

An Automated Detection Methodology for Dry Well-Mixed Layers

Stephen D. Nicholls*

*Joint Center for Earth Systems Technology, University of Maryland-Baltimore County, and
Mesoscale Atmospheric Processes Laboratory, NASA-Goddard Space Flight Center,
Greenbelt, Maryland*

Karen I. Mohr

*Earth Sciences Division-Atmospheres, NASA-Goddard Space Flight Center,
Greenbelt, Maryland*

Submitted to the *Journal of Atmospheric and Oceanic Technology*, August 2018

*Corresponding author

Stephen D. Nicholls

Mesoscale Atmospheric Processes Laboratory

NASA-Goddard Space Flight Center

Greenbelt, MD 20771

Voice: (301) 614-6337

Fax: (301) 614-6297

stephen.d.nicholls@nasa.gov

Abstract

The intense surface heating over arid land surfaces produces dry well-mixed layers (WML) via dry convection. These layers are characterized by nearly constant potential temperature and low, nearly constant water vapor mixing ratio. To further the study of dry WMLs, we created a detection methodology and supporting software to automate the identification and characterization of dry WMLs from multiple data sources including rawinsondes, remote sensing platforms, and model products. The software is a modular code written in Python, an open source language. Radiosondes from a network of synoptic stations in North Africa were used to develop and test the WML detection process. The detection involves an iterative decision tree that ingests a vertical profile from an input data file, performs a quality check for sufficient data density, and then searches upward through the column for successive points where the simultaneous changes in water vapor mixing ratio and potential temperature are less than the specified maxima. If points in the vertical profile meet the dry WML identification criteria, statistics are generated detailing the characteristics of each layer in the profile. At the end of the vertical profile analysis, there is an option to plot analyzed profiles in a variety of file formats. Initial results show that the detection methodology can be successfully applied across a wide variety of input data and North African environments and for all seasons. It is sensitive enough to identify dry WMLs from other types of isentropic phenomena such as subsidence layers and distinguish the current day's dry WML from previous days.

1. Introduction

Intense daytime surface heating over arid land surfaces such as deserts produces dry convective mixing in the boundary layer. In the absence of external influences capable of generating significant mechanical turbulence such as fronts or thunderstorm outflow, local dry convection can produce a deep, well-mixed, nearly isentropic boundary layer that later becomes a dry well-mixed residual layer in the evening (Bounoua and Krishnamurti 1991; Gamo 1996; Warner 2004; Messenger et al. 2010; Garcia-Carreras et al. 2015). These dry well-mixed layers (WMLs) can retain their unique mid-tropospheric thermal and humidity structure for several days and be transported thousands of kilometers downstream (Carlson and Prospero 1972; Prospero and Carlson 1981; Zhang and Pennington 2004; Cuesta et al. 2008; Cuesta et al. 2009). In general, low- and mid-level dry WMLs are detrimental to moist convection due to water loading and dry entrainment into rising moist eddies (Lucas et al. 2000; Wang and Sobel 2012; Takemi 2014; Honda and Kawano 2015). However, for environments with strong, low-level moist convergence and shear, a mid-level dry WML can indirectly promote stronger updrafts and deep, often intense convection initially at the expense of weaker moist eddies. When the first weak buoyant plumes attempt to traverse an elevated, dry WML, dry air entrainment, water loading, and evaporative cooling will eliminate the buoyancy of those rising eddies and promote the formation of downdrafts and low-level cold pools. Given sufficient low-level moist convergence and shear, feedbacks can develop in which these cold pools foster the formation of larger (1-km scale), stronger ($> 10 \text{ m s}^{-1}$) moist eddies more impervious to dry air entrainment and water loading. The undiluted ascent within the core of wide, strong eddies makes it possible for them to grow to the level of free convection, access mid-level CAPE, and release built-up instability (Carr and Millard 1985; Peckham et al. 2004; James and Markowski 2010; Trier et al. 2015). The transport of mid-

level dry WMLs from upstream deserts overrunning low-level moist airstreams has long been observed to coincide with severe local storm development in North America (e.g., Carlson and Ludlam 1968; Schaefer 1974; Carlson et al. 1983; Benjamin and Carlson 1986; Lanicci et al. 1987). Mid-level dry WMLs are also associated with the potential for intense convection in Asia (e.g., Medina et al. 2010; Romatschke and Jr. 2011; Yamane et al. 2013; Liu and Liu 2018), coastal Australia (Allen and Karoly 2014; Grundstein et al. 2017), Central Europe (e.g., Bissolli et al. 2007; Graf et al. 2011; Matsangouras et al. 2014; Púčik et al. 2015), South America (Romatschke et al. 2010; Rasmussen and Houze 2011; Zuluaga and Houze 2015), and West Africa (e.g., Roca et al. 2005; Nicholls and Mohr 2010; Vizu and Cook 2017).

The dry WML is an important element of dryland boundary layer meteorology as well as to the convective environments of more humid regions. Our research goal was to develop a quantitative, efficient, and automated detection algorithm to identify and characterize dry WMLs within the vertical column. The first step was to define what a dry WML is so that it could be coded into a software algorithm. Starting from the working hypothesis that a dry WML is generated by dry convection over arid land surfaces and has a distinctive thermodynamic structure, we adopt the following definition of a dry WML and analysis principles:

Definition of dry WML: A layer of nearly constant potential temperature and low, nearly constant water vapor mixing ratio, capped by an inversion, with a positive vertical gradient of relative humidity.

- The WMLs result primarily from dry convective mixing due to intense surface heating.
- The magnitude of the allowable water vapor mixing ratio is determined *a priori* from analysis of surface observations in the area of interest.
- The analysis covers the lower to mid-troposphere.

The use of the acronym “WML” in this paper will imply a dry WML, one with a water vapor mixing ratio below the *a priori* empirically determined allowable level. Any regionally or temporally extensive study of WMLs would require the inspection and analysis of multiple simultaneous temperature and humidity profiles from various available data sources. While radiosonde and rawinsonde observations (hereafter, RAOBs) have the best vertical resolution and accuracy, temperature and humidity profiles can also be extracted from model reanalysis datasets (e.g., Dee et al. 2011; Gelaro et al. 2017) or retrieved from spaceborne sounders such as the Atmospheric Infrared Sounder (AIRS), the Atmospheric Microwave Sounding Unit (AMSU), or the Geostationary Operational Environment Satellite (GOES) Sounder (Ma et al. 1999; Aumann et al. 2003; Weng et al. 2003; Menzel et al. 2018). As compared to RAOBs these alternative inputs have the advantage of improved spatial coverage and, in some cases, improved temporal resolution. Whether the analysis source is ground observations, remote sensing products, or model output, manual inspection and analysis of hundreds, perhaps thousands, of profiles would be prohibitively time consuming and difficult to quality control. To make an extensive and statistically consistent study of WMLs possible, we adapted our detection methodology to accept inputs from a variety of data sources including remote sensing and model products.

In this paper, we describe the WML detection methodology and show examples of analyzed WMLs in RAOBs and other data inputs from North Africa. In addition to containing the Sahara Desert, North Africa has several important features making it a useful study area, including RAOB sites with long observational records, multiple land surface types, and multiple dynamical influences such as synoptic waves, large-scale subsidence, moist monsoon inflow, and marine inversion layers that challenged us in creating a flexible and accurate detection methodology. Because the software is written in Python, it is open source, readably transferable, and usable by

potential users. Here we demonstrate its utility and its potential application in North African meteorological process studies and possibly in other arid regions.

2. Background

a. Study area description

Our study area in North Africa encompasses 18°W–45°E, 0°–37°N, containing all of the Sahara Desert. Figure 1 is a map of the 53 synoptic stations used in our analysis and the Bailey’s ecoregion provinces of North Africa (Bailey 2004; Bailey 2014). These provinces blend the analysis of climate and actual and potential dominant vegetation types to delineate relatively stable and continuous ecosystem boundaries. The yellow shades in Fig. 1 define the ecoregion provinces in the Sahara Desert. The mountain provinces (west to east) cover the Hoggar, Tibesti, and Red Sea Hills ranges. The ten synoptic stations marked with a black square denote inner Saharan stations, which are unaffected by coastal processes such as marine layers. During July and August, the profiles at the two “mixed” stations (black circle and square outline), Tessalit and Agadez, have a moist monsoon layer, reflecting their location at the far northern edge of the West African monsoon (Lélé and Lamb 2010). The 41 remaining stations (black circles) on the coasts or in Mediterranean (red shades), steppe (orange), savanna (light greens and blues), and rainforest (dark greens and blues) divisions are defined as non-Saharan because they are affected by coastal processes or are outside the Sahara Desert.

All references to seasons in North Africa henceforth imply boreal seasons. References to the West African monsoon imply a period from late April to mid- to-late November when the stations south of Tamanrasset (22°N) receive most of their annual precipitation due to a low-level, moist southwesterly inflow from the Gulf of Guinea (review in Fitzpatrick et al. 2015). Along the

Mediterranean coast, the wet season is in the winter, primarily from intruding transient extratropical waves and Atlantic depressions (Knippertz 2003; Knippertz et al. 2003; Born et al. 2010). The ecoregion gradients in Fig. 1 from desert to steppe to savanna/woodland/meadow to rainforest provinces usefully delineate a decreasing expected occurrence of WMLs generated in the local boundary layer by dry convection due to intense surface heating (Warner 2004; Nicholson 2011). Even at Saharan stations, it can be difficult to separate the signal from dry convection from dry intrusions from transient extratropical or tropical waves or from regional subsidence (Roca et al. 2005; Vizzy and Cook 2009; Raveh-Rubin 2017). Despite the challenge posed by multiple possible larger-scale dynamical influences on local profiles around North Africa (e.g., Nicholson 1996; Camberlin 2009; Fink 2017), we developed our detection methodology using temperature and humidity profiles from every month of the year and using all stations in Fig. 1.

b. Data description

The foundation of this work was RAOBs from 53 North African stations (Fig. 1) to build the detection and plotting modules and to validate analyses performed with model and remote sensing inputs. We obtained RAOB data from the Integrated Global Radiosonde Archive (IGRA), version 2, at NOAA's National Centers for Environmental Information (NCEI). The IGRA has radiosonde, rawinsonde, and pilot balloon data in text format from stations around the world, with several hundred stations having records that extend back to the early 20th Century. While the data records in IGRA are rigorously quality controlled for implausible values and internal inconsistencies, individual soundings are of varying data densities, depending on the station's ability to report significant and mandatory levels and those in between (Durre et al. 2006). The earliest North African RAOBs used in this study are from 1948. Many of the stations have temporally non-continuous records, although for the purposes of this work, temporal continuity

was not important. All North African RAOBs that had temperature and humidity values at the mandatory levels from the station surface up to and including 500 hPa were defined as “useful”. Figure 2 contains maps of the useful observations available at each North African station since a) 1948 and b) 2003. Some of the stations, particularly in Libya, Algeria, and Nigeria, stopped reporting before 2003, but had enough quality older observations to be included in the counts. Applying our “useful” observation criteria to all RAOB launches since 1948, only 35% (450,000 observations) were useful for methodological development and testing.

After we derived a robust detection methodology for analysis of RAOBs, we applied it to analyzing temperature and humidity products from the cross-track scanning IR and microwave sounders, AIRS and AMSU, on NASA’s polar orbiting Aqua satellite (Parkinson 2003). The mid-day Aqua overpasses over North Africa coincide with the development of deep daytime boundary layers over arid surfaces (Marsham et al. 2013; Garcia-Carreras et al. 2015). The AIRS is an infrared spectrometer/radiometer that covers the 3.7–15.4- μm spectral range with 2378 spectral channels, and AMSU is a 15-channel microwave radiometer operating between 23 and 89 GHz (Aumann et al. 2003). Our work used the version 6 sensor-resolution, ungridded Level 2 (L2) temperature and humidity support products at 50-km $\times$ 50-km horizontal resolution, and 28 standard pressure levels plus intermediate pressure levels for a total of 100 pressure levels across a 1650-km swath. Although there are twice daily observations over Africa, this work involved only the daytime passes. The version 6 collection of L2 products include AIRS+AMSU combined retrievals (AIRX2SUP) and an AIRS-only retrieval (AIRS2SUP) that, compared to previous versions, significantly improve cloud clearing and the retrieval of surface and cloud parameters (Susskind et al. 2012; Susskind et al. 2014; Stubenrauch et al. 2017). Vergados et al. (2018) evaluated global tropospheric specific humidity maps among multiple gridded humidity datasets,

including one derived from AIRS+AMSU, finding that the climatologies varied by less than 15% depending on region and pressure level, with AIRS+AMSU ranked as the driest of the datasets.

We also used profiles from two reanalysis datasets in this study, 1) the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis-Interim (ERA-I) at 6-h and $0.75^\circ \times 0.75^\circ \times 60$ model levels (Dee et al. 2011); 2) NASA Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2) at 1-h and $0.625^\circ \times 0.50^\circ \times 72$ model levels (Gelaro et al. 2017). Both of these modeling centers run data assimilation systems that incorporate AIRS and other atmospheric sounder radiances as well as observations from the operational RAOB network. Because global operational modeling systems are run at resolutions too coarse (> 10 -km) to resolve convective-scale phenomena explicitly in North Africa, multiple authors have demonstrated that the assimilation of RAOBs strongly influences forecasted atmospheric state variables to an approximate diameter of 300–500-km around the synoptic station, depending on variable, model resolution, and model physics (Agustí-Panareda et al. 2010; Fink et al. 2011; Roberts et al. 2015). Because validation of remote sensing and reanalysis inputs is grounded at past and present operational synoptic stations, our examples of reanalysis performance should be considered the *best possible* for those datasets, degrading with distance from synoptic stations.

Because major modeling centers assimilate radiances from the AIRS and AMSU sensors (e.g., Gelaro et al. 2017), it is important to point out the independence of Version 6 AIRS and AIRS+AMSU temperature and humidity products from model products, such as MERRA-2. The AIRS and AIRS+AMSU products are derived from a physical retrieval system using a neural network first guess. The neural network consists of interconnected non-linear function approximators trained to infer a statistical relationship between inputs (AIRS and AMSU radiances) and outputs (temperature and humidity profiles) without requiring direct knowledge of

the functional input/output relationship (Blackwell 2005; Blackwell 2012). For a given version of AIRS retrieved products, the neural network was trained once from co-located model analysis fields over many different months and years. The historically trained neural network coefficients are then used in all subsequent operational retrievals so that each day's temperature and humidity profiles are independent of any analysis fields from any modeling center (Blackwell 2012; Milstein and Blackwell 2016). Given a first guess, the AIRS retrieval performs a physically-based cloud clearing step estimate of what would be the cloud-free infrared radiance spectrum using 9 single field-of-view AIRS observations in a 3×3 IR field-of-regard coinciding with an AMSU footprint and, when available, the microwave measurements in same position (Cho and Staelin 2006; Susskind et al. 2014). The AIRS retrieval methodology finds iterative solutions, starting from the first guess, to the radiative transfer equations on a field-of-regard basis that best match the derived cloud cleared radiances determined for that field-of-regard. The combination of the AIRS neural network-based first guess and physically-based cloud clearing methodology makes robust, quality-controlled AIRS and AIRS+AMSU retrievals in almost all cloud conditions, including retrievals above overcast conditions and over storms (Blackwell and Milstein 2014; Milstein and Blackwell 2016; Reale et al. 2018).

3. WML Detection Methodology

a. Governing reference values

In the development phase, we made first guesses of the parameters defining the analysis column and what defines “dry”, the maximum allowable water vapor mixing ratio (w), and “nearly constant”, the simultaneous maximum allowable change in potential temperature (θ) and maximum allowable change in w from reported level to level. These reference (*ref*) values

emerged from multiple rounds of test analyses and inspections of analyzed skew-T diagrams versus spreadsheets of the input data. The following abbreviations and final values of our reference values reflect our preference for greater accuracy over sensitivity in the WML detection in North Africa:

- Top of the analysis column is 500 hPa (p_{ref}).
- Bottom of the analysis column (z_{bref}) depends on location.
- Maximum value of w (w_{ref}) is 0.007 kg kg^{-1} .
- Maximum change in θ with height $\left(\frac{\partial\theta}{\partial z}\right)_{ref}$ is 1.0 K km^{-1} .
- The reference maximum change in w with height $\left(\frac{\partial w}{\partial z}\right)_{ref}$ is $0.0008 \text{ kg kg}^{-1} \text{ km}^{-1}$.
- The actual maximum allowable change in w between two levels (l_2 and l_1) is scaled by the layer mean w ($\bar{w}_{l_2-l_1}$) and w_{ref} , thus $\left(\frac{\bar{w}_{l_2-l_1}}{w_{ref}}\right) \cdot \left(\frac{\partial w}{\partial z}\right)_{ref}$.

The choice to use a reference height above ground level (z_{bref}) rather than a pressure level acknowledges that analysis points in any land-based dataset may occur at elevations ranging from near sea-level (e.g., Dakar, Douala) to well above 1000 m (e.g., Tamanrasset, Asmara). In desert ecoregion provinces, the surface layer is often superadiabatic due to the intense surface sensible heat flux (Bounoua and Krishnamurti 1991; Warner 2004). Such a lapse rate is relatively easy to identify and to program an exclusion. For our barren, hyperarid Saharan-classified stations, the value of z_{bref} is 0. However, at our “non-Saharan” and for selected months at our “mixed” stations, there may be significant mechanical mixing due to surface roughness and/or low-level monsoon or sea breeze inflows (Parker et al. 2005; Gounou et al. 2012). At these stations, the value of z_{bref} is 500 m. It was easier to exclude the lowest part of the boundary layer (accuracy over sensitivity)

than to attempt to separate buoyant from mechanical mixing at analysis points in different environments.

Appropriate literature and the realities of data density in the IGRA RAOBs guided the choice of the top of the analysis column (p_{ref}). For this reference value, it was easier and more uniform across different types of data inputs to use a pressure level. Above 500 hPa, even for RAOBs containing both mandatory and significant levels, the vertical resolution of observations and the availability of valid dewpoint depression values greatly decrease. In studies using research flight data over the western Sahara, Saharan mixed and residual layers do not extend above 6-km/450-hPa (Messenger et al. 2010; Marsham et al. 2013). If only the mandatory radiosounding levels are reported below 500 hPa, a RAOB's data density is considered low. However, the vertical resolution of model and remote sensing products may not be much better than RAOBs with just mandatory levels. To maximize our ability to use the IGRA RAOBs and multiple other inputs, we decided to include RAOBs containing just mandatory levels in our algorithm development and to analyze only to a pressure level of 500 hPa.

We derived the maximum mixing ratio (w_{ref}), our “dry” value, from a cumulative frequency distribution analysis of w from a calendar year's set of RAOBs at In Salah, Casablanca, and Agadez. Choosing 2007 at random to analyze, we found that 0.007 kg kg^{-1} is the 90th percentile value for w at In Salah, occurring exclusively in the lowest 50 hPa of the atmospheric column. The 90th percentile value at In Salah is also the 90th percentile at Casablanca above the summertime subsidence inversion capping the marine layer and above the monsoon layer at Agadez, indicated by low-level southwesterlies. During testing using other years and stations, we did not find it necessary to adjust the value of w_{ref} from our first guess.

The first guess values for the changes in w and θ with height had to be within the sensitivity of radiosounding sensors and ground tracking stations (Office of the Federal Coordinator of Meteorology 1997) and potentially achievable by other inputs such as retrieved temperature and humidity from space-based sounders (Menzel et al. 2018). Our first guess for isentropic conditions, $\left(\frac{\partial\theta}{\partial z}\right)_{ref}$, was 1.0 K km^{-1} , derived from Bounoua and Krishnamurti (1991) who determined that $3 \text{ K } 3000 \text{ m}^{-1}$ was a representative lapse rate over an entire Saharan mixed layer. This first guess proved to be a widely useful reference value. For nearly constant w , $\left(\frac{\partial w}{\partial z}\right)_{ref}$, we started with a value estimated from hand analysis of several mid-summer RAOBs from In Salah, $0.002 \text{ kg kg}^{-1} \text{ km}^{-1}$.

Starting from these initial lapse rate values, we evaluated the sensitivity of analyzed WML detections to the allowable mixing ratio and potential temperature lapse rates for a selection of North African RAOB launch locations (Fig. 1). Figures 3 and 4 show analyzed WMLs (each WML is a different color) from two representative RAOB profiles from Tamanrasset, Algeria and Tunis, Tunisia, respectively. These figures show a subset of six tested configurations where $\left(\frac{\partial\theta}{\partial z}\right)_{ref}$ was constant and $\left(\frac{\partial w}{\partial z}\right)_{ref}$ was either constant (CON) or was scaled by $\left(\frac{\bar{w}_{l_2-l_1}}{w_{ref}}\right)$ (VAR). Figure 3 shows the algorithm to have greater sensitivity to changes in the maximum allowable mixing ratio lapse rate than the maximum allowable potential temperature lapse rate. Qualitative analysis of this and other RAOB profiles (See Fig. 4 and supplementary figures) show our initial guess for $\left(\frac{\partial\theta}{\partial z}\right)_{ref}$ works well, yet smaller values of $\left(\frac{\partial w}{\partial z}\right)_{ref}$ were needed to improve consistency with hand analysis. We found the best results for North Africa applied a maximum allowable potential temperature lapse value of 1.0 K km^{-1} and a maximum allowable mixing ratio lapse rate

value of $0.00035 \text{ kg kg}^{-1} \text{ km}^{-1}$ (constant) / $0.0008 \text{ kg kg}^{-1} \text{ km}^{-1}$ (scaled). Smaller and larger lapse rate values produced less accurate WML detections because the detection criteria were either too constraining or generous, respectively. Figure 4 reveals that applying a constant allowable mixing ratio lapse for WML analysis in North Africa has an inherent flaw. Specifically, if the mixing ratio profile approaches or falls below the allowed mixing ratio change over 1 km then any isentropic layer is identified as WML regardless of the dewpoint temperature profile (See Fig. 4, ~550 hPa). As shown in Figure 4, this flaw can be addressed by scaling the maximum allowable mixing ratio lapse rate by $\left(\frac{\bar{w}_{l_2-l_1}}{w_{ref}}\right)$. This adjustment is included in our study because it ensures that all analyzed WMLs contain dewpoint temperature profiles that are always roughly parallel to lines of constant mixing ratio. Properly tuning the allowed mixing ratio lapse rate is crucial for WML detection because isentropic layers in North Africa are generated by dry convection (our interest), large-scale subsidence, and shear-based turbulence, yet only layers associated with dry convection will also be well mixed (Moeng and Sullivan 1994; Conzemius and Fedorovich 2006).

In the development phase, we encountered inputs with marginally acceptable data densities below 500 hPa, such as soundings with only mandatory levels and at high altitude synoptic stations, such as Asmara (2325 m). We also had plenty of examples where the analyzed and plotted WMLs looked smaller than those that might be hand analyzed on a skew-T chart. For inputs of poor data densities or if the user requires a more generous interpretation of “well-mixed”, less restrictive values can be set, for example, $\left(\frac{\partial \theta}{\partial z}\right)_{ref}$ at 2.0 K km^{-1} , $\left(\frac{\partial w}{\partial z}\right)_{ref}$ at $0.00095 \text{ kg kg}^{-1} \text{ km}^{-1}$ (scaled allowable mixing ratio lapse rate), or $\left(\frac{\partial w}{\partial z}\right)_{ref}$ at $0.00035 \text{ kg kg}^{-1} \text{ km}^{-1}$ (constant allowable mixing ratio lapse rate). The adjustability of the software makes it possible to adapt the detection

methodology for user preferences, other types of WMLs (e.g., moist well mixed layers), and for other regions.

b. The WML analysis process

Figure 5 is the decision tree of the software. During the initial development of the decision tree, we used about 10% of the IGRA database, selecting a sample of soundings scattered across seasons, years, and stations (Fig. 1). The software is written in multiple Python modules that govern the basic information flow, set up (blue) to input data (yellow) to data evaluation (gray) to output (green). The execution (red) is controlled at both the top and the bottom of the tree, depending on both data input availability and quality. The set-up module (blue box) governs the setting of analysis period and location. “Additional information” can be interpreted as broadly as the user needs, to include anything needed for analysis that is not already in the input files. For example, if an analysis location is not contained within the NCEI IGRA station inventory, then the location name and position will have to be added manually. The set-up module is also the entry point where the user can group analysis points as we did into categories for different treatment of the base of the analysis column (z_{bref}). While we created individual sub-modules within the input module to read each type of data used in this study, the yellow boxes in Fig. 5 represent the generalized reading and quality checking of any type of input data. If the user’s work plan involves input from multiple data sources, the user can either preserve the native vertical resolution of each useful profile or linearly interpolate all useful profiles to the pressure levels associated with the least vertically dense input data. Since some dataset providers may already have θ and w available for download, calculating these variables may or may not be necessary.

The detection (gray boxes) begins with the search for the first level in the analysis column, i.e., between z_{bref} and p_{ref} , where w is at or less than w_{ref} . This level becomes $level_1$. Starting with a new analysis (column a), the next higher level in the profile is designated $level_2$, and the changes in θ and w are assessed ($l_2 - l_1$), providing the layer mean w is at or below w_{ref} . If all of these conditions are met (**Y**), a detection flag is set indicating that a WML exists (true, “T”), and a set of statistics on the WML starts to accumulate. Both the positive outcome and the negative outcome (**N**) below the analysis box return to the upward iteration of vertical levels in the analysis column. If the new level’s w is at or below w_{ref} and the WML flag is true, a multi-level WML exists, and column b is invoked. For a multi-level WML, not only must the current layer ($l_2 - l_1$) meet the θ and w change and mean w conditions, but the entire presumptive WML from the previously identified base to $level_2$, ($l_2 - base$), must meet those same conditions. Notably, the algorithm permits the user to apply either a scaled or constant maximum allowable mixing ratio lapse rate (i.e., $\left(\frac{\bar{w}_{l_2-l_1}}{w_{ref}}\right) = 1$). The cycle of iteration up the analysis column and the testing of change conditions continue until the change conditions are no longer satisfied, or the top of the analysis column is reached.

If the change conditions fail, a counter iterates to record that a WML has been detected in the analysis column, and the statistics for that WML stop accumulating. We now know its location and thickness in the analysis column and its mean θ and w . Resetting the flag to false (“F”) starts the search for another suitable layer for analysis. If the top of the analysis column (p_{ref}) is reached, the count of WMLs and a summary of the WML characteristics for that profile set are finalized in the output module (green box). The output module also has an option to plot an analyzed skew-T with different colors indicating each of the WMLs detected in the profile. The user can generate skew-T plots at any resolution or file format supported by Python including postscript (.ps),

encapsulated postscript (.eps), and portable network graphics (.png) images. Reading and quality checking input data continue until end-of-file conditions (red boxes) occur.

4. Detection Examples

a. RAOB examples

Figure 6 contains archetypal examples of analyzed IGRA RAOBs with WMLs. The WMLs (orange shaded areas) in Fig. 6 can originate from local dry convective mixing during the current day, residual boundary layers from previous days, or transport from an upstream desert where the dry convection occurred (e.g., Bounoua and Krishnamurti 1991; Messenger et al. 2010; Garcia-Carreras et al. 2015). In addition, soundings may contain radiation and subsidence inversions with or without WMLs. In Fig. 6a, In Salah has a superadiabatic surface layer below 972 hPa due to intense surface heating and a single, deep WML characteristic of inner Saharan locations during the summer months. In Fig. 6b, the coastal city of Tunis has a marine boundary layer capped by a deep subsidence inversion and a WML from 648–533 hPa. Figures 6c and 6d are from the mixed category station, Agadez, on the southern edge of the Sahara, demonstrating profiles before and after the onset of the West African monsoon. Figure 6c contains a typical late spring-early summer profile before monsoon onset. There is a surface layer below 949 hPa and a WML from 925–700 hPa, similar to In Salah (Fig. 6a), but shallower. Figure 6d is an example of Agadez during the West African monsoon season, containing a monsoon layer below 850 hPa, a WML from 817–700 hPa, and a subsidence inversion starting at 541 hPa.

Figure 7 has examples of more challenging environments to analyze. In early May, the surface heating at In Salah (Fig. 7a) is not as intense as in mid-summer (Fig. 6a). In Fig. 7a, there are two analyzed WMLs from 925 hPa to 500 hPa. There is a larger than allowable change in w

from 976 hPa to 925 hPa, so that the first WML (orange) starts at 925 hPa. It terminates at 850 hPa because of the larger than allowable changes in θ between 850–792 hPa. The second WML (gray) extends from 792 hPa to the top of the analysis column at 500 hPa. The probability that the mixing in today's boundary layer will erode through the previous night's residual layer, as in Figs. 6a and 6c, depends upon the intensity of the surface heating (Messenger et al. 2010; Garcia-Carreras et al. 2015). Growth of the current day's mixed layer can be limited by reductions in surface heating such as suppressed insolation due to dust (Marshall et al. 2008; Marshall et al. 2016), or to air mass changes, or to cloudiness from topographically induced circulations, transient extratropical waves, or organized convection associated with tropical waves (Knippertz et al. 2009; Vizi and Cook 2009; Knippertz and Todd 2010; Birch et al. 2012; Bou Karam et al. 2014). If the boundary layer growth is insufficient to mix out the remnant residual layers, then multiple WML layers can occur as in Fig. 7a. The slight cooling that takes place as the afternoon mixed layer transitions to the nighttime residual layer marks it as a different WML than the next day's mixed layer (Cuesta et al. 2008; Cuesta et al. 2009; Messenger et al. 2010). The ability to analyze, quantitatively and uniformly, multiple soundings from the previous days allows cross checking and confirmation of layered mid-tropospheric structure due to local diurnal boundary layer processes.

Tamanrasset (Fig. 7b) is an oasis city in the Hoggar Mountains with less intense surface heating and dry convection than In Salah (Cuesta et al. 2008; Wahab et al. 2010), yet it presents a complex profile with four analyzed WMLs. The two analyzed WMLs (orange and grey shading) between the surface layer and 700 hPa (834 to 825 hPa and 734 hPa to 700 hPa) demonstrate the capability of the decision tree to detect subtle variability in the profile mixing ratio lapse rate with height, which is not be easily attained from hand analysis. This is best illustrated in Figure 3 (same

case as Figure 7b), where altering the maximum allowable mixing ratio lapse rate by just 0.0002 kg kg⁻¹ km⁻¹, yields a vastly different WML analysis. Above 700 hPa, the blue shading indicates the third WML (700–541 hPa) and a fourth old residual layer (manila) between 536–500 hPa. The lower panels of Fig. 7 are from locations outside of the Sahara, Abidjan (Fig. 7c) and Ngaoundéré (Fig. 7d). Abidjan is in the Bailey’s ecoregion province “humid forests with short dry season” (Fig. 1), and the RAOB is from that short winter dry season. There is a WML from 689–627 hPa but not one above 600 hPa, as the change in θ is too large. At Ngaoundéré’s pre-monsoon onset RAOB, there one shallow WML analyzed between 850 and 735 hPa and a second, smaller layer extends from 711–700 hPa. Another nearly isentropic layer exists from 696 to 650 hPa, but it is not identified as a WML because the change in w exceeds the allowed limit. In both locations, shallow WMLs are identified despite both being seasonally humid locations. In future papers, we provide results from trajectory analysis that define whether these WMLs are from local boundary layer processes or transported from the Sahara.

b. Examples from other datasets

The development of the WML decision tree (Fig. 5) required a robust sample of RAOB observations throughout North Africa that did not need to have either temporal or spatial continuity. However, atmospheric process studies may require temporal and/or spatial continuity in temperature and humidity fields. The temporally and spatially sparse record of upper air observations over much of the world motivated the adaptation of the detection methodology and software to ingest remote sensing and model products. In this section, we show examples of WML analysis using AIRS+AMSU and model reanalysis (ERA-I and MERRA-2) at their native vertical resolution to demonstrate the utility and sensitivity of the detection analysis with multiple different inputs. Analysis using AIRS only and the Japanese 55-year Reanalysis (JRA-55) product inputs

are also possible but are not shown for sake of brevity. As in Nicholls and Mohr (2010), we assume RAOB and satellite observations to be valid if they are within 3 hours of the analysis time, whereas model reanalysis data is linearly interpolated to the analysis time. The principal intent of this section is to demonstrate and discuss how our detection algorithm analyzes WMLs across multiple data input sources under similar conditions. A more rigorous quantitative validation of the WML detection and its use in process studies of Saharan air intrusions will be treated in future papers.

Figures 8–11 are from different environments in North Africa and have four panels, with RAOB data in panel a, AIRS+AMSU products in panel b, ERA-I in panel c, and MERRA-2 in panel d. Figure 8 is a summer sounding from In Salah with a shallow WML from the top of the surface layer to 925 hPa and another, deeper WML extending from 850 hPa to 600 hPa (Fig. 8a). The soundings in panels b–d are largely similar to the RAOB in panel a, with WML detections occurring near the surface and into the mid-troposphere. The AIRS+AMSU profile (Fig. 8b) fails to capture the shallow WML near the surface, but it able to largely capture the height and thickness of the WML analyzed above 850 hPa in the ROAB profile. However, the mixing ratio lapse rate in the AIRS+AMSU profile does slightly exceed the allowed rate between 753 and 729 hPa, resulting in two analyzed WMLs rather than a single continuous layer. Corresponding ERA-I and MERRA-2 profiles (Figs. 8c and d) both show one deep, continuous WML extending from 950 to 650 hPa.

Unlike the RAOB, both reanalysis products in Fig. 8 are unable to represent the nearly isentropic, but not well mixed layer that exists between 925 and 850 hPa. RAOB winds from surface to 850 hPa had strength (up to 10 m s^{-1}) and shear (veering southeast to south), whereas in both reanalysis profiles the velocities were generally no more than 5 m s^{-1} and unidirectional. The presence of wind shear between 925 and 850 hPa suggests the potential of WML erosion by

mechanical turbulence from low-level winds increasing in velocity and/or shear throughout the day (Flamant et al. 2007; Cuesta et al. 2009; Birch et al. 2012). Global models may not be able to account for turbulent local wind patterns resulting from differential heating or surface/terrain effects due to their application of relatively coarse horizontal model grids (Garcia-Carreras et al. 2013; Roberts et al. 2015). The operational AIRS+AMSU retrieval is independent of any model wind patterns and was sensitive enough to account for the decrease in w with height to 850 hPa (Susskind et al. 2014; Milstein and Blackwell 2016). The model data assimilation schemes may not have been sensitive enough to retrieve realistic profiles of w below 850 hPa from AIRS+AMSU radiances nor allow this information to override information from other model fields (Faccani et al. 2009; Agustí-Panareda et al. 2010; Fink et al. 2011). Unlike both reanalysis models and the AIRS+AMSU retrieval, the RAOB shows an abrupt transition from WML to a subsidence layer at 600 hPa. Comparatively, the reanalysis models and the AIRS+AMSU retrieval have smoother, deeper transitions, as their vertical resolutions are becoming much larger at this point in the troposphere.

Dakar, on the west coast of West Africa, has a marine layer in the lowest portion of its RAOB in Fig. 9 and a single analyzed WML between 850 and 800 hPa. Model soundings (Figs. 9c and 9d) successfully replicate the presence of marine layer above Dakar due to their assimilation of RAOB data. Differences in horizontal grid spacing yield interpolated profiles that produce analyzed WMLs at notably different pressure levels (ERA-I: 650-700 hPa, MERRA-2: 850-825 hPa). The AIRS+AMSU profile (Fig 9b) is clearly hindered by its low sensitivity to abrupt atmospheric property changes in the vertical, yet is improved by its higher horizontal data resolution (13.5 km at nadir) as compared to model reanalysis products (60 km+). The resulting

profile is lacks a defined marine layer, but the AIRS+AMSU analyzed WML is largely consistent with RAOB data.

Similarly, the models have a more detailed, although not faultless, representation of the column above Ouagadougou in Fig. 10. On this June date in the RAOB (Fig. 10a), the West African monsoon is apparent below 925 hPa, and there is a WML from 700 hPa to 600 hPa. Both reanalysis products also have narrower WMLs analyzed in the same general region above a deep, dry subsidence layer, but the AIRS+AMSU profile lacks any analyzed WMLs between 600 to 700 hPa because missing ratio lapse rate slightly exceeds the maximum allowed value. In the example at Tunis (Fig. 11), the vertical structure of its column is less complex than the example at Ouagadougou. The ERA-I (Fig. 11c) best approximates the ROAB profile, yet no data product (retrieval or model) show more than a hint the marine layer present in the ROAB profile. The RAOB has a WML analyzed from 724 hPa to 593 hPa between subsidence layers. Both model products have WMLs analyzed around 700 hPa, with MERRA-2 exhibiting a smaller WML between 850 and 825 hPa, which is not present in the ROAB profile. The AIRS+AMSU and MERRA-2 profiles are largely similar to each other, but the steep mixing ratio lapse rates associated the AIRS+AMSU profile prevent any WML detections whatsoever. It cannot be determined just from the profiles in Fig. 11 why the MERRA-2 is less successful than ERA-I in capturing the general structure reflected in the RAOB, particularly the depth of the subsidence between the marine layer and 700 hPa. Information from other model fields (not explored here) may have contravened the information from the assimilation of both RAOB data and AIRS+AMSU radiances.

For the sake of brevity, additional examples will not be shown and additional analysis left to future papers. The multiple examples from around North Africa in this paper reveal how

complex vertical structures due to multiple processes from local-scale surface heating gradients to continental-scale subsidence can occur in any season. The WML detection methodology can identify these layers anywhere within the analysis column, even between subsidence layers (e.g., Fig. 11a), or if there are multiple WMLs (e.g., Fig. 7) due to memory from local boundary layer processes from previous days.

5. Conclusion

Intense surface heating over dry land surfaces can produce strong dry convection. The resulting WMLs are characterized by nearly constant potential temperature and low, nearly constant water vapor mixing ratio. The WML is a frequent feature of dryland boundary layer processes. Transported downstream over low-level moist airstreams, WMLs can affect the potential for intense deep convective clouds. To further the study of WMLs, we created a detection methodology and supporting software to automate the identification and characterization of WMLs from multiple data sources including RAOBs and remote sensing and model products. The software is modular code written in Python, an open source language.

The network of synoptic stations in North Africa was a challenging and varied testing ground for methodology development, containing stations in hyperarid desert to humid tropical forest ecoregions. We used RAOBs from all seasons to develop and then test our parameter values and layer identification process. Parameter values that required setting included the analysis column (top of the surface layer to 500 hPa), our maximum water vapor mixing ratio, and maximum allowable changes with height in water vapor mixing ratio and potential temperature. These parameters govern an iterative decision tree that ingests a sounding from an input data file, performs a quality check for sufficient data density, and then searches up through the column for

successive points where the changes in water vapor mixing ratio and potential temperature are less than specified maxima. If points in the sounding meet the identification criteria, statistics are generated on layer characteristics (e.g., layer base, top, mean potential temperature), and a count of layers in the profile is started. At the top of the analysis column, any layer accounting ceases and a summary of the WMLs in the sounding is generated. There is an option to plot analyzed soundings in a variety of file formats.

Because of temporally and spatially sparse upper air data records in many parts of the world, we adapted the software package to ingest remote sensing and gridded model products such as AIRS+AMSU temperature and humidity profiles and MERRA-2 reanalysis. The software has separate sub-modules for different types of input, and analysis parameter values can be adjusted as appropriate to the data source or the user's preferences for detection sensitivity. The examples of different input datasets shown in this paper used the same parameter values. While the remote sensing and model product profiles did not capture all of the complexity in the vertical columns above the synoptic stations used as examples, the detection methodology was sensitive enough to identify WMLs in those products in the same area of the mid-troposphere as the comparison RAOBs. We intend to examine more broadly and deeply the performance of remote sensing and model products versus RAOBs as the next step in using our detection methodology in process studies of desert-sourced dry well-mixed airstreams.

Acknowledgments

This work would not have been possible without the support of the Science of Terra and Aqua program, Atmospheric Science Measurement Team (Program Scientist, R. Kakar, NASA HQ) and the support of the Modeling, Analysis, and Prediction (Program Scientist, D. Constantine,

514 NASA HQ). Data came from multiple sources: NOAA NCEI, USDA Forest Service, NCAR DSS,
515 and NASA's GMAO and GES DISC. We had help and advice on AIRS+AMSU products from the
516 Goddard Sounder Research Team (J. Susskind, L. Iredell, J. Blaisdell) and the DISC (T. Hearty,
517 A. Savtchenko, C. Shie). We also had useful discussions with O. Reale, S. Braun, R. Shi, and S.
518 Lang. The software can be downloaded from the Goddard Cloud Library at [https://](https://portal.nccs.nasa.gov/cloudlibrary/)
519 portal.nccs.nasa.gov/cloudlibrary/. We thank our reviewers for greatly improving this manuscript
520 and suggesting useful changes to the analysis process.

References

- Agustí-Panareda, A., A. Beljaars, C. Cardinali, I. Genkova, and C. Thorncroft, 2010: Impact of assimilating AMMA soundings on ECMWF analyses and forecasts. *Wea. Forecasting*, **25**, 1142-1160.
- Allen, J. T., and D. J. Karoly, 2014: A climatology of Australian severe thunderstorm environments 1979–2011: inter-annual variability and ENSO influence. *Int. J. Climatol.*, **34**, 81-97.
- Aumann, H. H., and Coauthors, 2003: AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems. *IEEE Trans. Geosci. Remote Sens.*, **41**, 253-264.
- Bailey, R. G., 2004: Identifying ecoregion boundaries. *Environ. Management*, **34**, S14-S26.
- Bailey, R. G., 2014: *Ecoregions: The Ecosystem Geography of the Oceans and Continents 2nd Ed.* Springer, New York, 180 pp.
- Benjamin, S. G., and T. N. Carlson, 1986: Some effects of surface heating and topography on the regional severe storm environment. Part I: Three-dimensional simulations. *Mon. Wea. Rev.*, **114**, 330-343.
- Birch, C. E., D. J. Parker, J. H. Marsham, and G. M. Devine, 2012: The effect of orography and surface albedo on stratification in the summertime Saharan boundary layer: Dynamics and implications for dust transport. *J. Geophys. Res. Atmos.*, **117**, D05105, doi: 10.1029/2011JD015965.
- Bissolli, P., J. Grieser, N. Dotzek, and M. Welsch, 2007: Tornadoes in Germany 1950–2003 and their relation to particular weather conditions. *Global Planet. Change*, **57**, 124-138.

543 Blackwell, W. J., 2005: A neural-network technique for the retrieval of atmospheric temperature
544 and moisture profiles from high spectral resolution sounding data. *IEEE Trans. Geosci.*
545 *Remote Sens.*, **43**, 2535-2546.

546 Blackwell, W. J., 2012: Neural network Jacobian analysis for high-resolution profiling of the
547 atmosphere. *EURASIP J. Adv. Signal Proc.*, **2012**, 71-82, doi: 10.1186/1687-6180-2012-1171.

548 Blackwell, W. J., and A. B. Milstein, 2014: A Neural Network Retrieval Technique for High-
549 Resolution Profiling of Cloudy Atmospheres. *IEEE J. Sel. Topics Appl. Earth Obs. Remote*
550 *Sens.*, **7**, 1260-1270.

551 Born, K., A. H. Fink, and P. Knippertz, 2010: Meteorological processes influencing the weather
552 and climate of Morocco. *Impacts of Global Change on the Hydrological Cycle in West and*
553 *Northwest Africa*, P. Speth, and Coauthors, Eds., Springer, 150–163.

554 Bou Karam, D., and Coauthors, 2014: Synoptic-scale dust emissions over the Sahara Desert
555 initiated by a moist convective cold pool in early August 2006. *Quart. J. Roy. Meteor. Soc.*,
556 **140**, 2591-2607.

557 Bounoua, L., and T. N. Krishnamurti, 1991: Thermodynamic budget of the five day wave over
558 the Saharan Desert during summer. *Meteor. Atmos. Phys.*, **47**, 1-25.

559 Camberlin, P., 2009: Nile Basin climates. *The Nile : Origin, Environments, Limnology and*
560 *Human Use*, H. J. Dumont, Ed., Springer, 307-333.

561 Carlson, T. N., and F. H. Ludlam, 1968: Conditions for the occurrence of severe local storms.
562 *Tellus*, **20**, 206-223.

563 Carlson, T. N., and J. M. Prospero, 1972: The large-scale movement of Saharan air outbreaks
564 over the northern equatorial Atlantic. *J. Appl. Meteor.*, **11**, 283-297.

565 Carlson, T. N., S. G. Benjamin, G. S. Forbes, and Y.-F. Li, 1983: Elevated mixed layers in the
 566 regional severe storm environment: Conceptual model and case studies. *Mon. Wea. Rev.*, **111**,
 567 1453-1473.

568 Carr, F. H., and J. P. Millard, 1985: A composite study of comma clouds and their association
 569 with severe weather over the Great Plains. *Mon. Wea. Rev.*, **113**, 370-387.

570 Cho, C., and D. H. Staelin, 2006: Cloud clearing of Atmospheric Infrared Sounder hyperspectral
 571 infrared radiances using stochastic methods. *J. Geophys. Res. Atmos.*, **111**,
 572 doi:10.1029/2005JD006013.

573 Conzemius, R. J., and E. Fedorovich, 2006: Dynamics of sheared convective boundary layer
 574 entrainment. Part I: Methodological background and large-eddy simulations. *J. Atmos. Sci.*,
 575 **63**, 1151-1178.

576 Cuesta, J., J. H. Marsham, D. J. Parker, and C. Flamant, 2009: Dynamical mechanisms
 577 controlling the vertical redistribution of dust and the thermodynamic structure of the West
 578 Saharan atmospheric boundary layer during summer. *Atmos. Sci. Lett.*, **10**, 34-42.

579 Cuesta, J., and Coauthors, 2008: Multiplatform observations of the seasonal evolution of the
 580 Saharan atmospheric boundary layer in Tamanrasset, Algeria, in the framework of the African
 581 Monsoon Multidisciplinary Analysis field campaign conducted in 2006. *J. Geophys. Res.*
 582 *Atmos.*, **113**, DOI: 10.1029/2007JD009417.

583 Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance
 584 of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553-597.

585 Durre, I., R. S. Vose, and D. B. Wuertz, 2006: Overview of the Integrated Global Radiosonde
 586 Archive. *J. Climate*, **19**, 53-68.

587 Faccani, C., and Coauthors, 2009: The impacts of AMMA radiosonde data on the French global
588 assimilation and forecast system. *Wea. Forecasting*, **24**, 1268-1286.

589 Fink, A. H., 2017: Mean climate and seasonal cycle. *Meteorology of Tropical West Africa*, D. J.
590 Parker, and M. Diop-Kane, Eds., Wiley Blackwell, 1-39.

591 Fink, A. H., and Coauthors, 2011: Operational meteorology in West Africa: Observational
592 networks, weather analysis and forecasting. *Atmos. Sci. Lett.*, **12**, 135-141.

593 Fitzpatrick, R. G. J., C. L. Bain, P. Knippertz, J. H. Marsham, and D. J. Parker, 2015: The West
594 African monsoon onset: A concise comparison of definitions. *J. Climate*, **28**, 8673-8694.

595 Flamant, C., and Coauthors, 2007: Airborne observations of the impact of a convective system
596 on the planetary boundary layer thermodynamics and aerosol distribution in the inter-tropical
597 discontinuity region of the West African monsoon. *Quart. J. Roy. Meteor. Soc.*, **133**, 1175-
598 1189.

599 Gamo, M., 1996: Thickness of the dry convection and large-scale subsidence above deserts.
600 *Bound.-Layer Meteor.*, **79**, 265-278.

601 Garcia-Carreras, L., and Coauthors, 2015: The turbulent structure and diurnal growth of the
602 Saharan atmospheric boundary layer. *J. Atmos. Sci.*, **72**, 693-713.

603 Garcia-Carreras, L., and Coauthors, 2013: The impact of convective cold pool outflows on
604 model biases in the Sahara. *Geophys. Res. Lett.*, **40**, 1647-1652.

605 Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and
606 Applications, Version 2 (MERRA-2). *J. Climate*, **30**, 5419-5454.

607 Gounou, A., F. Guichard, and F. Couvreux, 2012: Observations of diurnal cycles over a West
608 African meridional transect: pre-monsoon and full-monsoon seasons. *Boundary-Layer
609 Meteorology*, **144**, 329-357.

610 Graf, M. A., M. Sprenger, and R. W. Moore, 2011: Central European tornado environments as
611 viewed from a potential vorticity and Lagrangian perspective. *Atmos. Res.*, **101**, 31-45.

612 Grundstein, A., M. Shepherd, P. Miller, and S. E. Sarnat, 2017: The role of mesoscale-
613 convective processes in explaining the 21 November 2016 epidemic thunderstorm asthma
614 event in Melbourne, Australia. *J. Appl. Meteor. Climatol.*, **56**, 1337-1343.

615 Honda, T., and T. Kawano, 2015: How does mid-tropospheric dry air affect the evolution of
616 supercellular convection? *Atmos. Res.*, **157**, 1-16.

617 James, R. P., and P. M. Markowski, 2010: A numerical investigation of the effects of dry air
618 aloft on deep convection. *Mon. Wea. Rev.*, **138**, 140-161.

619 Knippertz, P., 2003: Tropical–Extratropical Interactions Causing Precipitation in Northwest
620 Africa: Statistical Analysis and Seasonal Variations. *Mon. Wea. Rev.*, **131**, 3069-3076.

621 Knippertz, P., and M. C. Todd, 2010: The central west Saharan dust hot spot and its relation to
622 African easterly waves and extratropical disturbances. *J. Geophys. Res. Atmos.*, **115**, D12117,
623 doi:12110.11029/12009JD012819.

624 Knippertz, P., M. Christoph, and P. Speth, 2003: Long-term precipitation variability in Morocco
625 and the link to the large-scale circulation in recent and future climates. *Meteor. Atmos. Phys.*,
626 **83**, 67-88.

627 Knippertz, P., and Coauthors, 2009: Dust mobilization and transport in the northern Sahara
628 during SAMUM 2006 – a meteorological overview. *Tellus B: Chem. Phys. Meteor.*, **61**, 12-
629 31.

630 Lanicci, J. M., T. N. Carlson, and T. T. Warner, 1987: Sensitivity of the Great Plains severe-
631 storm environment to soil-moisture distribution. *Mon. Wea. Rev.*, **115**, 2660-2673.

632 Lélé, M. I., and P. J. Lamb, 2010: Variability of the Intertropical Front (ITF) and Rainfall over
 633 the West African Sudan–Sahel Zone. *J. Climate*, **23**, 3984-4004.

634 Liu, N., and C. Liu, 2018: Synoptic environments and characteristics of convection reaching the
 635 tropopause over Northeast China. *Mon. Wea. Rev.*, **146**, 745-759.

636 Lucas, C., E. J. Zipser, and B. S. Ferrier, 2000: Sensitivity of tropical West Pacific oceanic squall
 637 lines to tropospheric wind and moisture profiles. *J. Atmos. Sci.*, **57**, 2351-2373.

638 Ma, X. L., T. J. Schmit, and W. L. Smith, 1999: A nonlinear physical retrieval algorithm—Its
 639 application to the GOES-8/9 Sounder. *J. Appl. Meteor.*, **38**, 501-513.

640 Marsham, J. H., D. J. Parker, C. M. Grams, B. T. Johnson, W. M. F. Grey, and A. N. Ross, 2008:
 641 Observations of mesoscale and boundary-layer scale circulations affecting dust transport and
 642 uplift over the Sahara. *Atmos. Chem. Phys.*, **8**, 6979-6993.

643 Marsham, J. H., and Coauthors, 2016: The contrasting roles of water and dust in controlling daily
 644 variations in radiative heating of the summertime Saharan heat low. *Atmos. Chem. Phys.*, **16**,
 645 3563-3575.

646 Marsham, J. H., and Coauthors, 2013: Meteorology and dust in the central Sahara: Observations
 647 from Fennec supersite-1 during the June 2011 Intensive Observation Period. *J. Geophys. Res.*
 648 *Atmos.*, **118**, 4069-4089.

649 Matsangouras, I. T., P. T. Nastos, H. B. Bluestein, and M. V. Sioutas, 2014: A climatology of
 650 tornadic activity over Greece based on historical records. *Int. J. Climatol.*, **34**, 2538-2555.

651 Medina, S., R. A. Houze, A. Kumar, and D. Niyogi, 2010: Summer monsoon convection in the
 652 Himalayan region: terrain and land cover effects. *Quart. J. Roy. Meteor. Soc.*, **136**, 593-616.

653 Menzel, W. P., T. J. Schmit, P. Zhang, and J. Li, 2018: Satellite-based atmospheric infrared
 654 sounder development and applications. *Bull. Amer. Meteor. Soc.*, **99**, 583-603.

655 Messenger, C., D. J. Parker, O. Reitebuch, A. Agusti-Panareda, C. M. Taylor, and J. Cuesta, 2010:
 656 Structure and dynamics of the Saharan atmospheric boundary layer during the West African
 657 monsoon onset: Observations and analyses from the research flights of 14 and 17 July 2006.
 658 *Quart. J. Roy. Meteor. Soc.*, **136**, 107-124.

659 Milstein, A. B., and W. J. Blackwell, 2016: Neural network temperature and moisture retrieval
 660 algorithm validation for AIRS/AMSU and CrIS/ATMS. *J. Geophys. Res. Atmos.*, **121**, 1414-
 661 1430.

662 Moeng, C.-H., and P. P. Sullivan, 1994: A comparison of shear- and buoyancy-driven planetary
 663 boundary layer flows. *J. Atmos. Sci.*, **51**, 999-1022.

664 Nicholls, S. D., and K. I. Mohr, 2010: An analysis of the pre-storm environment of intense West
 665 African convective systems. *Mon. Wea. Rev.*, **138**, 3721–3739.

666 Nicholson, S. E., 1996: A review of climate dynamics and climate variability in Eastern Africa.
 667 *Limnology, Climatology and Paleoclimatology of the East African Lakes*, T. C. Johnson, and
 668 E. O. Odada, Eds., CRC Press, 25-56.

669 ———, 2011: *Dryland Climatology*. Cambridge University Press, 516 pp.

670 Office of the Federal Coordinator of Meteorology, 1997: Federal Meteorological Handbook No.3
 671 Rawinsonde and Pibal Observations. FCM-H3-1997, Dept. of Commerce, National
 672 Oceanographic and Atmospheric Administration, Washington, DC, 191 pp.

673 Parker, D. J., and Coauthors, 2005: The diurnal cycle of the West African monsoon circulation.
 674 *Quart. J. Roy. Meteor. Soc.*, **131**, 2839-2860.

675 Parkinson, C. L., 2003: Aqua: An Earth-Observing Satellite mission to examine water and other
 676 climate variables. *IEEE Trans. Geosci. Remote Sens.*, **41**, 173-183.

677 Peckham, S. E., R. B. Wilhelmson, L. J. Wicker, and C. L. Ziegler, 2004: Numerical simulation
678 of the interaction between the dryline and horizontal convective rolls. *Mon. Wea. Rev.*, **132**,
679 1792-1812.

680 Prospero, M. J., and T. N. Carlson, 1981: Saharan air outbreaks over the tropical North Atlantic.
681 *Pure Appl. Geophys.*, 677-691.

682 Púčik, T., P. Groenemeijer, D. Rýva, and M. Kolář, 2015: Proximity soundings of severe and
683 nonsevere thunderstorms in Central Europe. *Mon. Wea. Rev.*, **143**, 4805-4821.

684 Rasmussen, K. L., and R. A. Houze, 2011: Orographic convection in subtropical South America as
685 seen by the TRMM satellite. *Mon. Wea. Rev.*, **139**, 2399-2420.

686 Raveh-Rubin, S., 2017: Dry intrusions: Lagrangian climatology and dynamical impact on the
687 planetary boundary layer. *J. Climate*, **30**, 6661-6682.

688 Reale, O., E. L. McGrath-Spangler, W. McCarty, D. Holdaway, and R. Gelaro, 2018: Impact of
689 adaptively thinned AIRS cloud-cleared radiances on tropical cyclone representation in a
690 global data assimilation and forecast system. *Wea. Forecasting*, doi:10.1175/WAF-D-1117-
691 0175.1171, in press.

692 Roberts, A. J., J. H. Marsham, and P. Knippertz, 2015: Disagreements in low-level moisture
693 between (re)analyses over summertime West Africa. *Mon. Wea. Rev.*, **143**, 1193-1211.

694 Roca, R., J.-P. Lafore, C. Piriou, and J.-L. Redelsperger, 2005: Extratropical dry-air intrusions
695 into the West African monsoon midtroposphere: An important factor for the convective
696 activity over the Sahel. *J. Atmos. Sci.*, **62**, 390-407.

697 Romatschke, U., and R. A. H. Jr., 2011: Characteristics of precipitating convective systems in
698 the premonsoon season of South Asia. *J. Hydrometeor.*, **12**, 157-180.

699 Romatschke, U., S. Medina, and R. A. Houze, 2010: Regional, seasonal, and diurnal variations
 700 of extreme convection in the South Asian region. *J. Climate*, **23**, 419-439.
 701 Schaefer, J. T., 1974: The life cycle of the dryline. *J. Appl. Meteor.*, **13**, 444-449.
 702 Stubenrauch, C. J., A. G. Feofilov, S. E. Protopapadaki, and R. Armante, 2017: Cloud
 703 climatologies from the infrared sounders AIRS and IASI: strengths and applications. *Atmos.*
 704 *Chem. Phys.*, **17**, 13625-13644.
 705 Susskind, J., J. Blaisdell, and L. Iredell, 2012: Significant advances in the AIRS Science Team
 706 Version-6 retrieval algorithm. *SPIE Optical Engineering + Applications Conference*, San
 707 Diego, SPIE, doi: 10.1117/1112.929953.
 708 Susskind, J., J. M. Blaisdell, and L. Iredell, 2014: Improved methodology for surface and
 709 atmospheric soundings, error estimates, and quality control procedures: The Atmospheric
 710 Infrared Sounder Science Team version-6 retrieval algorithm. *J. Appl. Remote Sens.*, **8**, doi:
 711 10.1117/1111.JRS.1118.084994.
 712 Takemi, T., 2014: Convection and precipitation under various stability and shear conditions:
 713 Squall lines in tropical versus midlatitude environment. *Atmos. Res.*, **142**, 111-123.
 714 Trier, S. B., G. S. Romine, D. A. Ahijevych, R. J. Trapp, R. S. Schumacher, M. C. Coniglio, and
 715 D. J. Stensrud, 2015: Mesoscale thermodynamic influences on convection initiation near a
 716 surface dryline in a convection-permitting ensemble. *Mon. Wea. Rev.*, **143**, 3726-3753.
 717 Vergados, P., A. J. Mannucci, C. O. Ao, O. Verkhoglyadova, and B. Iijima, 2018: Comparisons
 718 of the tropospheric specific humidity from GPS radio occultations with ERA-Interim, NASA
 719 MERRA, and AIRS data. *Atmos. Meas. Tech.*, **11**, 1193-1206.
 720 Vizu, E. K., and K. H. Cook, 2009: A mechanism for African monsoon breaks: Mediterranean
 721 cold air surges. *J. Geophys. Res. Atmos.*, **114**, D01104.

722 ———, 2017: Mesoscale convective systems and nocturnal rainfall over the West African Sahel:
723 role of the Inter-tropical front. *Clim. Dyn.*, **42**, 1-28.

724 Wahab, M. A., M. El-Metwally, R. Hassan, M. Lefèvre, A. Oumbe, and L. Wald, 2010:
725 Assessing surface solar irradiance and its long-term variations in the northern Africa desert
726 climate using Meteosat images. *Int. J. Remote Sens.*, **31**, 261-280.

727 Wang, S., and A. H. Sobel, 2012: Impact of imposed drying on deep convection in a cloud-
728 resolving model. *J. Geophys. Res. Atmos.*, **117**, D02112, doi: 02110.01029/02011JD016847.

729 Warner, T. T., 2004: *Desert Meteorology*. Cambridge University Press, 595 pp.

730 Weng, F., L. Zhao, R. R. Ferraro, G. Poe, X. Li, and N. C. Grody, 2003: Advanced microwave
731 sounding unit cloud and precipitation algorithms. *Radio Science*, **38**, 8068, doi:
732 8010.1029/2002RS002679.

733 Yamane, Y., T. Hayashi, M. Kiguchi, F. Akter, and A. M. Dewan, 2013: Synoptic situations of
734 severe local convective storms during the pre-monsoon season in Bangladesh. *Int. J.*
735 *Climatol.*, **33**, 725-734.

736 Zhang, C., and J. Pennington, 2004: African dry air outbreaks. *J. Geophys. Res. Atmos.*, **109**,
737 doi:10.1029/2003JD003978.

738 Zuluaga, M. D., and R. A. Houze, 2015: Extreme convection of the near-equatorial Americas,
739 Africa, and adjoining oceans as seen by TRMM. *Mon. Wea. Rev.*, **143**, 298-316.

740

741

Figure Captions

Figure 1. A map of the Bailey’s ecoregion provinces for North Africa with the RAOB stations indicated by black squares (Saharan stations), black dots (non-Saharan stations), and black open squares around dots (mixed stations that function as Saharan stations during most but not all months). The color scheme refers to the ecoregion division and vegetation provinces within them (“Mtn” applies to mountain ecoregions). Red shades apply to the Mediterranean division, yellow to the tropical/subtropical desert division, and orange to the tropical/subtropical steppe division. Greens and blues apply to the savanna division (lighter shades) and rainforest division (darker shades).

Figure 2. Maps of the North African stations with the number of useful radiosonde launches resulting in a RAOB that could be used for algorithm development. a) Useful launches since the earliest observation in the IGRA, 1948. b) Useful launches since 2003, the first full year of operations of the Aqua satellite mission. The use of input from Aqua sounders and from model reanalysis starts with 2003, although model reanalysis can go further back in time. The stations marked with zero are stations that do not have observations during the Aqua era.

Figure 3. Algorithm-analyzed WML detections (shaded) from a Tamanrasset RAOB observation using several maximum allowable potential temperature and mixing ratio lapse rates parameterizations. Water vapor mixing ratio lapse rates either scale with the layer mean mixing ratio (VAR) or are constant (CON). For each panel, the potential temperature lapse rate is fixed at the rate shown. The first analyzed WML is shaded orange, the second gray, the third cyan, and the fourth manila. Shown wind barbs are in m/s. Additional lapse rate parameterization options for this case are shown in the supplemental material as Figure S1.

Figure 4. Same as Figure 3, except for Tunis. Additional lapse rate parameterization options for this case are shown in the supplemental material as Figure S2.

Figure 5. Decision tree of the WML detection software. The blue box is the analysis set up. The yellow boxes indicate data input. Gray boxes contain the analysis decisions. The green box contains output actions. Red boxes at the top and bottom control program execution.

Figure 6. Archetypal RAOBs of vertical structures found in North Africa. The red and green lines represent the temperature and dewpoint vertical profiles, respectively, and the orange shaded layer denotes algorithm-detected WMLs. The station elevation (m), number of useful levels in the analysis column, and the number of detected WMLs are indicated. a) In Salah is typical of summertime low elevation inner Saharan stations. b) The profile at Tunis is typical of northern coastal stations. c) Agadez is on the southern edge of the Sahara, but the West African monsoon reaches Agadez during July and August. This is a typical late spring-early summer profile at Agadez before monsoon onset. d) An example at Agadez during the monsoon season.

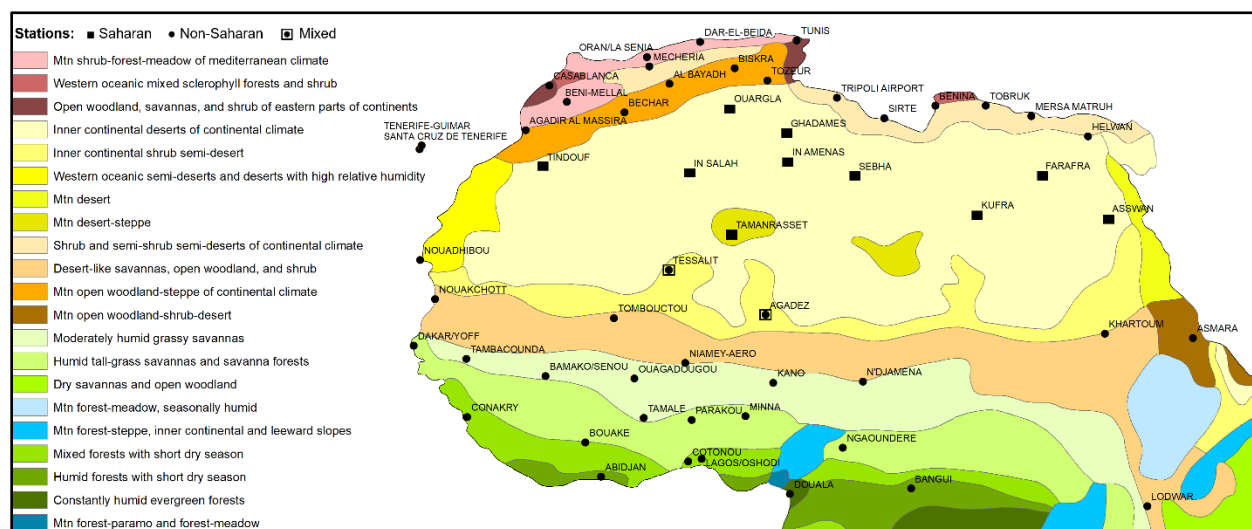
Figure 7. Examples of more challenging daytime environments from 2015 at a) In Salah and b) Tamanrasset in the spring; c) Abidjan in the winter; d) Ngaoundéré in the spring prior to the monsoon onset. The first analyzed WML is shaded orange, the second gray, the third cyan.

Figure 8. Four-panel figure comparing WML detections at In Salah in the summer from a) RAOB data, b) AIRS+AMSU temperature and humidity products, c) ERA-I reanalysis, d) MERRA-2 reanalysis. The orange shading in each panel is the first WML detected in that profile.

Figure 9. As in Figure 8, except for springtime in Dakar

Figure 10. As in Figure 8, except for springtime in Ouagadougou.

785 **Figure 11.** As in Figure 8, except for summertime in Tunis.



787

788 **Figure 1.** A map of the Bailey's ecoregion provinces for North Africa with the RAOB stations
 789 indicated by black squares (Saharan stations), black dots (non-Saharan stations), and black open
 790 squares around dots (mixed stations that function as Saharan stations during most but not all
 791 months). The color scheme refers to the ecoregion division and vegetation provinces within them
 792 ("Mtn" applies to mountain ecoregions). Red shades apply to the Mediterranean division, yellow
 793 to the tropical/subtropical desert division, and orange to the tropical/subtropical steppe division.
 794 Greens and blues apply to the savanna division (lighter shades) and rainforest division (darker
 795 shades).

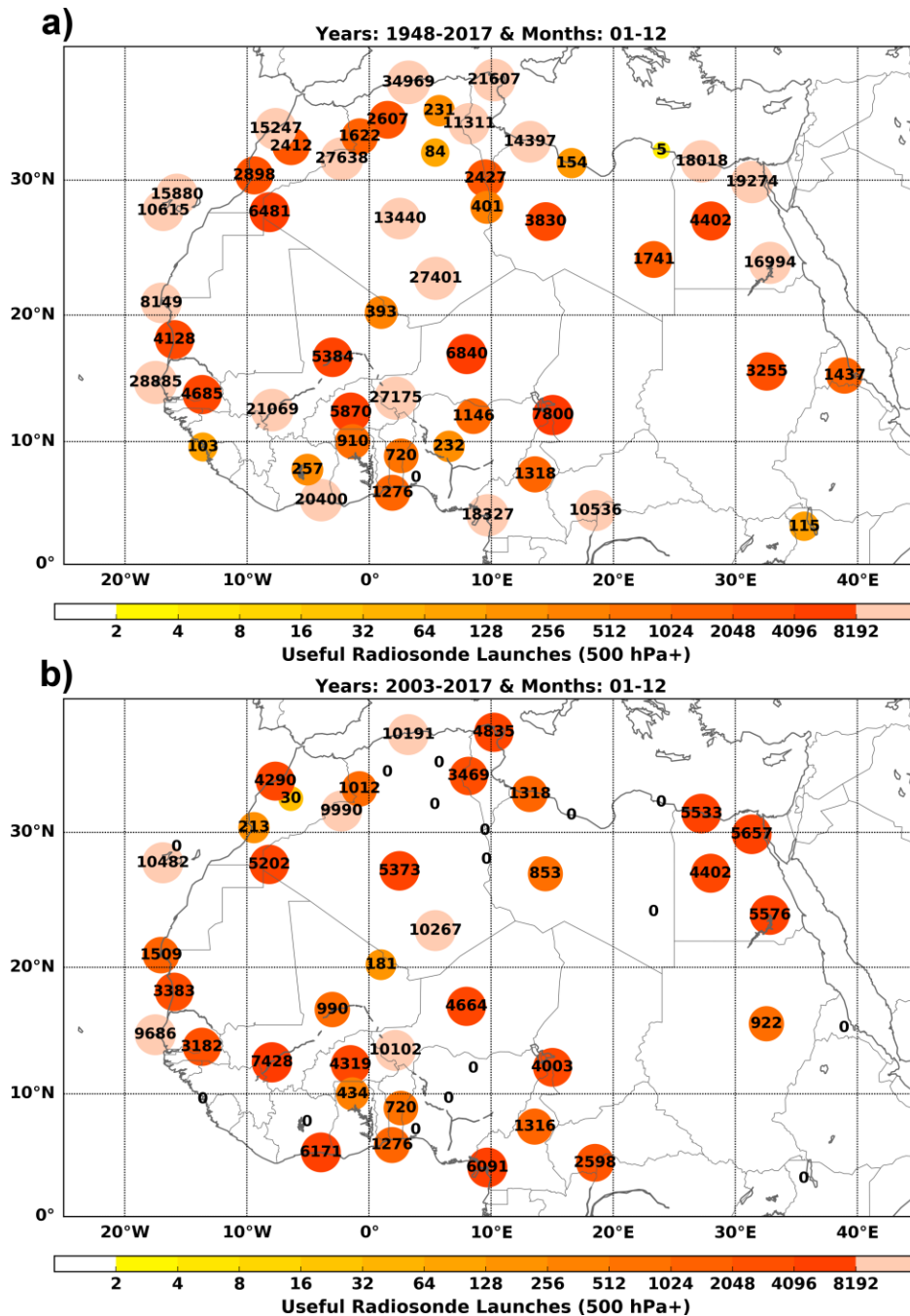


Figure 2. Maps of the North African stations with the number of useful radiosonde launches resulting in a RAOB that could be used for algorithm development. a) Useful launches since the earliest observation in the IGRA, 1948. b) Useful launches since 2003, the first full year of operations of the Aqua satellite mission. The use of input from Aqua sounders and from model reanalysis starts with 2003, although model reanalysis can go further back in time. The stations marked with zero are stations that do not have observations during the Aqua era.

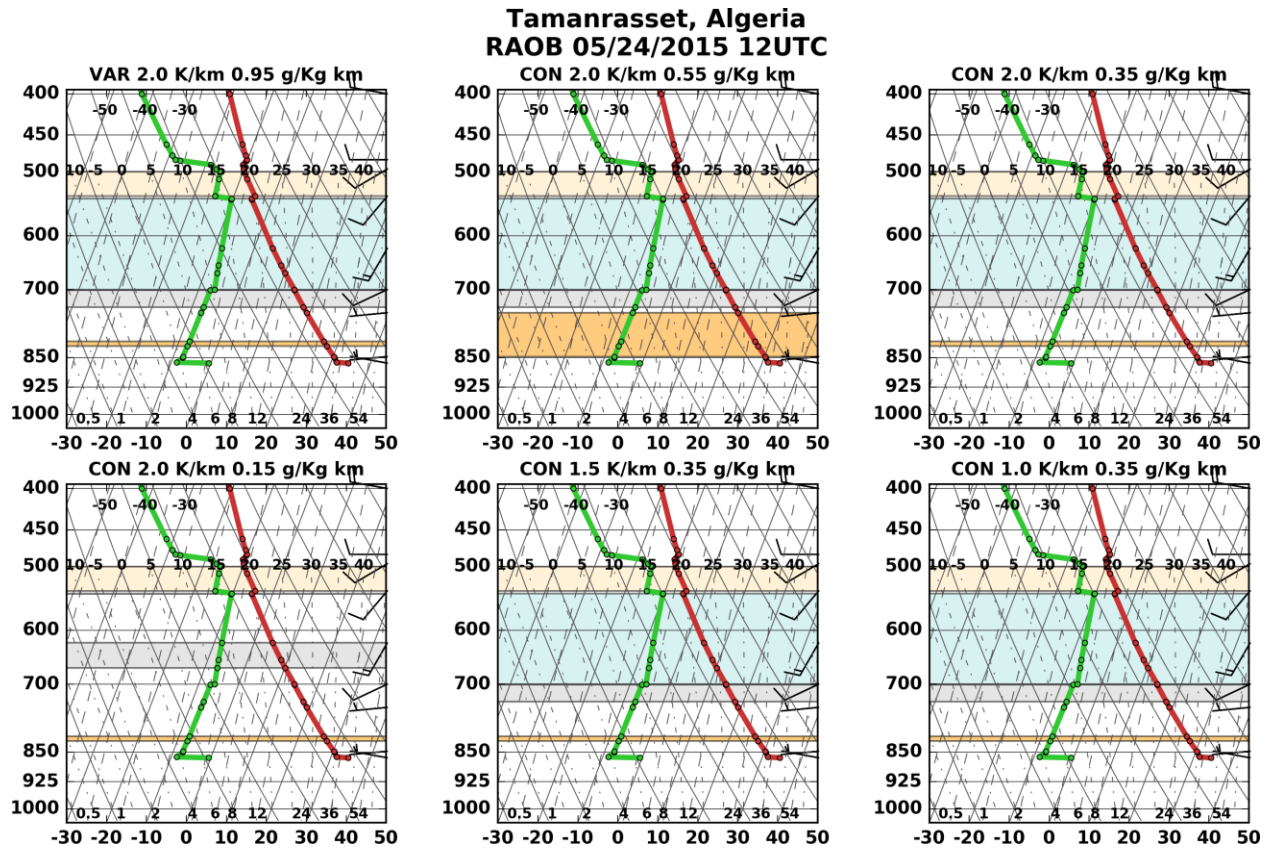


Figure 3. Algorithm-analyzed WML detections (shaded) from a Tamanrasset RAOB observation using several maximum allowable potential temperature and mixing ratio lapse rates parameterizations. Water vapor mixing ratio lapse rates either scale with the layer mean mixing ratio (VAR) or are constant (CON). For each panel, the potential temperature lapse rate is fixed at the rate shown. The first analyzed WML is shaded orange, the second gray, the third cyan, and the fourth manila. Shown wind barbs are in m/s. Additional lapse rate parameterization options for this case are shown in the supplemental material as Figure S1.

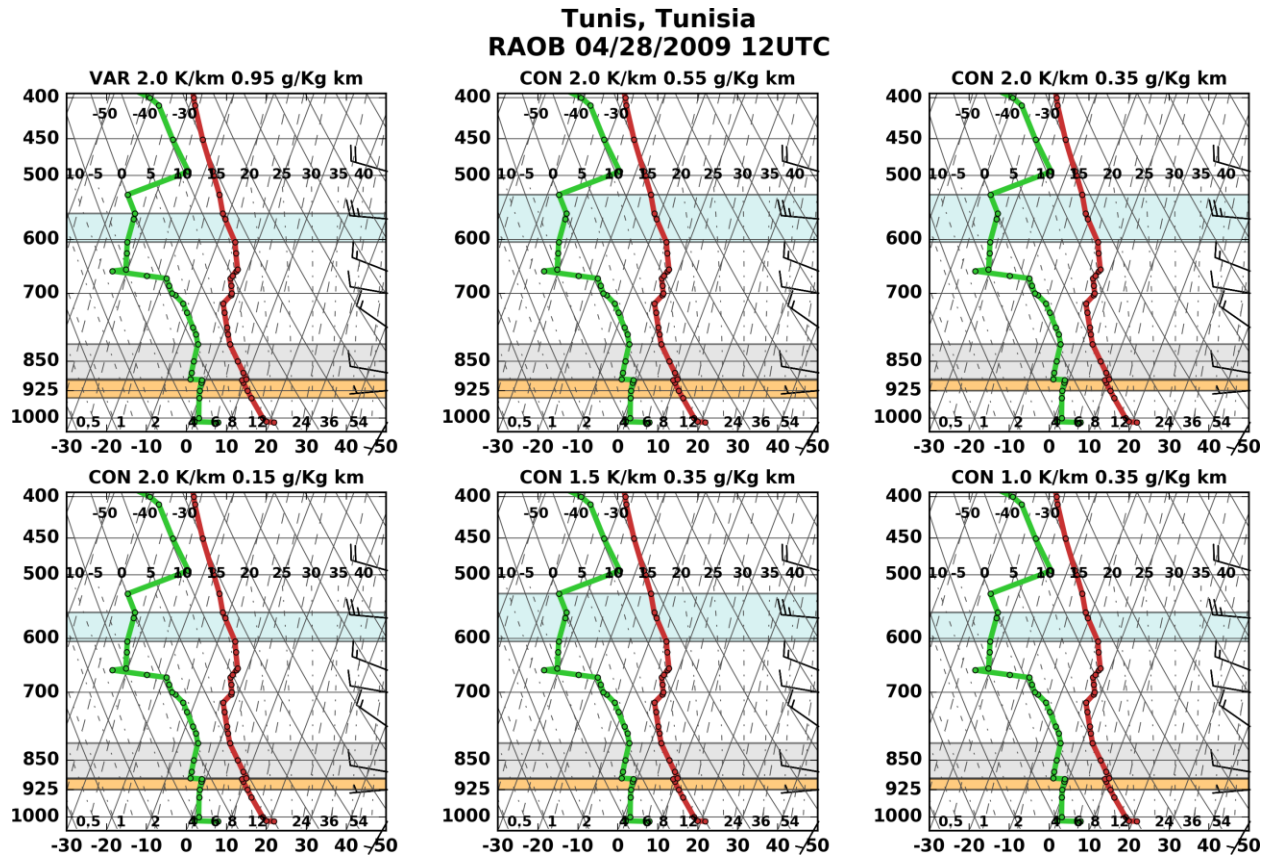


Figure 4. Same as Figure 3, except for Tunis. Additional lapse rate parameterization options for this case are shown in the supplemental material as Figure S2.

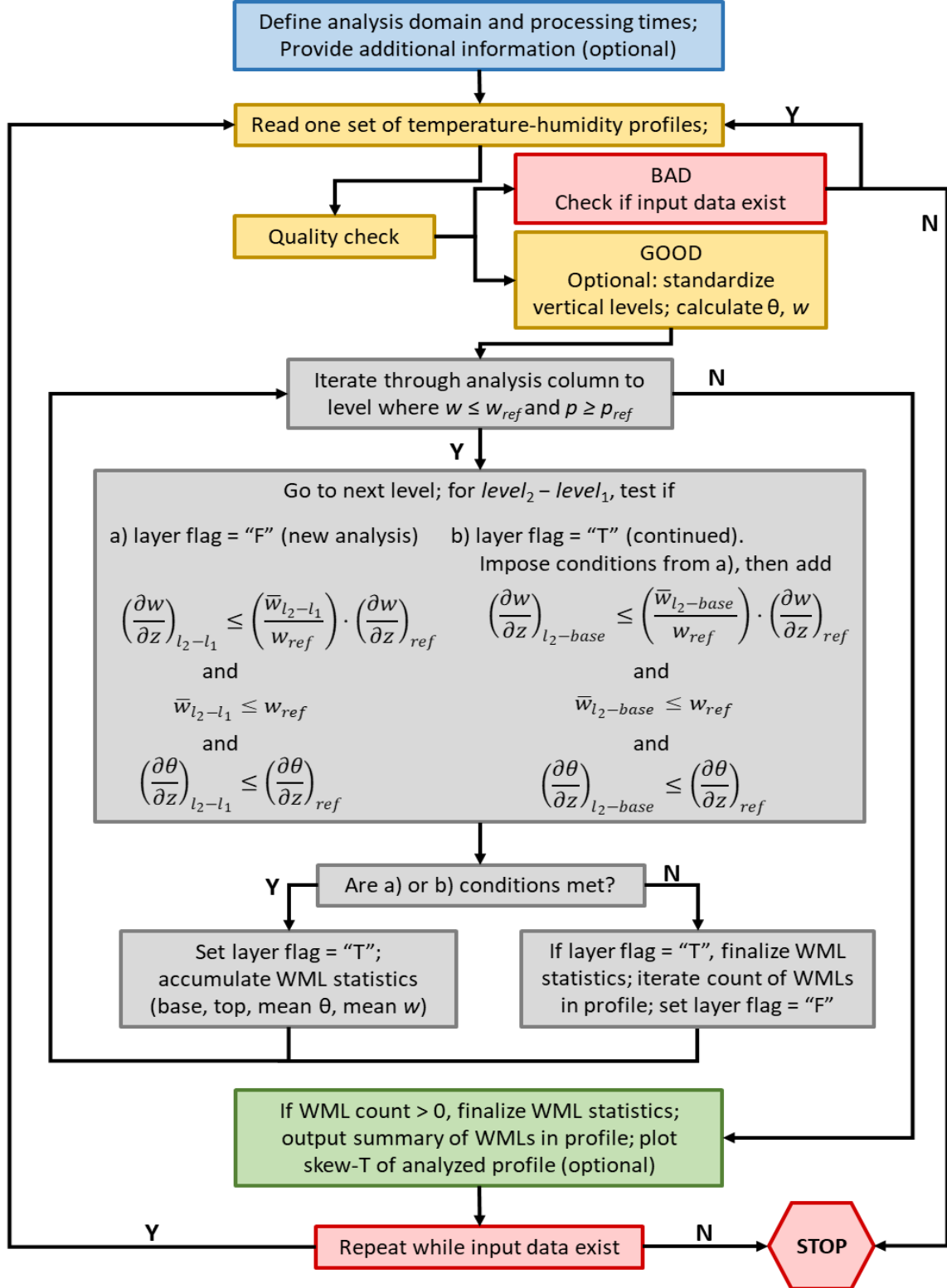


Figure 5. Decision tree of the WML detection software. The blue box is the analysis set up. The yellow boxes indicate data input. Gray boxes contain the analysis decisions. The green box contains output actions. Red boxes at the top and bottom control program execution.

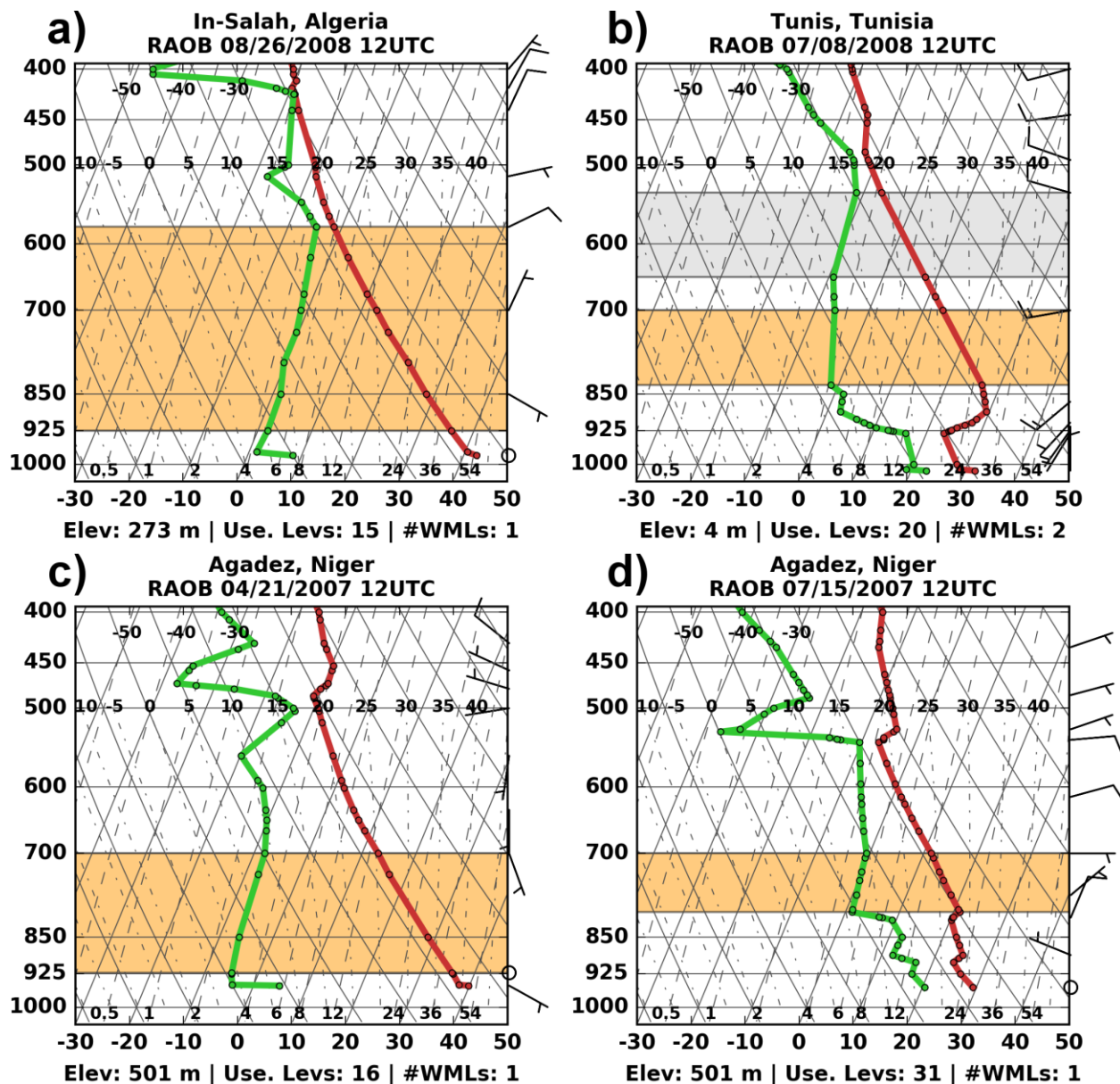


Figure 6. Archetypal RAOBs of vertical structures found in North Africa. The red and green lines represent the temperature and dewpoint vertical profiles, respectively, and the orange shaded layer denotes algorithm-detected WMLs. The station elevation (m), number of useful levels in the analysis column, and the number of detected WMLs are indicated. a) In Salah is typical of summertime low elevation inner Saharan stations. b) The profile at Tunis is typical of northern coastal stations. c) Agadez is on the southern edge of the Sahara, but the West African monsoon reaches Agadez during July and August. This is a typical late spring-early summer profile at Agadez before monsoon onset. d) An example at Agadez during the monsoon season. Shown wind barbs are in m/s.

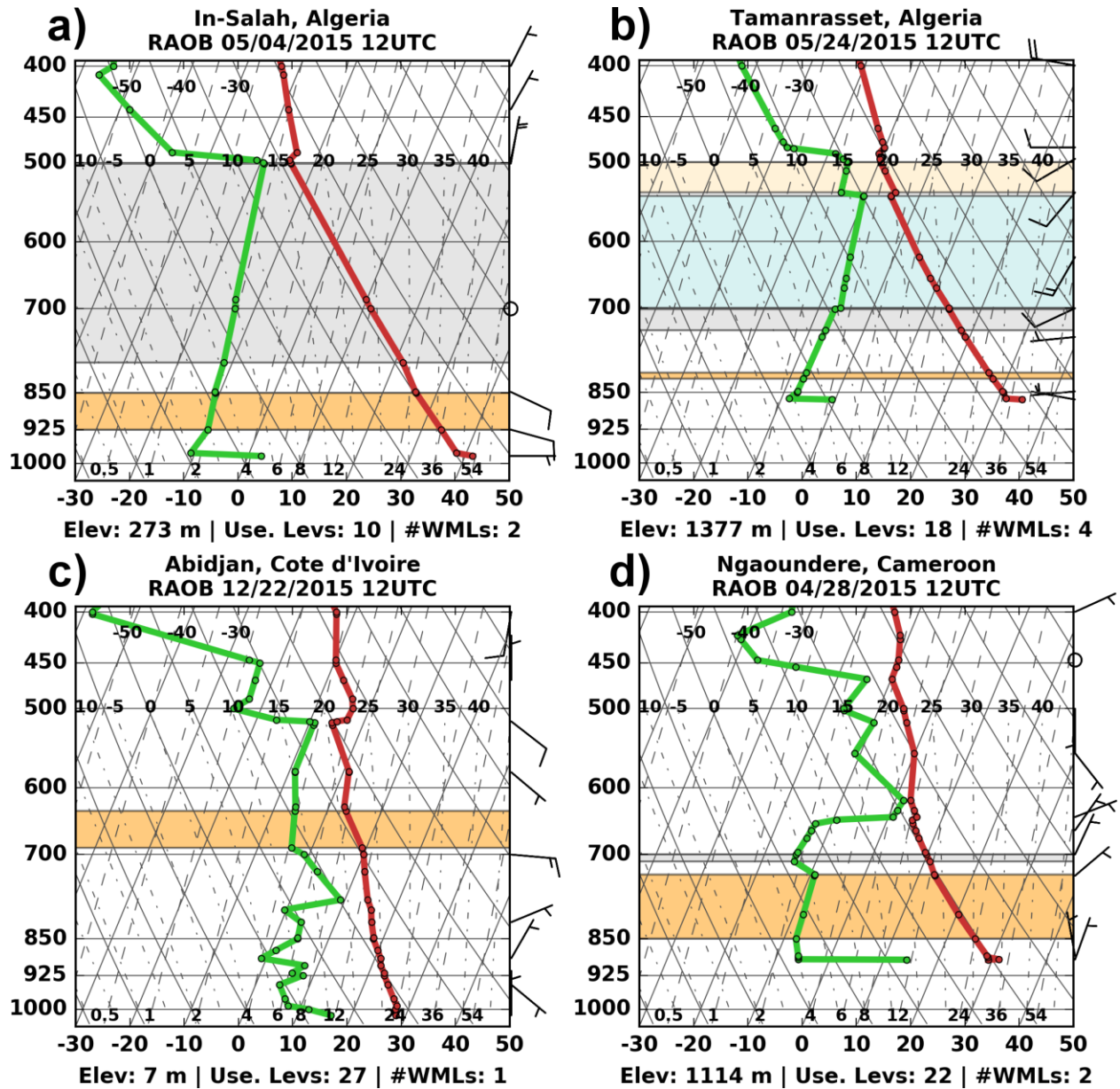


Figure 7. Examples of more challenging daytime environments from 2015 at a) In Salah and b) Tamanrasset in the spring; c) Abidjan in the winter; d) Ngaoundéré in the spring prior to the monsoon onset. The first analyzed WML is shaded orange, the second gray, the third cyan. Shown wind barbs are in m/s.

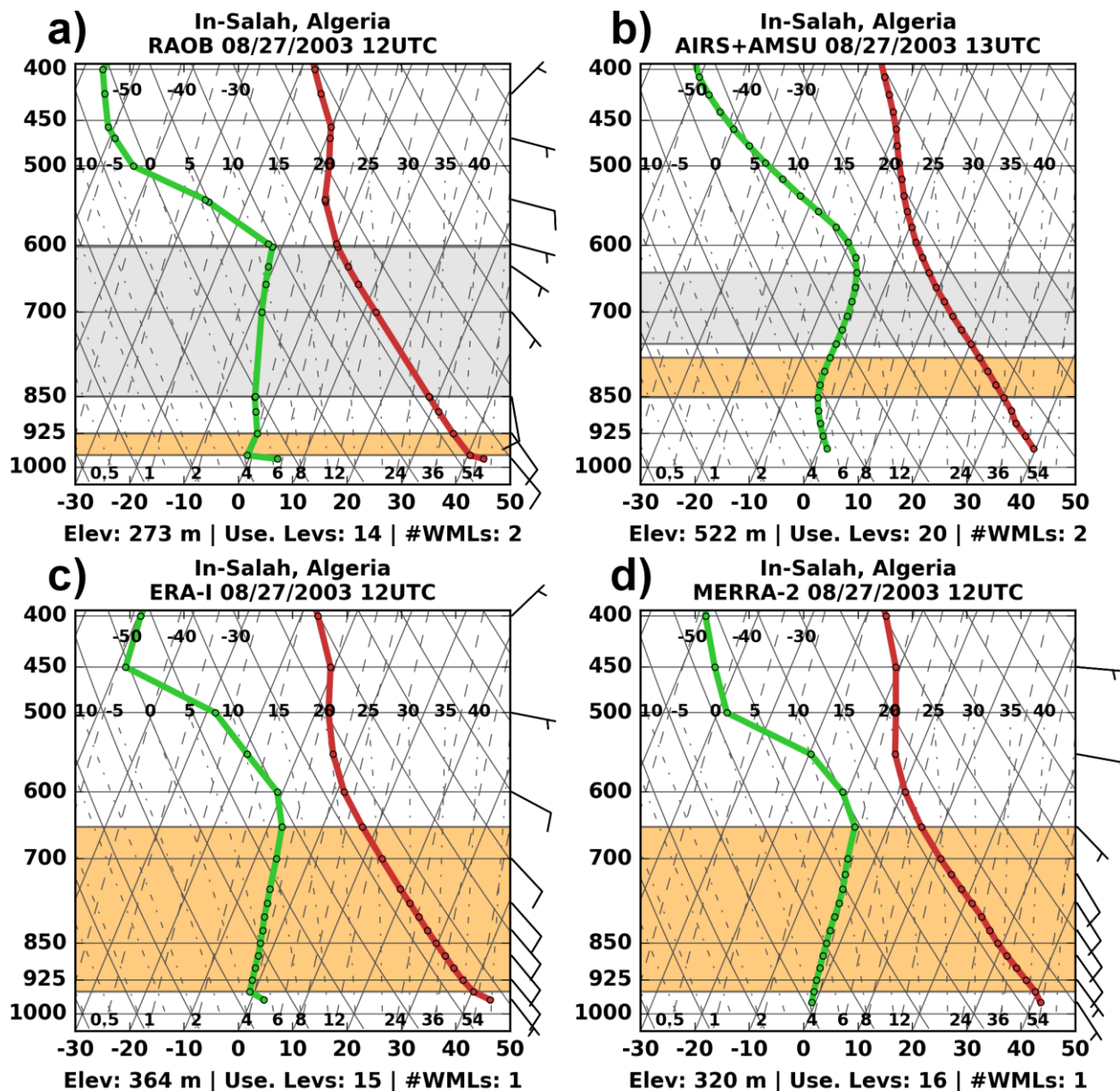


Figure 8. Four-panel figure comparing WML detections at In Salah in the summer from a) RAOB data, b) AIRS+AMSU temperature and humidity products, c) ERA-I reanalysis, d) MERRA-2 reanalysis. The orange shading in each panel is the first WML detected in that profile. Wind barbs are in m/s.

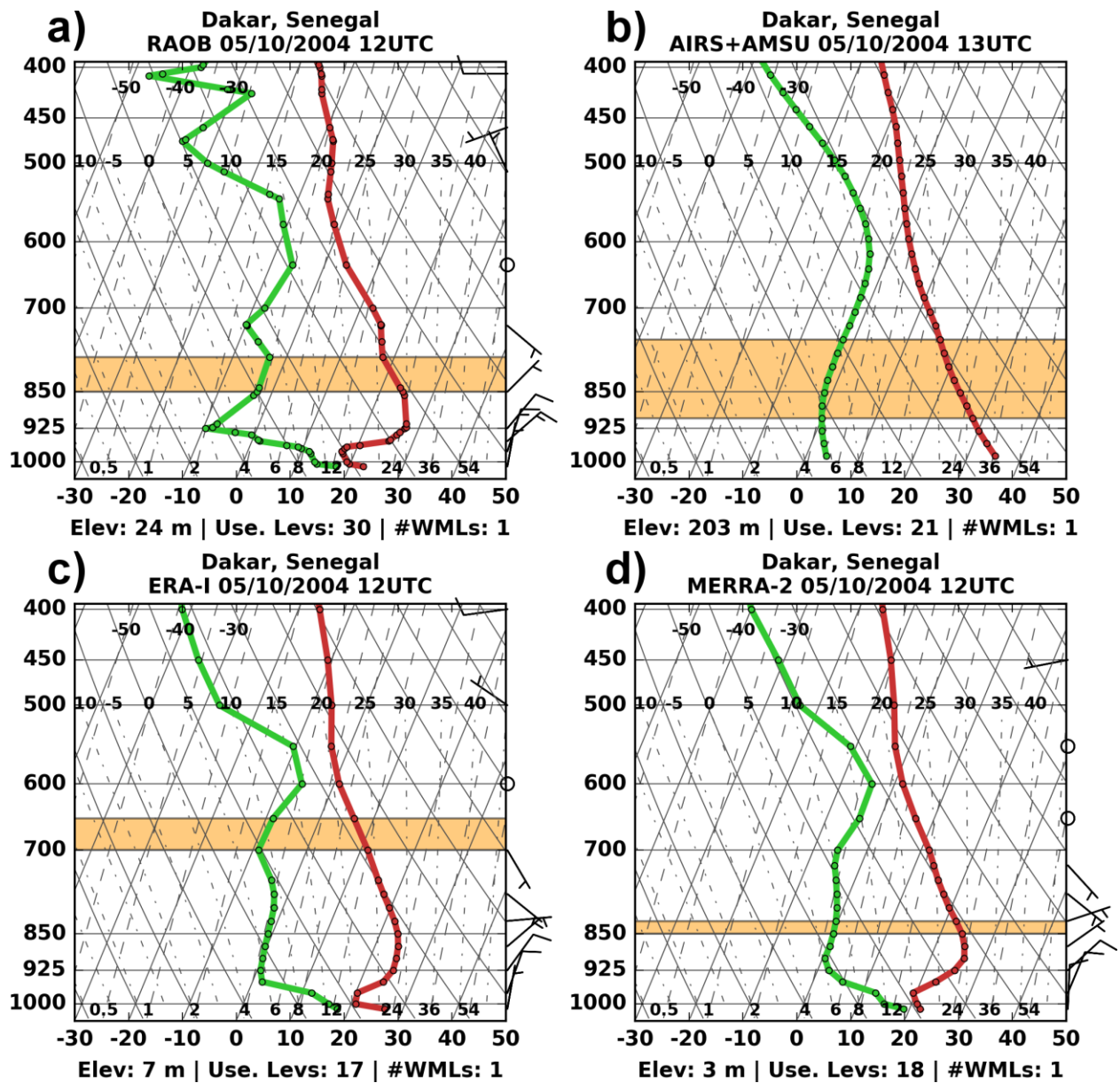


Figure 9. As in Figure 8, except for springtime in Dakar.

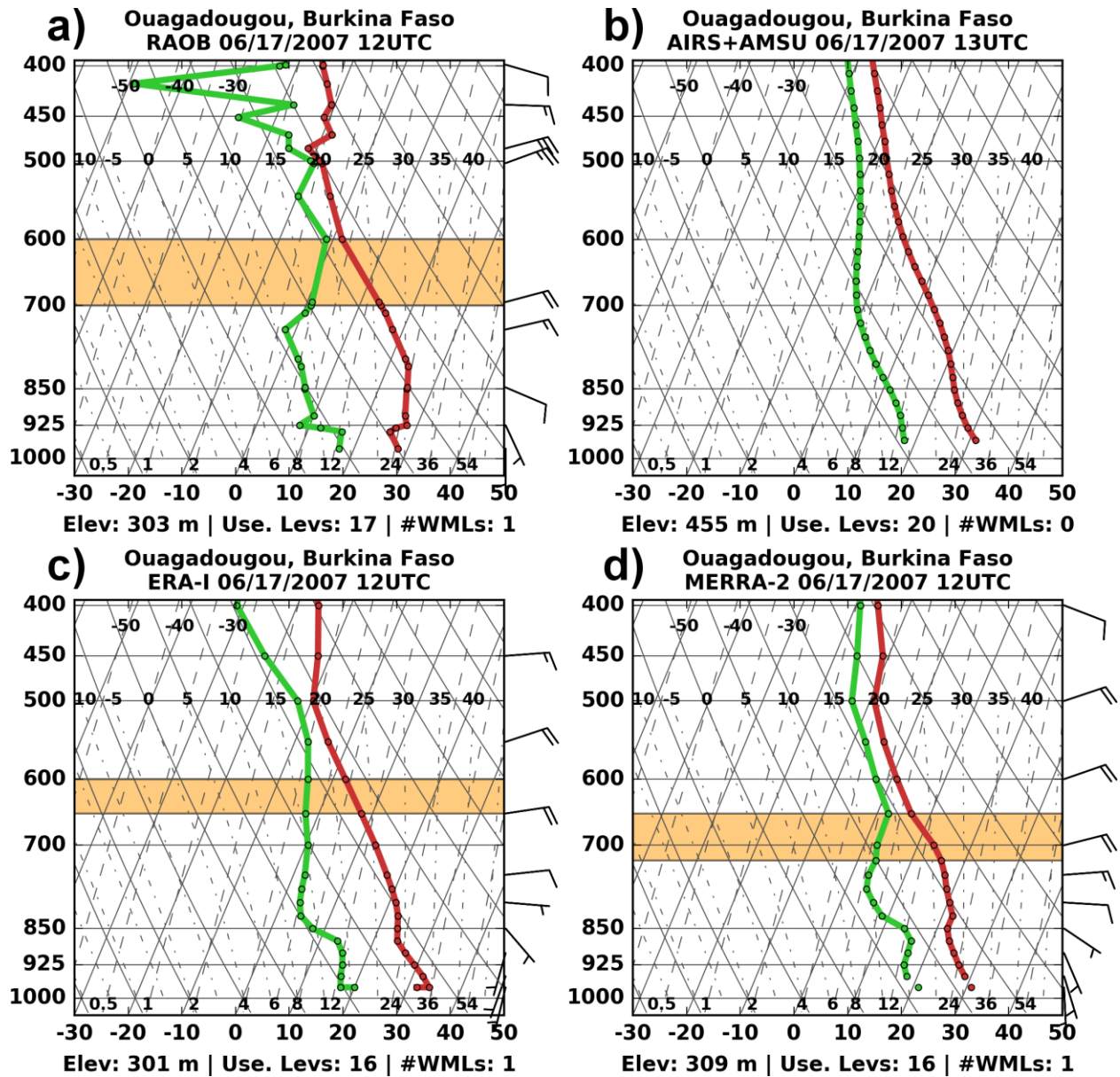


Figure 10. As in Figure 8, except for springtime in Ouagadougou.

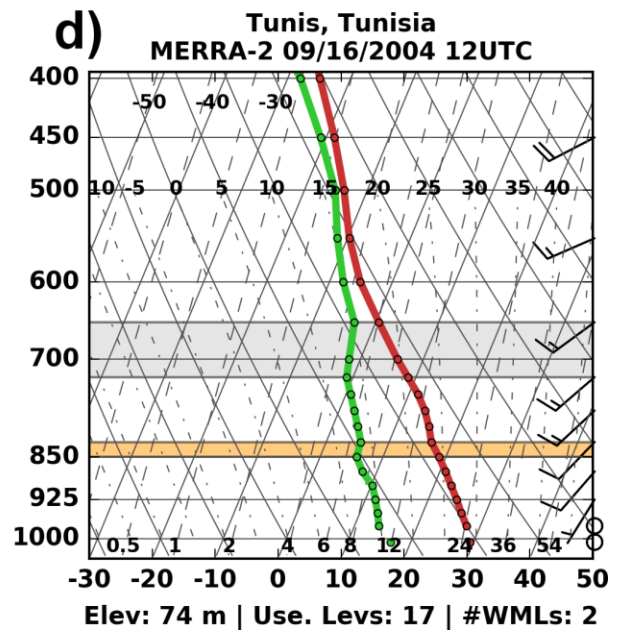
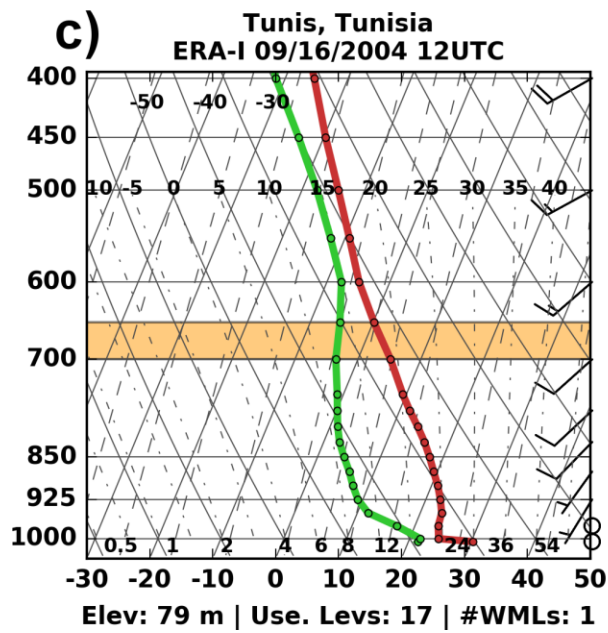
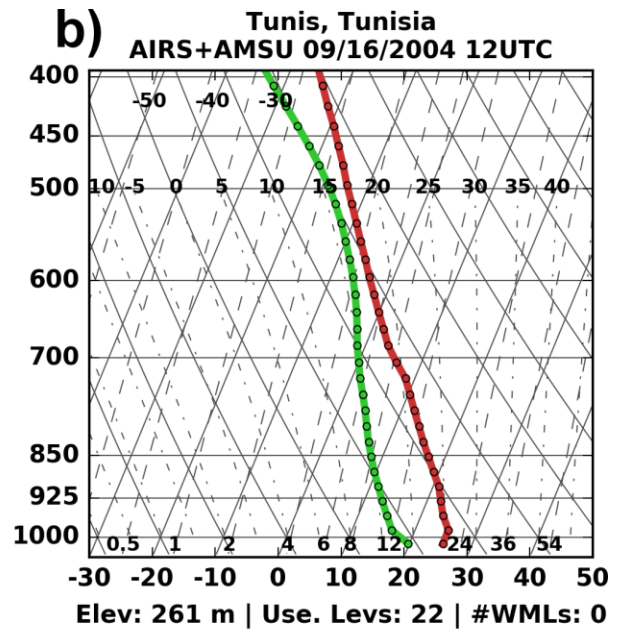
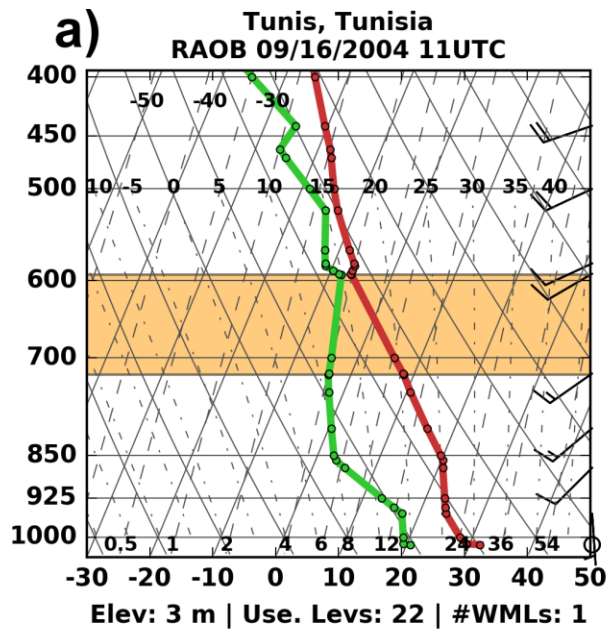


Figure 11. As in Figure 8, except for summertime in Tunis.