Improving thermodynamic profile retrievals from microwave radiometers by including Radio Acoustic Sounding System (RASS) observations

Irina V. Djalalova[1,2], David D. Turner[3], Laura Bianco[1,2],

James M. Wilczak[2], James Duncan[1,2]~~,~~[*], Bianca Adler[1,2] and Daniel Gottas[2]

[1] Cooperative Institute for Research in Environmental Sciences (CIRES), Boulder, CO, USA

[2] National Oceanic and Atmospheric Administration, Physical Sciences Laboratory, Boulder, CO, USA

~~[3]National~~[3] National Oceanic and Atmospheric Administration, Global Systems Laboratory, Boulder, CO USA

~~Draft: in preparation for Journal of Atmospheric Measurement Techniques (12/18/2020)~~

[*]Now at WindESCo, Burlington, MA

Formatted: *Font color: Auto*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Corresponding author address: Irina V. Djalalova (Irina.V.Djalalova@noaa.gov), NOAA/Physical Science Laboratory, 325 Broadway, mail stop: ~~-~~PSD3, Boulder, CO 80305. Tel.: 303-497-6238. Fax: 303-497-6181.

Outline

Abstract

1. **Introduction**
2. **XPIA dataset**
    - **2.1 MWR measurements**
    - **2.2 Radiosonde measurements**
    - **2.3 WPR-RASS measurements**
    - **2.4 BAO data**
3. **Physical retrievals**
    - **3.1 Iterative retrieval technique**
    - **3.2 ~~PR's~~Physical retrieval bias correction and temperature profiles**
    - **3.3 Averaging kernel**
4. **Results**
    - **4.1 ~~PRs~~Physical retrieval statistical analysis from Akernel**
    - **4.2 Statistical analysis of ~~PRs compared to NN~~physical retrievals up to 5 km AGL**

Formatted: *Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0" + Indent at: 0.25"*

Formatted: *Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0" + Indent at: 0.25"*

Formatted: *Outline numbered + Level: 2 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"*

Formatted: *Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0" + Indent at: 0.25"*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

4.3 Statistics for the least close to the climatological profiles

4.4 Virtual temperature statistics

5. Conclusions

Appendix A

Data availability

Author contribution

Acknowledgments

References

Formatted: *Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0" + Indent at: 0.25"*

Abstract

Thermodynamic profiles are often retrieved from the multi-wavelength brightness temperature observations made by microwave radiometers (MWRs) using regression methods (linear, quadratic approaches), artificial intelligence (neural networks), or physical-iterative methods. Regression and neural network methods are tuned to mean conditions derived from a climatological dataset of thermodynamic profiles collected nearby. In contrast, physical-iterative retrievals use a radiative transfer model starting from a climatologically reasonable value of temperature and water vapor, with the model run iteratively until the derived brightness temperatures match those observed by the MWR within a specified uncertainty.

In this study, a physical-iterative approach is used to retrieve temperature and humidity profiles from data collected during XPIA (eXperimental Planetary boundary layer Instrument Assessment), a field campaign held from March to May 2015 at NOAA's Boulder Atmospheric Observatory (BAO) facility. During the campaign, several passive and active remote sensing

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

instruments as well as in-situ platforms were deployed and evaluated to determine their suitability for the verification and validation of meteorological processes. Among the deployed remote sensing instruments was a multi-channel MWR, as well as two radio acoustic sounding systems (RASS), associated with 915-MHz and 449-MHz wind profiling radars.

Having the possibility to combine the information provided by the MWR and RASS systems, in this study the physical-iterative approach is tested with different observational inputs: first using data from surface sensors and the MWR in different configurations, and then including data from the RASS. These temperature retrievals are ~~also compared to those derived by a neural network method, assessing their relative accuracy~~assessed against 58 co-located radiosonde profiles. Results show that the combination of the MWR and RASS observations in the physical-iterative approach allows for a more accurate characterization of low-level temperature inversions, and that these retrieved temperature profiles match the radiosonde observations better than ~~all other approaches, including~~ the ~~neural network~~temperature profiles retrieved from only the MWR, in the ~~atmospheric~~ layer between the surface and 5 km above ground level (AGL~~.~~). Specifically, in this layer of the atmosphere, both root mean square errors and standard deviations of the difference between radiosonde and retrievals that combine MWR and RASS are improved by ~0.5 ^{o}C compared to the other methods. Pearson correlation coefficients are also improved.

Formatted: *Font: Calibri, 12 pt, Font color: Auto, Not Highlight*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

We provide the comparison of the temperature physical retrievals to the neural network retrievals in Appendix A.

1. Introduction

Formatted: Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"

To monitor the state of the atmosphere for process understanding and for model verification and validation, scientists rely on observations from a variety of instruments, each one having its set of advantages and disadvantages. Using several diverse instruments allows one to monitor different aspects of the atmosphere, while combining them in an optimized synergetic approach can improve the accuracy of the information we have on the state of the atmosphere.

During the eXperimental Planetary boundary layer Instrumentation Assessment (XPIA) campaign, an U.S. Department of Energy sponsored experiment held at the Boulder

Formatted: Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

Atmospheric Observatory (BAO) in Spring 2015, several instruments were deployed (Lundquist et al., 2017) with the goal of assessing their capability for measuring flow within the atmospheric boundary layer. XPIA investigated novel measurement approaches, and quantified uncertainties associated with these measurement methods. While the main interest of the XPIA campaign was on wind and turbulence, measurements of other important atmospheric variables were also collected, including temperature and humidity. Among the deployed instruments were two identical microwave radiometers (MWRs) and two radio acoustic sounding systems (RASS), as well as radiosondes launches that were used for verification.

MWRs are passive sensors, sensitive to atmospheric temperature and humidity content that allow for a high temporal observation of the state of the atmosphere, with some advantages and limitations. In order to estimate profiles of temperature and humidity~~, they observe atmospheric~~ from the observed brightness ~~temperature and apply radiative transfer equations (Rosenkranz, 1998) and~~temperatures (Tb), several methods could be applied such as regressions, neural network retrievals~~ (Solheimet al., ,1998a, and 1998b; Ware et al., 2003),~~ or physical retrieval methodologies ~~that can~~which include more information about the atmospheric state in the retrieval process~~ (Turner and Blumberg, 2019).~~. Radiative transfer equations (Rosenkranz, 1998) are commonly used to train statistical retrievals, or as forward models used within physical retrieval methods. Advantages of MWRs include their compact design, the relatively high temporal resolution of the measurements (2-3 minutes), the possibility to observe the vertical structure of both temperature and moisture~~,~~ through the ~~deep layer~~depth of the ~~atmosphere that can be monitored including~~troposphere during both clear and cloudy conditions, and their capability to operate in a standalone mode.

__Formatted:__ Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

Disadvantages include ~~the~~ limited accuracy~~, as the temperature and humidity profiles are not actively measured but retrieved, their lower accuracy~~ in the presence of rain because of scattering of radiation ~~due to~~from raindrops in the atmosphere (and because ~~some~~ water can ~~still~~ deposit on the radome, although the instruments use a hydrophobic radome and force airflow over the surface of the radome during rain to mitigate this impact), rather coarse vertical resolution, and for retrievals the necessity to have a site~~-~~-specific climatology. Other disadvantages include the challenges related to performing accurate calibrations (Küchler et al., 2016, and references within), radio frequency interference (RFI), and the low accuracy on the retrieved liquid water path (LWP) especially for values of LWP less than ~~50~~20 g/m^2.

RASS, in comparison, are active instruments that emit a longitudinal acoustic wave upward, causing a local compression and rarefaction of the ambient air. These density variations are tracked by the Doppler radar associated with the RASS, and the speed of the propagating sound wave is measured. The speed of sound is related to the virtual temperature (Tv) (North et al., 1973), and therefore, RASS are routinely used to remotely measure vertical profiles of virtual temperature in the boundary layer. Being an active instrument, the RASS is in general more accurate than a passive instrument (Bianco et al., 2017), but they also come with their sets of disadvantages. The main limitations of RASS for retrieval purposes are its low temporal resolution (typically a 5-min averaged RASS profile is measured once or twice per hour), and their limited altitude coverage. Recent studies (Adachi and Hashiguchi, 2019) have shown that to make them more suitable to operate in urban areas RASS could use parametric speakers to take advantage of their high directivity and very low side lobes. Nevertheless, the maximum height reached by the RASS is still limited, being a function of both radar frequency

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

and atmospheric conditions (May and Wilczak, 1993), and is determined both by the attenuation of the sound, which is a function of atmospheric temperature, humidity, and frequency of the sound source, and the advection of the propagating sound wave out of the radar's field-of-view. Therefore, data availability is usually limited to the lowest several ~~km, dependent~~kilometers, depending on the frequency of the radar. In addition, wintertime coverage is usually considerably lower than that in summer, due to a higher probability of stronger winds advecting the sound wave away from the radar in the winter.

To get a better picture of the state of the temperature and moisture structure of the atmosphere, it makes sense to try to combine the information obtained by both MWR and RASS. Integration of different instruments has been of scientific interest for several years (Han and Westwater 1995; Stankov et al. 1996; Bianco et al., 2005; Engelbart et al., 2009; Cimini et al., 2020~~,~~; Turner and Löhnert, 2020, to name some). In this study we particularly focus on the combination of the MWR and RASS observations in the retrievals to improve the accuracy of the temperature profiles in the lowest 5 km compared to ~~the standard MWR retrievals obtained through neural network (NN) processing, or compared to~~ physical retrieval approaches that do not include the information from RASS measurements. Some studies have used analyses from numerical weather prediction (NWP) models as an additional constraint in these variational retrievals (e.g., Hewison 2007; Cimini et al. 2005, 2011; Martinet et al. 2020); however, we have elected not to include model data in this study because we wanted to evaluate the impact of the RASS profiles on the retrievals from a purely observational perspective.

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

This paper is organized as follows: Section 2 summarizes the experimental dataset; Section 3 introduces the principles of the physical retrieval approaches used to obtain vertical profiles of the desired variables; Section 4 produces statistical analysis of the comparison between the different retrieval approaches and radiosonde measurement; finally, conclusions are presented in Section 5.

2. XPIA data

Formatted: *Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"*

The data used in our analysis were collected during the XPIA experiment, held in Spring 2015 (March-May) at the NOAA's Boulder Atmospheric Observatory (BAO) site, in Erie, Colorado (Lat.: 40.0451 N, Lon.: 105.0057 W, El.: 1584 m MSL). XPIA was the last experiment conducted at this facility, as after almost 40 years of operations the BAO 300-m tower was demolished at the end of 2016 (Wolfe and Lataitis, 2018). XPIA was designed to assess the capability of different remote sensing instruments for quantifying boundary layer structure, and was a preliminary study as many of these same instruments were later deployed, among other campaigns, for the second Wind Forecast Improvement Project WFIP2 (Shaw et al., 2019; Wilczak et al., 2019) which investigated flows in complex terrain for wind energy applications, and were for example used to study cold air pool and gap flow characteristics (Adler et al., ~~2020~~2021; Banta et al., 2020; Neiman et al., 2019). The list of the deployed instruments included active and passive remote-sensing devices, and in-situ instruments mounted on the BAO tower. Data collected during XPIA are publicly available at https://a2e.energy.gov/projects/xpia. A detailed description of the XPIA experiment can be

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)* found in Lundquist et al. (2017), while a specific look at the accuracy of the instruments used in this study can be found in Bianco et al. (2017).

2.1 MWR measurements

Two identical MWRs~~,~~ (Radiometrics MP-3000A) managed by NOAA (MWR-NOAA) and by the University of Colorado (MWR-CU), were deployed next to each other at the visitor center ~600 m south of the BAO tower (see Lundquist et al., 2017 for a detailed map of the study area). Prior to the experiment, both MWRs were calibrated using an external liquid nitrogen target and an internal ambient target and thoroughly serviced (sensor cleaning, radome replacement, etc.). MWRs are passive devices which record the natural microwave emission in the water vapor and oxygen absorption bands from the atmosphere, providing measurements of the brightness temperatures. Both MWRs have 35-channels spanning a range of frequencies, with 21 channels in the lower (22-30 GHz) K-band frequency band, of which 8 channels were used during XPIA: 22.234, 22.5, 23.034, 23.834, 25, 26.234, 28 and 30 GHz; and 14 channels in the higher (51-59 GHz) V-band frequency band, of which all were used in XPIA: 51.248, 51.76, 52.28, 52.804, 53.336, 53.848, 54.4, 54.94, 55.5, 56.02, 56.66, 57.288, 57.964 and 58.8 GHz. Frequencies in the K-band are more sensitive to water vapor and cloud liquid water, while frequencies in the V-band are sensitive to atmospheric temperature due to the absorption of atmospheric oxygen (Cadeddu et al., 2013). V-band frequencies or channels also can be divided in two categories: the opaque channels, 56.66 GHz and higher, that are more informative in the layer of the atmosphere from the surface to ~1 km AGL, and the transparent channels, 51-56 GHz, that are more informative above 1 km AGL. Both MWRs observed at the zenith and at 15-

Formatted: Font: 11 pt, Highlight

Formatted: Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

and 165-degree elevation angles in the north-south plane (referred to as oblique elevation scans hereafter; note zenith views have 90-degree elevation ~~angles~~angle). In addition, each MWR was provided with a separate surface sensor to measure pressure, temperature, and relative humidity at the installation level that was ~2.5 m ~~above ground level (AGL).~~ AGL. ~~MWRs are passive devices which record the natural microwave emission in the water vapor and oxygen absorption bands from the atmosphere, providing measurements of the brightness temperatures.~~ Vertical profiles of temperature (T), water vapor density (WVD), and relative humidity (RH) were retrieved in real-time during XPIA every 2-3 minutes using a neural network (NN) approach provided by the manufacturer of the radiometer, Radiometrics (Solheim et al. 1998a, and 1998b; Ware et al., 2003~~NN approach provided by the private manufacturing company Radiometrics (Solheim et al. 1998). The NN used a training dataset based on a 5-year climatology of profiles from radiosondes launched at the Denver International Airport, 35 miles south-east from the XPIA site. NN-based MWR vertical retrieval profiles were obtained using the zenith and an average of two oblique elevation scans, all extending for 58 levels up to 10 km, with nominal vertical levels depending on the height (every 50 m from the surface to 500 m, every 100 m from 500 m to 2 km, and every 250 m from 2 to 10 km, AGL). In this study we make use of the NN zenith and of the NN oblique, where the latter can average out small-scale horizontal inhomogeneities of the atmosphere~~). Although the physical retrieval configurations used in this study do not exactly match the MWR configurations used for NN retrievals, a comparison of both physical and neural network retrievals to the radiosonde temperature data is presented in Appendix A.

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

~~The MWR-CU~~ Both MWRs nominally operated from 9 March to 7 May 2015, ~~while~~although the MWR-NOAA was unavailable between 5-27 April 2015. For the overlapping dates, temperature profiles retrieved from the two MWRs showed very good agreement with less than 0.5 ~~K~~^{o}C bias and 0.994 correlation (Bianco et al., 2017). For this reason, and because the MWR-CU was available for a longer time period, we use only the MWR-CU (hereafter simply called MWR).

2.2 Radiosonde measurements

Between 9 March and 7 May 2015, while the MWR was operational, radiosondes were launched by the National Center for Atmospheric Research (NCAR) assisted by several students from the University of Colorado over three selected periods, one each in March, April, and May. There was a total of 59 launches, mostly four times per day, around 14:00, 18:00, 22:00 and 0200 UTC (8:00, 12:00, 16:00 and 20:00 local standard time, LST). All radiosondes were Vaisala RS92. The first 35 launches, between 9-19 March, were done from the visitor center, while the 11 launches, between 15-22 April, and 13 launches, between 1-4 May, were done from the water tank site, ~1000 meters apart (see Lundquist et al., 2017 for a detailed map of the study area). The radiosonde measurements included temperature, dewpoint temperature, and relative humidity, to altitudes usually higher than 10 km AGL, with measurements every few seconds.

2.3 WPR-RASS measurements

Formatted: *Indent: First line: 0.5"*

Formatted: *Font: Not Bold*

Formatted: *Indent: First line: 0.5"*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Two NOAA wind profiling radars (WPRs), operating at frequencies of 915-MHz and 449-MHz, were deployed at the visitor center (same location ~~of~~as the MWR) during XPIA. These systems are primarily designed to measure the vertical profile of the horizontal wind vector, but co-located RASS also observe profiles of virtual temperature in the lower atmosphere, with different resolutions and height coverages depending on the WPR. Thus, the RASS associated with the 915-MHz WPR (hereafter referred to as RASS 915) measured virtual temperature from 120 to 1618 m with a vertical resolution of 62 m, and the 449 MHz RASS (hereafter referred to as RASS 449) sampled the boundary layer from 217 to 2001 m with a vertical resolution of 105 m. The maximum height reached by the RASS is a function of both radar frequency and atmospheric conditions (May and Wilczak, 1993), and is usually lower for RASS 915 data, as will be shown later in the analysis.

The RASS data were processed using a radio frequency interference (RFI)-removal algorithm (performed on the RASS spectra), a consensus algorithm (Strauch et al. 1984) performed on the moment data using a 60% consensus threshold, a Weber-Wuertz outlier removal algorithm (Weber et al., 1993) performed on the consensus averages, and a RASS range-correction algorithm (Görsdorf and Lehmann, 2000) using an average relative humidity setting of 50% determined from the available observations.

2.4 BAO data

The BAO 300-m tower was built in 1977 to study the planetary boundary layer (Kaimal and Gaynor 1983). During XPIA, measurements were collected at the surface (2 m) and at six higher levels (50, 100, 150, 200, 250 and 300 m AGL). Each tower level was equipped with 2

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

sonic anemometers on orthogonal booms, and one sensor based on a Sensiron SHT75 solid-state sensor to measure temperature and relative humidity with a time resolution of 1 s, and averaged over five minutes.

The observational temperature and water vapor surface data are used from the more accurate observations at the BAO tower 2 m AGL level (Horst~~,~~ et al., 2016), to replace the data measured by the less accurate MWR inline surface sensor.

3. Physical retrievals

~~Other than NN approaches, a~~A physical retrieval (PR) iterative approach can be used to retrieve vertical profiles of thermodynamic properties from the MWR observations (Maahn et al 2020). In this case, using a radiative transfer model, the process starts with a climatologically reasonable value of temperature and water vapor, and is iteratively repeated until the computed brightness temperatures match those observed by the MWR within the uncertainty of the observed brightness temperatures (Rodgers, 2000; Turner and Löhnert, 2014; Maahn et al. 2020).

3.1 Iterative retrieval technique

For this study, the ~~physical retrieval (PR)~~PR uses a microwave radiative transfer model, MonoRTM (Clough et al., 2005), which is fully functional for the microwave region and was intensively evaluated previously on MWR measurements (Payne et al. 2008; 2011). We start with the state vector $\mathbf{X_a}$ = [**T, Q,** LWP]^T, where superscript T denotes transpose. **T** (K) and **Q**

Formatted: *Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"*

Formatted: *Space Before: 0 pt, After: 0 pt*

Formatted: *Space Before: 0 pt, After: 0 pt, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

(g/kg) are temperature and water vapor mixing ratio profiles at 55 vertical levels from the surface up to 17 km, with the distance between the levels increasing exponentially-like with height. LWP is the liquid water path in (g/m^2) that measures the integrated content of liquid water in the entire vertical column above the MWR, and is a scalar. For this study we have $\mathbf{X_a}$ with dimensions equal to 111 x 1 (two vectors **T** and **Q** with 55 levels each, and LWP). ~~-~~We are using the retrieval framework of Turner and Blumberg (2019), but only using MWR data (no spectral infrared) and will augment the retrieval to include RASS profiles of Tv.

The observation vector **Y** from the beginning includes temperature and water vapor mixing ratio measured at the surface, and ~~brightness temperature (~~**Tb**~~)~~ measured by the MWR. The MonoRTM model **F** is used as the forward model ~~to estimate the observation vector **Y**~~ from the current state vector **X**, ~~from~~ Eq. ~~(1),~~(1), and is then compared to the observation vector **Y**, iterating until the difference between **F(X)** and **Y** is small within a specified uncertainty~~:~~.

$$\cancel{X_{n+1} = X_a + (S_a^{-1} + K^T S_\varepsilon^{-1} K)^{-1} K^T S_\varepsilon^{-1} [Y - F(X_n) + K(X_n - X_a)]} \quad \cancel{(1)}$$

$$\boldsymbol{X_{n+1} = X_n + (S_a^{-1} + K^T S_\varepsilon^{-1} K)^{-1} K^T S_\varepsilon^{-1} [Y - F(X_n) + K(X_n - X_a)]} \quad (1)$$

with~~:~~

$$\boldsymbol{X_a} = \begin{bmatrix} \boldsymbol{T} \\ \boldsymbol{Q} \\ L \end{bmatrix} \qquad \boldsymbol{S_a} = \begin{bmatrix} \boldsymbol{\sigma^2_{TT}} & \boldsymbol{\sigma^2_{TQ}} & 0 \\ \boldsymbol{\sigma^2_{QT}} & \boldsymbol{\sigma^2_{QQ}} & 0 \\ 0 & 0 & \sigma^2_L \end{bmatrix} \boldsymbol{K_{ij}} = \frac{\partial \boldsymbol{F_i}}{\partial \boldsymbol{X_j}}$$

$$\boldsymbol{S_\varepsilon} = \begin{bmatrix} \sigma^2_{T\,sfc} & 0 & 0 & 0 & 0 \\ 0 & \sigma^2_{Q\,sfc} & 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{\sigma^2_{Tb_{zenith}}} & 0 & 0 \\ 0 & 0 & 0 & \boldsymbol{\sigma^2_{Tb_{zenith+oblique}}} & 0 \\ 0 & 0 & 0 & 0 & \boldsymbol{\sigma^2_{Tv_{RASS915(449)}}} \end{bmatrix}$$

Formatted: *Indent: First line: 0"*

Formatted: *Font: 11 pt*

Formatted: *Font: Not Bold*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

$$\boldsymbol{X}_a = \begin{bmatrix} \boldsymbol{T} \\ \boldsymbol{Q} \\ LWP \end{bmatrix} \quad \boldsymbol{S}_a = \begin{bmatrix} \boldsymbol{\sigma}^2_{TT} & \boldsymbol{\sigma}^2_{TQ} & 0 \\ \boldsymbol{\sigma}^2_{QT} & \boldsymbol{\sigma}^2_{QQ} & 0 \\ 0 & 0 & \sigma^2_{LWP} \end{bmatrix} \quad \boldsymbol{K}_{ij} = \frac{\partial \boldsymbol{F}_i}{\partial \boldsymbol{X}_j}$$

$$\boldsymbol{S}_\varepsilon = \begin{bmatrix} \sigma^2_{Tsfc} & 0 & 0 & 0 \\ 0 & \sigma^2_{Qsfc} & 0 & 0 \\ 0 & 0 & \boldsymbol{\sigma}^2_{Tb_{zenith}}\ ①\ \boldsymbol{or}\ \boldsymbol{\sigma}^2_{Tb_{zenith+oblique\ avrg}}\ ② & 0 \\ 0 & 0 & 0 & \boldsymbol{\sigma}^2_{Tv_{RASS915(449)}}\ \text{3 or 4} \end{bmatrix}$$

where i and j in the K_{ij} definition mark channel and vertical level respectively, and **Y**, depending on the configuration used, being equal to:

$$\boldsymbol{Y}_1 = \begin{bmatrix} T_{sfc} \\ Q_{sfc} \\ \boldsymbol{Tb}_{zenith} \end{bmatrix} \quad \boldsymbol{Y}_2 = \begin{bmatrix} T_{sfc} \\ Q_{sfc} \\ \boldsymbol{Tb}_{zenith} \\ \boldsymbol{Tb}_{zenith+oblique} \end{bmatrix}$$

$$\boldsymbol{Y}_3 = \begin{bmatrix} T_{sfc} \\ Q_{sfc} \\ \boldsymbol{Tb}_{zenith} \\ \boldsymbol{Tb}_{zenith+oblique} \\ \boldsymbol{Tv}_{RASS915} \end{bmatrix} \quad \boldsymbol{Y}_4 = \begin{bmatrix} T_{sfc} \\ Q_{sfc} \\ \boldsymbol{Tb}_{zenith} \\ \boldsymbol{Tb}_{zenith+oblique} \\ \boldsymbol{Tv}_{RASS449} \end{bmatrix}$$

$$\boldsymbol{Y}_1 = \begin{bmatrix} T_{sfc} \\ Q_{sfc} \\ \boldsymbol{Tb}_{zenith} \end{bmatrix} \quad \boldsymbol{Y}_2 = \begin{bmatrix} T_{sfc} \\ Q_{sfc} \\ \boldsymbol{Tb}_{zenith+oblique\ avrg} \end{bmatrix}$$

$$\boldsymbol{Y}_3 = \begin{bmatrix} T_{sfc} \\ Q_{sfc} \\ \boldsymbol{Tb}_{zenith+oblique\ avrg} \\ \boldsymbol{Tv}_{RASS915} \end{bmatrix} \quad \boldsymbol{Y}_4 = \begin{bmatrix} T_{sfc} \\ Q_{sfc} \\ \boldsymbol{Tb}_{zenith+oblique\ avrg} \\ \boldsymbol{Tv}_{RASS449} \end{bmatrix}$$

Formatted: *Indent: First line: 0"*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

The superscripts T and -1 in (1) indicate transpose or inverse matrix, respectively. Also, vectors and matrices are shown in bold.~~-~~ Note that we are including the 2-m surface-level observations of temperature and water vapor mixing ratio (Tsfc and Qsfc, respectively) as part of the observation vector **Y**, and thus the uncertainties in these observations are included in $\mathbf{S}_{\varepsilon}$.

Formatted: *Font: Bold*

The first guess of the state vector **X**, $\mathbf{X}_1$ in Eq. (1), is set to be equal to the mean state vector of climatological estimates, or a "prior" vector $\mathbf{X}_a$, which is calculated independently for each month of the year from climatological sounding profiles (using 10 years of data) in the Denver area.

$\mathbf{S}_a$ is the covariance matrix of the "prior" vector that includes not only temperature or water vapor variances but also the covariances between them. Using 3,000 radiosondes launched by the NWS in Denver, we interpolated each radiosonde profile to the vertical levels used in the retrieval, after which we computed the covariance of temperature and temperature, temperature and humidity, and humidity and humidity for different levels. **K** is the Jacobian matrix, computed using finite differences by perturbing the elements of **X** and rerunning the radiative transfer model.

Formatted: *Indent: First line: 0.5"*

Formatted: *Font: Bold*

We start with four configurations for the observational vector **Y** ($\mathbf{Y}_1$, $\mathbf{Y}_2$, $\mathbf{Y}_3$, and $\mathbf{Y}_4$). The MWR provides ~~the~~ **Tb** measurements ~~in all schemes,~~ from 22 channels from the zenith scan for the zenith only ~~in~~ configuration ($\mathbf{Y}_1$ ~~(~~, which also includes the 2-m in-situ observations of temperature and humidity), ~~and~~while when using the zenith ~~and~~plus oblique ~~in~~ Tb inputs ($\mathbf{Y}_2$, $\mathbf{Y}_3$, and $\mathbf{Y}_4$, also including the 2-m in-situ observations of temperature and humidity) the same 22 channels were used from the zenith scans together with only the four opaque channels (56.66, 57.288, 57.964 and 58.8 GHz) from the oblique scans. Using additional measurements

Formatted: *Font: Bold*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

from the co-located radar systems with RASS, we may further expand the observational vector with either RASS 915 (**Y**$_3$) or RASS 449 (**Y**$_4$) virtual temperature observations. The covariance matrix of the observed data, **S**$_\varepsilon$, depends on the chosen **Y**$_i$ as it is highlighted by the red numbers in the matrix description, with increasing dimensions from **Y**$_1$ to **Y**$_2$ and additional increasing dimensions to **Y**$_3$ ~~and~~or **Y**$_4$ through the multi-level measurements of the RASS (Turner and Blumberg, 2019). Table 1 summarizes the observational information included in these four different configurations of the PR.

	T_{sfc}	Q_{sfc}	Tb_{zenith}	~~$Tb_{zenith\text{-}oblique}$~~ $Tb_{oblique\_avrg}$	$Tv_{RASS915}$	$Tv_{RASS449}$
***Y**$_1$ = MWRz*	X	X	X			
***Y**$_2$ = MWRzo*	X	X	X	X		
***Y**$_3$ = MWRzo915*	X	X	X	X	X	
***Y**$_4$ = MWRzo449*	X	X	X	X		X

*Table 1. Four PR configurations corresponding to the four observational **Y**$_i$ vectors in Eq. (1).*

The uncertainty in the MWR Tb observations was set to the standard deviation from a detrended time-series analysis for each channel during cloud-free periods, which is described in detail in Section 3.2. The derived uncertainties ranged from 0.3 K to 0.4 K in the 22 to 30 GHz

Formatted: *Font: 14 pt, Italic*

Formatted: *Space After: 8 pt, Line spacing: Multiple 1.08 li*

Formatted: *Font: 11 pt, Italic*

Formatted: *Indent: First line: 0", Space After: 8 pt, Line spacing: Multiple 1.08 li*

Formatted Table

Formatted: *Font: 14 pt, Italic*

Formatted: *Font: 14 pt, Italic*

Formatted: *Space After: 8 pt, Line spacing: Multiple 1.08 li*

Formatted: *Font: 14 pt, Italic*

Formatted: *Space After: 8 pt, Line spacing: Multiple 1.08 li*

Formatted: *Font: 14 pt, Italic*

Formatted: *Space After: 8 pt, Line spacing: Multiple 1.08 li*

Formatted: *Font: 14 pt, Italic*

Formatted: *Space After: 8 pt, Line spacing: Multiple 1.08 li*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

channels, and 0.4 to 0.7 K in the 52 to 60 GHz channels. We ~~assume~~assumed that there ~~is~~was no ~~covariance~~correlated error between the different ~~instruments~~MWR channels.

For the RASS, collocated RASS and radiosonde profiles were compared and the standard deviation of the differences in Tv were determined as ~~well as~~a function of the radar's signal-to-noise ratio (SNR). This relationship resulted in uncertainties that ranged from 0.8 K at high SNR values to 1.5 K at low SNR values. Again, we assumed that there was no correlated error between different ~~channels (MWR) or height levels (RASS) of each instrument, therefore this~~RASS heights. Following all these assumptions, the covariance matrix $\mathbf{S}_\varepsilon$ is diagonal.

The Jacobian matrix, **K**, has dimensions m x 111, where m is the length of the vector $\mathbf{Y}_i$, therefore its ~~dimensions increase~~dimension increases correspondingly with the inclusion of more observational data. **K** makes the "connection" between the state vector and the observational data and should be calculated at every iteration.

3.2 ~~Bias~~Physical retrieval bias-correction and temperature profiles

Observational errors propagate through the retrieval into the derived profiles (i.e. the bias of the observed data will contribute to a bias in the retrievals.) For that, retrieval uncertainties in Eq. (1) from $\mathbf{Y} = \mathbf{Y}_1$ or $\mathbf{Y}_2$ derive only from uncertainties in surface and MWR data, while retrieval uncertainties from $\mathbf{Y} = \mathbf{Y}_3$ or $\mathbf{Y}_4$ are coming from uncertainties in surface, MWR, and RASS measurements.

While the bias of the retrieval depends on both the sensitivity of the forward model and the observational ~~uncertainty~~systematic offset, we can try to eliminate, or at least to reduce, the systematic error in the MWR observations. To this aim, we first looked for clear sky days (to

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

reduce the degrees of freedom associated with clouds) during the period of the measurements. One method to identify clear-sky times is to use ~~brightness temperature~~ Tb observations in the 30 GHz liquid water ~~vapor~~ sensitive channel. The random uncertainty in ~~brightness temperature was~~ Tb is calculated as ~~its~~an average of the Tb standard deviation ~~during clear sky times and for this channel is approximately 0.3 K (but during periods with liquid-bearing clouds overhead,~~ in a one-hour sliding window through all data points of a day. (It also could be computed as the standard deviation of the ~~30 GHz Tb is markedly higher than this threshold due~~difference between Tb and the smoothed Tb to ~~the non-homogeneous nature of clouds and thus their contribution to the downwelling microwave radiance).~~eliminate daily temperature variability.) Four clear-sky days ~~were selected,~~have been chosen using a criterion of 0.3 K uncertainty in the 30 GHz channel: March 10 and 30, and April 13 and 29~~.~~, 2015. During periods with liquid-bearing clouds overhead, this criterion is markedly higher (more than 0.7 K) and much higher for the rainy periods (> 4 K). While those calculations were applied on a daily basis, it is important to mention that the days are not uniform in terms of cloudiness or rain. Therefore, we used the data for 2-3 hours around the time of radiosonde launches to determine to which category a particular radiosonde profile belongs, clear-sky, cloudy or rain. In this way, we found that from 58 radiosonde launches used in our statistical analysis, 41 belong to the clear-sky category, 12 - to cloudy but non-precipitating conditions, and 5 - to rainy periods. For the four chosen clear-sky days not only were the daily uncertainties of 30 GHz Tb below 0.3 K, but both sets of uncertainties described above were extremely similar with the averaged difference less than 0.05 K.

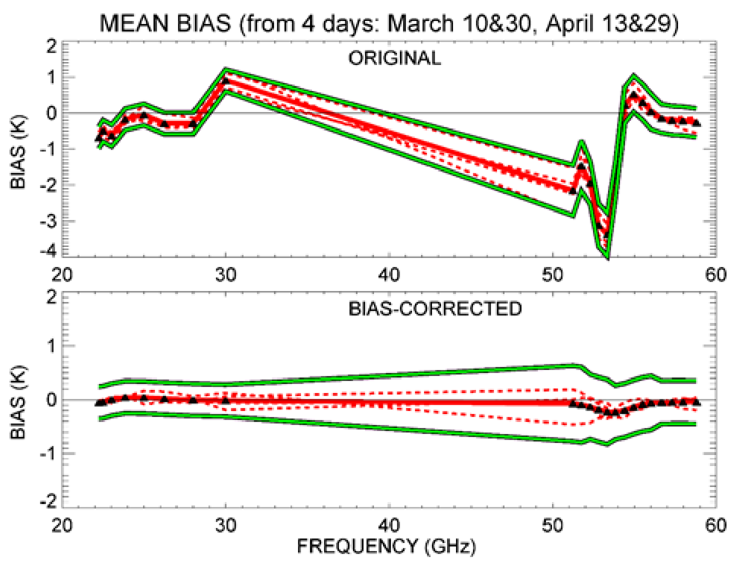
Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

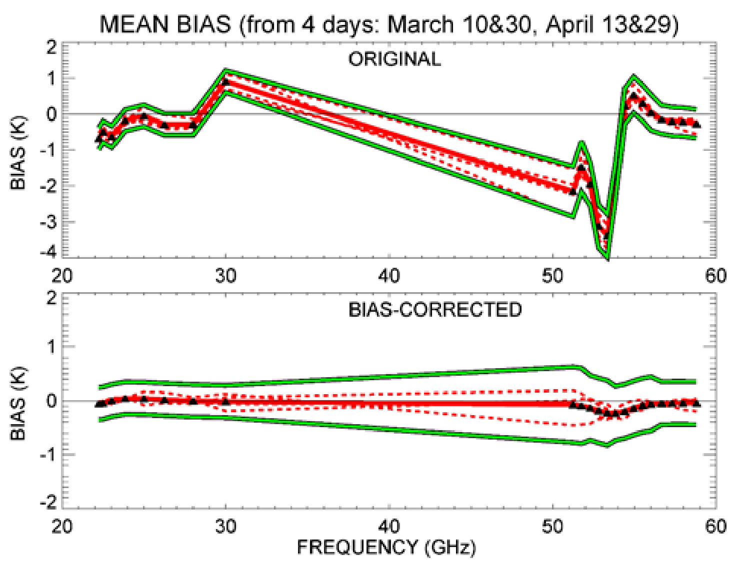
The bias was ~~then~~ computed ~~on all~~for each of the 22 channels as the averaged difference between the observed Tb from the MWR zenith observations, and the forward model calculation applied to the prior, over these selected clear-sky days, and then subsequently removed from all ~~measurements~~of the observations. We compute the bias in the bias-correction procedure only from the zenith scans, assuming that the same bias is suitable for other scans. Also, we assume that the true bias is an offset that is nearly independent of the scene, so that the sensitivity to the scene (e.g., clear or cloudy, zenith or off-zenith) is small. To investigate that, we eliminated the radiosondes launched during rainy periods (5 out of 58 cases) and found that the average temperature profiles were very little different than when all radiosonde profiles were used, with the maximum bias and RMSE absolute differences 0.12 K and 0.11 K respectively up to 5 km AGL. Fig. 1 shows the results of the bias-correction for the four chosen clear-sky days. The green lines on this figure indicate the MWR random errors~~ at each frequency calculated as the standard deviation of Tb averaged over one-hour sliding window~~; these are 0.3-0.4 K for K-band channels and 0.~~6~~4-0.7 K for V-band channels.

Formatted: *Font color: Black*

Formatted

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*





Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Fig.1. Bias for the four chosen clear-sky days (red-dashed lines) and their mean (red solid line) for the original observations in the top panel, and for the bias-corrected data in the bottom panel. Green lines are the uncertainty boundaries around the mean bias. Frequencies used in the PR algorithm are marked with black triangles in both panels.

Formatted: *Indent: First line: 0"*

~~This bias correction was applied to the brightness temperature used in the PR approach; however, the NN retrievals used the uncorrected brightness temperature, since it was non-trivial for us to reprocess those retrievals.~~

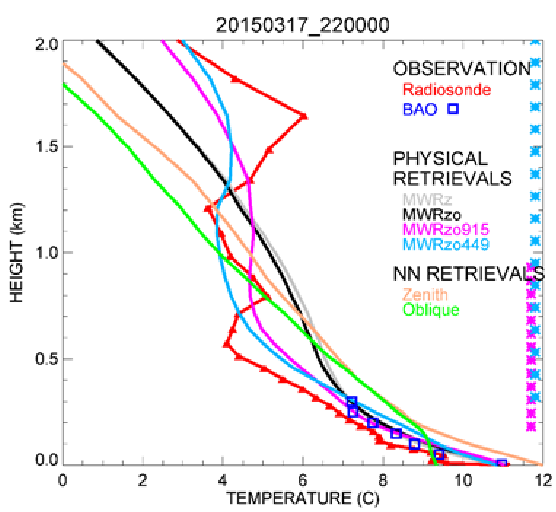
The retrieved profiles of the four different PR configurations presented in Table 1 (MWRz, MWRzo, MWRzo915, MWRzo449) were compared to the radiosonde profiles~~, as well as to the NN retrievals.~~. BAO tower temperature and mixing ratio data at the seven available levels were used as an additional validation dataset, without any interpolation.

To compare radiosonde observations against the PR ~~and NN retrieved~~ profiles, all these profiles were interpolated vertically to the same PR heights, and PR~~ and NN~~ profiles were averaged in the time window between 15 minutes before and 15 minutes after each radiosonde launch. Since the radiosonde ascends quite quickly in the lowest kilometers of the atmosphere (~15-20 min to reach 5 km), we estimated that the 30-minute temporal window is representative of the same volume of the atmosphere measured by the radiosonde.

An example of the different temperature retrievals and their relative performance, data obtained on 17 March 2015 at 2200 UTC is presented in Fig. 2. Temperature profiles up to 2 km

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

AGL from the four PR configurations (MWRz, MWRzo, MWRzo915, MWRzo449) are compared to the radiosonde data in red~~,~~ and to the BAO measurements in blue squares~~, and to the NN profiles (NN zenith in beige, and NN oblique in green).~~. The MWRz and MWRzo profiles~~, as well as those from the NNs,~~ are very smooth and depart quite substantially from the radiosonde measurements, being unable to reproduce the more detailed structure of the atmospheric temperature profile measured by the radiosonde, while the MWRzo449 profile (in light-blue) demonstrates a better agreement with both the radiosonde and BAO measurements (blue squares). Note that all four of the PRs match the BAO observations reasonably well~~, while the NN retrievals are warm-biased.~~ . The MWRzo915 profile (in magenta) also tries to follow the elevated temperature inversion observed by the radiosonde, successfully only in the lower part of the atmosphere (below 1 km AGL) where RASS 915 measurements are available. This behavior will be also addressed in the following section and in the statistical analysis presented later in the manuscript.



Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

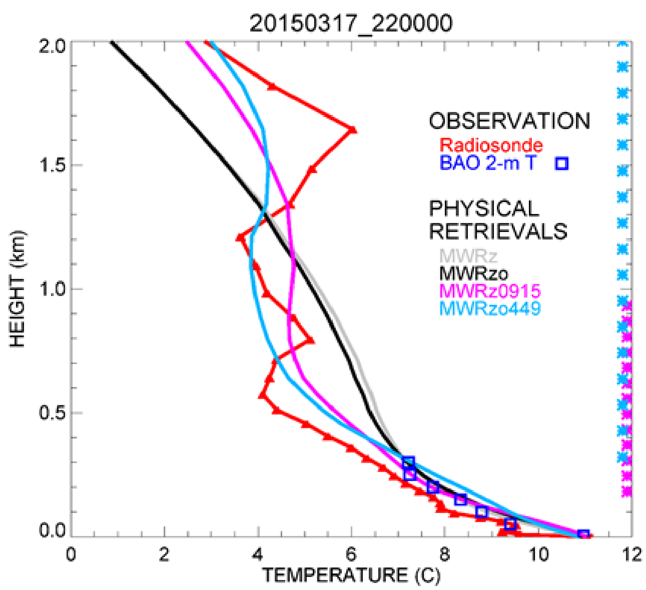


Fig. 2. Temperature profiles obtained by the four PR configurations: MWRz in gray, MWRzo in black, MWRzo915 in magenta, and MWRzo449 in light-blue~~; NN retrievals: NN zenith in beige, and NN averaged oblique in green.~~. These retrievals are compared to radiosonde measurements, in red, and BAO tower observations, in blue squares. The heights with available RASS virtual temperature measurements (RASS 915 in magenta and RASS 449 in light-blue), are marked by the asterisks on the right Y-axis.

3.3 Averaging kernel

The averaging kernel, **Akernel** (Masiello et al., 2012, Turner and Löhnert, 2014) from Eq. (1) can be calculated as:

$$\boldsymbol{Akernel} = \boldsymbol{B}^{-1}\,\boldsymbol{K}^{T}\,\boldsymbol{S}_{\varepsilon}^{-1}\,\boldsymbol{K} \qquad (2)$$

__Formatted:__ Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

where:

$$B = S_a^{-1} + K^T S_\varepsilon^{-1} K$$

Both matrices, **Akernel** and **B**, have dimensions 111 x 111 in our configuration. The **Akernel** ~~matrix has~~ provides useful information about the calculated retrievals, such as vertical resolution and degrees of freedom for signal at each level. Thus, the rows of **Akernel** provide the smoothing functions that have to be applied to the retrievals (Rodgers, 2000) to help minimize the vertical representativeness error in the comparison between the various retrievals and the radiosonde profiles due to very different vertical resolutions of these profiles.

Using the averaging kernel, the smoothed radiosonde observed profiles will be therefore computed as:

$$X_{smoothed\_sonde} = Akernel\ (X_{sonde} - X_a) + X_a \qquad (3)$$

The **Akernel** in Eq. (2) depends on the retrieval parameters (e.g., which datasets are used in the **Y** vector, the values assumed in the observation covariance matrix **S_ε**, and the sensitivity of the forward model (i.e., its Jacobian), etc.), so for our four PR configurations it is possible to calculate four different kernels: **A_MWRz, A_MWRzo, A_MWRzo915 and A_MWRzo449**, respectively.

While the top left corner of the **Akernel** matrix (1:55, 1:55) is devoted to temperature, and it will be called **AT_MWR** hereafter, the next (56:110, 56:110) elements are devoted to water vapor mixing ratio, and will be called **AQ_MWR**.

Formatted: *Font: Bold*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

For each of the four **Akernels**, a smoothed radiosonde profile can be computed for each radiosonde profile using Eq. (3). In the presence of temperature inversions or other particular structures in the atmosphere these smoothed profiles can be quite different from each other and also from the original unsmoothed radiosonde profile.

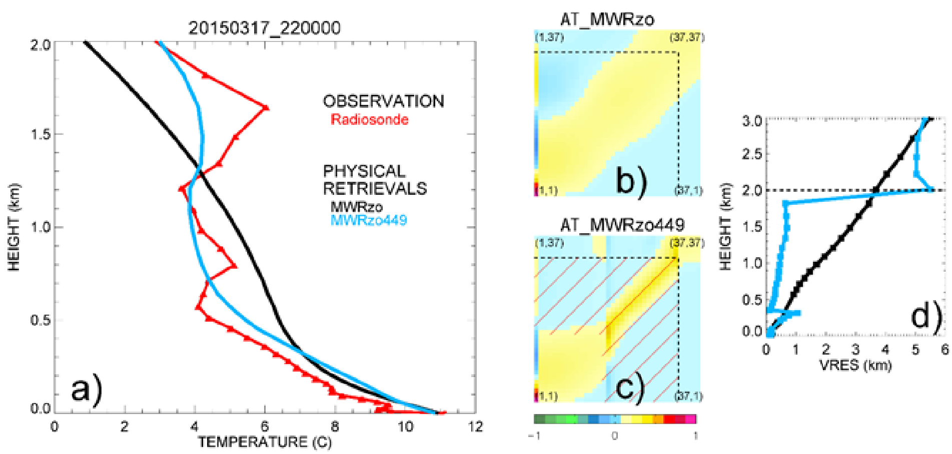
Therefore, in the statistical analysis presented later in the manuscript (in section 4.2), mean bias, root mean square error (RMSE), and Pearson correlation coefficients will be computed between the MWR's retrievals and both the unsmoothed and smoothed radiosonde profiles, where the latter were computed using their respective **Akernels**. Additional observational data help to resolve the atmospheric structure in more detail, therefore we would expect to obtain better statistical evaluations from the configurations including additional RASS observations compared to the runs without RASS data.

The improvement in the retrieved temperature profiles presented in Fig. 2 obtained using additional RASS data can be explained and clearly shown by the **ATkernel** itself. Figure 3 includes the temperature profiles of the radiosonde (unsmoothed and **ATkernel**'s smoothed) and PRs of MWRzo and MWRzo449 (panel a), and the **ATkernels** corresponding to these PRs in the color plots in the middle of the figure (panels b and c). These color plots are a schematic visualization of the 37 x 37 top left corner of the **ATkernel** matrix that illustrates the part of the **ATkernel** up to 3 km, for reference. Dash lines mark the 2 km vertical level.

The rows of the **ATkernel** provide a measure of the retrieval smoothing as a function of altitude, so the full-width half maximum of each **ATkernel** row estimates the vertical resolution of the retrieved solution at each vertical level (Merrelli and Turner, 2012). These plots of

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

temperature vertical resolution ~~vs~~versus height for MWRzo and MWRzo449 are included in Figure 3, panel d, for the same case presented in Fig. 2. Comparison of **ATkernel** color plots and vertical resolution plots of MWRzo vs MWRzo449 shows that additional observations from the RASS 449 significantly reduces the spread around the main diagonal from ~200m up to 2 km (in the layer of the atmosphere where RASS 449 measurements are available), thereby improving the vertical resolution of the retrievals (as clearly visible in panel d).



Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

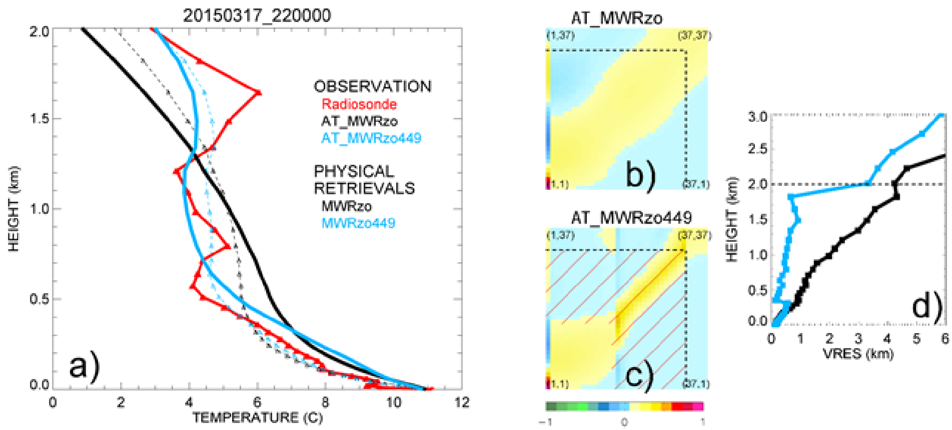


*Fig. 3. Panel a: observed temperature profiles from radiosonde, in red, from ATkernels smoothed radiosonde, **AT_MWRzo** in dashed black, and **AT_MWRzo449** in dashed light-blue; PRs from MWRzo PR in black, and from MWRzo449 PR in light-blue. Middle colored panels: 37x37 levels (surface to 3 km) of the Akernel matrix for temperature, b) **AT_MWRzo** and c) **AT_MWRzo449**. Right panel d: vertical resolution (VRES) as a function of the height for the MWRzo PR (black), and for the MWRzo449 PR (light-blue). ~~Dash~~Dashed lines on plots b)-d) mark 2 km AGL. Hatched area on panel c~~)~~ marks the RASS measurement heights.*

4. Results

PR~~ and NN retrieved~~ profiles have been evaluated against radiosonde observations. For additional verification, radiosonde data from 59 launches taken between 9 March and 4 May 2015 were first of all compared to the BAO tower measurements, up to 300 m AGL. These observed data sets match very well, with a correlation coefficient of 0.99 and a standard

Formatted: *Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

deviation of ~0.7 ~~°K.~~ °C. However, one radiosonde profile showed a large bias (> 5 ~~°K~~°C) against all seven levels of BAO temperature measurements and against all PRs ~~and NNs~~, therefore we decided to exclude this particular radiosonde profile from the statistical calculations.

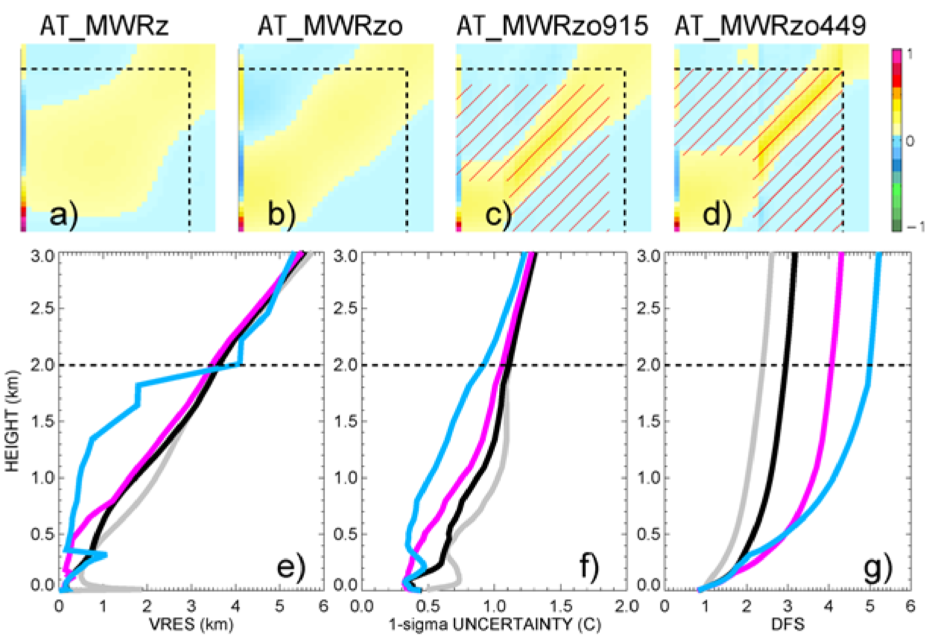
4.1 ~~PRs~~Physical retrieval statistical analysis from Akernel

To complete the analyses on the **ATkernel** changes and dependencies from different types of observational data used in the PRs, the **ATkernels**, averaged over all radiosonde events, are shown in Fig. 4, panels a-d, for the four PR configurations of Table 1, in the same way as shown in Fig. 3, b-c. A clearly visible gradual narrowing of the spread around the main diagonal is obtained by the usage of the additional observations, from MWR zenith only (panel a), to MWR zenith-oblique (panel b), to the larger impact obtained by the usage of RASS 915 (panel c) and RASS 449 (panel d) data.

Other statistically important features to analyze in the PRs, besides vertical resolution, are the retrieval uncertainty, and the degree of freedom for signal (DFS). These three features are also shown in Fig.4, panels e-g, at each of the heights of the retrieved solution, up to 3 km AGL, and averaged over all radiosonde events. While the vertical resolution (panel e) shows the width of the atmosphere layer used for each retrieval height (the vertical resolution is computed as the full-width half-maximum (FWHM; Maddy and Barnet, 2008) value of the averaging kernel), the uncertainty (panel f) gives a measure of the retrieval correctness (computed by propagating the uncertainty of the observations and the sensitivity of the forward model), and the DFS (panel g) is a measure of the number of independent pieces of

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

information used in the retrieved solution. For example, at the 1 km AGL level the vertical resolution of MWRzo449 equals 0.5 km, i.e. information from +/- 0.5 km around the retrieval height are considered in the retrieval, while all other retrievals use the information from +/- 2 km. Also, the uncertainty of the MWRzo449 retrieval up to ~~1km~~1 km AGL is around 0.5 ~~°K~~°C while the other retrievals have higher uncertainties of up to 1 ~~°K~~°C. The higher accuracy of the MWRzo449 retrievals is because they use more observational information compared to the other retrieval configurations.



Formatted: Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

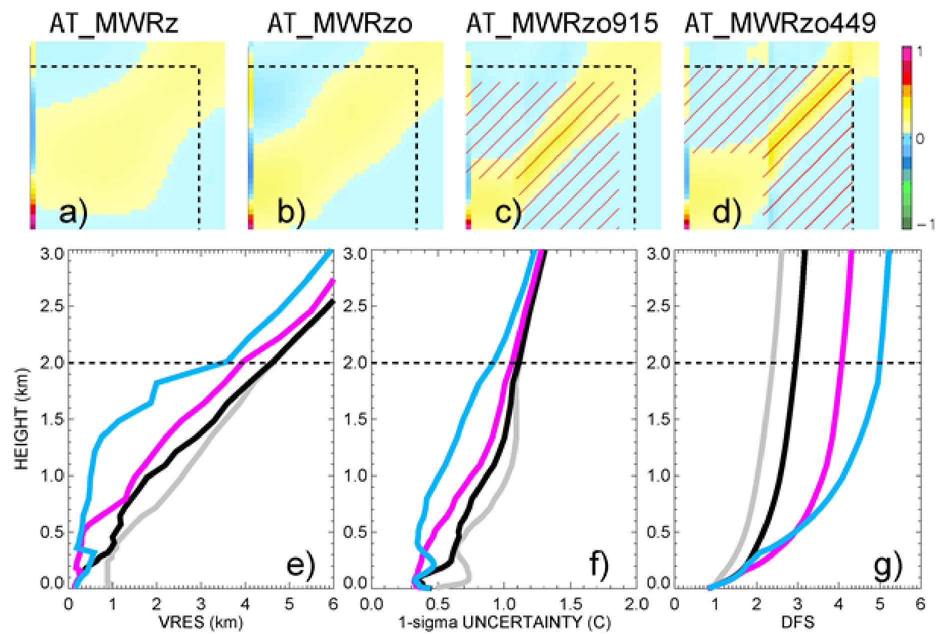


*Fig. 4. Top four-color images: **ATkernels** for MWRz (panel a), MWRzo (panel b), MWRzo915 (panel c) and MWRzo449 (panel d), averaged over all radiosonde events. Hatched area on panels c) and d) marks the RASS measurement heights. Bottom three panels from left to right: vertical resolution (VRES) in km (panel e), one-sigma uncertainty derived from the posterior covariance matrix in °C (panel f), and cumulative Degree of Freedom (DFS, panel g) as a function of height for temperature, averaged over all radiosonde events (MWRz is in gray, MWRzo is in black, MWRzo915 is in magenta, and MWRzo449 is in light-blue). ~~Dash~~Dashed lines mark 2 km AGL on all panels.*

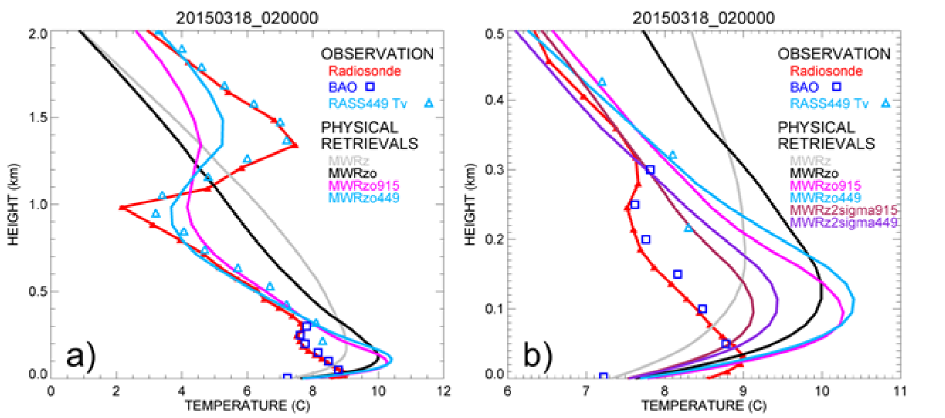
Formatted: Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

The improvements from MWRz (in gray) to MWRzo (in black), then to MWRzo915 (in magenta), and finally to MWRzo449 (in light-blue) are visible in all three panels (Fig 4 e-g), whereas MWRzo449 has the best statistical measures compared to the other PRs, particularly below 2 km AGL, where RASS 449 measurements are available. Finally, it is interesting that below 200 m AGL the MWRzo915 has slightly better statistics compared to the MWRzo449, as could be expected due to the first available height of the RASS 915 being lower (120 m AGL) than the first available height for the RASS 449 (217 m AGL) and due to the finer vertical resolution of the 915-MHz RASS. This suggests that if additional observations were available in the lowest several 100 m ~~layer~~ of the atmosphere where RASS measurements are not available, improvements might be even better closer to the surface, where temperature inversions, if present, are sometimes difficult to retrieve correctly.

As a matter of fact, we found several cases during XPIA when the temperature profile exhibits inversions, with the lowest happening in the surface layer. Figure 5a shows one of the most complex cases, with several temperature inversions visible in the temperature profile from the radiosonde (red line), in the temperature measurements from the BAO tower (blue squares), and in the virtual temperature measured by the RASS 449 (light blue triangles). We note that the virtual temperature profile is in close agreement with the temperature measured by radiosonde. Generally, the moisture contribution to the virtual temperature is less than a degree K, decreasing substantially for dryer air. Among the PR profiles, the PRs including RASS data show better agreement with the radiosonde in the atmospheric layer where RASS measurements are available, as ~~was already~~ shown in Fig. 2 for a different date. Unfortunately,

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

this better performance is not visible below the first available RASS measurement, i.e. from the surface up to ~200m AGL, where the PRs with additional RASS data have the largest positive bias compared to both radiosonde and BAO data in this layer. We ~~believe~~found that the MWR data, especially those from the oblique scans, in this case have a bias in the observed brightness temperatures that propagates through the retrieval calculations, and including other observational data is not enough to correct it in the layer between the surface data and the first available RASS measurement.



Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

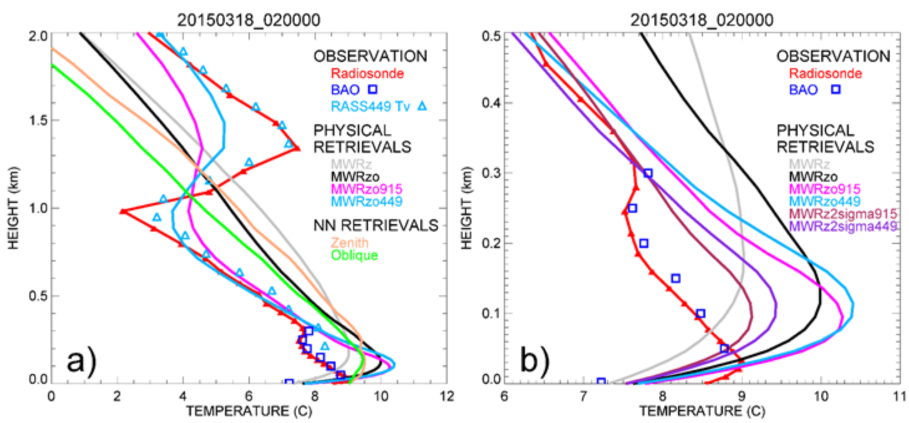


Fig. 5. Panel a~~)~~, as in Fig. 2 but for 18 March 2015 at 0200 UTC. The RASS 449 virtual temperature is included as light blue triangles. Panel b~~)~~ shows the same data ~~(except for the NN retrievals)~~ presented in panel a~~),~~, but only up to 500 m AGL, and includes PR profiles in which the MWR uncertainties were increased by a factor of two, MWRz915 in maroon and MWRz449 in violet.

After several trials, we found that when RASS measurements are included, temperature profiles in this and similar cases exhibiting inversions could be improved -by increasing the random uncertainty of MWR observations, and only using the zenith MWR measurements, because the oblique MWR brightness temperature measurements (which give more information in the lower layer of the atmosphere) seemingly have a bias that competes with the active and more accurate measurements from the RASS and surface observations. In this way, the PR approach is granted more freedom to get an optimal profile in the gap between the

Formatted: Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border) lowest RASS measurements and the surface measurement. Proof of this is presented in Figure 5b, that shows the same data as in 5a, but including the profiles obtained when increasing the assumed MWR Tb uncertainties by a factor of two, hereafter called MWRz2sigma915 and MWRz2sigma449, in maroon and violet respectively. The increased accuracy of these temperature profiles compared to MWRzo915 and MWRzo449 are obvious in the layer of atmosphere closer to the surface. Later we will show that these last two PR configurations demonstrate improved statistics over all 58 cases, and also through the layer of the atmosphere up to ~~5km~~5 km. We note that these last two PR configurations, that were found to work well for this dataset, might not be optimal for other datasets. During XPIA the RASS measurements impact (particularly those from the RASS 449) was important in the PR approach. This might not be the case for other datasets or over different seasons, when RASS coverage might not be as good as ~~that~~ during XPIA. For this reason, we think that attention has to be used to determine what is the best configuration to use when dealing with PR approaches. On the positive side, the advantage is that the user can determine and has control on what is the optimal configuration to use in his/her dataset, in terms of different inputs to employ and their relative uncertainty.

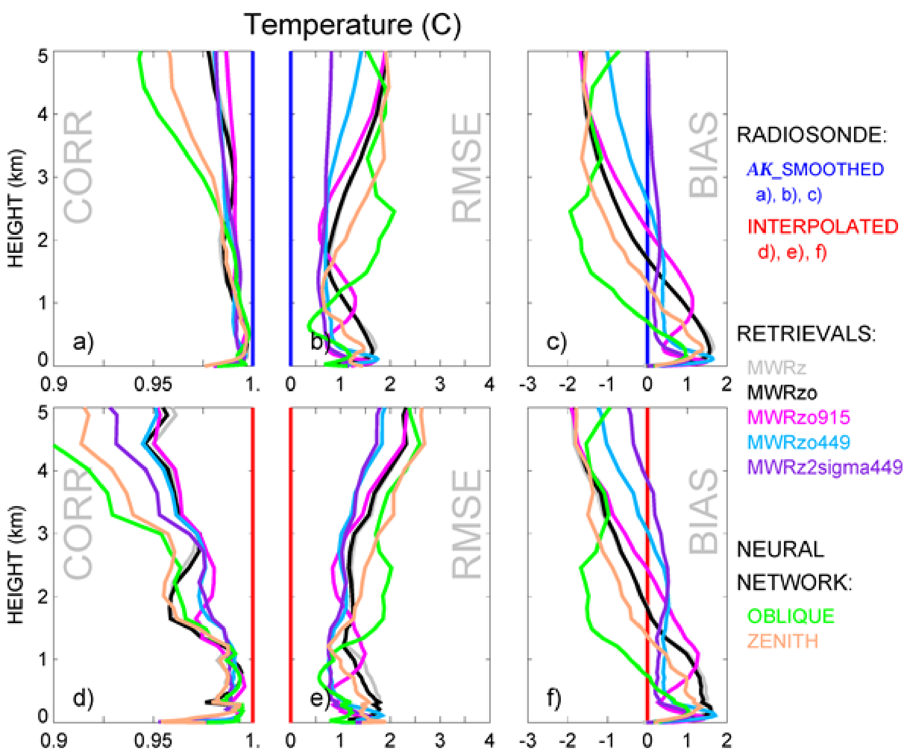
4.2 Statistical analysis of ~~PRs compared to NN~~physical retrievals up to 5km AGL

~~Since the iteratively~~ We calculated ~~PRs and the NN retrievals are obtained by very different approaches, we find it very important to compare their~~ the relative statistical behavior~~. We do this~~ of PRs for both ~~for~~ temperature and mixing ratio, providing ~~this~~the comparison in two ways: first ~~using~~to the ~~**Akernel**~~ smoothed radiosonde ~~data obtained~~ using

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

the averaging kernel matrix (as described in section 3.3~~,~~), and second comparing to the original, unsmoothed, radiosonde profiles, just interpolated to the 55 PR vertical levels.

Figure 6 shows the statistical results of these comparisons for temperature, in terms of Pearson correlation, RMSE, and mean bias, averaged over all radiosonde events.



Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

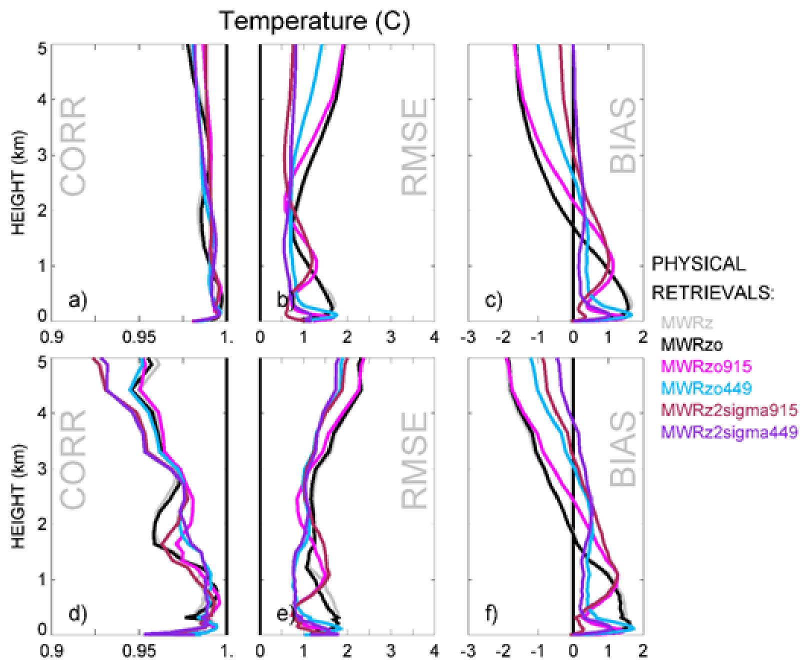


*Fig. 6. Pearson correlation, RMSE, and mean bias for temperature profiles of MWRz in gray, MWRzo in black, MWRzo915 in magenta, MWRzo449 in light-blue, MWRz2sigma915 in maroon and MWRz2sigma449 in violet, computed comparing to smoothed radiosonde data (using their relative **ATkernel**) in panels a-c, and against the original radiosonde measurements in panels d-f.* ~~*The same comparisons for NN profiles, with NN zenith in beige, and NN averaged oblique in green, are made against the corresponded smoothed radiosonde data in the top panel and against original radiosonde data in the bottom panel.*~~

__Formatted:__ Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

These results confirm the superiority of the MWRz2sigma449 temperature retrieval over the other PRs. While this is not true at all heights, this retrieval shows improved distribution of RMSE and bias for the atmospheric layer up to 5 km AGL. The differences between the MWRz2sigma915 ~~profile is not included in the figure to not overcrowd it, but its behaviour compared to~~and the MWRzo915 ~~is~~profiles are similar to ~~that of~~those between the MWRz2sigma449 ~~compared to~~and the MWRzo449 ~~profile~~profiles, reducing the drastic bias found in the layer closer to the ground. The differences between the two ways of comparison, against the smoothed **ATkernel** or the original radiosonde data, are small in terms of RMSE and bias, but more evident in terms of correlation as ~~it~~ can be expected because of the smoothing technique applied to the radiosonde profiles through Eq. (3). Above and below ~1.~~5~~6 km AGL the bias, RMSE, and correlation profiles of the PRs show very different behavior. While statistical ~~measures~~scores above ~1.~~5~~6 km AGL are very similar for the four PRs introduced in Table 1, they are better for the MWRz2sigma915 and MWRz2sigma449 ~~PR~~PRs, especially when compared to the smoothed radiosonde profiles. Differences between the profiles show more variability in the lowest ~~1.5 km. NN retrievals, both for zenith and averaged oblique, are very variable from height to height and generally have much larger RMSE and bias, and worse correlation coefficients compared to PRs.~~~1.6 km where most of the active RASS measurements are available. Also, while both PR profiles related to the RASS 449, MWRzo449 and MWRz2sigma449, have almost constant bias and RMSE from 200m up to at least 3 km, the RASS 915 based PR profiles, MWRzo915 and MWRz2sigma915, have biases and RMSEs that vary with height. Due to the lower first range gate of the RASS 915 measurements, the PR profile of MWRz2sigma915 has the smallest bias and RMSE compared to all other PR profiles in

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

the surface to 200 m layer. With quickly decreasing availability of RASS 915 measurement above this layer, the bias and RMSE of MWRzo915 and MWRz2sigma915 became larger, and in some higher layers even larger than the corresponding statistical measures of MWRz and MWRzo. This marks the importance of active measurements spanning a prominent vertical layer to provide a useful application of these data in a radiative transfer model.

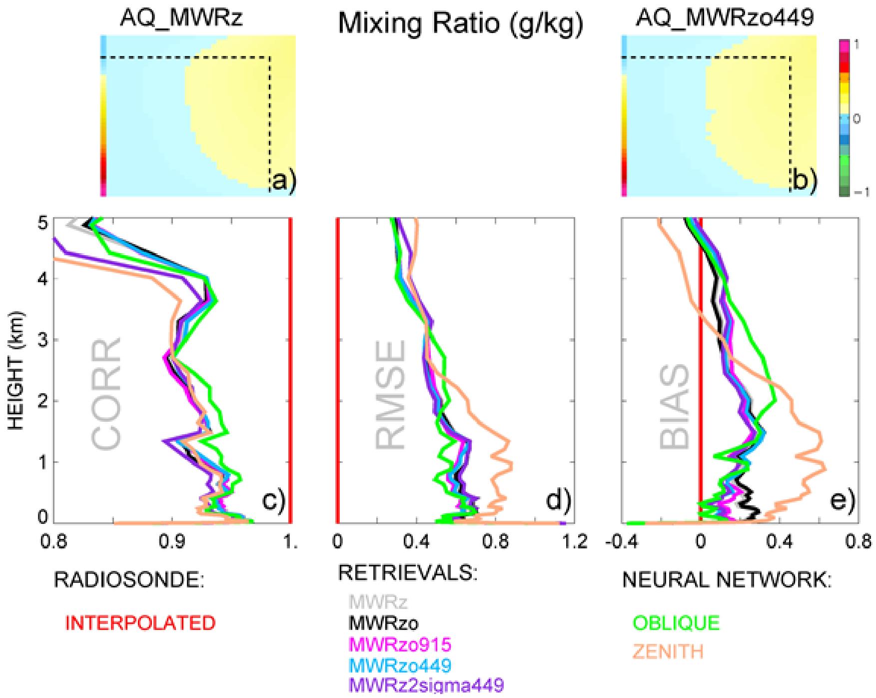
Besides temperature profiles, the ~~NN and~~ PR retrievals also provide water vapor mixing ratio profiles. It is understandable that the different configurations of PRs are not noticeably different from each other in relation to moisture, because the Tv observations from the RASS are dominated by the ambient temperature (not moisture), and thus have little impact on the water vapor retrievals.

Figure 7 includes the two **AQkernels** corresponding to the PRs MWRz and MWRzo449 in panels a and b, which are averaged over all radiosonde events and appear to be almost identical. More detailed statistical estimations of PRs mixing ratio in Fig 7 c-e, also averaged through all radiosonde events, show very similar correlations, RMSEs, and biases for all PRs included in the figure, meaning that the impact of including RASS observations is minimal on this variable. ~~These PR mixing ratio profiles are also statistically very close to the averaged oblique NN retrieval mixing ratio profiles, with the zenith NN retrieval mixing ratio profiles showing the worst statistics in terms of RMSE and bias. Overall, we conclude that the PR retrievals are not degraded on average compared to the NN moisture retrievals.~~

__Formatted:__ Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

700



Formatted: Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

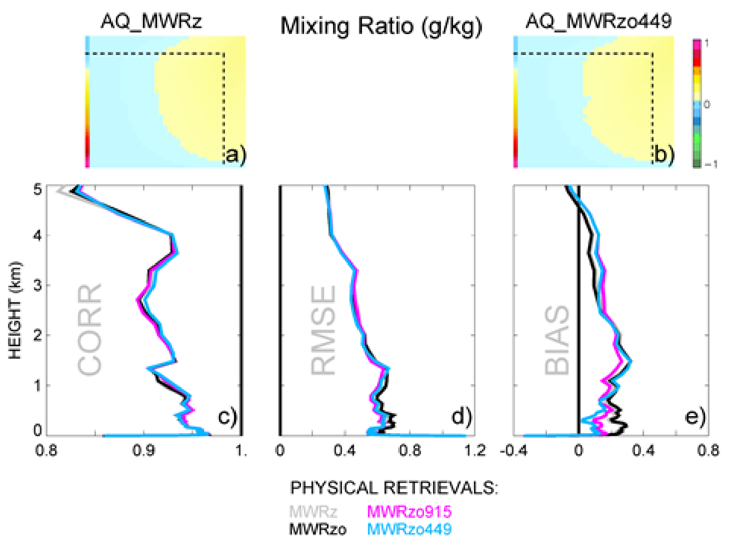


Fig. 7. Top two-color images: ***AQkernels*** *for MWRz (panel a) and MWRzo449 (panel b), averaged over all radiosonde events and shown up to 3 km AGL with dash lines mark 2 km AGL on both panels. Bottom three panels are the same as panels d-f in Figure 6, but for mixing ratio estimation.*

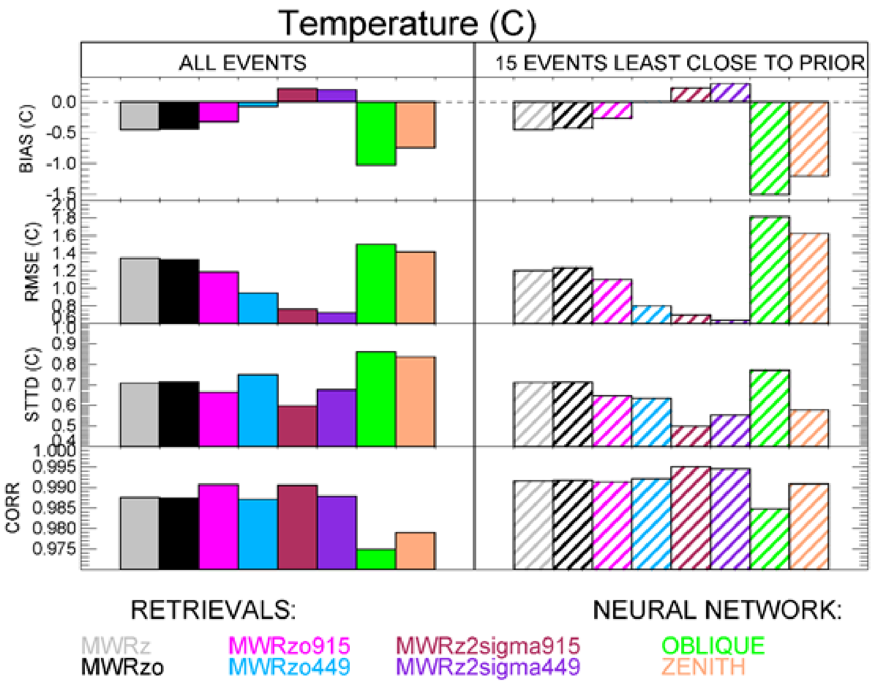
4.3 Statistics for cases far from the climatological mean

~~While both approaches, physical and neural network retrievals, are quite different,~~ ~~both~~Physical retrievals use climatological data as a constraint or for building the statistical relationships used in the retrieval. Statistically, the averaged profiles of both temperature and moisture variables are very close to the climatological averages. However, the most interesting

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

and difficult profiles to retrieve are the cases furthest from the climatology (Löhnert and Maier, 2012). To check the behavior of the retrieved data in such events, we first calculated the RMSE for each radiosonde profile relative to the prior profiles for 42 vertical levels from the surface up to 5 km AGL, and then we selected the 15 cases with the largest 0-5km layer averaged RMSEs compared to the prior. All comparisons are done against the corresponded smoothed **ATkernel** radiosonde data, using **AT_MWRz, AT_MWRzo, AT_MWRzo915, AT_MWRzo449, AT_MWRz2sigma915, AT_MWRz2sigma449** for all six PRs~~, and~~ ~~**AT_MWRz, AT_MWRzo**~~ ~~for NN zenith and NN oblique retrievals respectively~~.

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*



Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

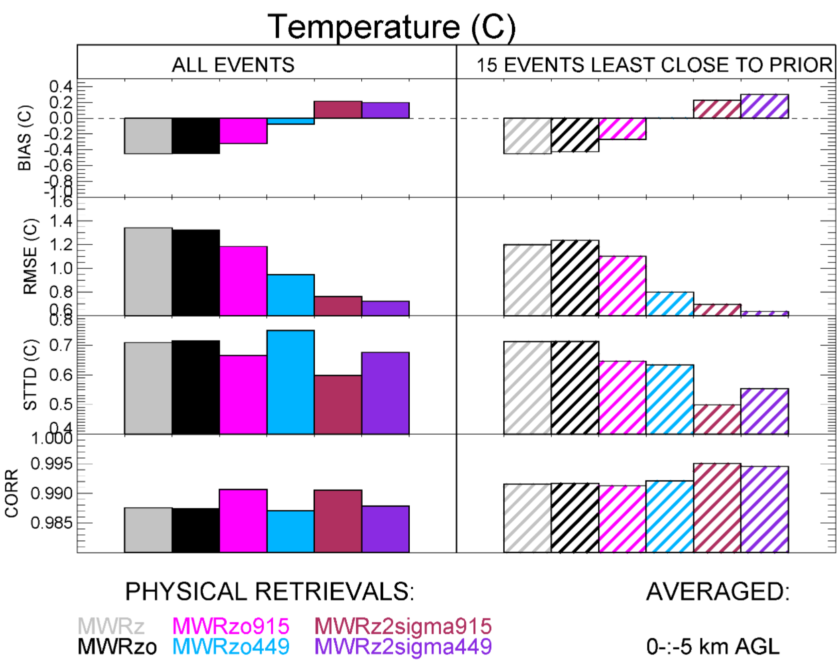


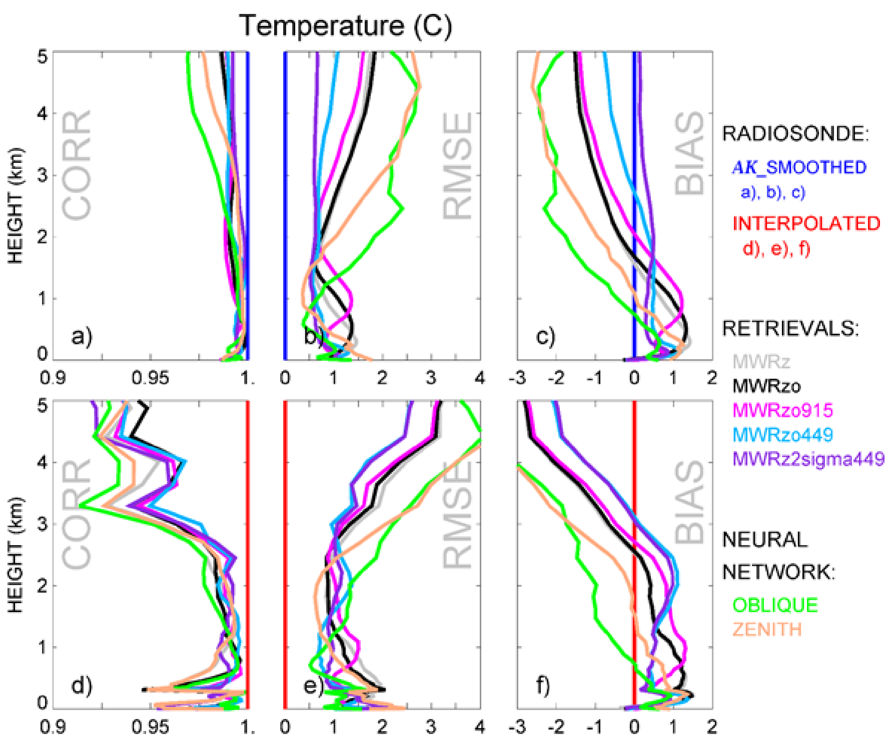
Fig. 8. From top to bottom: biases (retrievals minus ATkernel radiosonde), RMSEs, standard deviations of the difference between retrievals and ATkernel radiosonde, and Pearson correlations for the six PR configurations so far introduced ~~and both NN retrievals~~, averaged from the surface to 5 km AGL, averaged over all radiosonde data (solid boxes), and averaged over the 15 events furthest from the priors (hatched boxes).

Figure 8 shows the temperature statistical analysis for the entire radiosonde data set (solid boxes) and to just the fifteen chosen events (hatched boxes) for bias, RMSE, standard deviation of retrieval differences to the radiosonde data, and Pearson correlation, calculated as

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

the weighted averaged over the 42 vertical heights up to 5 km AGL. The vertical resolution of the Physical Retrievals is not uniform, with more frequent levels closer to the surface. If a simple average of the data from all levels is used, the near-surface layer will be weighted more compared to the upper levels of the retrievals. To avoid this, a vertical average over the lowest 5km AGL is performed using weights at each vertical level determined by the distance between the levels. Differences in the statistics when using the entire radiosonde data set or the fifteen profiles furthest from the prior are noticeable, especially for bias and RMSE, but also for the standard deviation. All PRs that include RASS observations show better performance compared to strictly MWR-only PR profiles (i.e., MWRz and MWRzo) for almost all statistical comparisons. Also, the statistical behavior of the MWRz2sigma915 and MWRz2sigma449 retrievals are the best in terms of RMSE and standard deviation for all events and for RMSE, standard deviation, and correlation coefficient, for the fifteen profiles furthest from the climatological average. Fig. ~~Finally, we note~~8 also shows that ~~the NN profiles are the least accurate retrievals for all of the statistics for the entire radiosonde data set,~~RMSE, standard deviation and ~~have the highest bias, RMSE and the lowest~~ correlation have improved scores for the 15 events furthest from the prior when compared to all temperature profiles for all PRs using active RASS measurements.

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*



749

Formatted: Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

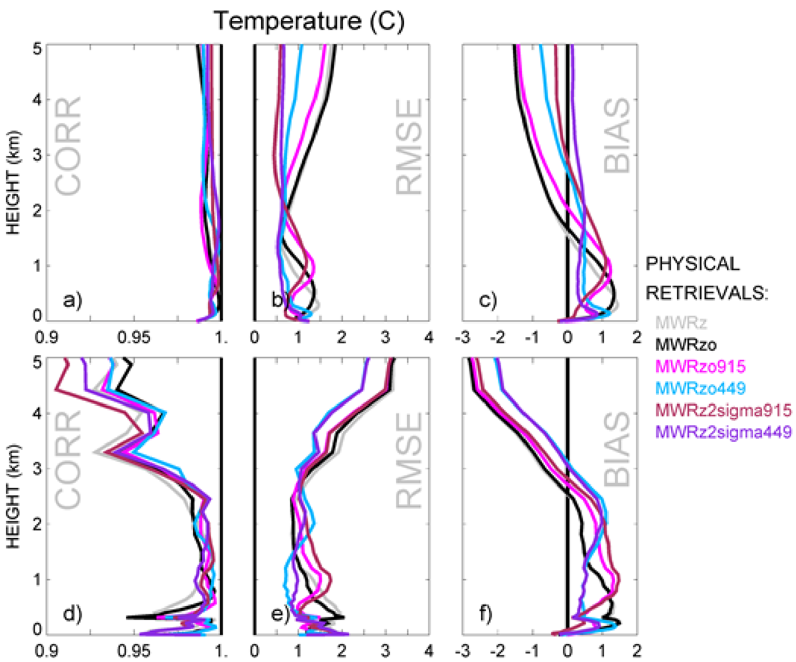


Fig. 9. The same as Fig. 6 but for the temperature over 15 furthest from prior radiosonde profiles.

To investigate the vertical structure of the error statistics for the 15 events furthest from the radiosonde climatology, profiles of correlation, RMSE and bias for these events are shown in Figure 9 for the layer 0-5 km. ~~-~~The ~~MWRz449~~MWRz2sigma915 and MWRz2sigma449 profiles, ~~which were seen in~~ having~~Fig. 8 to have~~ the best ~~layer~~ averaged statistics in Fig. 8, are seen ~~to be~~ as good as, or better, than the other methods for the 0-2 km layer.~~-~~ Importantly, for heights above 2km AGL, where there is no additional observational data from RASS, all of the PRs ~~are~~

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

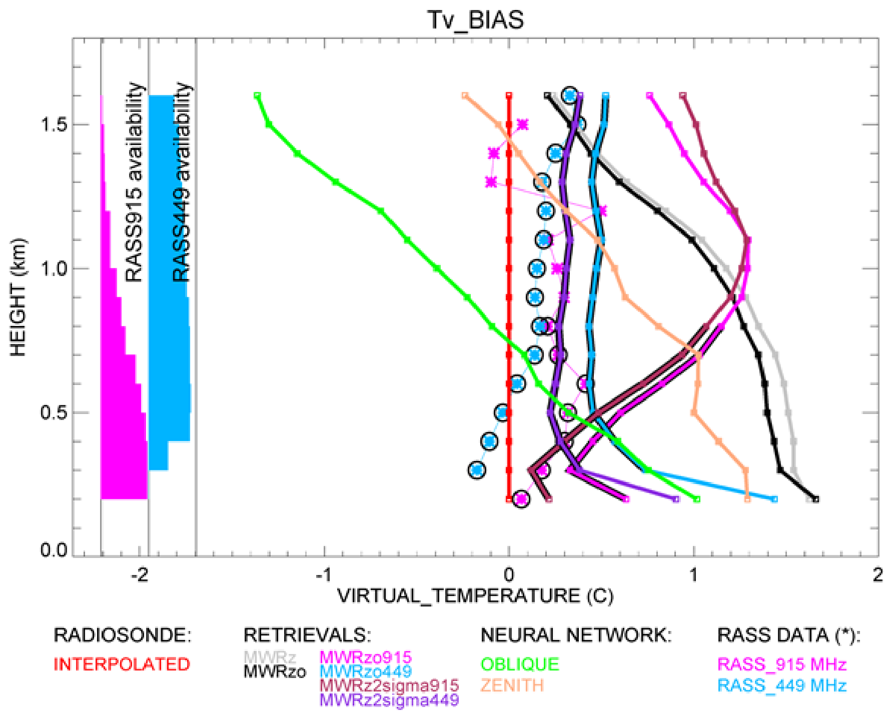
~~better than the NN profiles, with the MWRz2sigma449 and MWRz449 being the best. We note that the increased accuracy of the PRs relative to the NNs is more obvious in Fig. 9 for the 15 events when compared to the entire data set in~~ with RASS are closer to the "true" radiosonde temperature compared to the PRs without RASS.~~Fig. 6. Also, it can be seen that the NNs for the 15 events are worse than they are for the entire data set, especially in the 2-5km layer, which indicates (not surprisingly) that the NNs accuracy degrades when the atmosphere is far from its climatology.~~

4.4 Virtual temperature statistics

The above analysis confirms the superiority of MWRz2sigma915 and MWRz2sigma449 compared to the other PRs ~~and to the NN retrievals~~ for this dataset. In this section we show the direct comparison of the retrieved profiles to the original radiosonde and RASS virtual temperature profiles. Using temperature and moisture retrieval output, we calculated "retrieved virtual temperature profiles" and interpolated all profiles and RASS data on a regular vertical grid, going from 200 m to 1.6 km with 100 m range, for easy comparison.

Figure 10 shows Tv retrieved profile biases compared to the original radiosonde data as solid lines, and RASS 915 and RASS 449 Tv bias as asterisks. A zero bias is denoted by the red line. On the left side of the figure we show bar charts of the RASS measurement availability as a function of height. The widest part of these charts corresponds to 100% data availability. Heights with RASS availability greater than 50% are marked with additional circles over the asterisks.

Formatted: Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)



Formatted: Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

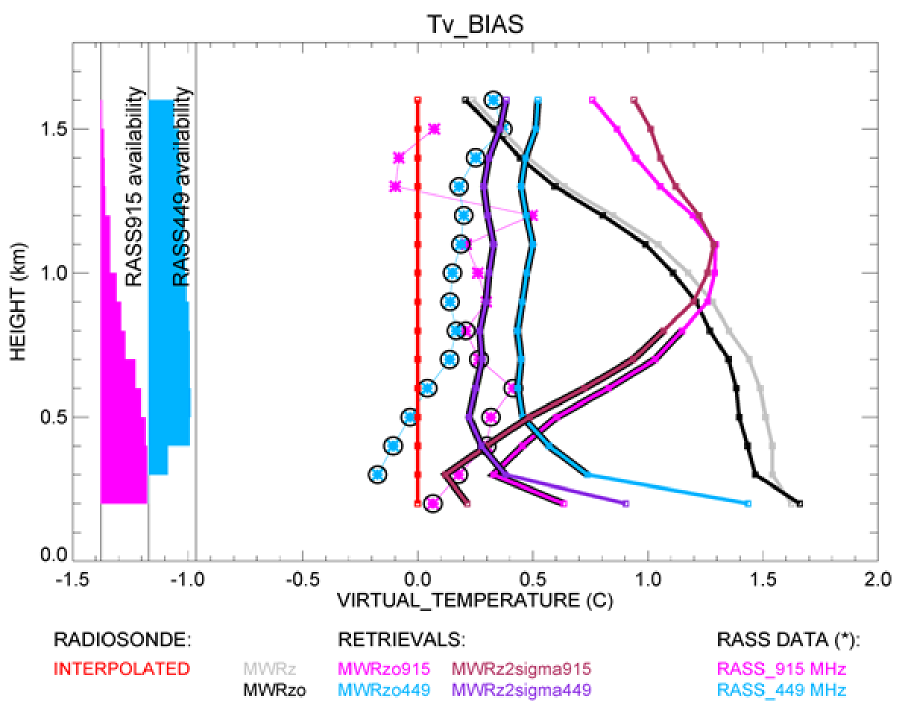


Fig. 10. Bias of virtual temperature for all six PR configurations ~~and both NN retrievals~~ compared to the original radiosonde measurements. RASS data are marked by asterisks and by additional circles for the RASS data with more than 50% availability, according to the availability bar charts on the left.

While RASS 449 data are available at almost all heights up to 1.6 km, the RASS 915 data availability decreases considerably with height, lowering to 50% availability around 800 m AGL. All PRs with input from RASS data, MWRzo915 and MWRzo449, and MWRz2sigma915 and MWRz2sigma449~~ with larger MWR uncertainties~~, are also marked with additional black lines at

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

the heights with at least 50% of relative RASS data availability. This figure clearly shows the superiority of MWRz2sigma449 and MWRz2sigma915 (in the layer with > 50% RASS 915 data availability) compared to MWRz and MWRzo configurations, which do not include RASS data, as well as to MWRzo915 and MWRzo449 which include RASS data and MWR zenith and oblique data. For MWRzo449 and MWRz2sigma449 profiles, RASS 449 data were almost always available, therefore it is easy to identify similar features between Tv bias profiles of the RASS 449 and the PRs including it. Thus, for the MWRzo449 and MWRz2sigma449 the Tv bias is more uniform through the heights compared to all other PRs that do not include RASS data~~, and to both NN retrievals.~~. Moreover, because MWRzo449 and MWRz2sigma449 Tv bias profiles follow tightly the trend of the RASS 449 with height, the difference between MWRzo449 and RASS 449 biases equals ~0.32 ^{o}C and the difference between MWRz2sigma449 and RASS 449 biases equals ~0.14 ^{o}C over the ~1.3 km atmospheric layer where most of RASS 449 measurements are available, uniformly distributed through the heights. Finally, the average differences between these MWRzo449 and MWRz2sigma449 Tv profiles and the radiosonde virtual temperature equal ~0.56 ^{o}C and ~0.34 ^{o}C respectively. ~~From these results we can assume that the final bias of the PRs that include additional RASS data derives from a combination of the RASS data bias itself, of the uncertainty of the retrieval model, and of the MWR brightness temperature biases, even though we tried to correct for the latter.~~

We note that as an alternative to using the PR ~~temperatures~~temperature profiles at all heights, one could consider replacing the PR temperatures with RASS observations up to the maximum height reached by the RASS, and then use the PR retrieval above that. ~~-~~To do this the

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

moisture contribution to the RASS virtual temperatures could be removed by using either the relative humidity measured by radiometer or by a climatology of the moisture term.

5. Conclusions

Formatted: *Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"*

In this study we used the data collected during the XPIA field campaign to test different configurations of a physical-iterative retrieval (PR) approach in the determination of temperature and humidity profiles from data collected by microwave radiometers, surface sensors, and RASS measurements. We tested the accuracy of several PR configurations, two that made use only of surface observations and MWR observed brightness temperature (zenith only, MWRz, and zenith plus oblique, MWRzo), and others that included the active observations available from two co-located RASS (one, RASS 915, associated with a 915-MHz, and the other, RASS 449, associated with a 449-MHz wind profiling radar). Radiosonde launches were used for verification of the retrieved profiles and Neural Network retrieved profiles were also used for comparison~~. The NN retrievals used in this study were obtained either using the zenith angle only, or the average of the oblique scans (based on the averaged Tb of 15- and 165-degree scans) without including the zenith. Other MWR systems (Rose et al., 2005) provide retrieved profiles that include the information from both oblique and zenith scans.~~ (see Appendix A).

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Inclusion of the observations from the active RASS instruments in the PR approach improves the accuracy of the temperature profiles, particularly when ~~low-level~~ temperature inversions are present. Of the PRs configurations tested, we find better statistical agreement with the radiosonde observations when the RASS 449 is used together with the surface observations and brightness temperature from only the zenith MWR observations ~~(MWRz2sigma449),~~ and doubling the random radiometric uncertainty on the MWR observations (MWRz2sigma449) relative to the uncertainty calculated over the selected clear-sky days ~~(Fig. 1).~~. This configuration is also more accurate compared to MWRzo915 or MWRz2sigma915 (which use RASS 915 observation), because of the deeper RASS 449 height coverage. ~~-~~The larger assumed radiometric uncertainty in the MWR Tb observations allows the retrieval to overcome both (a) the ~~(~~small~~)~~ systematic errors that exist between the MWR ~~(which could be in either the~~ observed Tb values ~~or in the MonoRTM used as the forward model)~~ and the RASS~~,~~ measurements and (b) the systematic errors that exist in forward microwave radiative models (Cimini et al. 2018).

We also selected 15 cases when temperature profiles from the radiosonde observations were the furthest from the mean climatological average, and reproduced the statistical comparison over this subset of cases. These are the cases usually the most difficult to retrieve and the most important to forecast; therefore, it is essential to improve the retrievals in these

Formatted: Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border) situations. Even for this subset of selected cases we find that MWRz2sigma449 produces better statistics, proving that the inclusion of active sensor observations in MWR passive observations would be beneficial for improving the accuracy of the retrieved temperature profiles also in the upper layer of the atmosphere where RASS measurements are not available (at least up to 5 km AGL). However, we note that this result may be dependent on the fact that our oblique measurements were taken at a 15-degree elevation angle, and that MWRs in locations with unobstructed views allowing for scans down to 5 degrees may provide similar improvements to the temperature profile accuracy in the lowest 0-1 or even 0-2 km AGL layers (Crewell and Löhnert, 2007).

Finally, we also considered the impact of the inclusion of RASS measurements on the retrieved humidity profiles, but in this case the inclusion of RASS observations did not produce significantly better results, compared to the configurations that do not include them. This was not a surprise as RASS measures virtual temperature, effectively adding very little extra information to the water vapor retrievals. In this case a better option would be to consider adding other active remote sensors such as water vapor differential absorption lidars (DIALs) to the PRs. Turner and Löhnert (2020) showed that including the partial profile of water vapor observed by the DIAL substantially increases the information content in the combined water

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

vapor retrievals. Consequently, to improve both temperature and humidity retrievals a synergy between MWR, RASS, and DIAL systems would likely be necessary.

Appendix A

The XPIA NN retrievals use a training dataset based on a 5-year climatology of profiles from radiosondes launched at the Denver International Airport, 35 miles south-east from the XPIA site. NN-based MWR vertical retrieval profiles were obtained using the zenith or an average of two oblique elevation scans, 15- and 165-degrees, all with 58 levels extending from the surface up to 10 km, with nominal vertical levels depending on the height (every 50 m from the surface to 500 m, every 100 m from 500 m to 2 km, and every 250 m from 2 to 10 km, AGL).

Fig. 1A shows composite NN vertical profiles of temperature (separately for the zenith and averaged obliques) calculated for radiosonde launch times, and the corresponding PR profiles already introduced in Fig. 6. As expected, the averaged oblique NN profile has lower bias and RMSE compared to the zenith NN profile below 1km AGL, while the zenith NN profile improved above this level.

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

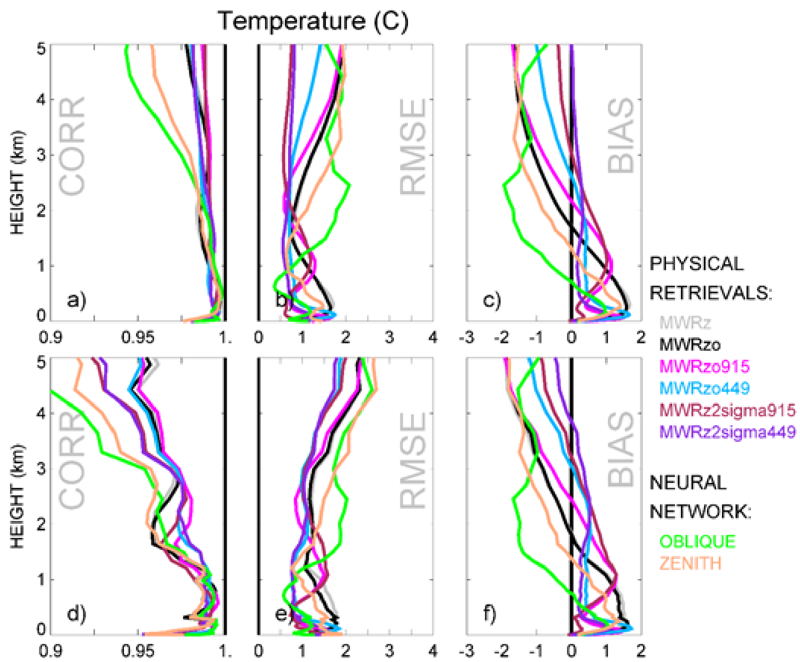


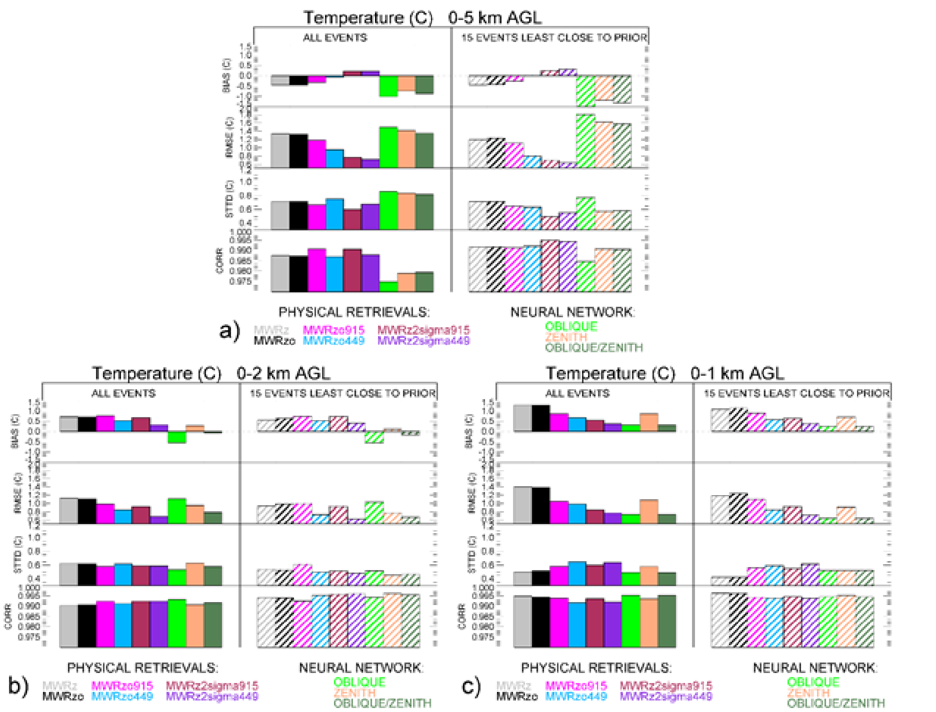
Fig. 1A. The same as Fig. 6 but with additional NN temperature profiles, from zenith in beige and from averaged oblique – in green.

We note that in this comparison the MWR Tb data have been bias-corrected before being used in the Physical Retrieval configurations, as discussed in Section 3.2, while the NN retrievals use the uncorrected Tb, since it was non-trivial for us to reprocess those retrievals. Zenith NN profiles have larger bias and RMSE and smaller correlation coefficient above 1 km

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

AGL compared to all PR profiles. This is possibly due to the Tb bias in the transparent channels of the V-band frequencies.

To optimize the use of the two types of NN scan data, we combined the NN retrieved profiles using only the averaged oblique scans up to 1 km AGL and the zenith scans above 1 km. Fig. 2A is the same as Fig. 8, now including also the three NN profiles (averaged oblique only, zenith only, and their combination) presenting the statistics in three different layers of atmosphere: from the surface to 5 km AGL, from the surface to 2 km AGL, and from the surface to 1 km AGL (a, b and c panels).



Formatted: Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

Fig. 2A. The same as Fig. 8 but including NN profile statistics from averaged oblique scans in beige, from zenith – in green, and from their combination – in spruce. Panels a, b, and c show the temperature statistics from the surface up to 5, 2 and 1 km AGL respectively.

Oblique only (and oblique and zenith combined) NN profiles show the best statistics in the layer closest to the surface, up to 1 km AGL, panel c, while in the deeper atmosphere layer up to 5 km all PR profiles have improved statistics compared to NNs, panel a. Panel b has mixed results: MWRz2sigma449 has the lowest RMSE, and the combined NN retrieved profiles show just slightly larger RMSE and almost the same standard deviation and correlation. It is important to admit that while potential NN bias-correction generally cannot change the oblique statistics, it may improve the zenith profiles, especially above 1 km AGL, therefore improving the combined NN profiles statistics.

Data availability

All data are publicly accessible at the DOE Atmosphere to Electrons Data Archive and Portal, found at https://a2e.energy.gov/projects/xpia (Lundquist et al., 2016).

Author contribution

Irina Djalalova completed the primary analysis with physical retrieval approach through MONORTM using XPIA data. Daniel Gottas contributed to the post-processing of the RASS data. Irina Djalalova prepared the manuscript with contributions from all co-authors.

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Acknowledgements

We thank all the people involved in XPIA for instrument deployment and maintenance, data collection, and data quality control, and particularly the University of Colorado Boulder for making the CU MWR data available. Funding for this study was provided by the NOAA/ESRL Atmospheric Science for Renewable Energy (ASRE) program.

Formatted: *Indent: First line: 0"*

References

Adachi, A. and H. Hashiguchi, 2019: Application of parametric speakers to radio acoustic sounding system. *Atmos. Meas. Tech.*, **12**, 5699–5715, 2019, https://doi.org/10.5194/amt-12-5699-2019.

Adler, B., J. M. Wilczak, L. Bianco, I. Djalalova, J. B. Duncan Jr., D. D. Turner, ~~2020~~2021: Observational case study of a persistent cold air pool and gap flow in the Columbia River Basin. *Under review to J. of Appl.* ~~*In preparation*~~*Meteor. Climatol*.

Banta, R. M., and coauthors, 2020: Characterizing NWP model errors using Doppler lidar measurements of recurrent regional diurnal flows: Marine-air intrusions into the Columbia River Basin. *Month. Wea. Rev.*, **148**, 927-953, ~~doi:10.1175/MWR-D-19-0188.1~~https://doi.org/10.1175/MWR-D-19-0188.1

Formatted: *Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

~~Blumberg, W. G., D. D. Turner, S. M. Cavallo, J. Gao, J. Basara, A. Shapiro, 2019: An Analysis of the Processes Affecting Rapid Near-Surface Water Vapor Increases during the Afternoon to Evening Transition in Oklahoma. *J. of Appl. Meteorol. and Climatol.*, **58(10)**, 2217–2234, https://doi.org/10.1175/JAMC-D-19-0062.1.~~

Bianco L., D. Cimini, F. S. Marzano, and R. Ware, 2005: Combining microwave radiometer and wind profiler radar measurements for high-resolution atmospheric humidity profiling, *J. Atmos. Ocean. Tech.*, **22**, 949–965, https://doi.org/10.1175/JTECH1771.1.

Bianco, L., K. Friedrich, J. M. Wilczak, D. Hazen, D. Wolfe, R. Delgado, S. Oncley, and J. K. Lundquist, 2017: Assessing the accuracy of microwave radiometers and radio acoustic sounding systems for wind energy applications. *Atmos. Meas. Tech.*, **10**, 1707-1721, https://doi.org/10.5194/amt-10-1707-2017.

Cadeddu, M. P., J. C. Liljegren, and D. D. Turner, 2013: The Atmospheric radiation measurement (ARM) program network of microwave radiometers: instrumentation, data, and retrievals, *Atmos. Meas. Tech.*, **6**, 2359–2372, https://doi.org/10.5194/amt-6-2359-2013.

Cimini, D., T. J. Hewison, L. Martin, J. Guldner, C. Gaffard, F. S. Marzano, 2006: Temperature and humidity profile retrievals from ground-based microwave radiometers during TUC, Meteorologische Zeitschrift, Vol. 15, No. 5, 45-56, DOI: 10.1127/09411-D-2948/2006/0099

Cimini, D., E. Campos, R. Ware, S. Albers, G. Giuliani, J. Oreamuno, P. Joe, S. E. Koch, S. Cober, and E. Westwater, 2011: Thermodynamic Atmospheric Profiling during the 2010 Winter

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Olympics Using Ground-based Microwave Radiometry, *IEEE Trans. Geosci. Remote Sens.*, **49**, 12, https://doi.org/10.1109/TGRS.2011.2154337.

Cimini, D., Rosenkranz, P. W., Tretyakov, M. Y., Koshelev, M. A., and Romano, F., 2018: Uncertainty of atmospheric microwave absorption model: impact on ground-based radiometer simulations and retrievals, *Atmos. Chem. Phys.*, **18**, 15231–15259, https://acp.copernicus.org/articles/18/15231/2018/.

Cimini, D., M. Haeffelin, S. Kotthaus, U. Löhnert, P. Martinet, E. O'Connor, C. Walden, M. Collaud Coen, and J. Preissler, 2020: Towards the profiling of the atmospheric boundary layer at European scale—introducing the COST Action PROBE. *Bulletin of Atmospheric Science and Technology*, **1**, 23–42, https://doi.org/10.1007/s42865-020-00003-8.

~~Cimini, D., Rosenkranz, P. W., Tretyakov, M. Y., Koshelev, M. A., and Romano, F., 2018: Uncertainty of atmospheric microwave absorption model: impact on ground-based radiometer simulations and retrievals, *Atmos. Chem. Phys.*, **18**, 15231–15259, https://acp.copernicus.org/articles/18/15231/2018/.~~

Clough, S.A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. Iacono, K. E. Cady-Pereira, S. Boukabara and P. D. Brown, 2005: Atmospheric radiative transfer modeling: A summary of the AER codes, *JQSRT*, vol 91, no. 2, pp 233-244, ~~2005,~~ https://doi.org/10.1016/j.jqsrt.2004.05.058.

Crewell, S., U. Löhnert, 2007: Accuracy of Boundary Layer Temperature Profiles Retrieved With Multifrequency Multiangle Microwave Radiometry, IEEE TGRS, VOL. 45, NO. 7, JULY 2007, DOI: 10.1109/TGRS.2006.888434.

Formatted: *Font color: Custom Color(RGB(17,85,204))*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Engelbart, D., W. Monna, J. Nash, 2009: Integrated Ground-Based Remote-Sensing Stations for Atmospheric Profiling, *COST Action 720 Final Report*, EUR 24172, https://doi.org/10.2831/10752.

Görsdorf, U., and V. Lehmann, 2000: Enhanced Accuracy of RASS-Measured Temperatures Due to an Improved Range Correction. J. Atmos. Oceanic Technol., **17 (4)**, 406–416, https://doi.org/10.1175/1520-0426(2000)017<0406:EAORMT>2.0.CO;2.

Han, Y., and E. R. Westwater, 1995: Remote sensing of tropospheric water vapor and cloud liquid water by integrated ground-based sensors. *J. Atmos. Oceanic Tech.*, **12**, 1050-1059~~.~~, DOI: https://doi.org/10.1175/1520-0426(1995)012<1050:RSOTWV>2.0.CO;2

~~Han, Y.,~~Hewison, T., 2007: 1D-VAR Retrieval of Temperature and ~~E. R. Westwater, 2000: Analysis and improvement of tipping calibration for ground-based microwave radiometers.~~Humidity Profiles From a Ground-Based Microwave Radiometer, *IEEE Trans. Geosci. Remote Sens.*, ~~**38,** 1260–1276,~~ **45(7)**, 2163–2168, https://doi.org/10.1109/TGRS.2007.898091. ~~https://doi.org/10.1109/36.843018.~~

Horst, T. W., S. R. Semmer, and I. Bogoev, 2016: Evaluation of Mechanically-Aspirated Temperature/Relative Humidity Radiation Shields, *18th Symposium on Meteorological Observation and Instrumentation, AMS Annual Meeting*, New Orleans, LA, 10-15 January 2016, https://ams.confex.com/ams/96Annual/webprogram/Paper286839.html.

Kaimal, J. C., and J. E. Gaynor, 1983: The Boulder Atmospheric Observatory. *J. Climate Appl. Meteor.*, **22**, 863–880, https://doi.org/10.1175/1520-0450(1983)022<0863:TBAO>2.0.CO;2.

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Küchler, N., D. D. Turner, U. Löhnert,and S. Crewell, 2016: Calibrating ground-based microwave radiometers: Uncertainty and drifts, *Radio Sci.*, **51**, 311–327,doi:10.1002/2015RS005826.

Löhnert U. and O. Maier, 2012: Operational profiling of temperature using ground-based microwave radiometry at Payerne: prospects and challenges. *Atmos. Meas. Tech.*, **5**, 1121–1134, https://doi.org/10.5194/amt-5-1121-2012.

Lundquist, J. K., J. M. Wilczak, R. Ashton, L. Bianco, W. A. Brewer, A. Choukulkar, A. Clifton, M. Debnath, R. Delgado, K. Friedrich, S. Gunter, A. Hamidi, G. V. Iungo, A. Kaushik, B. Kosović, P. Langan, A. Lass, E. Lavin, J. C.-Y. Lee, K. L. McCaffrey, R. K. Newsom, D. C. Noone, S. P. Oncley, P. T. Quelet, S. P. Sandberg, J. L. Schroeder, W. J. Shaw, L. Sparling, C. St. Martin, A. St. Pe, E. Strobach, K. Tay, B. J. Vanderwende, A. Weickmann, D. Wolfe, and R. Worsnop, 2017: Assessing state-of-the-art capabilities for probing the atmospheric boundary layer: the XPIA field campaign. *Bull. Am. Meteor. Soc.*, **98**, 289–314, https://doi.org/10.1175/BAMS-D-15-00151.1.

Maahn, M., D. D. Turner, U. Löhnert, D. J. Posselt, K. Ebell, G. G. Mace, and J. M. Comstock, 2020: Optimal estimation retrievals and their uncertainties: What every atmospheric scientist should know. *Bull. Amer. Meteor. Soc.*, **101**, E1512-E1523, https://doi~~:~~.org/10.1175/BAMS-D-19-0027.1

~~May P. T., Moran, K. P., and Strauch, R. G., 1989: The Accuracy of RASS Temperature Measurements, *J. Appl. Meteorol.*, **28**, 1329–1335.~~

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Maddy, E. S. and C. D. Barnet, 2008: Vertical Resolution Estimates in Version 5 of AIRS Operational Retrievals. *IEEE TGRS*, VOL. **46**, NO. 8, AUGUST 2008, https://doi.org/10.1109/TGRS.2008.917498

Martinet, P., D. Cimini, F. Burnet, B. Ménétrier, Y. Michel, and V. Unger, 2020: Improvement of numerical weather prediction model analysis during fog conditions through the assimilation of ground-based microwave radiometer observations: a 1D-Var study, *Atmos. Meas. Tech.*, **13**, 6593–6611, https://doi.org/10.5194/amt-13-6593-2020.

May, P. T. and J. M. Wilczak, 1993: Diurnal and Seasonal Variations of Boundary-Layer Structure Observed with a Radar Wind Profiler and RASS. Mon. Wea. Rev., **121**, 673–682, https://doi.org/10.1175/1520-0493(1993)121<0673:DASVOB>2.0.CO;2.

Masiello, G., C. Serio, and P. Antonelli, 2012: Inversion for atmospheric thermodynamical parameters of IASI data in the principal components space. *Quart. J. Roy. Meteor. Soc.*, **138**, 103–117, https://doi.org/10.1002/qj.909.

Merrelli, A. M., and D. D. Turner, 2012: Comparing information content of upwelling far infrared and midinfrared radiance spectra for clear atmosphere profiling. *J. Atmos. Oceanic Technol.*, **29**, 510–526, https://doi.org/10.1175/JTECH-D-11-00113.1.

Neiman, P. J., D. J. Gottas, and A. B. White, 2019: A Two-Cool-Season Wind Profiler–Based Analysis of Westward-Directed Gap Flow through the Columbia River Gorge. *Month. Wea. Rev.*, **147**, 4653-4680, https://doi.org/10.1175/MWR-D-19-0026.1.

North, E. M., A. M.Peterson, and H. D. Parry , 1973: RASS, a remote sensing system for measuring low-level temperature profiles. *Bull. Am. Meteor. Soc.*, **54**, 912–919.

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Payne, V. H., J. S. Delamere, K. E. Cady-Pereira, R. R. Gamache, J.-L. Moncet, E. J. Mlawer, and S. A. Clough, 2008: Air-broadened half-widths of the 22- and 183-GHz water-vapor lines. *IEEE Trans. Geosci. Remote Sens.*, **46**, 3601-3617, ~~doi:10.1109/TGRS.2008.2002435.~~https://doi.org/10.1109/TGRS.2008.2002435.

Payne, V. H., E. J. Mlawer, K. E. Cady-Pereira, and J.-L. Moncet, 2011: Water vapor continuum absorption in the microwave. *IEEE Trans. Geosci. Remote Sens.*, **49**, 2194-2208, ~~doi:10.1109/TGRS.2010.2091416.~~https://doi.org/10.1109/TGRS.2010.2091416.

Rodgers, C. D., 2000: Inverse Methods for Atmospheric Sounding: Theory and Practice. *Series on Atmospheric, Oceanic and Planetary Physics*, Vol. 2, World Scientific, 238 pp.

~~Rose, T., S. Crewell, U. Löhnert, C. Simmer, 2005: A network suitable microwave radiometer for operational monitoring of the cloudy atmosphere, *Atmospheric Research*, **75**, Issue 3, 183–200, https://doi.org/10.1016/j.atmosres.2004.12.005.~~

Rosenkranz, P. W., 1998: Water vapour microwave continuum absorption: A comparison of measurements and models. *Radio Science*, **33**, 919–928~~.~~, https://doi.org/10.1029/98RS01182.

Shaw, W., and Coauthors, 2019: The Second Wind Forecast Improvement Project (WFIP 2): General Overview. *Bull. Am. Meteor. Soc.*, **100(9)**, 1687–1699, https://doi.org/10.1175/BAMS-D-18-0036.1.

Formatted: *Space After: 12 pt, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Solheim, F., J. R. Godwin, J., and R. Ware, 1998a: Passive ground-based remote sensing of atmospheric temperature, water vapor, and cloud liquid profiles by a frequency synthesized microwave radiometer. *Meteorol. Z.*, **7**, 370–376.

Solheim F., J. R. Godwin, E. R. Westwater, Y. Han, S. J. Keihm, K. Marsh, R. Ware, 1998b: Radiometric profiling of temperature, water vapor and cloud liquid water using various inversion methods. *Radio Science*, **33**, 393–404, https://doi.org/10.1029/97RS03656.

Stankov, B. B., E. R. Westwater, and E. E. Gossard, 1996: Use of wind profiler estimates of significant moisture gradients to improve humidity profile retrieval. *J. Atmos. Oceanic Tech.*, **13**, 1285-1290~~.~~, DOI: https://doi.org/10.1175/1520-0426(1996)013<1285:UOWPEO>2.0.CO;2.

Strauch, R. G. , D. A. Merritt, K. P. Moran, K. B. Earnshaw, and D. V. De Kamp, 1983: The Colorado wind-profiling network. *J. Atmos. Oceanic Technol.*, **1**, 37–49, https://doi.org/10.1175/1520-0426(1984)001<0037:tcwpn>2.0.co;2.

Turner, D. D., and U. Löhnert, 2014: Information content and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based Atmospheric Emitted Radiance Interferometer (AERI). *J. Appl. Meteor. Clim.*, **53**, 752–771, https://doi.org/10.1175/JAMC-D-13-0126.1.

Turner, D. D., and W. G. Blumberg, 2019: Improvements to the AERIoe thermodynamic profile retrieval algorithm. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **12(5)**, 1339–1354, https://doi.org/10.1109/JSTARS.2018.2874968.

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*

Turner, D. D., and U. Löhnert, 2020: Ground-based Temperature and Humidity Profiling: Combining Active and Passive Remote Sensors. *In revision to Atmos. Meas. Tech.* ~~*Discuss.*, https://doi.org/10.5194/amt-2020-352~~*Discuss.*, https://doi.org/10.5194/amt-2020-352.

Ware R., Solheim F., Carpenter R., and Coauthors, 2003: A multi-channel radiometric profiler of temperature, humidity and cloud liquid. *Radio Science*, **38**, No. 4, 8079, https://doi.org/10.1029/2002RS002856.

Weber, B. L~~.,~~., D. B. Wuertz, D. C. Welsh, and R. Mcpeek, ~~1992~~1993: Quality controls for profiler measurements of winds and RASS temperatures. *J. Atmos. Oceanic Technol.*, **10**, 452–464, https://doi.org/10.1175/1520-0426(1993)010<0452:qcfpmo>2.0.co;2

Wilczak, J. M., and Coauthors, 2019: The Second Wind Forecast Improvement Project (WFIP2): Observational Field Campaign. *Bull. Am. Meteor. Soc.*, **100(9)**, 1701–1723, https://doi.org/10.1175/BAMS-D-18-0035.1.

Wolfe, D. E. and R. J. Lataitis, 2018: Boulder Atmospheric Observatory: 1977–2016: The end of an era and lessons learned. *Bull. Am. Meteor. Soc.*, **99**, 1345–1358, https://doi.org/10.1175/BAMS-D-17-0054.1.

Formatted: *Space After: 12 pt, Line spacing: Double, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)*