A New Method to Correct Radiosonde Temperature Biases Using Radio Occultation Data

JORDIS S. TRADOWSKY

Bodeker Scientific, Alexandra, New Zealand, and National Institute of Water and Atmospheric Research, Lauder, New Zealand, and Freie Universität Berlin, Berlin, Germany

CHRIS P. BURROWS

Met Office, Exeter, United Kingdom

SEAN B. HEALY

European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

JOHN R. EYRE

Met Office, Exeter, United Kingdom

(Manuscript received 31 March 2016, in final form 22 February 2017)

ABSTRACT

A new method to estimate radiosonde temperature biases using radio occultation measurements as a reference has been developed. The bias is estimated as the difference between mean radio occultation and mean radiosonde departures from collocated profiles extracted from the Met Office global numerical weather prediction (NWP) system. Using NWP background profiles reduces the impact of spatial and temporal collocation errors. The use of NWP output also permits determination of the lowest level at which the atmosphere is sufficiently dry to analyze radio occultation dry temperature retrievals. The authors demonstrate the advantages of using a new tangent linear version of the dry temperature retrieval algorithm to propagate bending angle departures to dry temperature departures. This reduces the influence of a priori assumptions compared to a nonlinear retrieval. Radiosonde temperature biases, which depend on altitude and the solar elevation angle, are presented for five carefully chosen upper-air sites and show strong intersite differences, with biases exceeding 2 K at one of the sites. If implemented in NWP models to correct radiosonde temperature biases prior to assimilation, this method could aid the need for consistent anchor measurements in the assimilation system. The method presented here is therefore relevant to NWP centers, and the results will be of interest to the radiosonde community by providing site-specific temperature bias profiles. The new tangent linear version of the linear Abel transform and the hydrostatic integration are described in the interests of the radio occultation community.

1. Introduction

For decades, radiosonde (RS) profiles have been assimilated into numerical weather prediction (NWP) systems, and since 2006, radio occultation (RO) data have been assimilated, demonstrating a positive impact on weather forecasts (see Healy 2008a; Poli et al. 2010; Rennie 2010b). However, the impact of the high-quality RO data may be limited by opposing biases between the observation types. Furthermore, local and regional variations in RS temperature biases have the potential to cause false horizontal temperature gradients in NWP analyses, which would lead to spurious features in the wind field. Satellite radiance measurements require bias corrections when assimilated into NWP models, and these corrections are computed relative to the model background or analysis, either statically (Eyre 1992; Harris and Kelly 2001) or via variational methods (Derber and Wu 1998; Dee 2004; Auligné et al. 2007). This can only be done consistently if sufficient ''anchor'' measurements are present in the assimilation system.

Corresponding author: Jordis S. Tradowsky, jordis@bodekerscientific.com

DOI: 10.1175/JAMC-D-16-0136.1

Among these, RO and RS are key contributors, and therefore, ensuring consistency between the bias characteristics of these observation types is important for the stability of the assimilation systems.

This study describes a method to calculate the RS temperature bias correction on a station-by-station basis using high-quality RO profiles (see, e.g., Anthes 2011) as a reference. A forecast impact study using the method described here to correct the RS temperature biases prior to assimilation into an NWP system is planned, and the results will be published separately.

The RO variable used here is the bending angle (BA) as a function of the impact parameter (native RO coordinate), but some assumptions (see section 3) allow the retrieval of a dry temperature (Tdry) profile, which is a valid approximation of the physical temperature when water vapor effects are negligible. In contrast to other studies (see, e.g., He et al. 2009; Sun et al. 2010, 2013; Ladstädter et al. 2015), which analyze the difference between RS and RO profiles based on spatial and temporal collocations, here, short-range global forecast (i.e., background) fields from the Met Office global NWP system (Davies et al. 2005; Rawlins et al. 2007) are used as a transfer medium; we calculate the background departures [observation minus model background (O-B)] for RO and RS, respectively, and compare the mean departures of the two observation types. The BA background fields are computed from the model fields using the forward model described in Healy and Thépaut (2006) and Burrows et al. (2014). The method applied here has three advantages: (i) compared to direct observation-to-observation collocations, the influence of differences in time and space is minimized, since every measurement (both RO and RS) has a collocated model background profile; that is, the model is interpolated to the position and time of each measurement; (ii) the lowest level at which the model humidity is negligible can be determined for each RO profile, which enables the use of the retrieved Tdry departures as low in the atmosphere as is reasonable; and (iii) a tangent linear (TL) retrieval can be used. The knowledge of atmospheric humidity facilitates the investigation of the RS bias using Tdry from 10 hPa to a location-dependent pressure level of about 100 hPa in the tropics and 300–400 hPa in the high latitudes (see also Ladstädter et al. 2015). The bias calculated at the lowest altitudes might, however, not be representative for all atmospheric conditions, as it is calculated from a subset of RO profiles that are sampled in especially dry conditions. To avoid including a priori knowledge about the humidity in the RO retrieval chain, we calculate the bias corrections only at altitudes for which water vapor effects are negligible. For operational use in NWP systems, a gradual transition of the bias corrections below this altitude must be applied (see Tradowsky 2016) to avoid discontinuities in the assimilated temperature profiles. As the radiosonde bias tends to be largest in the stratosphere, we expect this approach to be reasonable. In contrast, for example, Sun et al. (2013) use Tdry down to 150 hPa globally and wet temperature retrievals, which include a priori knowledge from the National Centers for Environmental Prediction (NCEP) 12-h forecasts farther down in the atmosphere.

This paper is based on the results of the Radio Occultation Meteorology Satellite Application Facility (ROM SAF) visiting scientist projects number 26 [see Tradowsky (2015) for more details on the method] and number 31 [see Tradowsky (2016) for details on the preparation of the forecast impact study and on the structural uncertainty].

2. Observational data and the NWP system

a. Radiosonde data

More than 800 upper-air sites launch RSs on weather balloons to measure vertical profiles of temperature, humidity, and, depending on the RS type, pressure. Many manufacturers correct, among other things, the radiation bias in the RS temperature and humidity profiles before the data are released. Even after this correction, the temperature profiles can have substantial biases that vary with the solar elevation angle (SEA) but also depend on the postflight processing applied at the ground station. The RS temperature bias is commonly calculated per sonde type (see, e.g., He et al. 2009; Sun et al. 2010, 2013), though Milan and Haimberger (2015) found variations of the temperature bias for stations launching the same sonde type, supporting our approach to calculate the bias separately for each station. A thorough evaluation of sources of biases in RS measurements can be found in Dirksen et al. (2014). Here, we provide a method to correct the remaining temperature bias prior to assimilation into NWP models. A bias correction is calculated for most (for technical reasons, not all) land-based RS launch sites that regularly disseminated operational data in 2014. The results for five example sites (see Table 1) comprising different climate regimes are analyzed here.

The temperature bias is calculated at those standard pressure levels of RS profiles submitted in the alphanumeric "TEMP" (Ingleby and Edwards 2015) format (i.e., 1000, 925, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, and 10 hPa) that have negligible humidity. In the Met Office NWP system, TEMP profiles have a

TABLE 1. Country, latitude, longitude, World Meteorological Organization (WMO) station identifier (ID), and RS/radar type for the example upper-air sites analyzed in this study. The RS type ''unknown, not specified'' actually includes three Russian RS types, namely, I-2012, MRZ-3MK, and AK2m (B. Ingleby 2015, personal communication).

Country	Lat	Lon	ID	RS/radar type
Germany	52.22°	14.12°	10393	Vaisala RS92
Russia (west)	69.35°	88.27°	23078	AVK-AK2–02, MARL-A/Vektor-M AK2–02, MARL-A/Vector-M-BAR
Russia (east)	59.55°	150.78°	25913	MARL-A/Vektor-M AK2–02, AVK-BAR, MARL-A/Vector-M-BAR, unknown
Indonesia	−1.18°	136.12°	97560	Meisei
Antarctica	−69.0°	39.58°	89532	Meisei

cold bias of −0.05 K that is caused by the conversion from degree Celsius to kelvins (see Ingleby and Edwards 2015). In addition, the encoding/decoding of the RS profiles in TEMP format can cause an offset, for example, −0.095 K for the RS92 with DigiCORA III processing (see Ingleby and Edwards 2015), leading to a combined bias of almost −0.15 K in the RS92 temperature data within the Met Office NWP system.

In this study, only those RS profiles that passed the operational quality control of the Met Office observation processing system in 2014 are analyzed. Outliers in the RS dataset are rejected based on the median absolute deviation (MAD). To calculate the MAD, first, the RS departures for all profiles at one upper-air site are calculated. Then, the absolute deviation is calculated by subtracting the median of all RS O-Bs from each RS departure profile. The median value of the absolute deviation at one standard pressure level is the MAD for the given level. Based on Leys et al. (2013), a moderately conservative threshold of 2.5 is chosen to reject outliers, such that the RS temperature at a certain level is rejected if an RS O-B is more than 2.5 × MAD away from the median of all RS departures at that level. While mean and standard deviation (SD) are especially sensitive to outliers and are therefore problematic for their detection, the MAD presents a robust measure to reveal outliers.

b. Radio occultation data

The signal transmitted by the Global Navigation Satellite System (GNSS) at about 20 000-km altitude is received by a low-Earth-orbit (LEO) satellite. The measurements are made during the radio occultation event, that is, when the GNSS satellite rises above or sets behind the horizon. In this configuration, the signal passes through the limb of the atmosphere, where it is bent and delayed before it is received by the LEO satellite. The measured phase shift of the received signal allows the retrieval of atmospheric variables, mainly temperature and pressure, and given additional a priori knowledge about the atmospheric state, water vapor. Since the temperature and pressure values are retrieved assuming that water vapor is negligible, these quantities are often referred to as Tdry and dry pressure, respectively, by the RO community. A description of the RO technique can be found in Kursinski et al. (1997, 2000). Since the RO measurement is based on precise timing available from atomic clocks, it offers the possibility to be traceable to the international system of units [Système International d'Unitès (SI)] standard of time (Leroy et al. 2006). This ensures the long-term stability of RO data and makes them valuable for climate studies. In contrast to other remote sensing techniques, RO measurements are nearly independent of rain and clouds, and the profiles are retrieved from the higher atmosphere down into the boundary layer, with a vertical resolution of 0.1–1 km and a horizontal resolution of 100–300 km around the tangent point [see Fig. 3 in Anthes (2011)]. As the noise in the RO profile increases at high altitudes, the measured BA profile is usually merged with a climatological BA profile using statistical optimization in order to retrieve refractivity. This is applied above a certain altitude, for example, above 40 km in Healy (2001). Thus, the influence of the RO measurement in the retrieval of refractivity and temperature decreases with altitude, while the influence of the climatology increases.

This study uses the near-real-time BA data of the U.S.–Taiwanese Constellation Observing System for Meteorology, Ionosphere and Climate–Formosa Satellite Mission 3 (COSMIC–FORMOSAT-3, hereinafter COSMIC; see, e.g., Anthes et al. 2008). The profiles from both rising and setting occultations are used, as the influence of separating them becomes negligible given an appropriate sample size (see Tradowsky 2015). In a preprocessing step prior to assimilation into the Met Office global data assimilation system, the BA profiles are assessed for their quality. Central to this procedure is a one-dimensional variational (1D-Var) algorithm that uses collocated model background information to obtain an optimal solution for each observation profile. Complete profiles are flagged for rejection if (i) the 1D-Var assimilation fails to converge within 20 iterations; (ii) the initial cost function over all levels is greater than 2.5; or (iii) the final cost function over all levels, that is, at convergence, is greater than 2.0. The values in the

(conservative) rejection criteria and the number of iterations required were chosen based on experience. Approximately 10% of the RO profiles are rejected. Furthermore, BAs are rejected on a level-by-level basis if the absolute value of the observation minus the 1D-Var solution is greater than the assumed observation error multiplied by 5.

THE RO TEMPERATURE NULL SPACE

It should be emphasized that RO is not a direct measurement of temperature, and some a priori information is required to make a temperature retrieval well posed. The need for a priori information implies that there is a RO measurement "null space," meaning there are some atmospheric profile perturbations that do not affect the measured values. Conversely, information about these perturbations cannot be retrieved directly from the measurement alone (see Rodgers 2000, their section 2.2.1).

RO measurements in the dry atmosphere are sensitive to density ρ as a function of altitude h. The density, however, is proportional to the quotient of pressure and temperature ($\rho \propto P/T$). Therefore, a priori information is needed to split into the influence caused by $T(h)$ and $P(h)$.

As a consequence, temperature perturbations ΔT for which the amplitude grows exponentially with the density scale height ($H = RT/g$) as

$$\Delta T(h) = k \exp^{(h/H)}, \tag{1}$$

where k is a constant small perturbation, are difficult to detect with RO.

For example, assume two altitude levels h_1 (lower) and h_2 (upper) separated by Δh with a temperature variation $T(h)$ between the two levels. Perturbing the temperature based on Eq. (1) leads to increased temperature at h_2 and, therefore, the thickness of the layer increases, which leads to an increase of the pressure at level h_2. Thus, while $P(h_2)$ and $T(h_2)$ are changing, the ratio $P(h_2)/T(h_2)$ stays approximately constant, and thus, the density as a function of altitude is not significantly affected.

This temperature null space is a fundamental limitation of the RO technique and is distinct from the exponentially decreasing influence of the top-level pressure (see, e.g., Kursinski et al. 1997), which is calculated from an a priori temperature and the refractivity using Eq. (9).

Temperature perturbations that grow exponentially based on Eq. (1) quickly produce unphysical temperature values on altitude levels. However, more subtle perturbation patterns—partly composed of this exponential growth—remain potentially problematic. For example, the temperature bias highlighted by Steiner et al. (1999, see their Fig. 8b), produced by perturbing the a priori information used in their geophysical retrieval, is in the RO measurement null space.

Because of this RO null space, the technique presented here will only be able to estimate the contribution to the RS biases that the BA observations can determine uniquely. However, the assimilation of RO in NWP systems (Healy 2008b) and reanalysis (Poli et al. 2010; Simmons et al. 2014; Kobayashi et al. 2015) is seen to anchor the temperatures at around 100 hPa, which is an indication that RO is able to provide useful bias information at these levels.

c. *NWP system*

In this study, the Met Office global NWP system provides the model backgrounds for the RO and RS profiles. Using model fields as transfer medium reduces differences that would be caused by imperfect collocations.

During the investigated year, 2014, the model was due for an update of the dynamical core from the version described in Davies et al. (2005) to a version described in Walters et al. (2014). As a result, the resolution changed on 15 July from N512L70 to N768L70. This corresponds to a decrease in the grid length from about 25 km in midlatitudes to about 17 km. Also, the time step decreased from 10 to 7.5 min.

3. Method

The RS temperature bias corrections are calculated on a station-by-station basis, giving a vertical bias correction profile for each of the 762 studied sites ("site" and "station" are used synonymously). In this paper, a carefully chosen subset of five sites comprising different climate regimes and different sonde types is presented. The results for all sites are available as supplementary material to Tradowsky (2015). The bias correction profiles extend from 10 to at least 100 hPa and, for most of the nontropical stations, considerably lower into the atmosphere. The lower boundary is determined by the humidity in the background profile as described in section 3c. Since the bias corrections at the lowest levels are calculated from a subset of RO profiles sampled in especially dry air masses, these bias corrections might not be representative for all atmospheric conditions.

For each upper-air site, the RO BA profiles within 500 km of the site are selected, and the following are calculated: (i) the mean RO BA profile, (ii) the mean NWP background BA profile, and (iii) the mean BA departure (observation minus background) profile. Thus, instead of using only those RS and RO profiles that are closely collocated in time and space, all

occultations within 500 km of the site and all RSs launched at the site, independent of the sonde type, are analyzed. Increasing the vicinity radius has only a minor influence on the departure statistics [radii between 300 and 2000 km have been studied in Tradowsky (2015); see also Sun et al. (2010)]. There is only a slight tendency to increase the SD as the radius is increased. A radius of 500 km is therefore used to obtain a sufficient sample size while keeping the SD low.

The RS temperature biases depend on the solar elevation, as can be seen in, for example, Philipona et al. (2013), Sun et al. (2013), and Dirksen et al. (2014), and mean RO Tdry departures also show a slight dependence on the SEA, as can be seen in Tradowsky (2015). The small influence of solar elevation on RO BAs, which is caused by the ionosphere, propagates through to refractivity and temperature in the retrieval chain, as described by Healy and Culverwell (2015). Therefore, both datasets are divided into four SEA ranges, that is, high (SEA $> 22.5°$), low ($7.5° <$ SEA $< 22.5°$), dusk ($-7.5° <$ SEA $< 7.5°$), and night (SEA $< -7.5°$). RSs are typically launched at 0000 and 1200 UTC (at some stations, also 0600 and 1800 UTC), and the SEA at these times depends on geographic location and time of the year. Thus, for a given site, the separation by SEA may include preferential sampling for particular seasons.

To estimate biases in RS temperature profiles, the mean BA departure profile is propagated to a Tdry departure profile. First, the linear Abel transform is used to obtain inferred refractivity departures (see section 3a), and then, the hydrostatic integration of the refractivities gives Tdry departures as described in section 3b.

The RS temperature bias correction is calculated as the difference between the mean RO Tdry departures and the mean RS temperature departures, that is,

$$\overline{O_{\mathrm{RO}} - O_{\mathrm{RS}}} \simeq \overline{O_{\mathrm{RO}} - B_{\mathrm{RO}}} - \overline{O_{\mathrm{RS}} - B_{\mathrm{RS}}}, \qquad (2)$$

where O is the observation and B is the background, forward modeled into observation space. The $\simeq$ denotes that the assumption has been made that B_{RO} and B_{RS} are equally representative of the true values at the RO and RS locations, respectively (i.e., the central assumption is that the NWP forecast bias does not vary between the RO and RS locations). This is a more robust assumption compared to direct collocations between measurements, which are made assuming that the atmosphere does not vary over the separation distance. A similar double-differencing approach is used by Haimberger et al. (2012) to homogenize radiosonde temperature records.

In this investigation, a new TL Tdry retrieval, which calculates the Tdry departures from BA departures, is used. This retrieval is described here for the first time in the peer-reviewed literature, though it is based on the linear calculations developed in Syndergaard (1999), who originally proposed them to assess error propagation in the retrieval chain. The TL retrieval makes it possible to estimate the Tdry departures implied by any subset of BA departures. Here, the BA departures up to an impact height of 35 km are used as this is the altitude range of RSs. The BA departures above this height are set to zero. In practice, this enables us to reduce the influence of a priori information on our results. More detail about the retrieval can be found in Tradowsky (2015). Up to now, all Tdry retrievals use a combination of RO BA and a climatological/model BA at altitudes above 30–40 km (Ho et al. 2012, their Table 1) to minimize the effect of increasing noise in high-altitude RO BAs. Usually, this is done with "statistical optimization," which blends RO and climatological BAs, decreasing the weight of the RO BAs with increasing altitude. Depending on the processing center, different high-level initializations are applied, as it is described in Ho et al. (2012, their appendix A). Thus, no consensus about how best to initialize upper-level BAs exists in the community, but the spread between the different retrievals can be used to estimate the structural uncertainty (e.g., Ho et al. 2012; Steiner et al. 2013). The influence of the statistically optimized upper-level BAs propagates down into lower levels when refractivity and Tdry are retrieved. Thus, the whole retrieval depends on the climatological BA profile chosen by the processing center [see estimate of differences in, e.g., Ho et al. (2012) and Steiner et al. (2013)], although the influence of climatology decreases with decreasing altitude. The method proposed here eliminates the use of a climatological BA profile by using a cutoff of the BA departures above an impact height of 35 km as described in section 3d, which can be seen as an extreme case of the statistical optimization. As noted above, this altitude range is also a better match to the altitudes sampled by RSs.

a. (Tangent) linear Abel transform

BA profiles are retrieved from the measured phase shift of the GNSS signal, assuming local spherical symmetry of the atmospheric refractive index (see Kursinski et al. 2000). Under this assumption, the relation between the BA, α, and the refractive index, n as a function of tangent radius r, is described by the Abel transform (Fjeldbo et al. 1971):

$$n(r) = \exp\left[\frac{1}{\pi}\int_x^\infty \frac{\alpha(a)}{\sqrt{a^2 - x^2}}\, da\right], \qquad (3)$$

where $x = nr$ equals the impact parameter a for the ray with tangent radius r. The impact parameter for a given

ray represents the distance of closest approach of the undeflected part of the ray from the local center of curvature of Earth at the position of the ray tangent point [see, e.g., Fig. 1 in Kursinski et al. (2000) for occultation geometry].

For convenience, the refractivity N is defined as $N = 10^6(n-1)$. Assuming linear variation of the BA between successive observation levels, Syndergaard (1999) linearized Eq. (3), giving the refractivity as a function of the tangent radius, which is implicitly related to the refractive index as $r = a/n(r)$,

$$N(r) \approx \frac{10^6}{\pi} \int_x^\infty \frac{\alpha(a)}{\sqrt{a^2 - x^2}}\, da. \tag{4}$$

Discretizing Eq. (4) [see appendix A in Syndergaard (1999)] gives

$$\mathbf{N} = \mathbf{A}\boldsymbol{\alpha}, \tag{5}$$

with $\mathbf{A}$ [see Eq. (3.33) in Syndergaard (1999)] being an $m \times m$ triangular matrix, where m is the dimension of the profile.

A priori information is required to solve the Abel transform, as the upper limit of the integral is infinity, but the BAs are only available up to ≈60 km. Within this project, for each profile, the BA value at the highest level and the mean NWP background temperature at this level are used to initialize a (fairly arbitrary) extrapolation allowing the computation of the Abel integral up to infinity (for details, see Burrows and Healy 2016). The refractivity at the highest level is then used (again, with the background temperature at this level) to initialize the hydrostatic integral to obtain Tdry at all levels below. As a consequence of these two integrals, each retrieved value of Tdry contains information from every value of refractivity above, and each of these refractivity values contains information from the BA at the respective level and above, including the uppermost BA value and those extrapolated values defined by the uppermost BA. There is, therefore, considerable sensitivity of Tdry to the highest level of the BA profile. To minimize the influence of a priori, TL calculations as described below are used here.

As long as the departure from the linearization state is small, the linear approximation of a nonlinear function provides a good approximation of the function for small perturbations. TL models are used extensively in data assimilation systems (Hoffman et al. 1992), for example, to estimate the development of the atmospheric state at short time scales, as well as in observation operators. To build the TL retrieval code, the partial derivative of each equation is calculated in sequence.

In this study, observed and NWP background BAs are closely collocated, as the model is interpolated to the measurement position and time, and, therefore, are sufficiently similar to consider the matrix $\mathbf{A}$ to be equal for RO and model BA. Hence, the matrix $\mathbf{A}$ can also be used to propagate BA departures, since $\mathbf{N}_2 - \mathbf{N}_1 \simeq \mathbf{A}\boldsymbol{\alpha}_2 - \mathbf{A}\boldsymbol{\alpha}_1 = \mathbf{A}(\boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_1)$ This TL version of the Abel transform, where the BA departures $\delta\boldsymbol{\alpha}$ are used to calculate the refractivity departures $\boldsymbol{\delta}\mathbf{N}$, can be expressed in the same notation as Eq. (5):

$$\delta\mathbf{N} = \mathbf{A}\delta\boldsymbol{\alpha}. \tag{6}$$

The TL Abel transform, together with the TL version of the Tdry calculation, is used throughout this paper unless stated otherwise.

b. Tdry calculation

The relation between refractivity N, temperature T, pressure P, and water vapor pressure P_w is described in the Smith–Weintraub equation (Smith and Weintraub 1953):

$$N = c_1\frac{P}{T} + c_2\frac{P_w}{T^2}, \tag{7}$$

where $c_1 = 77.6\,\mathrm{K\,hPa^{-1}}$ and $c_1 = 3.73 \times 10^5\,\mathrm{K^2\,hPa^{-1}}$ are empirical constants.

To compute the pressure from the refractivity assuming a dry atmosphere, the hydrostatic equation is integrated using a gravitational acceleration g that varies with latitude and altitude. With the ideal gas law and assuming the refractivity varies exponentially with the geopotential height z, the contribution to the pressure from each discrete layer can be calculated as (for details, see Tradowsky 2015)

$$\Delta P = \frac{g}{Rc_1}\frac{N_i - N_{i+1}}{\ln(N_i/N_{i+1})}(z_{i+1} - z_i), \tag{8}$$

where R is the specific gas constant. The calculation of the pressure at each impact height level [impact height IH is the native coordinate for BAs and is related to the geometric height (altitude) h via IH $\approx h + 6.3 \times N(h)$; see Healy et al. 2007] is initiated at the highest level, which requires a priori knowledge about the temperature at this level. The a priori temperature used here is a mean value of the NWP background temperature at 59.06 km geopotential height, which is closest to the highest impact height level. The average includes background temperatures at 59.06 km for all those profiles that are used to calculate the RO O-B at a given site.

With this temperature, the top-level pressure is calculated as

$$P(z_{\text{top}}) = \frac{N(z_{\text{top}})T(z_{\text{top}})}{c_1}. \quad (9)$$

At altitudes where the water vapor contribution is negligible, Tdry is calculated from the pressure and Abel-derived refractivity with the dry term of Eq. (7) as

$$T_{\text{dry}}(z) = c_1\frac{P(z)}{N(z)}. \quad (10)$$

A TL Tdry calculation, comprising the Abel transform, the hydrostatic integration, and the Smith–Weintraub equation, is used here to calculate the Tdry departures (δT_{dry}) and can be expressed with the linear Tdry operator $\mathbf{K}$ and the BA departure as

$$\delta\mathbf{T} = \frac{\partial\mathbf{T}}{\partial\boldsymbol{\alpha}}\delta\boldsymbol{\alpha} = \mathbf{K}\delta\boldsymbol{\alpha}, \quad (11)$$

where $\mathbf{K}$ is the matrix of partial derivatives of Tdry with respect to the BA, assuming all a priori information required in the retrieval is fixed.

A calculation with the original nonlinear equations is done to provide linearization states, for the purpose of comparison, and to study the performance of the Tdry calculation (see appendix A).

c. Determining the lowest dry level

Since output from an NWP model is used as a transfer medium, it is possible to determine the lowest level of negligible humidity in the background fields separately for each RO profile, enabling us to use the Tdry profile as low in the atmosphere as is reasonable. The metric used to determine the deviation of the Tdry from the actual temperature is described in Scherllin-Pirscher et al. (2011) as $T_{\text{dry}} - T \approx -4/5 \times c_{\text{q2T}} \times q$, where q is the model specific humidity and $c_{\text{q2T}} = (c_2/c_1)/a_w = 7728$. The quantity a_w is the ratio of the dry air to water vapor gas constants, and c_1 and c_2 are the constants from Eq. (7). The level with lowest altitude above which the difference in physical and dry temperature remains below 0.09 K is determined for each profile. This value is chosen based on the suggestions in Scherllin-Pirscher et al. (2011), and also Fig. A1 (and analogous figures for different profiles, not shown here) supports this decision, since the model temperature and the model Tdry (computed from forward-modeled BAs) agree well in the lower levels that satisfy the criterion.

The BA departure statistics are calculated using BA values at all dry levels. If fewer than 10 dry profiles contributed to these statistics, then no subsequent analysis is performed as the averages may be unreliable.

d. Upper cutoff criterion for BA departures

The GNSS-RO Tdry retrieval is an example of an "exact retrieval" (Rodgers 2000), because it does not consider the impact of measurement noise and tries to fit the measured values exactly, including the noise they contain. As the noise in the RO BAs increases with altitude in the upper stratosphere and above, conventional retrievals of refractivity, and hence Tdry, apply statistical optimization, which blends the RO BA profile with a climatological BA profile, and the resulting profile is smoother than the original observed profile (see appendix B for a brief description of statistical optimization). Therefore, when statistical optimization is used in the retrieval algorithm, the measured BAs contribute less information to the retrieval at higher altitudes, where the climatology dominates. As stated above, the Abel transform and the hydrostatic integral both serve to propagate BA information down through the whole profile. Therefore, if statistical optimization is used, the climatology will influence the resulting Tdry retrieval even at levels below which the climatology has a finite contribution. This is not desirable when estimating biases of RS, so here, we intend to use the "exact" Tdry retrieval described above in a way that minimizes the impact of noisy or biased data. Therefore, a linear approach is used, which makes it possible to restrict the vertical range of BAs that contribute to the Tdry departures and, furthermore, avoids the use of prior information. By setting the BAs to zero above 35 km, the method bears similarity to an extreme case of statistical optimization where, in the BA departures, the observed BA is replaced with the background values above 35 km (or, equivalently, both observation and background are replaced by an independent climatology above this impact height). Here, Tdry departures that have been calculated from BA departures below 35 km are analyzed, which eliminates the influence of a priori assumptions and model bias above this impact height. To calculate RS bias corrections, we use a double-differencing approach, which is valid under the assumption that the biases in the model background profiles that are collocated with the RS observations are the same as those collocated with the RO observations. Therefore, it is important to use model information in the same altitude range for both observing systems. To reiterate, using the BA departures up to an altitude far above the RS balloon burst altitude would invalidate this central assumption, because the impact of any model bias at these uppermost levels would contribute to the Tdry departures below (see Fig. 1 and appendix B), while these

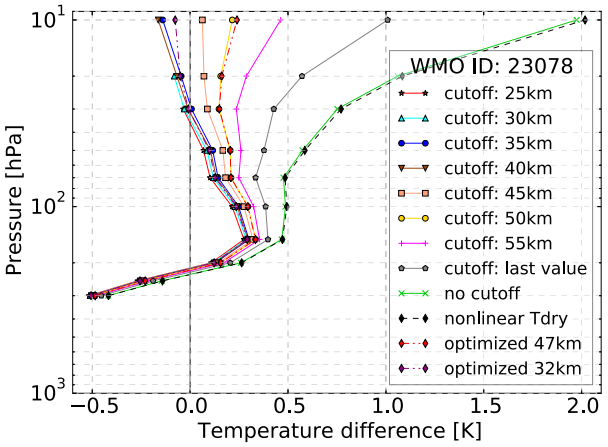

FIG. 1. Sensitivity of the mean Tdry departure to different upper cutoff impact heights. The mean Tdry departure is calculated from up to 842 RO profiles within a radius of 500 km around the example site 23078 in western Russia. Also shown is the Tdry departure calculated with the nonlinear retrieval (black dashed) and the Tdry departures calculated using an approach similar to statistical optimization with 50% of the background BA used at 32 (purple dash–dotted) and 47 km (red dash–dotted), respectively.

high-level biases will *not* contribute to RS temperature departures.

Burrows and Healy (2016, their Fig. 8) show how the global Tdry departures for different cutoff impact heights differ when using either the Met Office or the ECMWF BA background fields, highlighting the importance of minimizing any influence from the model, which will not be removed by the double differencing. When no cutoff is applied, the difference between the Tdry departures is >1 K at 10 hPa. The BA departure statistics of Met Office and ECMWF are very similar up to an impact height of about 35 km. This implies that the differences in the Tdry departures are caused by model BA biases of opposite sign at higher altitudes (see Fig. 7 in Burrows and Healy 2016) that are propagated down in the retrieval chain.

In Fig. 1, the TL Tdry departures for different cutoff impact heights (including no cutoff) are compared with Tdry departures calculated using BA profiles that are blended with climatology (similar to statistical optimization). In addition, Fig. 1 shows the nonlinear Tdry departures for the example site in western Russia. Tradowsky (2015) shows similar figures (excluding the statistical optimization) for all five example sites, showing that the departures for most stations tend to become more positive with increasing cutoff impact height [Burrows and Healy (2016) show that this is caused by the model background fields]. Note that the similarity of the green line (no cutoff) and the dashed black line (nonlinear) indicates the accuracy of the linearization. A major impact on the Tdry departures results from setting solely the BA departure value at the highest level to zero prior to the TL Tdry calculation (cf. green and gray lines). This large sensitivity is an artifact of using the RO BA value at the highest level to extrapolate the observation vector to infinity. In appendix B, this sensitivity is evaluated in detail by analyzing covariance and Jacobian matrices.

The dash–dotted curves show Tdry departures calculated using model/RO BAs that are blended with a smooth climatological BA (in this case, the background profiles are used), with 50% weight on a priori information at 47 km (red) and 32 km (violet), respectively. Conventional RO retrievals typically use statistical optimization to smooth the profiles, and we aim to illustrate what effect different implementations of statistical optimization can have on the retrieval (see appendix B for a description of the applied smoothing). When using 50% a priori at 32 km, the Tdry departures closely follow the line for a cutoff of 35 km (blue), with the biggest difference of about 0.1 K at the top level. Similarly, using 50% a priori at 47 km closely resembles a cutoff at 50 km. Thus, setting the BA departures above a certain altitude to zero gives results similar to blending the RO and model BA profile with climatology but does not require a priori information.

Like every assumption that is made in a retrieval, reasonable variations in the choice of the upper cutoff impact height contribute to the structural uncertainty. This is analogous to the structural uncertainty related to choosing a climatology/weighting to be used in the calculation of statistically optimized BAs (see difference in dash–dotted lines in Fig. 1). Ho et al. (2012) estimate the structural uncertainty resulting from different processing schemes as the differences and standard deviations between individual centers and the intercenter mean. For global RO Tdry calculated from different processing centers, Ho et al. (2012) find the mean differences to be between −0.27 and 0.15 K in an altitude range of 8–30 km. To give an estimate of the structural uncertainty caused by the choice of the cutoff impact height, the range (largest minus smallest value) of the departures calculated for cutoff impact heights between 35 and 55 km is shown in Fig. 2. The black line is calculated from ≈77 000 globally distributed COSMIC-6 BA profiles measured during 2014, while the other lines are calculated from all COSMIC BA profiles within 500 km around the respective upper-air site [the station identifier (ID) is given in the legend]. In general, the structural uncertainty increases with decreasing pressure. The structural uncertainty of the global statistics is calculated to estimate the typical behavior. It stays below 0.2 K at pressures above 50 hPa and increases to about 0.47 K at 10 hPa. Our estimation of the global structural uncertainty resulting from choosing an upper cutoff is thus similar to the structural uncertainty resulting from

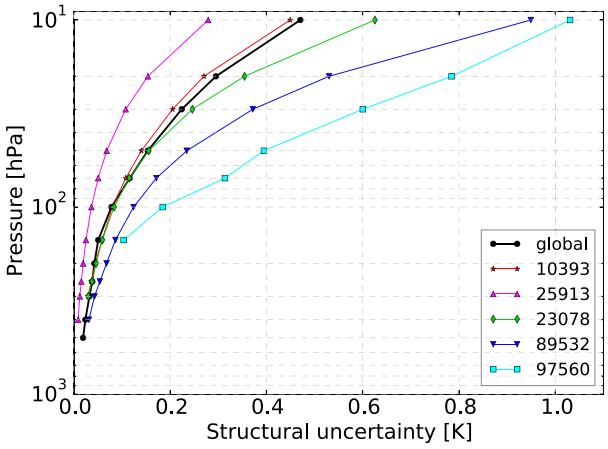

FIG. 2. Estimate of the structural uncertainty in Tdry departures for five example stations and for the mean global Tdry departures calculated from ≈77 000 COSMIC-6 profiles. The structural uncertainty is calculated as the range (largest departure − smallest departure) of Tdry departures inferred from BA departures with different upper impact height cutoffs between 35 and 55 km.

different processing as presented in Ho et al. (2012). The structural uncertainty is appreciably lower at the Russian upper-air site with the station ID 25913 and appreciably larger at the Indonesian site (ID 97560). Strong vertical gradients, caused by gravity waves above this tropical site, could explain the strong dependence on the cutoff impact height, but also, the comparably small sample size (≤242 ROs, as compared with ≤1093 for site 25913) could affect the results. Estimating the structural uncertainty is of major importance for climate studies. For the purpose of estimating RS biases, the spread of the departures between 35 and 55 km might, however, be an overestimation of the uncertainty, as a cutoff impact height of 55 km would invalidate the central assumption for the double-differencing technique; that is, the model bias is constant over the separation distance (see also Burrows and Healy 2016).

In summary, the application of the cutoff at 35 km is justified for two reasons: First, this reduces the effect of noisy and arbitrarily extrapolated RO data from the upper stratosphere and mesosphere that would otherwise affect all calculated bias corrections. Second, it avoids the influence of model biases above the RS altitude range, which, if the cutoff was not applied, would result in biased RO departures whose influence would persist even after the double differencing.

4. Comparison of departure statistics for RS and RO

The RO Tdry departures interpolated to standard pressure levels (blue), the RS temperature departures (pink), and the bias corrections (green) are presented in Figs. 3–7 for five example upper-air sites (see Table 1). Since the RS and RO departure statistics depend on the SEA, the bias and, hence, the bias correction profile (the bias correction is shown in the plots, identical to the bias but opposite sign) is calculated separately for high, low, dusk, and night launches as defined in section 3. The departure statistics and the bias corrections are displayed with error bars representing the associated sampling uncertainty [standard error (SE) of the mean, see Eqs. (B2) and (B3)], taking into account the sample size (dashed blue and pink lines) and the SD of the means. An estimate of the structural uncertainty, which is not included in the error bars, is given in Fig. 2. The horizontal cyan-colored line in Figs. 3–7 marks the highest standard pressure level (lowest altitude) where at least 95% of the RO profiles are used; that is, not more than 5% of the profiles are excluded because of atmospheric humidity exceeding the threshold defined in section 3c or because they failed the quality check. Below this altitude, the bias corrections are not representative for all atmospheric conditions at the respective site, as the selected RO profiles are sampled in especially dry conditions. This could lead to particular bias characteristics in the RO departures, probably originating from the model, which would invalidate our central assumption as no masking is applied for RS. Thus, while the RO departures below the horizontal cyan line in Figs. 3–7 are valid for certain atmospheric conditions, the resulting bias correction cannot be used to correct all RS profiles for the site.

From the five example sites shown here, the RS temperature bias correction is smallest at the German site (Fig. 3; bias < 0.5 K), which uses a Vaisala RS92 sonde, and at the Antarctic station (Fig. 7; bias < 1 K) using Meisei sondes. At the German example site, we find a small warm bias at 10 hPa for high SEAs. Also, Ladstädter et al. (2015) found this warm bias, however, with a stronger extent. Between ≈200 and 30 hPa, our analysis shows a small negative RS bias, which is not obvious in Ladstädter et al. (2015), who analyzed the differences in RS and RO for the years 2002–13. The differences in the results between this study and Ladstädter et al. (2015) might be caused by the choice of analyzed years, as the vendor correction in the RS92 profiles might have changed. Furthermore, as described in section 2a, RS92 TEMP profiles are almost 0.15 K too cold in the Met Office global NWP system, which partly explains the negative bias found in our analysis. Although the Indonesian site (Fig. 6) launches Meisei sondes like the Antarctic example site, the bias corrections are larger, reaching about 1.3 K. This indicates that the RS bias does not depend solely on the RS type (see also Milan and Haimberger 2015) but also on the station, possibly related to varying ground station software, different climate regimes [Sun et al. (2013) analyze the RS bias for different latitude bands], and the season in

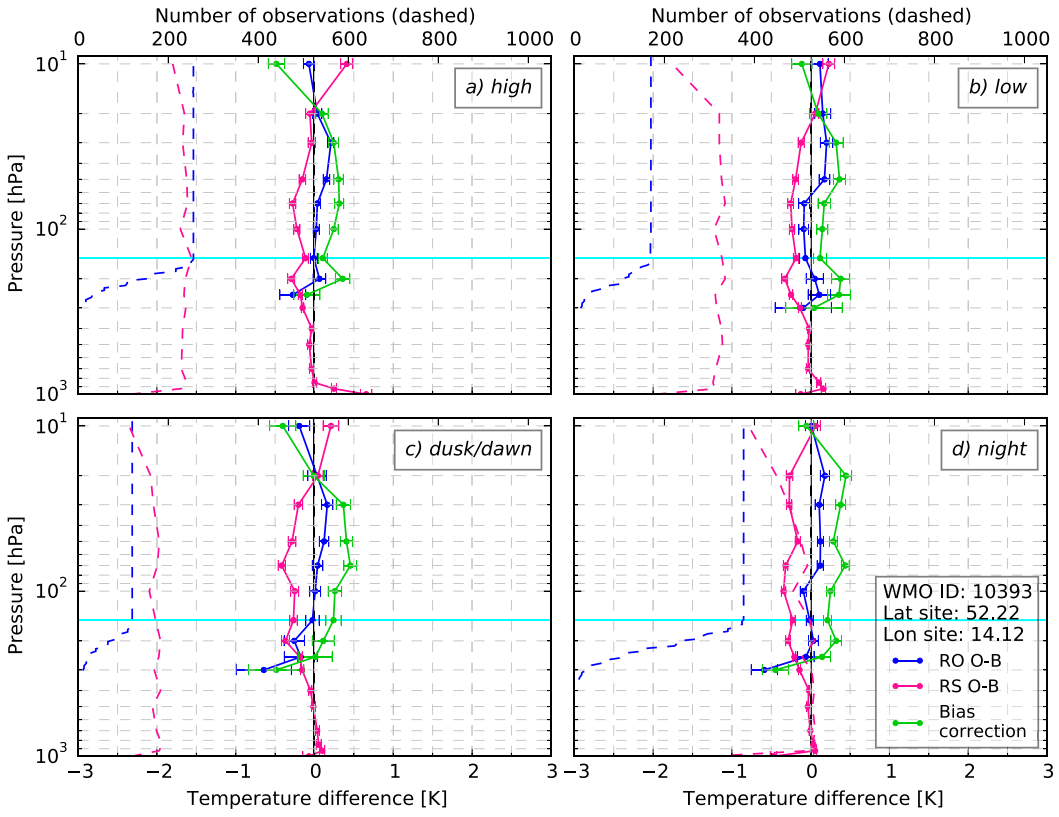

FIG. 3. Mean RO Tdry departures (blue), mean RS temperature departures (pink), and bias correction (RO O-B − RS O-B; green) at the example site in Germany for different SEA ranges: (a) high, (b) low, (c) dusk/dawn, and (d) night. The horizontal cyan line indicates the highest standard pressure level (lowest altitude) where at least 95% of the RO profiles are included. The error bars represent the SEs (sampling uncertainty) as calculated in Eqs. (B2) and (B3). An estimate of the structural uncertainty is provided in Fig. 2.

which the profiles are sampled. This finding supports our approach to calculate the bias correction on a station-by-station basis rather than based on the RS type.

While the RS bias correction at the site in western Russia (Fig. 4) stays below 0.7 K for dusk and night and is only larger at the lowest pressure level for high and low SEAs, a larger bias correction is needed at the site in eastern Russia (Fig. 5). Here, the bias correction is largest for dusk and night, reaching values up to about 2.5 K. Interestingly, the RS temperatures at many Russian sites show cold biases, leading to a positive bias correction being required for all SEAs and at most levels. This is in contrast with the theoretically expected radiation bias, that is, a warm bias during daytime and a cold bias during nighttime. Our results, however, agree with the findings by Rennie (2010a) for RSs tracked with the Russian "AVK" radar. The cold bias that prevails for all SEAs could be caused by the correction of biases in the ground system software. In general, the raw RS profiles are not disseminated, but instead vendor-corrected profiles are supplied for which the applied corrections are not always traceable.

The temperature bias correction is calculated here on a station-by-station basis, disregarding which type of sonde/radar is used. This is a compromise made to achieve a sufficient sample size for statistical significance. Though it may lead to problems in the operational use for individual cases (e.g., if a Vaisala sonde would be launched at a site with a bias correction of +2.5 K), we expect that applying the bias correction operationally will improve the RS temperature profiles on average. The RS bias corrections are calculated from the highest dry standard pressure level (lowest altitude) to the lowest standard pressure level (highest altitude) where enough profiles are available. Thus, bias correction profiles are calculated to a minimum pressure level of 10 hPa. For radiosondes that reach lower pressures, Tradowsky (2016) describes how the bias correction is extended upward for the use in a forecast impact study.

5. Summary and conclusions

A method to estimate the RS temperature bias based on a double-differencing approach of RS and RO

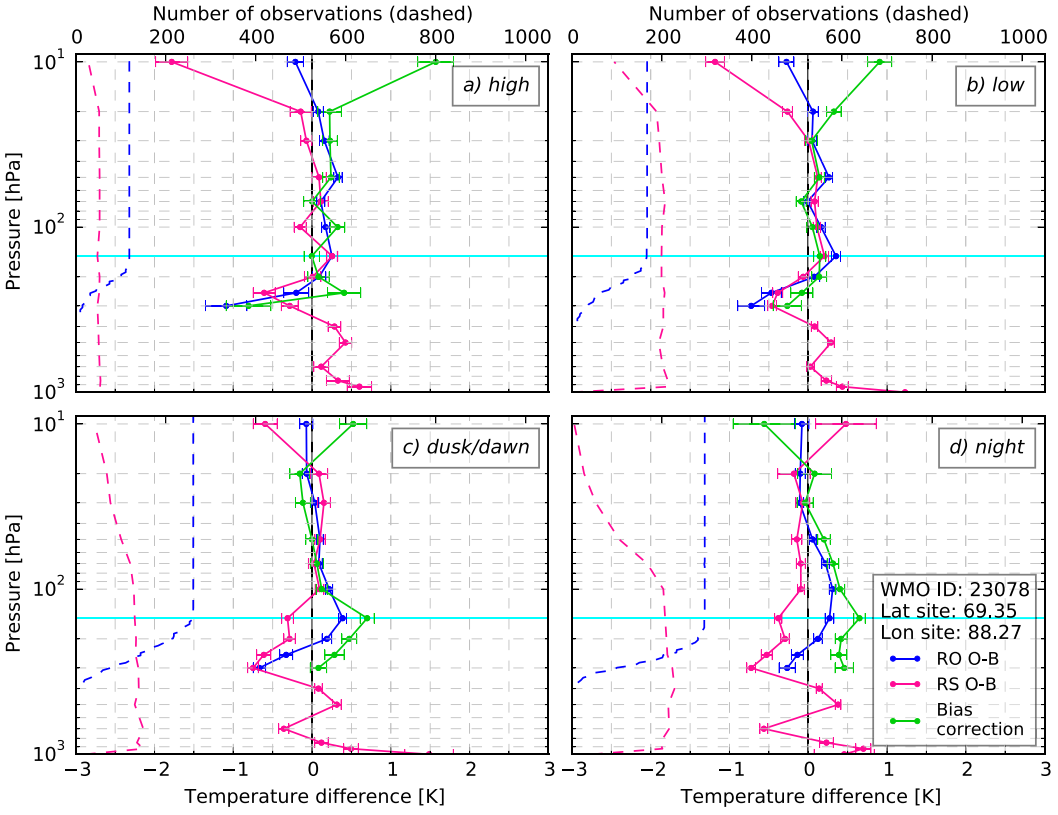

FIG. 4. As in Fig. 3, but for the example site in western Russia.

background departure statistics is presented, and the bias corrections are shown for five carefully chosen example sites.

First, RS departure statistics for a given site are calculated as the mean difference between the RS temperature and the NWP system background temperature. Similarly, the RO Tdry departure statistics, including all profiles within 500 km of the launch site, are calculated from the BA departures with a linear retrieval algorithm. Then, the difference between the RO and RS departure statistics (RO O-B − RS O-B) estimates the temperature bias correction to be applied to RS temperature profiles, thereby using the RO measurement as an unbiased reference. In this method, the NWP fields serve as a transfer medium, and this reduces the errors caused by imperfect collocation. The approach is based on the assumption that the bias in the NWP system does not vary within the vicinity radius of 500 km. As compared with the assumption of a nonvarying atmosphere, which is implicitly used for direct observation-to-observation collocations, this assumption leads to relatively small SDs. The double-differencing technique has two further advantages: (i) model humidity information allows the determination of the lowest level where RO Tdry can be used; and (ii) it is possible to use a TL retrieval of Tdry departures from BA departures, which reduces the sensitivity to a priori information.

In contrast to the nonlinear calculation of Tdry profiles from BA profiles, the TL version calculates Tdry departures from BA departures. As the aim is to use model background information from the same vertical range for both observation types, BAs are used from the lowest sufficiently dry level in the atmosphere to an impact height of 35 km. The BA departures above 35 km are set to zero, which limits the influence of a priori information and model biases at higher levels. The applied cutoff is comparable to an extreme case of statistical optimization, which uses 100% measured/forward-modeled BA values below 35 km and 100% climatological BA above 35 km in both the measured and the forward-modeled profiles. As the mean Tdry departures and their covariances depend on the choice of upper impact height cutoff (which is similar to how the conventional RO retrieval depends on the weighting in the statistical optimization), a detailed investigation of the cutoff is shown in appendix B.

In theory, a positive RS temperature bias is expected during daytime from solar radiation, while the emission of radiation by the RS can cause a negative bias during nighttime. Most RS manufacturers account for these

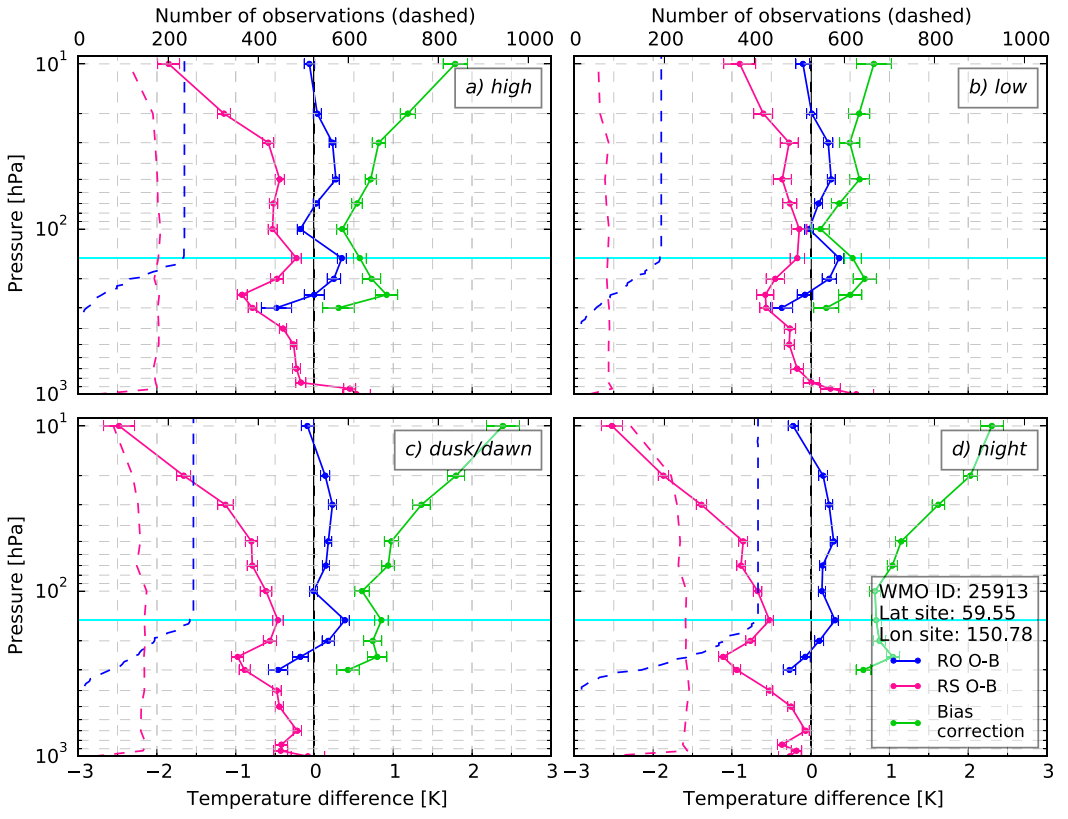

FIG. 5. As in Fig. 3, but for the example site in eastern Russia.

biases in the ground station software and disseminate a corrected RS profile. Such corrected profiles are analyzed here, which might explain why the sign of the bias does not always agree with the expectation based on theory. This is especially clear for the analyzed Russian upper-air sites where the temperature bias tends to be negative for all SEA ranges and can exceed −2 K at 10 hPa.

Depending on the station, the RS type, and the SEA, the magnitude of the bias varies. For some stations, the estimated temperature bias stays below ±0.5 K throughout the whole profile, while it exceeds ±2 K at others stations. An increase of the bias with increasing altitude is found for many stations, especially those that show large biases. Sites launching the Vaisala RS92 sonde, which is often used as reference (see, e.g., Agustí-Panareda et al. 2009), tend to show a small negative bias in the lower levels and a positive bias in the highest level during daytime and a slightly negative bias at most levels during nighttime. The bias is calculated for each upper-air site separately, and differences in the bias characteristic occur even if the same sonde type is launched, as was also found by Milan and Haimberger (2015). This can be caused, for example, by different versions of the vendor software. As described in Ingleby and Edwards (2015) and in section 2a, the bias correction is not fully independent of the NWP system, as the TEMP encoding–decoding and the temperature conversion introduces a cold bias for RS in the Met Office NWP system.

The technique presented here provides the basis for developing bias corrections for all operational RS stations (see Tradowsky 2016), which will then be applied in a forecast impact study using the Met Office global NWP system. Within the forecast impact study, for each observed RS profile, the bias correction for the corresponding site and SEA would be interpolated onto the observation levels. The interpolated bias correction can then be added to the observed temperatures prior to assimilation. The bias corrections described here are only available in the stratosphere down to an altitude where humidity starts to become significant. Tradowsky (2016) describes how the bias correction below this altitude is reduced smoothly to a Met Office specific default value in the preparation of a forecast impact study.

The method presented here can also be used to investigate how the reference-quality RS profiles from the Global Climate Observing System Reference Upper-Air Network (GRUAN; Global Climate Observing System 2007) compare to RO measurements (see Tradowsky 2016).

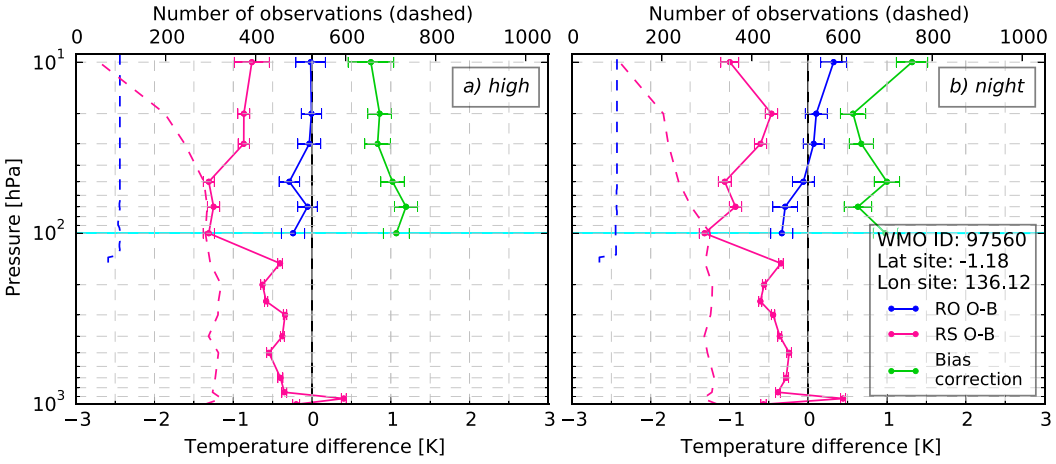

FIG. 6. As in Fig. 3, but for the example site in Indonesia.

In operational weather forecasts, the method presented here could be used to calculate the bias corrections on a regular basis. It could replace the Hawson correction (Hawson and Caton 1961) that is currently applied for some RS stations in the Met Office NWP system. The correction of RS temperatures prior to assimilation should serve the need for consistent measurements to anchor NWP models. Anchor measurements are not bias corrected as part of the data assimilation cycle, and Eyre (2016) shows the importance of having a sufficient number of observations to anchor the system. There are two aspects of anchoring measurements within the scope of NWP; that is, they provide the mean anchor to which the mean analysis is pulled, and they provide local anchors. Using the proposed method, the mean anchor is provided by the mean

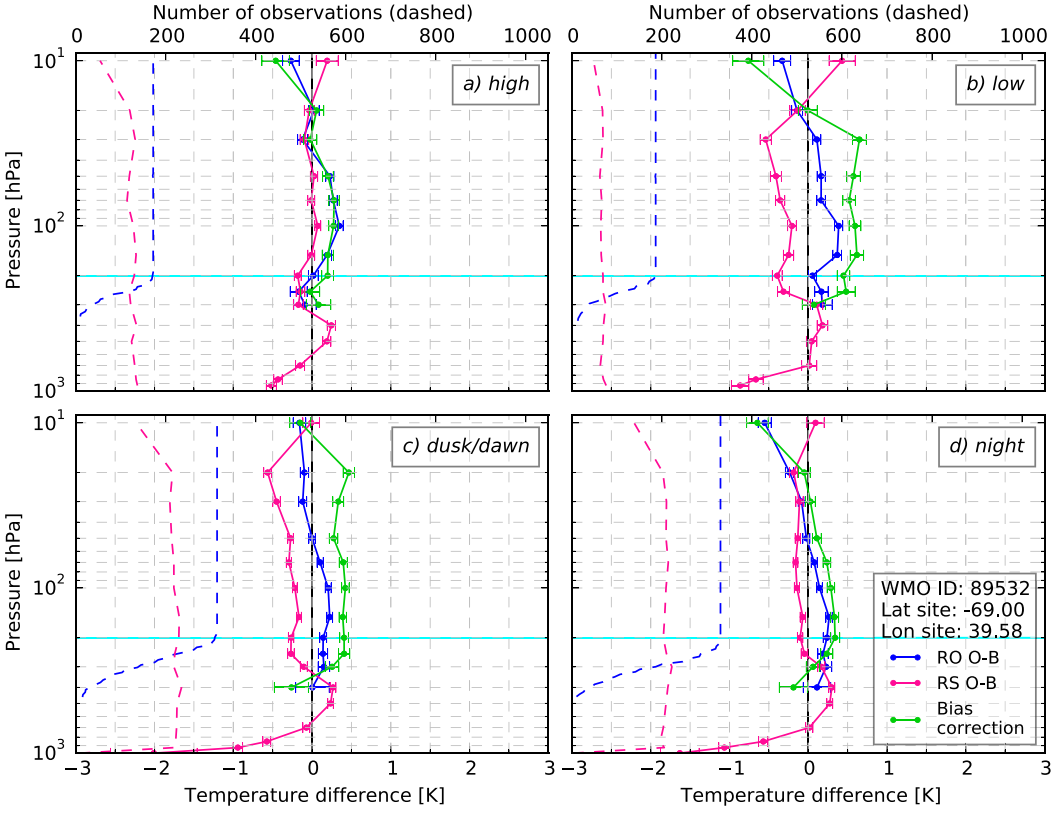

FIG. 7. As in Fig. 3, but for the example site in Antarctica.

RO observations to which the mean RS temperatures, averaged over all stations, are corrected. The local anchor, however, is provided by both the RO and RS observations. Enhancing the consistency between the observation types that anchor the temperature in a data assimilation system has the potential to improve the NWP skill, and further investigation will evaluate the performance of the proposed method.

Acknowledgments. This study was performed as a ROM SAF Visiting Scientist project, with J. S. Tradowsky as Visiting Scientist. C. B. Burrows and S. B. Healy are members of the ROM SAF, which is a decentralized facility of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). The authors are grateful that the ROM SAF enabled this study. We are thankful to Bruce Ingleby, who provided knowledge about RS; Axel von Engeln, who reviewed the ROM SAF Visiting Scientist report; Greg Bodeker, who provided helpful comments about the draft paper; and Ian Culverwell, who supported the study with his knowledge about statistics. The first author expresses her gratitude to Peter Builtjes, Jürgen Fischer, Greg Bodeker, and Richard Querel, who supervise her work, and to the German Academic Exchange Service for supporting her with a doctoral research grant. Furthermore, the authors are thankful to three anonymous reviewers whose comments improved this paper.

APPENDIX A

Performance of the Tdry Calculation

Since the model BA is calculated from the model variables as described in Burrows et al. (2014), it is possible to test the performance of the nonlinear Tdry calculation by comparing one model background temperature profile with the Tdry profile retrieved from the associated forward-modeled BA profile. If the assumptions in the forward (temperature to BA) and inverse (BA to temperature) calculations were identical, the differences of model temperature and model Tdry would be negligible in the dry atmosphere. But, because of differing assumptions, differences that vary with altitude are expected.

Figure A1 shows one typical profile of the background temperature (blue) and the Tdry (pink) retrieved from the associated background BA profile in the vicinity of the German example site (see Table 1). The profiles comprise pressure levels with negligible humidity, as determined following section 3c. Temperature differences of about 5 K around 1 hPa occur with decreasing magnitude toward lower altitudes (<2 K at 10 hPa).

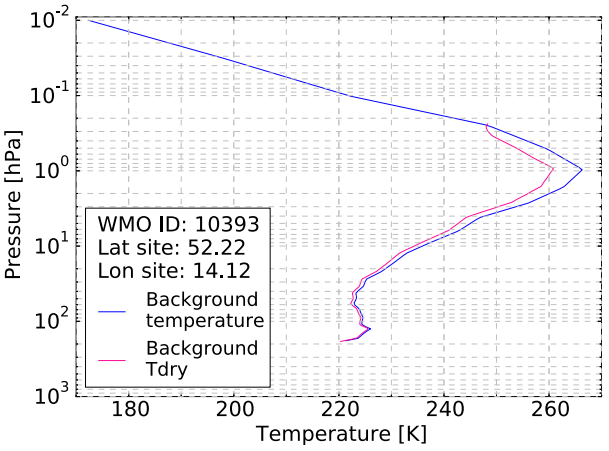

FIG. A1. Background temperature (blue) and Tdry (pink) for one individual profile from around the German example site. Tdry is calculated from the background BA using the nonlinear Tdry calculation.

The differences are caused by different assumptions in the forward and inverse calculations, especially about the variation of the quantities between levels [cf. Burrows et al. (2014) with section 3a] and above the highest level. In this project, the BA is assumed to fall exponentially with altitude above the uppermost observation in the Tdry retrieval (which is equivalent with the assumption of an isothermal atmosphere above an impact height of 60 km).

The highest level Tdry is plotted at a slightly too-high dry pressure compared to the model pressure, which is caused by different assumptions in the calculation of pressure in the model and Tdry calculations. Otherwise, the top-level Tdry agrees well with the background temperature, emphasizing the value of initializing the hydrostatic integration at the highest level with the model temperature as described in section 3b. Although some differences in the model temperature and Tdry are present, resulting from differences in the assumptions, it can be concluded that the nonlinear Tdry calculation as described in section 3 and further discussed in Burrows and Healy (2016) performs well.

APPENDIX B

Analysis of Sensitivity to the Upper Cutoff Height

a. Tdry departure covariance matrices for different upper cutoff impact heights

The method presented here uses a double-differencing approach that relies on the assumption of a constant model bias within the separation distance of the RO and RS measurements. Therefore, the profile of BA departures in the approximate altitude range of

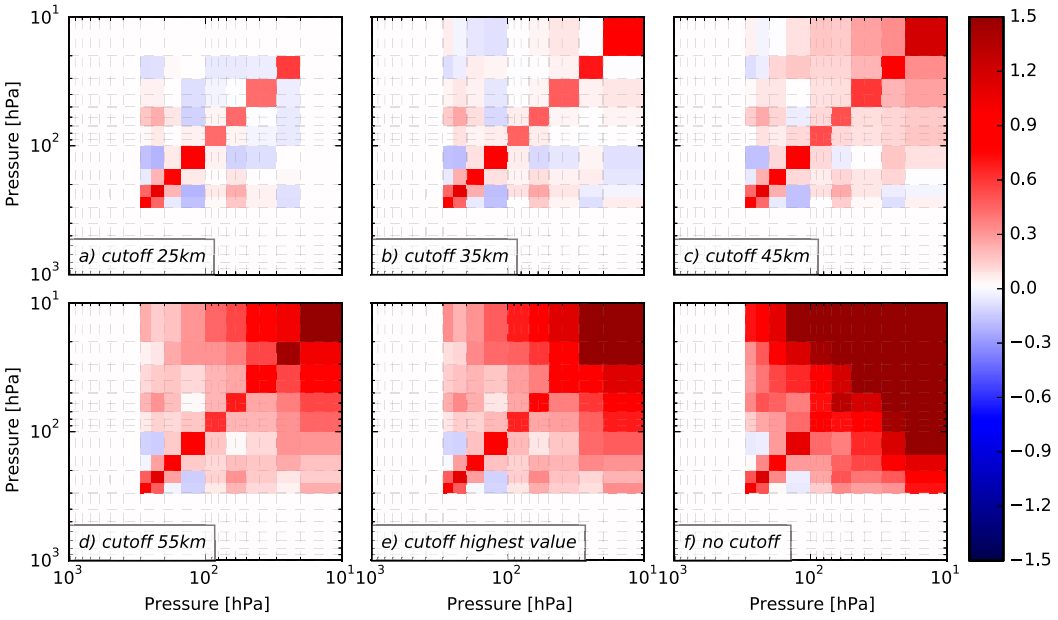

FIG. B1. Tdry departure covariance matrices for different upper cutoff impact heights; example site in western Russia. Color scale for temperature squared (K^2) is adjusted to a cutoff at an impact height of 35 km, which means that the colors are saturated for the higher cutoffs (they reach values of ≈36 K^2).

the RS departures must be propagated into Tdry departures, thus using model background information within the same altitude range for both observation types. The BA departures above the approximate burst altitude of RSs are set to zero to achieve this. To ensure an optimal choice of this upper cutoff impact height and to understand the effect of setting high-level departure values to zero, the respective covariance matrices are analyzed here. An estimate of the population covariance using a finite sample can be calculated with Eq. (B1), where x and y are variables, $\overline{x}$ and $\overline{y}$ are the sample means, and n is the sample size:

$$\mathrm{Cov}(x,y) = \frac{\sum_{i=1}^{n}(x-\overline{x})(y-\overline{y})}{n-1}. \qquad \text{(B1)}$$

The departure covariance is first calculated in BA space and is then propagated through the TL Tdry calculation to derive the Tdry departure covariances $\mathbf{C}_{\mathrm{Tdry}}$ as $\mathbf{C}_{\mathrm{Tdry}} = \mathbf{K}\mathbf{C}_{\alpha}\mathbf{K}^{\mathrm{T}}$, where $\mathbf{K}$ is the linear Tdry operator in matrix form, $\mathbf{C}_{\alpha}$ is the covariance matrix of BA departures, and $\mathbf{C}_{\mathrm{Tdry}}$ is the inferred covariance matrix of Tdry departures. The diagonal elements of the covariance matrix are the variances (squared SDs), while the nondiagonal elements indicate the covariation of the Tdry departures at one level with those at other levels.

Figure B1 shows the Tdry departure covariance matrices calculated from different subsamples of the BA departure profile (i.e., different upper cutoff impact heights). The color scale suits a cutoff at 35 km, which leads to saturation of the color for the panels with higher impact height cutoffs. The variances and covariances increase as the cutoff impact height is raised (from Fig. B1a to Fig. B1f) and a large influence is caused by setting the highest BA departure value to zero (cf. Fig. B1e with Fig. B1f). The sensitivity to the BA at the highest observation level is caused by arbitrary assumptions about the behavior of the BAs above the measured profile. The sensitivity to other BAs close to the top is mainly caused by noisy BA departures at high altitudes, which potentially contain significant biases from the NWP backgrounds. The Tdry departures at each level depend on the refractivity departures at *all* levels above, and each of these refractivity departures depend on *all* BA departures above. This causes the sensitivity to the (potentially large) BA departures at the top of the original profile; see also section 3a. Decreasing the upper cutoff impact height diminishes this dependence on the uppermost departures and therewith reduces the variance and covariance.

Using the TL Tdry calculation enables us to calculate the Tdry departures from a subset of BA departures, while the nonlinear version would not allow this flexibility. Since RSs only reach pressure levels of about 10 hPa, the influence of the NWP system and a priori information from altitudes above 10 hPa should be minimized. While a cutoff at 25 or 30 km is too low to allow comparisons to be made at 10 hPa, the impact

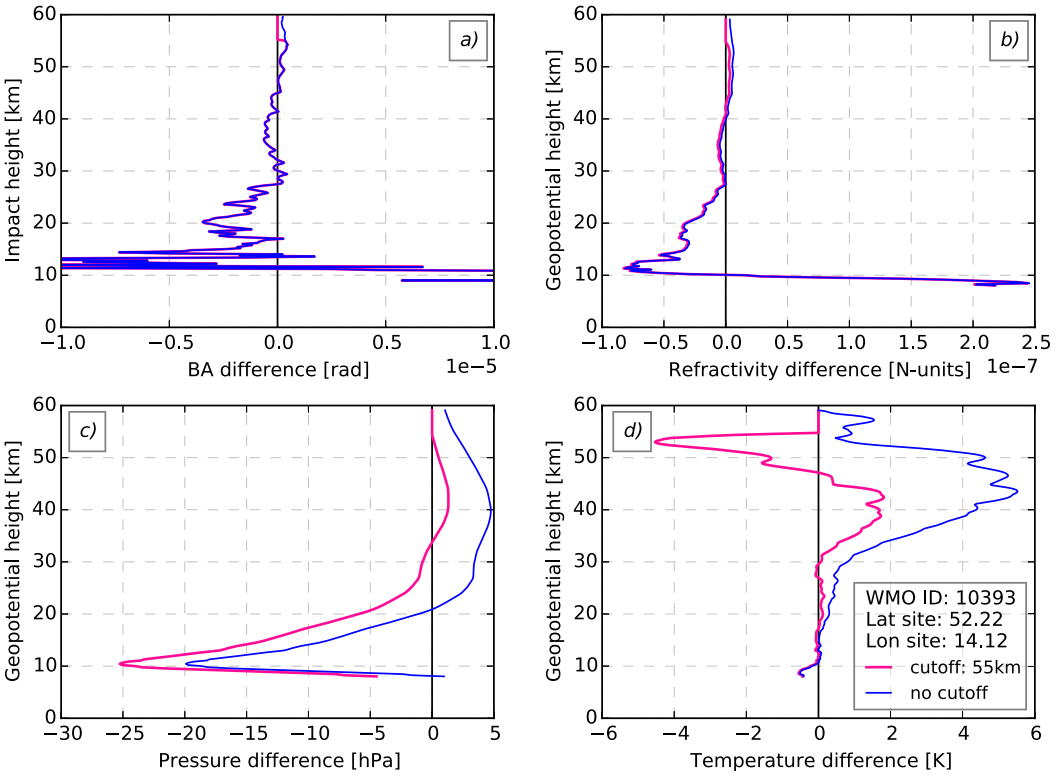

FIG. B2. The different steps in TL Tdry calculation for no cutoff (blue) and 55 km cutoff (pink), German example site. (a) The BA departures, (b) the inferred (TL) refractivity departures, (c) the inferred (TL) dry pressure departures, and (d) the inferred (TL) Tdry departures.

height of 35 km appears to be a good choice as variance and covariance are low and, importantly, these statistics are not significantly sensitive to small changes in the cutoff impact height at around 35 km (see also Burrows and Healy 2016). Therefore, the subset of the BA departures below the impact height of 35 km is used in this study, and the BA departures at higher levels are set to zero. The influence caused by choosing an upper cutoff is comparable to the impact of the choice for a high-level initialization (i.e., choice of climatology and smoothing) in the conventional retrieval chain based on statistical optimization (see Ho et al. 2009, 2012; Steiner et al. 2013).

The nonlinear Tdry calculation would, at all but the lowest levels, be highly dependent on the uppermost BAs and lead to large SDs, which reach up to ≈7 K at the upper-air sites shown in Table 1 (figures not shown here). Setting the highest BA departure to zero at the least halves the SD at the 10-hPa level, and further reducing the cutoff impact height decreases the SD as the Jacobians are unable to propagate information from noisy high-level BA departures downward. Given a cutoff of 35 km, the SD remains under 1 K for most example stations and at most levels. Reducing the SD (σ) positively affects the SE calculated as

$$\mathrm{SE} = \sigma/\sqrt{(n-1)}. \tag{B2}$$

The SE indicates the quality of the estimation of the mean. Knowing the SE of the bias correction is important to decide whether a bias correction is needed. The SE of the bias correction is calculated following Burns and Dobson (1981), as

$$\mathrm{SE}_{\mathrm{bias\_correction}} = \sqrt{\frac{\sigma_{\mathrm{RO}}^2}{n_{\mathrm{RO}}-1} + \frac{\sigma_{\mathrm{RS}}^2}{n_{\mathrm{RS}}-1}} = \sqrt{\mathrm{SE}_{\mathrm{RO}}^2 + \mathrm{SE}_{\mathrm{RS}}^2}. \tag{B3}$$

b. *Dependence on the cutoff impact height*

To understand the strong dependence of the Tdry departures to the cutoff impact height, Fig. B2 shows the different steps in the TL Tdry calculation: BA departures (Fig. B2a), inferred refractivity departures (Fig. B2b), inferred dry pressure departures (Fig. B2c), and inferred Tdry departures (Fig. B2d). The variables are plotted for a cutoff impact height of 55 km (pink) and no cutoff (blue).

Above 55 km, the BA departures without cutoff are higher than those with cutoff at 55 km (Fig. B2a), which

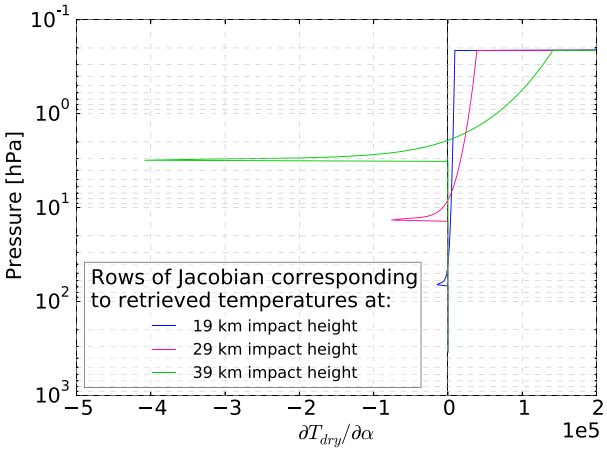

FIG. B3. Subset of a Jacobian matrix (partial derivative) for Tdry with respect to the BAs. The spike at the highest level is off scale, with the Jacobians reaching values of 3.8×10^5 (blue), 1.65×10^6 (pink), and 6.45×10^6 (green).

leads to a systematic difference in the TL refractivity departures $\delta N_{\text{no cut-off}}$ (Fig. B2b) calculated as $\delta\mathbf{N} = \mathbf{A}\delta\boldsymbol{\alpha}$. The pressure departure profile in Fig. B2c, which is calculated as the sum of the top-level pressure and the integral of the refractivity as described in section 3b, has a systematic difference of $\delta P_{\text{no cut-off}} > \delta P_{55\text{km}}$. Notably, the differences in refractivity and pressure are present farther down into the atmosphere than the differences in BA departures. Thus, a change in the upper cutoff impact height results in a different Tdry departure at all levels as can be seen in Fig. B2d.

How differences in the highest-level BA departures propagate farther down into the atmosphere can be illustrated with the Jacobian matrix of Tdry with respect to the BA [Burrows and Healy (2016) additionally show the Jacobian matrices of refractivity and pressure with respect to the BA]. Figure B3 shows a subset of three typical Tdry Jacobians. Each Jacobian shows a sharp spike at the pressure level corresponding to the impact height level of the Tdry retrievals (19, 29, and 39 km); that is, Tdry at a certain pressure level is highly sensitive to the BA at the same pressure level and those just above. The sharpness of this spike is a feature of the Abel integral transform. However, the Tdry is also influenced by BAs higher up in the atmosphere and especially by the highest BA, as the long tail in the Jacobians (a feature of the hydrostatic integration in the Tdry computation) and the spike at the highest level demonstrate.

c. *Optimization of high-level BAs*

In the conventional RO retrieval, the RO BAs are blended with a smooth a priori profile using optimization techniques. In general, the statistical optimization processing step can be written in a matrix/vector form as

$$\boldsymbol{\alpha}_{\text{smooth}} = \boldsymbol{\alpha}_{\text{clim}} + \mathbf{K}(\boldsymbol{\alpha}_{\text{RO}} - \boldsymbol{\alpha}_{\text{clim}}), \tag{B4}$$

where $\boldsymbol{\alpha}_{\text{smooth}}$, $\boldsymbol{\alpha}_{\text{clim}}$, and $\boldsymbol{\alpha}_{\text{RO}}$ are the vectors of optimized, climatological (a priori), and observed BA profiles, respectively, and $\mathbf{K}$ is the gain matrix. The gain matrix can be written as $\mathbf{K} = \mathbf{B}(\mathbf{B} + \mathbf{R})^{-1}$, where $\mathbf{R}$ and $\mathbf{B}$ are the assumed covariance matrices for the observed and climatological BA profiles, respectively. These matrices are often assumed to be diagonal, and typically, the observation error statistics are ~2 microradians, and the climatological errors are written as a scalar (~0.1) times the simulated BA value. This formulation ensures a smooth transition from observation to the climatological model in the upper stratosphere.

The smoothed RO and model BAs, which are plotted in Fig. 1, are calculated using a simplified smoothing as $\boldsymbol{\alpha}_{\text{RO}_{\text{smooth}}} = wt \cdot \boldsymbol{\alpha}_{\text{RO}} + (1 - wt) \cdot \boldsymbol{\alpha}_{\text{clim}}$, where the weight wt is calculated as $wt = \sigma^2_{\text{clim}}/(\sigma^2_{\text{clim}} + \sigma^2_{\text{RO}})$. The assumed observation error is set to $\sigma_{\text{RO}} = 2$ microradians, and the error in the climatology is assumed to be (for two example cases) $\sigma_{\text{clim}} = 0.01\alpha_{\text{clim}}$ (violet dash–dotted line in Fig. 1) and $\sigma_{\text{clim}} = 0.1\alpha_{\text{clim}}$ (red dash–dotted line in Fig. 1). With typically used values of climatological BAs at an impact height of 55 km (≈10 microradians) and an assumed 10% error on the climatology, the weight would be 0.2, which means that, at 55 km, the RO observation provides 20% of the information and the climatology 80%. Therefore, the method presented here does not neglect much more observational information than other more conventional approaches.

REFERENCES

Agustí-Panareda, A., and Coauthors, 2009: Radiosonde humidity bias correction over the West African region for the special AMMA reanalysis at ECMWF. *Quart. J. Roy. Meteor. Soc.*, **135**, 595–617, doi:10.1002/qj.396.

Anthes, R. A., 2011: Exploring Earth's atmosphere with radio occultation: Contributions to weather, climate and space weather. *Atmos. Meas. Tech.*, **4**, 1077–1103, doi:10.5194/amt-4-1077-2011.

——, and Coauthors, 2008: The COSMIC/FORMOSAT-3 mission: Early results. *Bull. Amer. Meteor. Soc.*, **89**, 313–333, doi:10.1175/BAMS-89-3-313.

Auligné, T., A. P. McNally, and D. P. Dee, 2007: Adaptive bias correction for satellite data in a numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, **133**, 631–642, doi:10.1002/qj.56.

Burns, R. B., and C. B. Dobson, 1981: Standard error of the difference between means. *Experimental Psychology: Research Methods and Statistics*, Springer Verlag, 151–157.

Burrows, C., and S. Healy, 2016: Sensitivity of radio occultation-based dry temperature retrievals to upper-level information and its relevance to radiosonde bias corrections. Met Office Forecasting Research Tech. Rep. 615, 17 pp. [Available online at http://www.metoffice.gov.uk/binaries/content/assets/mohippo/pdf/library/frtr_615_2016_2p.pdf.]

——, ——, and I. D. Culverwell, 2014: Improving the bias characteristics of the ROPP refractivity and bending angle operators. *Atmos. Meas. Tech.*, **7**, 3445–3458, doi:10.5194/amt-7-3445-2014.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood, 2005: A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Quart. J. Roy. Meteor. Soc.*, **131**, 1759–1782, doi:10.1256/qj.04.101.

Dee, D. P., 2004: Variational bias correction of radiance data in the ECMWF system. *Proc. ECMWF Workshop on Assimilation of High Spectral Resolution Sounders in NWP*, Reading, United Kingdom, ECMWF, 97–112. [Available online at http://www.ecmwf.int/sites/default/files/elibrary/2004/8930-variational-bias-correction-radiance-data-ecmwf-system.pdf.]

Derber, J. C., and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299, doi:10.1175/1520-0493(1998)126<2287:TUOTCC>2.0.CO;2.

Dirksen, R. J., M. Sommer, F. J. Immler, D. F. Hurst, R. Kivi, and H. Vömel, 2014: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde. *Atmos. Meas. Tech.*, **7**, 4463–4490, doi:10.5194/amt-7-4463-2014.

Eyre, J. R., 1992: A bias correction scheme for simulated TOVS brightness temperatures. ECMWF Tech. Memo. 186, 28 pp. [Available online at http://www.ecmwf.int/sites/default/files/elibrary/1992/9330-bias-correction-scheme-simulated-tovs-brightness-temperatures.pdf.]

——, 2016: Observation bias correction schemes in data assimilation systems: A theoretical study of some of their properties. *Quart. J. Roy. Meteor. Soc.*, **142**, 2284–2291, doi:10.1002/qj.2819.

Fjeldbo, G., A. J. Kliore, and V. R. Eshleman, 1971: The neutral atmosphere of Venus as studied with the Mariner V radio occultation experiments. *Astron. J.*, **76**, 123–140, doi:10.1086/111096.

Global Climate Observing System, 2007: GCOS Reference Upper-Air Network (GRUAN): Justification, requirements, siting and instrumentation options. WMO/TD Tech. Rep. 1379, 25 pp. [Available online at http://www.wmo.int/pages/prog/gcos/Publications/gcos-112.pdf.]

Haimberger, L., C. Tavolato, and S. Sperka, 2012: Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations. *J. Climate*, **25**, 8108–8131, doi:10.1175/JCLI-D-11-00668.1.

Harris, B. A., and G. Kelly, 2001: A satellite radiance-bias correction scheme for data assimilation. *Quart. J. Roy. Meteor. Soc.*, **127**, 1453–1468, doi:10.1002/qj.49712757418.

Hawson, C. L., and P. G. F. Caton, 1961: A synoptic method for the international comparison of geopotential observations. *Meteor. Mag.*, **90**, 333–364.

He, W., S.-P. Ho, H. Chen, X. Zhou, D. Hunt, and Y.-H. Kuo, 2009: Assessment of radiosonde temperature measurements in the upper troposphere and lower stratosphere using COSMIC radio occultation data. *Geophys. Res. Lett.*, **36**, L17807, doi:10.1029/2009GL038712.

Healy, S. B., 2001: Smoothing radio occultation bending angles above 40 km. *Ann. Geophys.*, **19**, 459–468, doi:10.5194/angeo-19-459-2001.

——, 2008a: Forecast impact experiment with a constellation of GPS radio occultation receivers. *Atmos. Sci. Lett.*, **9**, 111–118, doi:10.1002/asl.169.

——, 2008b: Assimilation of GPS radio occultation measurements at ECMWF. *GRAS SAF Workshop on Applications of GPSRO Measurements*, Reading, United Kingdom, ECMWF, 99–109. [Available online at http://www.ecmwf.int/sites/default/files/elibrary/2008/9809-assimilation-gps-radio-occultation-measurements-ecmwf.pdf.]

——, and J.-N. Thépaut, 2006: Assimilation experiments with CHAMP GPS radio occultation measurements. *Quart. J. Roy. Meteor. Soc.*, **132**, 605–623, doi:10.1256/qj.04.182.

——, and I. D. Culverwell, 2015: A modification to the standard ionospheric correction method used in GPS radio occultation. *Atmos. Meas. Tech.*, **8**, 3385–3393, doi:10.5194/amt-8-3385-2015.

——, J. Wickert, G. Michalak, T. Schmidt, and G. Beyerle, 2007: Combined forecast impact of GRACE-A and CHAMP GPS radio occultation bending angle profiles. *Atmos. Sci. Lett.*, **8**, 43–50, doi:10.1002/asl.149.

Ho, S.-P., and Coauthors, 2009: Estimating the uncertainty of using GPS radio occultation data for climate monitoring: Intercomparison of CHAMP refractivity climate records from 2002 to 2006 from different data centers. *J. Geophys. Res.*, **114**, D23107, doi:10.1029/2009JD011969.

——, and Coauthors, 2012: Reproducibility of GPS radio occultation data for climate monitoring: Profile-to-profile intercomparison of CHAMP climate records 2002 to 2008 from six data centers. *J. Geophys. Res.*, **117**, D18111, doi:10.1029/2012JD017665.

Hoffman, R. N., J.-F. Louis, and T. Nehrkorn, 1992: A method for implementing adjoint calculations in the discrete case. ECMWF Tech. Memo. 184, 20 pp.

Ingleby, B., and D. Edwards, 2015: Changes to radiosonde reports and their processing for numerical weather prediction. *Atmos. Sci. Lett.*, **16**, 44–49, doi:10.1002/asl2.518.

Kobayashi, S., and Coauthors, 2015: The JRA-55 Reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48, doi:10.2151/jmsj.2015-001.

Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield, and K. R. Hardy, 1997: Observing Earth's atmosphere with radio occultation measurements using the global positioning system. *J. Geophys. Res.*, **102**, 23 429–23 465, doi:10.1029/97JD01569.

——, ——, S. S. Leroy, and B. Herman, 2000: The GPS radio occultation technique. *Terr. Atmos. Ocean. Sci.*, **11**, 53–114, doi:10.3319/TAO.2000.11.1.53(COSMIC).

Ladstädter, F., A. K. Steiner, M. Schwärz, and G. Kirchengast, 2015: Climate intercomparison of GPS radio occultation, RS90/92 radiosondes and GRUAN from 2002 to 2013. *Atmos. Meas. Tech.*, **8**, 1819–1834, doi:10.5194/amt-8-1819-2015.

Leroy, S. S., J. A. Dykema, and J. G. Anderson, 2006: Climate benchmarking using GNSS occultation. *Atmosphere and Climate: Studies by Occultation Methods*, U. Foelsche, G. Kirchengast, and A. Steiner, Eds., Springer-Verlag, 287–301, doi:10.1007/3-540-34121-8_24.

Leys, C., C. Ley, O. Klein, P. Bernard, and L. Licata, 2013: Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psych.*, **49**, 764–766, doi:10.1016/j.jesp.2013.03.013.

Milan, M., and L. Haimberger, 2015: Predictors and grouping for bias correction of radiosonde temperature observations. *J. Geophys. Res. Atmos.*, **120**, 10 736–10 766, doi:10.1002/2015JD023635.

Philipona, R., and Coauthors, 2013: Solar and thermal radiation errors on upper-air radiosonde temperature measurements. *J. Atmos. Oceanic Technol.*, **30**, 2382–2392, doi:10.1175/JTECH-D-13-00047.1.

Poli, P., S. B. Healy, and D. P. Dee, 2010: Assimilation of global positioning system radio occultation data in the ECMWF ERA-Interim reanalysis. *Quart. J. Roy. Meteor. Soc.*, **136**, 1972–1990, doi:10.1002/qj.722.

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne, 2007: The Met Office global four-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **133**, 347–362, doi:10.1002/qj.32.

Rennie, M. P., 2010a: Investigation into the consistency of radiosonde and GPSRO observations and their application in NWP at the Met Office. Met Office Forecasting R&D Tech. Rep. 541, 40 pp. [Available online at http://research.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/reports/541.pdf.]

——, 2010b: The impact of GPS radio occultation assimilation at the Met Office. *Quart. J. Roy. Meteor. Soc.*, **136**, 116–131, doi:10.1002/qj.521.

Rodgers, C. D., 2000: *Inverse Methods for Atmospheric Sounding: Theory and Practice.* World Scientific, 238 pp.

Scherllin-Pirscher, B., G. Kirchengast, A. K. Steiner, Y.-H. Kuo, and U. Foelsche, 2011: Quantifying uncertainty in climatological fields from GPS radio occultation: An empirical-analytical error model. *Atmos. Meas. Tech.*, **4**, 2019–2034, doi:10.5194/amt-4-2019-2011.

Simmons, A. J., P. Poli, D. P. Dee, P. Berrisford, H. Hersbach, S. Kobayashi, and C. Peubey, 2014: Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim. *Quart. J. Roy. Meteor. Soc.*, **140**, 329–353, doi:10.1002/qj.2317.

Smith, E. K., and S. Weintraub, 1953: The constants in the equation for atmospheric refractive index at radio frequencies. *J. Res. Natl. Bur. Stand.*, **50**, 39–41, doi:10.6028/jres.050.006.

Steiner, A. K., G. Kirchengast, and H. P. Ladreiter, 1999: Inversion, error analysis, and validation of GPS/MET occultation data. *Ann. Geophys.*, **17**, 122–138, doi:10.1007/s00585-999-0122-5.

——, and Coauthors, 2013: Quantification of structural uncertainty in climate data records from GPS radio occultation. *Atmos. Chem. Phys.*, **13**, 1469–1484, doi:10.5194/acp-13-1469-2013.

Sun, B., A. Reale, D. J. Seidel, and D. C. Hunt, 2010: Comparing radiosonde and COSMIC atmospheric profile data to quantify differences among radiosonde types and the effects of imperfect collocation on comparison statistics. *J. Geophys. Res.*, **115**, D23104, doi:10.1029/2010JD014457.

——, ——, S. Schroeder, D. J. Seidel, and B. Ballish, 2013: Toward improved corrections for radiation-induced biases in radiosonde temperature observations. *J. Geophys. Res. Atmos.*, **118**, 4231–4243, doi:10.1002/jgrd.50369.

Syndergaard, S., 1999: Retrieval analysis and methodologies in atmospheric limb sounding using the GNSS radio occultation technique. Danish Meteorological Institute Scientific Rep. 99-6, 76 pp.

Tradowsky, J. S., 2015: Characterisation of radiosonde temperature biases and errors using radio occultation measurements. Radio Occultation Meteorology Satellite Application Facility Visiting Scientist Rep. 26, 56 pp. [Available online at http://www.romsaf.org/visiting_scientist.php#y2015.]

——, 2016: Radiosonde temperature bias corrections using radio occultation bending angles as reference. Radio Occultation Meteorology Satellite Application Facility Visiting Scientist Rep. 31, 63 pp. [Available at http://www.romsaf.org/visiting_scientist.php#y2016.]

Walters, D., N. Wood, S. Vosper, and S. Milton, 2014: ENDGame: A new dynamical core for seamless atmospheric prediction. Met Office Rep., 27 pp. [Available online at http://www.metoffice.gov.uk/binaries/content/assets/mohippo/pdf/s/h/endgamegovsci_v2.0.pdf.]