ROM SAF Report 44 Ref: SAF/ROM/METO/REP/RSR/44 Web: https://rom-saf.eumetsat.int Date: 30 June 2025



ROM SAF Report 44

# An investigation into the impacts of the vertical smoothing of GNSS-RO bending angle observations on Met Office NWP forecasts

Neill E. Bowler<sup>1</sup>, Christian Marquardt<sup>2</sup>

<sup>1</sup> Met Office, Fitzroy Road, Exeter, EX1 3PB, UK <sup>2</sup> EUMETSAT, Eumetsat Allee, 64295 Darmstadt, Germany

#### DOCUMENT AUTHOR TABLE

	Author(s)	Function	Date
Prepared by:	Neill Bowler	ROM SAF Local Manager	24/04 2025
Reviewed by:	Mary Forsythe	ROM SAF Steering Group member	12/05 2025
Reviewed by:	Sean Healy	ROM SAF Science Coordinator	19/06 2025
Approved by:	Kent B. Lauritsen	ROM SAF Project Manager	30/06 2025

#### DOCUMENT CHANGE RECORD

Version	Date	Ву	Description
1.0	30/06 2025	Neill Bowler	Final version following internal reviews

#### ROM SAF

The Radio Occultation Meteorology Satellite Application Facility (ROM SAF) is a decentralised processing centre under EUMETSAT which is responsible for operational processing of radio occultation (RO) data from the Metop, Metop-SG and Sentinel-6 satellites and radio occultation data from other missions. The ROM SAF delivers bending angle, refractivity, temperature, pressure, humidity, and other geophysical variables in near real-time for NWP users, as well as reprocessed Climate Data Records (CDRs) and Interim Climate Data Records (ICDRs) for users requiring a higher degree of homogeneity of the RO data sets. The CDRs and ICDRs are further processed into globally gridded monthly-mean data for use in climate monitoring and climate science applications.

The ROM SAF also maintains the Radio Occultation Processing Package (ROPP) which contains software modules that aid users wishing to process, quality-control and assimilate radio occultation data from any radio occultation mission into NWP and other models.

The ROM SAF Leading Entity is the Danish Meteorological Institute (DMI), with Cooperating Entities: i) European Centre for Medium-Range Weather Forecasts (ECMWF) in Reading, United Kingdom, ii) Institut D'Estudis Espacials de Catalunya (IEEC) in Barcelona, Spain, iii) Met Office in Exeter, United Kingdom, and iv) and Wegener Center, University of Graz, in Graz, Austria. To get access to our products or to read more about the ROM SAF please go to: https://rom-saf.eumetsat.int.

#### Intellectual Property Rights

All intellectual property rights of the ROM SAF products belong to EUMETSAT. The use of these products is granted to every interested user, free of charge. If you wish to use these products, EUMETSAT's copyright credit must be shown by displaying the words "copyright (year) EUMETSAT" on each of the products used.



#### Abstract

Motivated by apparent forecast improvements in the Met Office system with Spire-processed GNSS radio occultation bending angle observations, various experiments have been run to test the effect of increased vertical smoothing on forecast quality.

The initial experiments were run with additional smoothing applied to Spire's observations as part of the EUMETSAT Secretariat processing. In these experiments it was seen that increasing the smoothing decreased the standard deviation of the observation departures, but also increased the vertical correlation length-scales. These observations with additional smoothing were then ingested within a low-resolution version of the Met Office NWP system, and the forecast quality was seen to be improved with the observations using additional smoothing compared with the observations using the operational processing.

A second set of experiments were run which applied additional smoothing as a pre-processing step within the Met Office system. The smoothing is thus applied to the low-resolution BUFR observations which are normally assimilated operationally. This method has the advantage that it is applied to the whole observations dataset, but a disadvantage that it is applied to the low-resolution observations which posed some technical challenges. It also meant that it was possible to make the smoothing lengthscale proportional to the spacing between vertical levels in the Met Office model. Tests with applying the additional smoothing in this way demonstrated improved forecast performance over a wide range of variables. However, using a large smoothing length-scale produced degraded results, and the degradation was seen first in the tropical region, suggesting that less smoothing is beneficial there.

Further experimentation is planned which would demonstrate the impact of additional smoothing on a second NWP system. If these experiments show that additional, or different, smoothing is optimal in that system then it would demonstrate that this method provides an opportunity for all EUMETSAT members that assimilate bending angles to improve their forecasting systems. To achieve this would require data processing centres, such as EUMETSAT, make their observations available at high-resolution. It would also need tools to be made available for NWP centres to process these high-resolution observations in a manner which is best suited to their system.



# Contents

Ał	bstract	3					
1	Introduction						
2	Applying smoothing within EUMETSAT Secretariat's processing   2.1 Local Polynomial Kernel Regression filters   2.2 Bandwidths applied in the RO processing   2.2.1 The effect of different smoothing bandwidths   2.3 Results within the Met Office NWP system	7 9 10 11					
3	Method for applying smoothing as a pre-processing step 1   3.1 The smoothing algorithm 1   3.2 Power spectra of smoothed data 1						
4	Results with applying smoothing in pre-processing 20   4.1 Data assimilation statistics 20   4.2 Smoothing the logarithm of bending angle 21						
5	5 Conclusion 27						
Re	References 2 <sup>o</sup>						
R	ROM SAF (and earlier GRAS SAF) Reports 31						



# 1 Introduction

Numerical weather prediction (NWP) relies on establishing a good estimate of the current state of the atmosphere via the process of data assimilation (Clayton, Lorenc, and Barker, 2013). Data assimilation relies on having large quantities of high-quality observations, and one of the most impactful observation types is global navigation satellite systems (GNSS) radio occultation (RO) (Samrat et al., 2025). The raw observations of GNSS-RO are measurements of the Doppler shift of the radio signal and the time delay when the signal arrives at the satellite in low-earth orbit. These measurements are processed to calculate bending angle, which is smoothed and thinned relative to the original measurements. The quality of the processing of the observations is essential to the effective use of the observations, and one aspect of that processing is examined within this study.

During 2021, the Met Office and the European Centre for Medium-range Weather Forecasts (ECMWF) conducted a study which looked at the impact of assimilating GNSS radio occultation observations produced by Spire into their NWP systems (Lonitz et al., 2021). This study found that there were clear benefits from assimilating the additional data. Further, it also compared the benefit in the ECMWF system from assimilating observations processed by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) with the same observations processed by Spire. It was found that there was a slightly larger benefit when assimilating observations which had been processed by EUMETSAT Secretariat. Those same tests were repeated with the Met Office system, but they were not completed in time to be included in the final report. Those tests indicated that in the Met Office system there was slightly greater benefit from assimilating the observations when processed by Spire. Figure 1.1 shows the verification scorecard comparing an experiment with Spire-processed observations to an experiment with EUMETSAT Secretariat-processed observations. This shows that the forecasts for extra-tropical temperature and geopotential heights are mostly improved when using Spire-processed observations.

There are three main areas where the processing performed by Spire differed from that performed by EUMETSAT Secretariat. These are

- Spire uses an increased level of smoothing, meaning that the observations have smaller standard deviations and larger vertical correlations.
- The bias within the troposphere is smaller for EUMETSAT Secretariat processed observations.
- Above 35km the EUMETSAT Secretariat-processed observations have smaller standard deviations.

The first of these differences is mostly likely to be the source of the improvement seen within the Met Office system, leading to the suggestion of running experiments testing the level of smoothing applied in the observation processing.



							mi	ax =	20						
NH_W250															
NH_W500															
NH_W850												•		٠	
NH_W10m														•	
NH_T250															
NH_T500													•	•	
NH_T850			٠								•	•	٠		
NH_T_2m											•	٠	٠	٠	
NH_Z250														•	
NH_Z500														•	
NH_Z850										÷				٠	
TR_W250															
TR_W500															
TR_W850															
TR_W10m															
TR_T250															
TR_T500			٠		۷	٠	۷	۷	•	•	۷	۲	۲	٠	
TR_T850			٠	٠	۷	٠									
TR_T_2m				•	•		•		•						
SH_W250							•								
SH_W500				•			•								
SH_W850				•	•		•	ŀ	•					•	
SH_W10m		•	•	٠	۳	٠	٠	ŀ	٠	•		ŀ	•	•	
SH_T250															
SH_T500															
SH_T850														•	
SH_T_2m														•	
SH_Z250							4								
SH_Z500					4										
SH_Z850			۷	۷	•	۷	۷	۷	۳	·	÷	•	•	•	
	0+1	T+6	T+12	T+24	T+36	T+48	T+60	T+72	T+84	T+96	T+108	T+120	T+132	T+144	T+168

#### % Difference (Add Spire-processed vs. Add Spire reprocessed) - overall 0.4%, RMSE against ecanal for Equalized, 20211023 00:00 to 20220115 12:00

**Figure 1.1:** Forecast verification scorecard comparing the impact of assimilating Spire-processed observations against EUMETSAT Secretariat-processed observations. The area of the triangles is proportional to the change in root-mean-square error (RMSE), relative to the system with EUMETSAT Secretariat-processed observations, measured against ECMWF analyses. Green upwards-pointing (purple-blue downwards-pointing) triangles indicate better (worse) performance for the experiment with Spire processing. The y-axis denotes the forecast variable being considered. NH, TR and SH denote the northern hemisphere > 20 degrees, tropics and southern hemisphere < -20 degrees, respectively. The letter following the underscore indicates the weather parameter, with W, T and Z signifying vector wind field, temperature and geopotential height, respectively. The numbers after this indicate the height of the observation in hPa, except for 2m and 10m that are heights above the surface. The x-axis shows the forecast lead time in hours.



# 2 Applying smoothing within EUMETSAT Secretariat's processing

EUMETSAT Secretariat's level 1b processing of GNSS Radio Occultation observations consists of a fast version of the Phase Matching (or Phase Transform) algorithm; see Jensen et al. (2004) and Gorbunov and Lauritsen (2004) for details. This Fast Phase Transform (FPT) consists of a Fourier integral of a function of the (mathematical) complex observed signal using the impact parameter as independent variable. Technically, the Phase Transform is carried out as a Fast Fourier Transform over a regular impact parameter grid with a step size of 0.5m. The input data is upsampled from the observed complex signal, without having been subject to any form of smoothing. However, data at the bottom of an occultation consisting of noise only is cut off before the FPT is carried out. After the transform has been calculated, the bending angle retrieval continues with the following steps (for each of the two GNSS frequencies):

- 1. The transformed signal is subject to radio-holographic filter with a (one-sided) bandwith of 200m to address small-scale observational noise;
- Bending angles at the 0.5m impact grid are calculated by numerically differentiating the phase of the transformed complex signal using a Kay's weighted phase average estimator (Kay (1989)) applied over a sliding window of 75m;
- 3. Bending angles are smoothed using an adaptive height-dependent bandwidth (see below for more details on the bandwidth selection for this first explicit smoothing) and interpolated to a regular 25m impact parameter grid. Bending angles from the two GNSS frequencies are linearly combined to form the neutral atmospheric bending angle profile on the same "high-resolution" vertical impact parameter grid.
- 4. Finally, a second smoothing step, again with height-dependent bandwidths, is applied to the high-resolution bending angle profiles at a nominal resolution of 25m to produce a "thinned" set of bending angle profiles (one for each of the two GNSS frequencies, along with a profile characterising the neutral atmosphere). After applying the smoothing, the smoothed bending angles are interpolated to a standard non-equidistant grid of 247 impact parameter levels. These thinned bending angle profiles are disseminated to NWP users.

With their comparatively small filter bandwidths of 100 or 200m, the first two filter steps described above only result in a moderate smoothing of the final bending angle profiles. By far the largest amount of smoothing is introduced by the third and fourth step, which we refer to as the first and second bending angle smoothing in the following.

## 2.1 Local Polynomial Kernel Regression filters

For all but the second smoothing step described above, we apply a Local Polynomial or Kernel Regression smoother. In this type of smoother, a weighted low-order polynomial fit to the data in a short running kernel window is evaluated at its centre to obtain a smoothed value of the data at this point. Thus, for a given data set  $\{(x_i, y_i), i = 1, n\}$ , the relationship between the (potentially multi-variate) predictor variable x and a response variable y is assumed to be of the form

$$y_i = \boldsymbol{\mu}(x_i) + \boldsymbol{\varepsilon}_i \; ,$$

where  $\mu$  denotes an (unknown) function to be estimated, and  $\varepsilon_i$  a normally distributed and zero-mean random error. Further assuming that  $\mu$  can be locally Taylor expanded in the vicinity of a point x up to order p, e.g.,



$$\mu(z) \approx \sum_{j=0}^{p} \frac{\mu^{(j)}(x)}{j!} (z-x)^{j} + \ldots = \sum_{j=0}^{p} \beta_{i} (z-x)^{j} + \ldots$$

where  $\mu^{(j)}(x)$  denotes the *j*-th derivative of the model  $\mu$  at *x*. A local estimate of the model parameters  $\beta_i$  is then obtained by minimising the weighted least squares regression cost function

$$\min_{\beta_0\dots\beta_p} J = \sum_{i=1}^n K_h(x_i - x) \cdot \left( y_i - \sum_{j=0}^p \beta_j (x_i - x)^j \right)^2.$$

Here,  $K_h$  denotes a kernel (or weighting) function which is typically chosen as a symmetric probability density with a given bandwidth h. Common choices are the normal (or Gaussian) kernel width standard deviation h,

$$K_h^n(x) = \frac{1}{\sqrt{2\pi h^2}} e^{-\frac{x^2}{2h^2}}$$

or a uniform (boxcar) kernel with a one-sided width *h*:

$$K_h^u(x) = \begin{cases} rac{1}{2h} & ext{if } |x| \le h \\ 0 & ext{otherwise} \end{cases}$$

Local Polynomial Kernel Regression was studied extensively (e.g., Wand and Jones (1995), Fan and Gijbels (1996), Loader (1999)) and comes with a rich body of theoretical statistical results. Compared to ordinary linear filters, kernel regression filters provide improved performance near the boundaries, especially for odd polynomial orders; the local regression implementation also lends itself to process irregularly distributed data points and allows for varying bandwidths, thus providing a highly flexible smoothing methodology. The local fitting of a low order polynomial also means that non-stationary time series can be smoothed without introducing biases, as long as the kernel is not too wide compared to the length scale of the non-stationary variations in the data.

The earliest version of local kernel regression was proposed by Savitzky and Golay (1964), providing a fast method for equidistant data though without proper handling of data points near the boundary. Note that fitting a zeroth-order polynomial - i.e., a constant - is equivalent to an ordinary finite impulse-response (FIR) filter. Schafer (2011) also demonstrated the similarity in the frequency responses for certain traditional FIR filters with very flat passbands and Savitzy-Golay filters. Thus, local kernel regression can also be seen as one specific way to design ordinary linear filters, but providing more flexibility regarding local bandwidths and non-equidistant designs.

For computational efficiency, the EUMETSAT Secretariat implementation is based on binning the original data, saving the repeated calculation of kernel weights (see Wand and Jones (1995) and references therein). The drawback is that for low data densities (with respect to the chosen bins), the smoothing process might fail because some of the bins cannot be filled with data samples any more. Thus, applying a local polynomial regression smoother to a data set for a second time may fail if the data has been thinned or subsampled after the first application of the filter.

The amount of smoothing resulting from the application of a kernel regression filter depends on the kernel's width and shape as well as the polynomial order of the fit. In general, wider kernels will provide stronger smoothing. However, kernels with different shapes but the same nominal bandwidth h might provide significantly different amounts of smoothing. For example, a Gaussian kernel defined via its standard deviation h as nominal (one-sided) bandwidth used over an interval [-4h, 4h] is far wider and





**Figure 2.1:** Vertical bandwidths for the first (left) and second bending angle smoothing (middle) in EUMETSAT Secretariat's RO processing; also shown is the new bandwidth of the combined smoothing (right). Colours in the centre and right plot indicate different smoothing bandwidths applied in the thinning step. The legend indicates the bandwidth applied in the lower troposphere only.

results in more smoothing than a boxcar window with the same nominal one-sided bandwidth h used over the interval [-h,h]. Higher polynomial orders fitted over the same kernel provide less smoothing than lower order polynomials, or require larger bandwidths to provide similar amounts of smoothing. It is thus difficult to directly compare filter bandwidths of kernel regression smoothers without taking the kernel's shape and the order of the polynomial fit into account.

However, Marron and Nolan (1988) demonstrated that after a kernel-specific rescaling of different kernel windows, a common "canonical bandwidth" used with the rescaled kernels leads to a similar amount of smoothing regardless of the kernel applied to the data. In practice, this makes the amount of smoothing independent from the chosen kernel, and thus simplifies the interpretation of smoothing bandwidths considerably.

The smoothing in EUMETSAT Secretariat's processing is based on canonical bandwidths and kernels. However, to allow an intuitive assessment, we provide all smoothing bandwidths as one-sided bandwidths of an equivalent boxcar kernel. The actual implementation uses third-order local polynomial regression over a Gaussian kernel applied over the interval [-4h, 4h] where *h* denotes the local canonical bandwidth; the Gaussian kernel suppresses high-frequency noise more efficiently than an equivalent boxcar filter would do.

### 2.2 Bandwidths applied in the RO processing

Figure 2.1 shows the vertical bandwidth profiles applied during the first (high-resolution) and second (thinning) bending angle smoothing step, along with an estimated net bandwidth of the combined smoothing steps. The lines denote the one-sided equivalent boxcar bandwidths.

For the first bending angle smoothing step, the bandwidth is around 100m at altitudes below 5 km, to avoid the oversmoothing of fine-scale vertical structure retrieved by the wave optics processing. Around 45 km height, i.e. the upper stratosphere, the bandwidth typically reaches 1500m and remains constant above. This choice ensures that the smoothing of upper-level bending angle values is consistent with the smoothing applied to earlier, geometrical optics based retrieval schemes. Note that the lower bandwidth is fixed and identical for all occultations. On the other hand, the upper asymptotic bandwidth represents the vertical range covered by a straight-line tangent point during 0.75 seconds; it is calculated for each occultation based on the actual tangent point descent or ascent rate between 45 and 80 km height and





**Figure 2.2:** Robust estimates of the Observation-minus-Background bias (left) and standard deviation (right) for various smoothing bandwidths choices in the tropical troposphere. Statistics for the data originally provided by Spire is also shown (in grey). See text for details.

thus depends on the orbit geometry. The smoothing bandwidth for the secondary frequency is further adapted to minimise the occurrence of small-scale oscillations between 40 and 80 km in order to ensure smooth upper-level *neutral* bending angle characteristics.

In the second bending angle processing step, the vertical shape is similar, but bandwidths were varied to simulate the impact of different amounts of vertical smoothing; colours denote the different setting applied to reproduce a similar smoothing as the one likely applied to the original Spire data. The default (or baseline) bandwidth for the second bending angle smoothing step is denoted by the magenta line in the centre plot of Figure 2.1; the values of increasing bandwidths at both lower and upper end of the bandwidth profiles are denoted by the different colours. Note that the numbers in the legend indicate the setting for the lower bandwidth only.

#### 2.2.1 The effect of different smoothing bandwidths

Not surprisingly, increasing the amount of smoothing reduces the random uncertainty introduced by measurement noise. Thus, increased bandwidths in the thinning step of the processing reduce the standard deviations of bending angle retrievals against NWP data, as shown in Figure 2.2 for the tropical troposphere. Statistics were calculated using a robust estimator of the Observation-minus-Background covariance matrix initially proposed by Gnanadesikan and Kettenring (1972) and later improved by Maronna and Zamar (2002).

The figure also shows the statistics of the original Spire data. Concerning standard deviations, the figure suggests that the strongest smoothing (labelled as "900m" in both Figure 2.1 and Figure 2.2) generates similar statistical characteristics as found in the original Spire data. The same conclusion can be drawn from the correlation length-scales resulting from the various amounts of smoothing: Figure 2.3 shows the Full Width at Half Maximum (FWHM) of the vertical correlation function at each level. For the operational smoothing (labelled as "Baseline" in in Figure 2.3), the FWHM is approximately 500m in the troposphere, and then gradually increases to around 2km in the upper stratosphere. Using a smoothing bandwidth of 900m in the lower troposphere increases the correlation length-scale to be similar to the Spire-processed bending angle (labelled as "Original" in Figure 2.3).





**Figure 2.3:** Full-Width at Half-Maximum (FWHM) of the tropical O-B correlation matrix for Spire and Metop data. The FWHM is a proxy for the amount of vertical smoothing applied in the various retrievals.

Note that the original Spire-processed data exhibits significantly larger positive biases in the lower troposphere (see Figure 2.2), although this particular aspect of the Spire data is outside the scope of this report and will be addressed elsewhere.

## 2.3 Results within the Met Office NWP system

Experiments were run with the Met Office's NWP system, using the observations from Spire processed with the extra smoothing of 300, 600 and 900 m in the troposphere. The experiments were based on a low-resolution version of the Met Office NWP system, including coupling to the ensemble forecasting system to provide updated covariances to the data assimilation. The experiments used a weather forecasting model run with 640 by 480 grid points in the horizontal and 70 vertical levels, stretching from near the surface to 80 km altitude. The GNSS-RO observations are modelled in the NWP system using a one-dimensional operator using interpolation to pseudo-levels located between the model levels (Burrows, Healy, and I. D. Culverwell, 2014). The observation-error covariance matrix (Bowler, 2020) is assumed to be diagonal, i.e. observation errors are uncorrelated. Thus the correlations shown in Figure 2.3 are not modelled. Assuming that observation errors are uncorrelated is a common assumption across NWP centres, and therefore there has often been a desire to avoid introducing correlations through additional smoothing.

These experiments were run for a period between 22 October 2021 and 22 January 2022. During this period EUMETSAT purchased a large number of observations from Spire as part of a pre-operational test, for use only within EUMETSAT's member states. They assimilated all observations which were used operationally, plus the additional Spire observations. The model settings were based on the system which became operational in Parallel Suite 44, which was the operational configuration between December 2020 and May 2022.

In each case only the source of the input observations was changed, no attempt was made to alter the quality control or the observation uncertainties applied within the NWP system. Although the smoothed observations have smaller differences when compared with the NWP system, no account was made for this as the extra smoothing introduces vertical correlations which are not accounted for in the data assimilation system. Therefore, it was considered that the best course of action is to leave the observation uncertainties unchanged.

Figures 2.4 and 2.5 show the verification scorecards for the experiments using 300, 600 and 900 m smoothing, respectively, compared with the experiment which uses the operational smoothing. These

EUMETSAT

ROM SAF







**Figure 2.4:** Forecast verification scorecard showing the impact of using smoothing length-scales of 300 m (left) and 600 m (right) in the troposphere compared with the operational smoothing. Scores calculated using the RMSE, measured against ECMWF operational analyses.

As an example of the forecast improvement possible, Figure 2.6 shows the forecast verification scores for forecasts of temperature at various heights in the southern extra-tropics. At short forecast lead times there are large reductions in the RMSE below 600 hPa for the experiments using 600 and 900 m smoothing. At longer lead times the reduction in the RMSE in the lower troposphere reduces, but the reduction in the RMSE for 600 m smoothing is less than for the 900 m smoothing at T+96h. At longer lead times the RMSE for the 300 m smoothing experiment becomes similar to that for the experiment with operational smoothing. Whilst the RMSE for temperature is similar between 600 and 900 m smoothing in the short range, the benefits of 600 m smoothing appear to be lost into the medium range.







**Figure 2.5:** Forecast verification scorecard showing the impact of using a smoothing length-scale of 900m in the troposphere compared with the operational smoothing. Scores calculated using the RMSE, measured against ECMWF operational analyses.





**Figure 2.6:** Reduction in the forecast RMSE of temperature for the different smoothing experiments compared to the control, as a function of height in the southern extra-tropics measured against ECMWF operational analyses. Shown are results for 300 m smoothing (blue), 600 m smoothing (green) and 900 m smoothing (pink). The forecasts are verified at T+12h (top) and at T+96h (bottom).



## 3 Method for applying smoothing as a pre-processing step

The previous experiments considered the effect of differing levels of vertical smoothing within EUMET-SAT Secretariat's processing of Spire observations. This approach has the advantage that it is modifying an existing step within the processing chain. However, it only applied the smoothing to Spire's observations, not the whole set of GNSS-RO observations being assimilated. In this section we consider applying the additional smoothing as a pre-processing step within the Met Office's NWP system.

## 3.1 The smoothing algorithm

In this next step of experiments, the smoothing is applied to the low-resolution observations normally assimilated by the data assimilation system. This includes the BUFR files (Binary Universal Form for the Representation of meteorological data) which are exchanged over the Global Telecommunication System (GTS), as well as the high-volume observations from Spire which were the subject of the previous experiments. The smoothing that is applied to bending angle profiles is designed to be similar to that used in the processing of EUMETSAT Secretariat. Since bending angles typically vary exponentially with height, there will be a natural tendency for the smoothing to introduce a bias in the bending angles. Therefore, this is an aspect which will be considered explicitly.

Let us define a profile of bending angles as  $\alpha_i$  where *i* is the index of the observation within the profile. We will assume that the observations have been processed in such a way that any missing or corrupt data have been removed. Crucially, we will not assume that  $\alpha_i$  is positive for all *i* since the received bending angles can often be negative (I. Culverwell, 2021). We define the altitude of a given observation via the impact height ( $x_i$  for observation *i*). The impact height is defined as the impact parameter at tangent point, minus the earth's radius of curvature at this point. The impact parameter at the tangent point is defined as the distance of the ray from the centre of the earth times the refractive index of the air.

To begin with a normalised value for the impact height is calculated using

$$\tilde{x}_i = \frac{x_i - \langle x \rangle}{\lambda_i} \tag{3.1}$$

where  $\langle \cdot \rangle$  denotes the average over the profile and  $\lambda_i$  is the length-scale of the smoothing, which varies between experiments and with the impact height of the observation. The weight given to observation j when calculating a smoothed value for observation i is given by

$$w_i = \sqrt{e^{-\frac{(\bar{x}_i - \bar{x}_j)^2}{2}}}.$$
 (3.2)

Using these weights we then calculate coefficients for a cubic polynomial regression on the bending angles. The value of this polynomial at the observation location gives the value of the smoothed observation. This process is then repeated for all observations within the profile. As in section 2, the length scales reported here are reported as one-sided bandwidths of an equivalent boxcar kernel. Following Marron and Nolan (1988) this is equivalent to approximately  $1.74\lambda_i$ .

The smoothing is applied to the observations within the BUFR files exchanged over the GTS. The observations within these files are typically provided on 247 levels in the vertical, with separations ranging from 118 m in the lower troposphere to 308 m in the upper stratosphere. Since a cubic polynomial was used in all instances it was not possible to apply the method with a smoothing kernel whose length-scale is too short. Hence a minimum smoothing length-scale of 261 m was applied in all instances. Similarly, early results demonstrated that applying a large smoothing length-scale towards the top of the RO profile led to biases. Therefore, it was also chosen to limit the maximum smoothing length-scale to be 2610 m.

Within those constraints the aim of this section of work was to see if a smoothing length-scale proportional to the separation between vertical levels in the Met Office model performs well. Figure 3.1 shows the set of smoothing length-scales which have been used in these tests. The length-scales used away from the minimum and maximum have been slightly adjusted so that the minimum and maximum are approached smoothly. Using a smoothing scale of one half the model-level separation means that at 10 km impact height a smoothing scale of 458 m is used. With a smoothing scale of 1.5 times the model-level separation a smoothing scale of 1209 m is used at 10 km impact height. For reference, Figure 3.1 also shows the separation between adjacent levels in the Met Office model. As will be noted, these values are much smaller near the surface and much larger near the top of the model, due to the threshold being used on the kernel size.



**Figure 3.1:** Spatial scale used in for the smoothing kernel as a function of impact height of the observation — the impact height here is calculated based on an average atmosphere. Also shown are the differences between adjacent levels in the Met Office model when a point above the ocean is used, for comparison. Note that the spatial scale for the smoothing kernel is reported as the one-sided bandwidth of an equivalent boxcar kernel, so is multiplied by approximately 1.74 relative to the model level differences.

Smoothing in the presence of sharp gradients can lead to biases in the smoothed quantity, as was noted by Liu et al. (2020). Figure 3.2 shows the mean and standard deviation of the normalised departures between the observations and background with different levels of smoothing applied to the bending angles. From this we see that as the level of smoothing increases the standard deviation of the departures decreases. The largest decreases are seen when applying smoothing proportional to 0.5 or 0.75 times the model level spacing. Applying more smoothing than this results in only modest further reductions in the standard deviation. At high altitudes the mean departure for all the curves is the same, because they are all limited to 2610 m smoothing at these levels. Between 10 and 35 km the mean of the unsmoothed bending angle departures is close to zero. Using smoothing of 1.5 times the model level spacing results in a positive-negative change to the mean departure between 15 and 20 km. This is presumably because an aggressive smoothing strategy is being applied across the tropical tropopause. Above 20 km the 1.5 times smoothing has a negative bending angle departure, and this persists all the way to 35 km. Using smaller levels of smoothing results in a smaller bias, but it is apparent that a bias is created when applying all of the smoothing levels tested here. Some of the larger variations in the mean O-B statistics at high altitude, such as the increases at 45 and 49 km correspond to the location of model levels. Between the layers the model variables are interpolated (Burrows, Healy, and I. D. Culverwell, 2014) which is presumably creating a small bias in the forward-modelled bending angles.

Rather than smooth the bending angles directly, it's possible to apply the smoothing to the logarithm of the bending angle. To avoid problems with negative bending angles, the logarithm of the following





**Figure 3.2:** Mean and standard deviation of the normalised difference between observations and background bending angle forecasts ((O-B)/B) with various levels of smoothing, when the smoothing is applied to the bending angles. Data calculated by smoothing all operational bending angles for the 00 UTC cycle on the first day of each month in 2023.

quantity is calculated

$$\begin{aligned} \tilde{\alpha}_i &= \alpha_i + \left| \min_{\forall i} \alpha_i \right| + 10^{-7} & \text{for } \min_{\forall i} \alpha_i < 0 \\ &= \alpha_i & \text{for } \min \alpha_i \ge 0 \end{aligned}$$
 (3.3)

where  $\alpha$  denotes the bending angle. The modified bending angle is guaranteed to be non-negative, but ideally the conversion would ensure that the bending angles are sufficiently far from zero that the nonlinearity of the logarithm doesn't adversely alter the data. The smoothing is applied to the logarithm of  $\tilde{\alpha}_i$ , and from these the smoothed bending angles are calculated. Figure 3.3 shows the mean and standard deviation of the normalised bending angle depatures when the smoothing is applied to the logarithm of the bending angles. The reduction in the standard deviations of the departures is similar to what is seen in Figure 3.2. The mean departures are noticeably closer to each other between 5 and 35 km than when applying the smoothing to the bending angles directly. Therefore, the smoothing is changing the bias of the observations less when the smoothing is applied to the logarithm of the bending angles.

### 3.2 Power spectra of smoothed data

One way to help understand the impact of vertical smoothing of the profiles is to examine the power spectra of the smoothed and unsmoothed profiles. This is challenging since the observations are not on a regular vertical grid. To calculate the power spectrum, the observations are first interpolated to a regular 200 m grid in the vertical, and any missing data are removed. Values of zero bending angle are then added to the top of the profile, to ensure that each profile is the same length, ensuring a total of 301 entries within each zero-padded profile. Since the fast Fourier transform (FFT) is a cyclic algorithm, a copy of the zero-padded profile in reverse order is added to the start of the profile. In this way the FFT is performed upon the profile and its mirror image, padded with zeros at either end to ensure the calculation is performed on a cyclic dataset.





**Figure 3.3:** Mean and standard deviation of (O-B)/B with various levels of smoothing, when the smoothing is applied to the logarithm of the bending angles. Data calculated by smoothing the logarithm of all operational bending angles for the 00 UTC cycle on the first day of each month in 2023.

Figure 3.4 shows the power spectrum for the observations and the smoothed observations (using a length-scale equal to 0.75 times the model-level spacing). There is a clear difference between power spectra for the smoothed and original observations, particularly between frequencies of 0.0007 and  $0.002 \text{ m}^{-1}$ , which corresponds to a length-scales of 1429 and 500 m. This is consistent with the level of smoothing which is applied in the stratosphere and above.



**Figure 3.4:** Vertical power spectrum for the observations (blue) and the observations after smoothing (orange) with a smoothing length-scale proportional to 0.75 times the model-level spacing. Observations are taken from the ROMEX experiment during September 2022.

Although the smoothing is applied to the observations, it is principally the effect on the analysis increments which is of interest as this is how the NWP forecasts are influenced. To test this a 1D-Var assimilation was run to produce an analysis increment for each profile (both with and without the additional smoothing). Power spectra for these analysis increments, once again in terms of bending angles and using the same procedure described above, are shown in Figure 3.5. The analysis increments are filtered through the background-error covariance and are therefore much smoother than the observa-



tions themselves, which is reflected in the values of the power spectrum being much lower than for the observations. The difference between the power spectra with and without the additional smoothing is also much less and largely confined to lower frequencies than for the observations. These results are consistent with the expectation that much of the smoothing seen in the analysis increments derives from the background-error covariance used, rather than from the filtering applied to the observations.



**Figure 3.5:** Vertical power spectrum for the analysis increments without additional smoothing (blue) and with additional smoothing (orange) with a smoothing length-scale proportional to 0.75 times the model-level spacing. Observations are taken from the ROMEX experiment during September 2022.



# 4 Results with applying smoothing in pre-processing

As with the experiments which tested the effect of adding additional smoothing to EUMETSAT Secretariat's processing, low-resolution NWP experiments were run to test the effect of adding additional smoothing to the pre-processing of the observations within the Met Office system. These experiments were run for the same period as the previous experiments with the same model settings.

In order to understand the impact on the forecast of the additional smoothing we consider the change in the RMSE of the forecast, compared to ECMWF analyses for various quantities and lead times. Figures 4.1, 4.2 and 4.3 show the verification scorecards for experiments using a smoothing scale between one quarter of the model-level spacing and one and a half times the model-level spacing.

For smoothing of one quarter of the model-level spacing we note that there is a degradation in the forecast performance in the northern extra-tropics at medium-range. Given that the smoothing length-scale is rather modest, it seems likely that this is connected to the scheme's performance in the lower troposphere where the smoothing kernel is restricted to be at least 261 m without gaining the benefits of greater smoothing at higher altitudes. Using smoothing which is proportional to one half the model-level spacing (Figure 4.1, right) is beneficial for almost all variables and lead times. Increasing the level of smoothing to three-quarters the model-level spacing (Figure 4.2, left) further increases the forecast performance, especially in the extra-tropics. In the tropics the increased level of smoothing has a mixed effect. Whilst the scores are improved in the short range, it is noticeable that the temperature forecasts in the tropics at medium range perform better with the lower level of smoothing. This trend continues with smoothing proportional to (one times) the model-level spacing (Figure 4.2, right) where the forecast performance in the tropics is degraded by the use of additional smoothing. In the extra-tropics the forecasts are still improved by using this level of smoothing. If one increases the level of smoothing further (Figure 4.3) then the forecast performance is degraded, especially in the tropics.

## 4.1 Data assimilation statistics

One way to better understand the impact of the additional smoothing is to look at the statistics from the data assimilation step. These can both explain how the minimisation is affected by the smoothed observations, but also reflect its impact on other observations used by the assimilation. Given that 0.5 and 0.75 times the model-level spacing appear to be the most promising smoothing length-scales we focus on these in the following.

Figure 4.4 shows the average value of the data assimilation cost function during the minimisation. With the additional smoothing, the costfunction values are generally lower, due to the observations GNSS-RO being less noisy. Since GNSS-RO observations consitute a substantial fraction of the total cost function, the reduction of the GNSS-RO component adjusts the total cost function by a large amount. Figure 4.4 also shows the ratio of the cost function with and without additional smoothing when GNSS-RO observations are excluded from the calculation. The cost function is very similar in all cases, indicating that the additional smoothing does not bring the background or analysis much closer to these independent observations. For this statistic the experiment which smoothes the logarithm of the bending angle has a slightly lower cost function than the other observations. It should be noted that the minimisation is run as two separate steps: one at N108 ( $\sim$ 120 km resolution) and one at N216 ( $\sim$ 60 km resolution) (Payne, 2011). Hence, there is a jump in the statistics after 50 iterations as this is the point at which the resolution in increased.

To further understand the impact of the experiments, we can examine the differences between the observations and background forecast or analyses for individual observation types. Figure 4.5 shows the ratio of the root mean square difference between observations of potential temperature from sondes and the model estimates of the same. The plots are shown as ratios between three experiments with different smoothing schemes and the control. Points which are less than one indicate that the experiment has



					m	ах	= 2	20						_				max	= 2	!0			
NH W50										•		•	NH W50										<b>.</b> 11
NH W100									•		٠		NH W100										
NH W250								٠	۷		•	۷	NH W250									1	
NH W500				. 1				•	•	•	•		NH W500					1	1.				•
NH W850								٠		•	•	Ŧ	NH W850			1			1		-	-	
NH T50			. 1							1		۲	NH T50			¥ İ		1.					
NH T100		1	. 1	¥1				1	•	•	•	v	NH T100					1	1			1	
NH T250									٠	1		۲	NH T250							1da	1		
NH T500		1	1	. 1				•		•		•	NH T500			Ϋ́		÷.	1.		1		•
NH T850							٠			•			NH T850					1.	1.			-	
NH Z50				1	•	۲	۲	٠	•				NH Z50						1		1		
NH Z100									•	•	•		NH Z100		۸								
NH Z250		۸						•	۷		•	۷	NH Z250	•	۸		۰.					-	
NH Z500				. :					۷	•			NH Z500	•	۸							1	• • •
NH Z850	1		•	•				۲	۷	٧	۲	•	NH Z850						1		۲	•	• •
TR W50				1									TR W50	•	۸								
TR W100			•	1		•	٠	1	1				TR W100					•	1.	•			
TR W250													TR W250										
TR W500				. :								÷ .	TR W500										
TR W850				. ]					1	•	•	•	TR W850									1	1.1.1
TR T50				•			٠				•		TR T50	4	۸						•		•
TR T100				1					:	1	-	:	TR T100	4	۸								
TR T250				•	4	۸	۸						TR T250	4	۸	Δ.	Δ Δ		4	•			A A
TR T500			•	•	•	•				1			TR T500	•			• •	•	•		•		•
TR T850				•	•	•							TR T850			•							
SH W50									•				SH W50		۸								
SH W100	-							•	•				SH W100										
SH W250								٠					SH W250										• •
SH W500			- 1	1									SH W500					1					A A
SH W850								•					SH W850										
SH T50				•				•	۲	•			SH T50	4	۸								
SH T100										1.			SH T100	▲	۸								A A
SH T250													SH T250		۸								A A
SH T500				• ]					1	1.		1	SH T500								- 1		
SH T850			•						1	1	1		SH T850										
SH Z50		۸	٠					•	•	۷	٠	•	SH Z50		۸					•	•	1	
SH Z100	۸	۸	۸	•	۸					1			SH Z100		Δ	Δ.	▲ ▲		۸				A A
SH Z250	4	4							-				SH Z250		Δ	<b>A</b>	A .						<b>A A</b>
SH Z500	4	۸											SH Z500		۸							. ,	Δ Δ
SH Z850				•	1	1	1		1		-		SH Z850				•	1	1		1		<b>A A</b>
	12	24	90	20	00	72	84	96	80	20	32	44		12	24	99	9 0	12	34	90	80	2 6	4 2
	÷	÷.	ť	÷.	É	÷	Ť	Ť	+1	H H	Ŧ	÷		÷	÷	± i	έŤ	÷±	ž	Ê	41	75	
									÷	÷	÷	÷.								1.1	÷ι	÷μ	÷ É.

**Figure 4.1:** Scorecards showing the change in the RMSE for an experiment using smoothing proportional to one quarter the spacing between model levels (left) and proportional to one half the model-level spacing (right). Verification is performed against ECMWF analyses.

a smaller RMS difference than the control (i.e. an improved forecast or analysis). The statistics for all experiments are significantly below one above model level 40 (around 11 km altitude). This might be expected since GNSS-RO observations have most impact in the stratosphere and the improvements caused by smoothing will therefore be most visible in this region. The statistics when comparing the observations to the analysis (Figure 4.5, right) are generally smaller than those against the background forecast (Figure 4.5, left) which suggests that the additional smoothing is helping the minimisation to fit the sonde observations better. It is also notable that the statistics for smoothing proportional to 0.75 the model level spacing often performs better than smoothing proportional to 0.5 the model level spacing.

Figure 4.6 shows the same statistics as Figure 4.5, but for IASI observations on Metop-B. The Met Office assimilates only 130 of the many channels from this instrument, so only those channels are shown here. For this instrument the RMS statistics compared to the background are generally smaller than the statistics against the analysis, indicating that the additional smoothing is not directly helping the observations to be fit better. For some of the channels (around channel number 3000) the change in the RMS is relatively large. For these the smoothing using the logarithm of the bending angle appears to perform best, whereas they are generally similar for other channels.

## 4.2 Smoothing the logarithm of bending angle

As discussed in section 3.1 it is possible to apply the additional smoothing to either the bending angle or to the logarithm of the bending angles. Smoothing based on the logarithm of the bending angle affects



	max = 20	max = 20
NH W50		NH W50 🛕 🗛 🗛 🗛 🗛 🗛 🗛 🗛
NH W100	<b>A A A A A A A A </b>	
NH W250		NH W250
NH W500		NH W500
NH W850		NH W850
NH T50		
NH T100		
NH T250		NH T250
NH T500	· · · · · · · · · · · · · · · · · · ·	NH T500
NH T850		NH T850
NH 750		NH 750 A A A A A A A A A A A A A A A A A A A
NH 7100		NH 7100
NH 7250		NH 7250
NH 7500		NH 2500
NH 7850		NH 2850
TR W50		TR W50
TR W100		
TP W250	A	TP W250
TP W500	1999 - 1997 - 19	
TP W850		TP W850
TP T50		
TP T100		TP T100 Y Y Y
TP T250		TP T250
TP T500		
TP T950		TP 1950 7 7 7 7 7 7 7 7 7
SH W50		SH W50
SH W100		
SH W250		SH W250
SH W900		SH W950
SH T50		
SH 130		
SH T250		
CH T500		
SH T950		
SH 750		
SH 2100		
CH 7250		
SH ZZOU		
SH Z300		
o⊓ ∠850		
	-12 -24 -24 -24 -24 -24 -24 -24 -24 -24 -2	
	こここさささささ エココココー	

**Figure 4.2:** Scorecards showing the change in the RMSE for an experiment using smoothing proportional to three quarters the spacing between model levels (left) and proportional to the model-level spacing (right). Verification is performed against ECMWF analyses.

the mean bending angle less than applying the smoothing to the bending angle directly, so we expect that this will perform better.

Figures 4.7 shows the verification scorecard for forecasts which used smoothing of the logarithm of the bending angles proportional to one half and three quarters of the model level spacing. Comparing these graphs with the equivalent graphs in Figures 4.1 and 4.2 it appears that the benefit from smoothing with the logarithm of the bending angles is less than using the bending angle directly. For smoothing proportional to one half the model level spacing the largest differences are in the southern extra-tropics at medium range. For smoothing proportional to three quarters the model level spacing the largest differences are in the troposphere in the northern extra-tropics. This is contrary to expectations, since smoothing the logarithm affected the bias of the bending angles less than smoothing the bending angles directly. At higher levels of smoothing (not shown) there are some indications that the experiments smoothing the logarithm of the bending angle perform better, although in this case both experiments perform worse than the control. Thus it seems that the biases introduced by the moderate levels of smoothing are insignificant.



	max = 20	max = 20
NH W50		NH W50 🔻 🔺 🖌 🖌 🖌 🖌 🖌
NH W100		NH W100 A A A T T T T T T T T T T T T T T T T
NH W250		NH W250 🔻 🔹 🔺
NH W500		NH W500
NH W850		NH W850
NH T50	A A A A Y Y Y Y V V V	NH T50 V V V V V V V V V V V V V V V V V V V
NH T100	Δ <u> </u>	NH T100 V V V V V V V V V V V V V V V V V V
NH T250	V V V V	NH T250 🔍 🔍 🔍 🗸 🔹
NH T500	<b>A A A A A A A A A A</b>	NH T500 A A A A
NH T850		NH T850 7 7 7
NH Z50	<b>A A A A A A A A A</b>	NH Z50 🗛
NH Z100		NH Z100 🔻 🔒 👘 👘
NH Z250	<b>Δ</b> Δ Δ Δ Δ Δ Δ Τ Τ Τ	NH Z250 🗛 🗛 🗛 🗛 🗛 🗛 🗛
NH Z500	<b>A A A A A A A A A A</b>	NH Z500 🛕 🛕 🖌 🖌 🔺 🖌
NH Z850		NH Z850
TR W50	<b>\V \V \V \V \V \V \V \V</b>	TR W50 🗸 🗸 🗸 🛪 🛪 🛪 🛪 🛪 🛪 🛪 🛪
TR W100	· · · · · · · · · · · ·	TR W100 🛛 🗸 🗸 🖉 🗸 🛪 🛪 🔹 🔹
TR W250	* * * * * *	TR W250 v v v v v v v v v v v v v v v v v v v
TR W500	<b>V V V V V V V V V V</b>	TR W500 🔍 🔍 🔍 🔍 🔍 🔍 🔺 🔺 🔹
TR W850	<u>.</u>	TR W850
TR T50	$\nabla  \nabla  \nabla  \nabla  \nabla  \nabla  \nabla  \nabla  \nabla  \nabla $	TR T50 🗸 🗸 🛆 🖉 🖉 🖉 🖉 🖉 🖉 🖉 🖉 🖉
TR T100	$\bigtriangledown \bigtriangledown \bigtriangledown \bigtriangledown \bigtriangledown \lor \lor \lor \lor \lor \lor \lor \lor \lor$	TR T100 🔽 🤍 🏹 🗸 🗸 🖉 🖉 🖉 🖉 🖉 🖉
TR T250		TR T250
TR T500	<b>v v v v v v v v v</b>	TR T500 🛛 🗸 🖉 🖉 🖉 🖉 🖉 🖉 🖉
TR T850	<b>v v v v v v v</b>	TR T850 V V V V V V V V V V V V V V V V V V V
SH W50	×	SH W50 🗸 🕇 🔹 🔹
SH W100	<b>A A A A</b>	SH W100 🔺 🔺 📑 🔽
SH W250	A A A	SH W250 • • • • •
SH W500		SH W500 1 1 1 1 1 1
SH W850	• • · · · · · · · • • • •	SH W850
SH T50	<b>A A A A A A A A</b>	SH T50 🛕 🗛 🗛 🗛 🗛 🗛 🗛 🗛 🗛 🗛
SH T100	<u>A</u> A T T T T T A A	SH T100 🛛 🖓 🖓 🖓 🖓 🗖 🔍 🔻 🔻 🔻
SH T250	V V A A A A	SH T250 🛛 🔍 🔻 🔻 🔻 📩
SH T500	A A A A A A A A	SH T500 🔒 🕐 😗
SH T850		SH T850 🛕 🛕 🛕 🛕 🔺 🔺 🔺
SH Z50	<b>A A A A A A</b>	SH Z50 🔺 🗴 😗 😗 😗 👘
SH Z100	• • • • • • • • • • • •	SH Z100 🔻 🔺 🔺 🔺
SH Z250	<b>A A A A A A A A</b>	SH Z250 🛕 🛕 🔺 🔺 👘
SH Z500		SH Z500
SH Z850		SH Z850
	112 224 336 60 60 84 84 84 84 84 84 84 84 84 84 84 84 84	1 2 2 8 2 2 8 2 7 2 1 2 7 2 7 2 8 2 7 2 7
	+++++++++++++++++++++++++++++++++++++++	

**Figure 4.3:** Scorecards showing the change in the RMSE for an experiment using smoothing proportional to one and a quarter times the spacing between model levels (left) and proportional to one and a half times the model-level spacing (right). Verification is performed against ECMWF analyses.









**Figure 4.5:** Root mean square (RMS) difference between observations of potential temperature and the model background (top) or analyses (bottom). Statistics are shown as the ratio of the RMS difference for experiments with additional smoothing of GNSS-RO observations and the control (without additional smoothing). Error bars indicate 90% confidence intervals estimated using bootstrap resampling.





**Figure 4.6:** Root mean square (RMS) difference between observations from the IASI instrument on Metop-B and the model background (top) or analyses (bottom). Statistics are shown as the ratio of the RMS difference for experiments with additional smoothing of GNSS-RO observations and the control (without additional smoothing). Error bars indicate 90% confidence intervals estimated using bootstrap resampling. Experiments with smoothing proportional to one-half the model-level spacing (blue)), three-quarters the model-level spacing (orange) and three-quarters the model-level spacing, using the logarithm of the bending angle (green) are shown. The y-axis shows the channel number of those assimilated by the Met Office. Shaded regions correspond to channels that are principally sensitive to: temperature (red), the surface (yellow), ozone (green), water vapour (blue) and solar radiation (cyan).





					n	nax	= 2	20				
NH W50	۸	۸	۸	۸			٨					
NH W100	4	۸	۸									1
NH W250									۲		•	•
NH W500								•			•	•
NH W850								٠			•	
NH T50	4	۸	۸	۸								
NH T100		۸										
NH T250		٨	۸									•
NH T500									•			•
NH T850												
NH Z50	4	4	4	4								
NH Z100	4	Δ	۸									
NH Z250	4	4	۸						•		•	•
NH Z500	4								•			•
NH Z850								•	•	•	•	۲
TR W50		Δ	Δ	•		4		۸				
TR W100				1			•					
TR W250												
TR W500										-		
TR W850									•	•	•	•
TR T50						٠	٠	۲	۲		٠	
TR T100												
TR T250	-											
TR T500			1								3	
TR T850					-			-				
SH W50												
SH W100	-											
SH W250	1		-		-		-	-	-	-		
SH W500			1		-							
SH W850					-					-		
SH T50												
SH T100		-	-	1	-			·		-		-
SH T250	1	1						÷				
SH T500		-	1					-				
SH T850					÷	-		÷	t.	-	÷	•
SH 750	-											
SH 7100	-	A	A	<b>^</b>	<b>.</b>	<b>^</b>		· · · · ·	<u>.</u>			
SH 7250	<b>A</b>	^	<b>A</b>	<b>A</b>	<b>A</b>	<b>A</b>	•	•	•			•
SH 7500	<b>A</b>	^	^		-			÷	÷		÷.	
SH 7950	<b>.</b>	•	•		÷				÷	÷.	t.	•
JH 2000	4	*		<u> </u>	:	<u>.</u>			<u> </u>	· ·	A	A
	E E	F24	FЗ	+48	160	17.2	F82	96-	108	120	132	142
	Ę,	÷	ŕ	÷	÷	Ę,	Ę.	ŕ	Ě	Ť	÷	÷.

**Figure 4.7:** Scorecards showing the change in the RMSE for an experiment using smoothing of the logarithm of the bending angle proportional to half the spacing between model levels (left) and proportional to three-quarters the model-level spacing (right). Verification is performed against ECMWF analyses.



# 5 Conclusion

The motivation for this work originated in noting that observations that were processed by Spire provided improved forecasts in the Met Office system, compared to those that were processed by EUMETSAT Secretariat. The reason for this was felt likely to be connected to the increased level of smoothing that Spire applied to the profiles of bending angle. So a variety of experiments have been run which test the effect of additional smoothing being applied to the observations in order to understand whether this is the case.

The first experiments were performed by increasing the level of smoothing which was applied within EUMETSAT Secretariat's processing of bending angle observations made by Spire satellites. The aim here was to test increased levels of smoothing, with the largest smoothing being aimed at mimicking the smoothing applied within Spire's processing. The three levels of smoothing tested were referred to as 300, 600 and 900 m smoothing, which indicates the length-scale of the smoothing kernel applied to observations within the troposphere. It was found that the forecasts improved when using the additional smoothing and the largest improvements were seen with the 900 m smoothing.

The second set of experiments examined whether benefit would be seen when applying additional smoothing as a pre-processing step in the Met Office system. This has the advantage that it can be applied to all assimilated observations, not just those which have been processed by EUMETSAT Secretariat and allows the level of smoothing applied to be customised to the spacing between the levels in the Met Office's NWP model. The disadvantage of this method is that it is applied to the low-resolution bending angles, meaning that the local polynomial regression method will fail if the width of the smoothing kernel is too small. These experiments also demonstrated a positive impact of the additional smoothing, with an optimal level of smoothing being approximately three quarters the model-level spacing. With more smoothing the results in the tropics became worse, and with very high levels of smoothing the forecasts were degraded in all regions.

Analysis of the data assimilation statistics indicated that the analysis with additional smoothing gives improved fits to independent observations, which is leading to improved forecasts and therefore the background forecast also has a better fit to the independent observations. Thus it appears that the additional smoothing is removing noise from the GNSS-RO observations which allows the data assimilation system to perform better.

Although this study has investigated a number of areas around the impact of the smoothing of GNSS-RO observations, there remain a number of areas which would warrant further investigation. Chief among these questions is whether these results are unique to the Met Office's NWP system, or whether they generalise more widely. When applying the smoothing as a pre-processing step it was seen that the higher levels of smoothing gave negative results in the tropics, but a similar degradation was not seen in the initial study with EUMETSAT Secretariat's processing. Does this indicate that the limitations of applying the smoothing to low-resolution bending angle files has particular problems in the tropics, or does this indicate some other issue with the smoothing method used? The experiments which have been run thus far have concentrated on a period in the (northern hemisphere) autumn and winter. It would be of interest to test these results in a different period.

Further experimentation is planned to test the effect of additional smoothing on the ECMWF model. Given that the vertical resolution of the ECMWF model is approximately twice that of the Met Office model, it is expected that a lesser amount of smoothing would be optimal in that system. If this proves to be correct, then it would demonstrate that there is a need for flexibility in the smoothing of observations with different NWP centres able to customise the level of smoothing that they use. This could be achieved if data providers, such as EUMETSAT, make their observations available at high resolution. It would also require NWP centres to have the tools required to smooth and thin these high-resolution observations. Such tools may be something which is best provided by the ROM SAF.



## Acknowledgments

We would like to thank Ying-Hwa Kuo, Katrin Lonitz and Sean Healy for their helpful comments and suggestions during the course of this work.



# References

- Bowler, Neill E. (2020). "Revised GNSS-RO observation uncertainties in the Met Office NWP system". In: *QJRMS* 146, pp. 2274–2296. ISSN: 0035-9009. DOI: 10.1002/qj.3791.
- Burrows, C. P., S. B. Healy, and I. D. Culverwell (2014). "Improving the bias characteristics of the ROPP refractivity and bending angle operators". In: *AMT* 7, pp. 3445–3458. ISSN: 1867-1381. DOI: 10.5194/amt-7-3445-2014.
- Clayton, A. M., A. C. Lorenc, and D. M. Barker (2013). "Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office". In: *QJRMS* 139, pp. 1445–1461. DOI: 10.1002/qj.2054.
- Culverwell, I. (2021). "Anomalous GRAS radio occultations". In: *EUMETSAT ROM SAF Report* 40, pp. 1–23.
- Fan, J. and I. Gijbels (1996). *Local polynomial modelling and its applications*. Chapman and Hall / CRC.
- Gnanadesikan, R. and J. Kettenring (1972). "Robust estimates, residuals, and outlier detection with multiresponse data". In: *Biometrics* 28, pp. 81–124.
- Gorbunov, M. E. and K. B. Lauritsen (2004). "Analysis of wave fields by Fourier Integral Operators and their application for radio occultations". In: *Radio Sci.* 39. DOI: 10.1029/2003RS002971.
- Jensen, A. S. et al. (2004). "Geometrical optics phase matching of radio occultation signals". In: *Radio Sci.* accepted.
- Kay, S. (1989). "A fast and accurate single frequency estimator". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37.12, pp. 1987–1990. DOI: 10.1109/29.45547.
- Liu, Hui et al. (2020). "Analysis bias induced in assimilation of the radio occultation bending angle with complex structures in the tropical troposphere". In: *Quarterly Journal of the Royal Meteorological Society* 146.733, pp. 4030–4037. DOI: https://doi.org/10.1002/qj.3887.
- Loader, C. (1999). Local regression and likelihood. New York: Springer.
- Lonitz, K. et al. (2021). "Final Technical Note of "Impact assessment of commercial GNSS-RO data"". In: ESA Contract Report 4000131086/20/NL/FF/a, pp. 1–72. DOI: 10.21957/wrh6voyyi.
- Maronna, Ricardo A. and Ruben H. Zamar (2002). "Robust Estimates of Location and Dispersion for High-Dimensional Datasets". In: *Technometrics* 44.4, pp. 307–317. ISSN: 0040-1706. DOI: 10. 1198/004017002188618509.
- Marron, J. S. and D. Nolan (1988). "Canonical kernels for density estimation". In: *Statistics and Probability Letters* 7.3, pp. 195–199. ISSN: 01677152. DOI: 10.1016/0167-7152(88)90050-8.
- Payne, T. (2011). "Conjugate Gradient Algorithm and Hessian Eigenvector Preconditioning". In: Internal Met Office Report VSDP5, pp. 1–13.
- Samrat, Nahidul Hoque et al. (2025). "Observation impact evaluation through data denial experiments in the Met Office global numerical weather prediction system". In: *QJRMS*.
- Savitzky, A. and M. J. E. Golay (1964). "Smoothing and Differentiation of Data by Simplified Least Squares Procedures." In: Analytical Chemistry 36.8, pp. 1627–1639. ISSN: 0003-2700. DOI: 10. 1021/ac60214a047.
- Schafer, Ronald. W. (2011). "What is a Savitzky-Golay Filter?" In: *IEEE Signal Processing Magazine* 116.July, pp. 111–117.



Wand, M. P. and M. C. Jones (1995). Kernel smoothing. Chapman & Hall / CRC.



Reference	Report title
SAF/GRAS/METO/REP/GSR/001	Mono-dimensional thinning for GPS Radio Occulation
SAF/GRAS/METO/REP/GSR/002	Geodesy calculations in ROPP
SAF/GRAS/METO/REP/GSR/003	ROPP minimiser - minROPP
SAF/GRAS/METO/REP/GSR/004	Error function calculation in ROPP
SAF/GRAS/METO/REP/GSR/005	Refractivity calculations in ROPP
SAF/GRAS/METO/REP/GSR/006	Levenberg-Marquardt minimisation in ROPP
SAF/GRAS/METO/REP/GSR/007	Abel integral calculations in ROPP
SAF/GRAS/METO/REP/GSR/008	ROPP thinner algorithm
SAF/GRAS/METO/REP/GSR/009	Refractivity coefficients used in the assimilation of GPS radio occultation measurements
SAF/GRAS/METO/REP/GSR/010	Latitudinal Binning and Area-Weighted Averaging of Irregu- larly Distributed Radio Occultation Data
SAF/GRAS/METO/REP/GSR/011	ROPP 1dVar validation
SAF/GRAS/METO/REP/GSR/012	Assimilation of Global Positioning System Radio Occultation Data in the ECMWF ERA-Interim Re-analysis
SAF/GRAS/METO/REP/GSR/013	ROPP PP validation
SAF/ROM/METO/REP/RSR/014	A review of the geodesy calculations in ROPP
SAF/ROM/METO/REP/RSR/015	Improvements to the ROPP refractivity and bending angle operators
SAF/ROM/METO/REP/RSR/016	Simplifying EGM96 undulation calculations in ROPP
SAF/ROM/METO/REP/RSR/017	Simulation of L1 and L2 bending angles with a model iono-sphere
SAF/ROM/METO/REP/RSR/018	Single Frequency Radio Occultation Retrievals: Impact on Nu- merical Weather Prediction
SAF/ROM/METO/REP/RSR/019	Implementation of the ROPP two-dimensional bending angle observation operator in an NWP system
SAF/ROM/METO/REP/RSR/020	Interpolation artefact in ECMWF monthly standard deviation plots
SAF/ROM/METO/REP/RSR/021	5th ROM SAF User Workshop on Applications of GPS radio occultation measurements
SAF/ROM/METO/REP/RSR/022	The use of the GPS radio occultation reflection flag for NWP applications
SAF/ROM/METO/REP/RSR/023	Assessment of a potential reflection flag product
SAF/ROM/METO/REP/RSR/024	The calculation of planetary boundary layer heights in ROPP
SAF/ROM/METO/REP/RSR/025	Survey on user requirements for potential ionospheric products from EPS-SG radio occultation measurements
SAF/ROM/METO/REP/RSR/026	Estimates of GNSS radio occultation bending angle and re- fractivity error statistics
continued on next page	

### ROM SAF (and earlier GRAS SAF) Reports



continued

Reference	Report title
SAF/ROM/METO/REP/RSR/027	Recent forecast impact experiments with GPS radio occultation measurements
SAF/ROM/METO/REP/RSR/028	Description of wave optics modelling in ROPP-9 and suggested improvements for ROPP-9.1
SAF/ROM/METO/REP/RSR/029	Testing reprocessed GPS radio occultation datasets in a re- analysis system
SAF/ROM/METO/REP/RSR/030	A first look at the feasibility of assimilating single and dual frequency bending angles
SAF/ROM/METO/REP/RSR/031	Sensitivity of some RO measurements to the shape of the iono- spheric electron density profile
SAF/ROM/METO/REP/RSR/032	An initial assessment of the quality of RO data from KOMPSAT-5
SAF/ROM/METO/REP/RSR/033	Some science changes in ROPP-9.1
SAF/ROM/METO/REP/RSR/034	An initial assessment of the quality of RO data from Metop-C
SAF/ROM/METO/REP/RSR/035	An initial assessment of the quality of RO data from FY-3D
SAF/ROM/METO/REP/RSR/036	An initial assessment of the quality of RO data from PAZ
SAF/ROM/METO/REP/RSR/037	6th ROM SAF User Workshop
SAF/ROM/METO/REP/RSR/038	An initial assessment of the quality of RO data from COSMIC-2
SAF/ROM/METO/REP/RSR/039	Impacts of RO mission differences on trends in multi-mission data records
SAF/ROM/METO/REP/RSR/040	Anomalous GRAS radio occultations
SAF/ROM/METO/REP/RSR/041	Assessment of sensitivity of the ROM SAF 1D-Var solutions to various error covariance choices
SAF/ROM/METO/REP/RSR/042	A one-dimensional variational ionospheric retrieval for trun- cated GNSS Radio Occultation measurements
SAF/ROM/METO/REP/RSR/043	Applying the ROPP ionospheric 1D-Var retrieval to Metop extension data
SAF/ROM/METO/REP/RSR/044	An investigation into the impacts of the vertical smoothing of GNSS-RO bending angle observations on Met Office NWP forecasts
SAF/ROM/METO/REP/RSR/045	ТВА
SAF/ROM/METO/REP/RSR/046	8th EUMETSAT ROM SAF user workshop on GNSS radio occultation measurements
SAF/ROM/METO/REP/RSR/047	An initial assessment of the quality of RO data from FY-3E

ROM SAF Reports are accessible via the ROM SAF website: https://rom-saf.eumetsat.int.