holm. On the use of EDA background error variances in the ECMWF 4D-Var. Technical Memorandum 664, ECMWF, 2012.

N. Bormann, H. Lawrence, and J. Farnan. Global observing system experiments in the ECMWF assimilation system. Technical Memorandum 839, ECMWF, 2019.

F. Bouttier and G. Kelly. Observing-system experiments in the ECMWF 4D-Var data assimilation system. *Q.J.R. Meteorol. Soc.*, 127:1469—-1488, 2001.

N. E. Bowler. An assessment of GNSS radio occultation data produced by Spire. *Q.J.R. Meteorol. Soc.*, 146:3772–3788, 2020.

N. E. Bowler. Revised GNSS-RO observation uncertainties in the Met Office NWP system. *Q.J.R. Meteorol. Soc.*, 146(730):2274–2296, 2020. doi: {10.1002/qj.3791}.

A. M. Clayton, A. C. Lorenc, and D. M. Barker. Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Q.J.R. Meteorol. Soc.*, 139:1445–1461, 2013.

P Craven and G Wahba. Smoothing noisy data with spline functions. *Numer. Math.*, 31:377–403, 1979.

L. Cucurull, J. C. Derber, R. Treadon, and R. J. Purser. Assimilation of Global Positioning System Radio Occultation Observations into NCEP's Global Data Assimilation System. *Mon. Wea. Rev.*, 135: 3174–3193, 2007.

I. D. Culverwell, H. W. Lewis, D. Offiler, C. Marquardt, and C. P. Burrows. The Radio Occultation Processing Package, ROPP. *Atmospheric Measurement Techniques*, 8(4):1887–1899, 2015. ISSN 1867-8548. doi: 10.5194/amt-8-1887-2015.

T Davies, M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood. A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Q.J.R. Meteorol. Soc.*, 131:1759–1782, 2005.

G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. *Q.J.R. Meteorol. Soc.*, 131:3385–3396, 2005.

A. Geer, P. Bauer, and P. Lopez. Direct 4D-Var assimilation of all-sky radiances. Part II: Assessment. *Q.J.R. Meteorol. Soc.*, 136(652):1886–1905, 2010.

A.J. Geer. Significance of changes in medium-range forecast scores. *Tellus A: Dynamic Meteorology and Oceanography*, 1(68):30229, 2016. doi: 10.3402/tellusa.v68.30229.

M. E. Gorbunov and K. B. Lauritsen. Analysis of wave fields by Fourier integral operators and their application for radio occultations. *Radio Science*, 39(4), August 2004. ISSN 00486604. doi: 10.1029/2003RS002971.

W. Han and N. Bormann. Constrained adaptive bias correction for satellite radiance assimilation in the ECMWF 4D-Var system. Technical Memorandum 783, ECMWF, 2016.

F. Harnisch, S. B. Healy, P. Bauer, and S. J. English. Scaling of GNSS radio occultation impact with observation number using an ensemble of data assimilations. *Mon. Wea. Rev.*, 141:4395–4413, 2013.

S. B. Healy and J.-N. Thépaut. Assimilation experiments with CHAMP GPS radio occultation measurements. *Q. J. R. Meteorol. Soc.*, 132:606–623, 2006.

S-p. B. Ho and X. Zhou. Quality assessment of the commercial weather radio occultation data for GeoOptics and SPIRE and inter-comparison to COSMIC-2. Paper presented at the IROWG-8, 2021.

B. Ingleby, B. Candy, J. Eyre, T. Haiden, C. Hill, L. Isaksen, D. Kleist, F. Smith, P. Steinle, S. Taylor, W. Tennant, and C. Tingwell. The impact of COVID-19 on weather forecasts: A balanced view. *Geogphyscial Research Letters*, 48:e2020GL090699, 2021.

L. Isaksen, M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher, and L. Raynaud. Ensemble of Data Assimilations at ECMWF. Technical Memorandum 636, ECMWF, 2010.

Arne Skov Jensen, M S Lohmann, H-H Benzon, and A S Nielsen. Full Spectrum Inversion of radio occultation signals. *Radio Science*, 38(3), 2003. doi: 10.1029/2002RS002763.

E. R. Kursinski, G. A. Hajj, W. I. Bertiger, S. S. Leroy, T. K. Meehan, L. J. Romans, J. T. Schofield, D. J. McCleese, W. G. Melbourne, C. Thornton, T. P. Yunck, J. R. Eyre, and R. N. Nagatani. Initial results of radio occultation observations of earth's atmosphere using the Global Positioning System. *Science*, 271:1107–1100, 1996.

E. R. Kursinski, G. A. Hajj, J. T. Schofield, R. P. Linfield, and K. R. Hardy. Observing earth's atmosphere with radio occultation measurements using the Global Positioning System. *J. Geophys. Res.*, 102: 23.429–23.465, 1997a.

E R Kursinski, G A Hajj, J T Schofield, R P Linfield, and K R Hardy. Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *J. Geophys. Res.*, 102 (D19):23.429–23.465, 1997b.

A. Lorenc. Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, 112:1177–1194, 1986.

Ricardo A. Maronna, R. Douglas Martin, Victor J. Yohai, and Matías Salibián-Barrera. *Robust Statistics. Theory and Methods*. John Wiley & Sons, Ltd., second edition, 2019.

C. Marquardt, A. von Engeln, A. Nardo, and Y. Andres. EUMETSAT Assessment of Spire Commercial RO Data. Final Report ESA Contract 40001257000, EUMETSAT, 2020.

D. Masters, V. Irisov, V. Nguyen, T. Duly, O. Nogués-Correig, L. Tan, T. Yuasa, J. Ringer, R. Sikarin, M. Gorbunov, and C. Rocken. Status and plans for Spire's growing commercial constellation of GNSS science cubesats. Work presented at the IROWG Workshop 2019, Danish Meteorological Institute, Copenhagen, Elsinore, Denmark, 2019.

S. Migliorini and B. Candy. All-sky satellite data assimilation of microwave temperature sounding channels at the Met Office. *Q.J.R. Meteorol. Soc.*, 145(719, B):867–883, 2019. doi: {10.1002/qj.3470}.

T. Palmer, R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. Shutts, M. Steinheimer, and A. Weisheimer. Stochastic parametrization and model uncertainty. Technical Memorandum 598, ECMWF, 2009.

P. Poli, P. Moll, D. Puech, F. Rabier, and S. B. Healy. Quality control, error analysis, and impact assessment of FORMOSAT-3/COSMIC in numerical weather prediction. *Terr. Atmos. Ocean*, 20: 101–113, 2009.

F. Rabier, H. Jäarvinen, E. Klinker, J.-J. Mahfouf, and A. J. Simmons. The ECMWF operational implementation of four dimensional variational assimilation. Part I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.*, 126:1143–1170, 2000.

F. Rawlins, S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne. The Met Office global four-dimensional variational data assimilation scheme. *Q.J.R. Meteorol. Soc.*, 133:347–362, 2007.

M. P. Rennie. The impact of GPS radio occultation assimilation at the Met Office. *Q.J.R. Meteorol. Soc.*, 136:116–131, 2010.

C. Rocken, R. Anthes, M. Exner, D. Hunt, S. Sokolovsky, R. Ware, M. Gorbunov, W. Schreiner, D. Feng, B. Herman, Y.-H. Kuo, and X. Zou. Analysis and validation of GPS/MET data in the neutral atmosphere. *J. Geophys. Res.*, 102:29.849–29.866, 1997.

W.S. Schreiner, J.P. Weiss, R.A. Anthes, J. Braun, V. Chu, J. Fong, D. Hunt, Y.-H. Kuo, T. Meehan, W. Serafino, J. Sjoberg, S. Sokolovskiy, E. Talaat, T.K. Wee, and Z. Zeng. COSMIC-2 Radio Occultation Constellation: First Results. *Geophysical Research Letters*, 47(4), February 2020. ISSN 0094-8276. doi: 10.1029/2019GL086841.

Sergey V. Sokolovskiy, C. Rocken, W. Schreiner, and D. Hunt. On the uncertainty of radio occultation inversions in the lower troposphere. *Journal of Geophysical Research*, 115(D22):1–19, November 2010. ISSN 0148-0227. doi: 10.1029/2010JD014058.

D. G. H. Tan, E. Andersson, M. Fisher, and I. Isaksen. Observing-system impact assessment using a data assimilation ensemble technique: application to the ADM-Aeolus wind profiling mission. *Q.J.R. Meteorol. Soc.*, 133:381—-390, 2007.

N. Žagar, E. Andersson, and M. Fisher. Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. *Q.J.R. Meteorol. Soc.*, 131:987–1011, 2005.

G Wahba. Smoothing noisy data with spline functions. *Numer. Math.*, 24:383–393, 1975.

D. Walters, I. Boutle, M. Brooks, T. Melvin, R. Stratton, S. Vosper, H. Wells, K. Williams, N. Wood, T. Allen, A. Bushell, D. Copsey, P. Earnshaw, J. Edwards, M. Gross, S. Hardiman, C. Harris, J. Heming, N. Klingaman, R. Levine, J. Manners, G. Martin, S. Milton, M. Mittermaier, C. Morcrette, T. Riddick, M. Roberts, C. Sanchez, P. Selwood, A. Stirling, C. Smith, D. Suri, W. Tennant, P.L. Vidale, J. Wilkinson, M. Willett, S. Woolnough, and P. Xavier. The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. *Geoscientific Model Development*, 10, 2017.

Rand R. Wilcox. *Introduction to Robust Estimation and Hypothesis Testing*. Academic Press, fourth edition, 2017.