

Comparison of Convective Parameters Derived from ERA5 and MERRA-2 with Rawinsonde Data over Europe and North America

MATEUSZ TASZAREK,^{a,b,c} NATALIA PILGUJ,^d JOHN T. ALLEN,^e VICTOR GENSINI,^f HAROLD E. BROOKS,^{b,g} AND PIOTR SZUSTER^{h,i}

^a Department of Meteorology and Climatology, Adam Mickiewicz University, Poznań, Poland

^b National Severe Storms Laboratory, Norman, Oklahoma

^c Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

^d Department of Climatology and Atmosphere Protection, University of Wrocław, Wrocław, Poland

^e Central Michigan University, Mount Pleasant, Michigan

^f Department of Geographic and Atmospheric Sciences, Northern Illinois University, DeKalb, Illinois

^g School of Meteorology, University of Oklahoma, Norman, Oklahoma

^h Department of Computer Science, Cracow University of Technology, Krakow, Poland

ⁱ AGH Cracow University of Science and Technology, Krakow, Poland

(Manuscript received 25 June 2020, in final form 7 December 2020)

ABSTRACT: In this study we compared 3.7 million rawinsonde observations from 232 stations over Europe and North America with proximal vertical profiles from ERA5 and MERRA-2 to examine how well reanalysis depicts observed convective parameters. Larger differences between soundings and reanalysis are found for thermodynamic theoretical parcel parameters, low-level lapse rates, and low-level wind shear. In contrast, reanalysis best represents temperature and moisture variables, midtropospheric lapse rates, and mean wind. Both reanalyses underestimate CAPE, low-level moisture, and wind shear, particularly when considering extreme values. Overestimation is observed for low-level lapse rates, midtropospheric moisture, and the level of free convection. Mixed-layer parcels have overall better accuracy when compared to most-unstable parcels, especially considering convective inhibition and lifted condensation level. Mean absolute error for both reanalyses has been steadily decreasing over the last 39 years for almost every analyzed variable. Compared to MERRA-2, ERA5 has higher correlations and lower mean absolute errors. MERRA-2 is typically drier and less unstable over central Europe and the Balkans, with the opposite pattern over western Russia. Both reanalyses underestimate CAPE and CIN over the Great Plains. Reanalyses are more reliable for lower elevation stations and struggle along boundaries such as coastal zones and mountains. Based on the results from this and prior studies we suggest that ERA5 is likely one of the most reliable available reanalyses for exploration of convective environments, mainly due to its improved resolution. For future studies we also recommend that computation of convective variables should use model levels that provide more accurate sampling of the boundary layer conditions compared to less numerous pressure levels.


KEYWORDS: Severe storms; Thunderstorms; Climatology; Soundings; Forecast verification/skill; Reanalysis data

1. Introduction

Understanding the spatiotemporal climatology of severe convective storm environments has primarily been achieved through atmospheric reanalyses (Brooks et al. 2003, 2007; Gensini and Ashley 2011; Grams et al. 2012; Tippet et al. 2012; Allen et al. 2015; Rädler et al. 2018; Chen et al. 2020; Taszarek et al. 2020). Reanalysis products provide a four-dimensional (x, y, z, t) best guess of the atmospheric state, incorporating rawinsonde measurements, satellite retrievals, surface observations, and other sources assimilated with short-term forecasts from a frozen state model. Data from atmospheric reanalyses have a finer horizontal grid spacing and temporal frequency when compared to sparsely distributed rawinsonde observations (Potvin et al. 2010) and thus offer significant advantages for understanding convective

profiles and the past atmospheric state. This is especially important for parameters that are sensitive to the small spatial and temporal variations in atmospheric profile, such as storm-relative helicity (SRH) or CAPE (Markowski et al. 1998; Thompson et al. 2003). The gridded nature of reanalysis data also makes for an excellent source of initial and lateral boundary condition information for use in regional weather and climate modeling (Trapp et al. 2007, 2011; Robinson et al. 2013; Gensini and Mote 2014). However, given the limitations of the advanced assimilation and modeling process, it is unclear whether all reanalyses are capable of capturing reliable thermodynamic and kinematic profiles critical to understanding local convective environments. Thus, evaluation of reanalysis products is needed to provide researchers guidance as to the strengths and limitations of these datasets.

Although reanalyses have been commonly used in a large number of studies, only a limited number of elaborations

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Mateusz Taszarek, mateusz.taszarek@amu.edu.pl



This article is licensed under a Creative Commons Attribution 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

DOI: 10.1175/JCLI-D-20-0484.1

evaluated the accuracy of reanalysis for studying convective environments (Lee 2002; Allen and Karoly 2014; Gensini et al. 2014; Taszarek et al. 2018; King and Kennedy 2019; Li et al. 2020). Lee (2002) showed that NCEP–NCAR proximity profiles even at coarse resolutions provide a reasonable approximation of the convective environment when compared with collocated soundings. Kinematic variables were found to be best represented by reanalysis whereas thermodynamic parameters contained large differences that resulted from errors in low-level moisture fields. Thompson et al. (2003) compared 149 observed soundings with collocated Rapid Update Cycle 2 (RUC2; Benjamin et al. 2004) profiles in regional supercell environments, which revealed the largest errors in temperature and mixing ratios near the surface. Allen and Karoly (2014) examined ECMWF ERA-Interim reanalysis and compared proximity profiles with 3700 soundings over Australia. They found similar weaknesses in low-level moisture fields to Lee (2002) and noted issues in low-level wind profiles. At higher resolutions, the North American Regional Reanalysis (NARR) had higher correlations but similar limitations, along with a sensitivity of profiles to shallow convective schemes (Gensini et al. 2014). Taszarek et al. (2018) examined over 1 million soundings across Europe in comparison to ERA-Interim and found that vertical wind shear was consistently underestimated, especially in the 0–1-km layer. In addition, they also found that biases in CAPE demonstrate regional patterns with underestimation over eastern Europe and overestimation across Mediterranean area. Intercomparison by King and Kennedy (2019) evaluated profiles from several reanalyses against RUC2 proximity soundings considering only supercell thunderstorm events. Each of the evaluated reanalyses showed strengths and weaknesses, but broadly, they were found to adequately replicate the background convective environment for severe weather climatological research.

Although convective parameters from the recently released ERA5 have not been comprehensively evaluated yet, a few studies have already confirmed this reanalysis seems to perform well. Tarek et al. (2020) assessed the use of ERA5 as a potential reference dataset for hydrological modeling using two hydrological models over 3138 North American catchments. They concluded that ERA5 temperature and precipitation biases are reduced consistently, and provide the accuracy needed for hydrologic modeling, as compared to ERA-Interim. There is also a significant improvement from ERA-Interim to ERA5 in surface fields as compared to remotely sensed and in situ observations (Balsamo et al. 2018). Lei et al. (2020) concluded that ERA5 has a considerable improvement compared to ERA-Interim in sampling cloud cover over the Tibetan Plateau and eastern China. The performance of ERA5 and MERRA-2 in modeling wind generation in five different countries has also been tested by Olauson (2018). Comparison to measured wind speed values suggests that ERA5 performs better than MERRA-2 with higher correlations, 20% lower mean absolute error, and a better representation of diurnal cycle. Li et al. (2020) compared climatological aspects of a few sounding-derived variables with Community Atmosphere Model version 6 (CAM6) and ERA5 over a period of 1980–2014 and concluded that both CAM6 and ERA5 successfully reproduced environmental climatology, but with strong spatiotemporal correlations, and overall lower biases

in ERA5. Coffey et al. (2020) also concluded that ERA5 was able to represent tornado environments across the United States with a similar skill to the RUC2. A detailed evaluation of ERA5 accuracy has been performed by Hersbach et al. (2020). Comparison of ERA5 with gauge measurements, SYNOP, METAR, radar, upper-air, and buoy data yielded improved correlation to tropospheric temperature, wind, humidity, ocean wave height, and precipitation versus prior ECMWF reanalyses. These results suggest that the enhanced temporal and spatial resolution of ERA5 allows for an excellent representation of weather systems; however, these have yet to be explored for convective environments.

To assess which parameters should be evaluated, earlier studies using representative upper-air soundings and prior climatologies provide a range of candidates. Past research has examined environmental conditions favorable for severe convective storms by mining historical observed upper-air soundings representative of the near-storm and/or preconvective environment. Rasmussen and Blanchard (1998) developed a baseline climatology of environmental parameters favorable for the development of deep moist convection. At climatological aggregations, CAPE and vertical wind shear were found to be prudent discriminators between significant severe and severe thunderstorms, a result later confirmed by Brooks et al. (2003), Thompson et al. (2003), Craven and Brooks (2004), Thompson et al. (2007), and Trapp et al. (2007) for the United States; by Allen et al. (2011) for Australia; and by Púčik et al. (2015) and Taszarek et al. (2017, 2020) for Europe. Research on thunderstorms producing tornadoes yielded the additional importance of lifting condensation level (LCL) and low-level wind shear (e.g., Brooks et al. 2003; Rasmussen 2003; Thompson et al. 2003; Craven and Brooks 2004; Groenemeijer and van Delden 2007; Thompson et al. 2007, 2012; Coffey et al. 2019, 2020; Coniglio and Parker 2020). These results suggest that in the context of tornadic storms, performance for CAPE, LCL, and low-level wind shear is important, while deep layer shear is also relevant for storm mode and severe hazards more generally.

Convective inhibition (CIN) has also been found useful for describing the potential for convective initiation with low (high) values being associated with higher (lower) chances for severe thunderstorm development (Rasmussen and Blanchard 1998; Bunkers et al. 2010; Westermayer et al. 2017; Hoogewind et al. 2017; Taszarek et al. 2020). Isobaric level data can also provide effective discriminants; for example, Grams et al. (2012) found that wind speeds at 500 and 850 hPa, along with 500-hPa geopotential or surface to 850-hPa dewpoint, are useful provided the storm mode is known. Other authors found that the combination of individual convective parameters into composite indices, such as the significant tornado parameter (STP; Thompson et al. 2003, 2007; Coffey et al. 2019, 2020), significant hail parameter (SHIP; NOAA Storm Prediction Center; <https://www.spc.noaa.gov/>), supercell composite parameter (SCP; Thompson et al. 2003, 2007; Groppe and Davenport 2018), energy helicity index (EHI; Hart and Korotky 1991; Davies 1993), or the product of CAPE and 0–6-km wind shear (Brooks et al. 2003; Allen et al. 2011; Taszarek et al. 2017, 2020; Rodríguez and Bech 2021) can be useful in recognizing and forecasting specific types of convective storms. The combinative

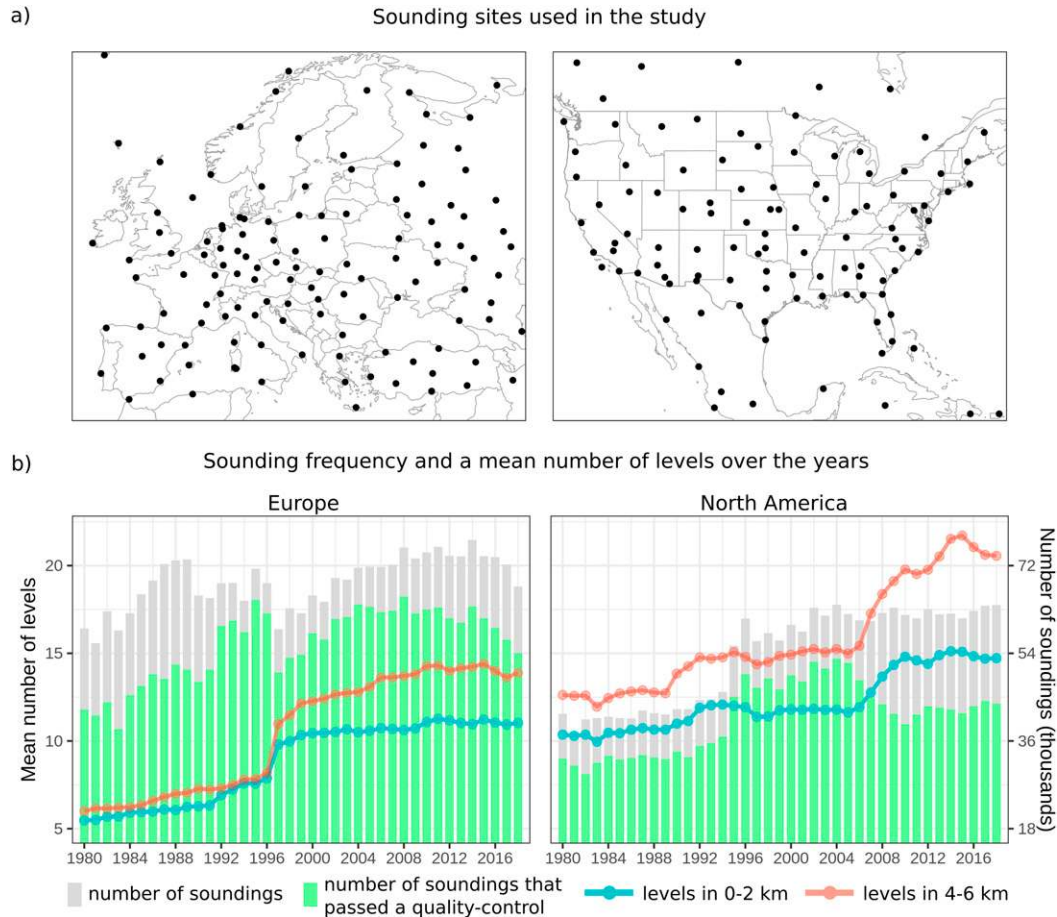


FIG. 1. (a) Sounding stations used in the study and (b) number of available sounding measurements over the years (including those that passed the quality-control phase) and a mean number of levels in 0–2- and 4–6-km layers.

effect of these parameters can have both desirable and undesirable properties (Doswell and Schultz 2006); given that their covariate nature multiplicatively magnifies the errors of individual components, these also warrant investigation.

The main motivation of this study is to evaluate convective parameters derived from hybrid-sigma model levels from ERA5 and MERRA-2 with a large number of high-quality sounding measurements and examine how well reanalysis can reconstruct convective climatologies. To better evaluate this skill, we use hybrid-sigma terrain-following model levels for computation of convective parameters instead of interpolated pressure levels. From the initial database of 5.1 million soundings derived from 232 European and North American stations from years 1980–2018, we use 3.7 million cases that meet strict quality-control criteria (Fig. 1; see appendix A). The main focus is on parameters and ingredients (Johns and Doswell 1992; Brooks et al. 2003; Doswell and Schultz 2006) supporting development of severe convective storms. For each sounding parameter we look at statistics that describe biases, variations in performance as a function of annual cycle, and changes in quality over the length of the dataset. We also present spatial climatologies of selected convective ingredients

to assess how they vary between the datasets. Results presented in this study may increase awareness of limitations associated with applying reanalysis datasets for modeling, case studies, and climate research. In addition, the ERA5 dataset evaluated in this study was previously used in Tazarek et al. (2020, 2021) for analyzing climatological aspects of severe convective storms and their trends across Europe and the United States.

2. Dataset and methodology

a. Sounding database and quality-control assumptions

Vertical atmospheric profiles were derived from the University of Wyoming upper-air observations database (<http://weather.uwyo.edu/upperair/>). For years 1980–2018 all available measurements were downloaded from 232 stations from Europe and North America that accounted for 5.1 million soundings. However, to ensure that biases obtained by comparing sounding and reanalysis data are not originating from poor-quality sounding data (appendix B), strict quality-control assumptions were applied. Each measurement was evaluated by 48 different

functions designed to detect various data issues within the profile. Focusing on the most important of these filters, as a first step we excluded all cases that had a first measurement higher than 10 m or a highest measurement lower than 6000 m (justified by a possibility of computation 0–6-km wind shear and mean wind) and fewer than 10 levels over the depth of the entire sounding. Up to 6000 m all levels needed to have complete measurements of u and v wind, temperature, moisture, height, and pressure. In higher levels, only missing moisture measurements were deemed acceptable (a common problem for soundings in the 1980s and 1990s) as it was nonessential for parcel parameter computations. If any other variable (height, pressure, temperature, u and v wind) was missing, then that level was removed from the sounding. If a profile had CAPE exceeding 0 J kg^{-1} , this required that a sounding have a definable equilibrium level (EL); that is, the CAPE calculation was completed.

Other evaluations were aimed at detecting unrealistic changes in temperature, moisture, and the wind profile (appendix B). Differences between sounding levels with temperature gradients exceeding 12 K km^{-1} between 0 and 2000 m and 10 K km^{-1} above 2000 m, were removed from the sounding to correct for potentially erroneous levels within a sounding. Levels consisting of saturated air (dewpoint depressions lower than 1°C) while also having dry-adiabatic gradients (at least three consecutive measurements) were also removed (e.g., appendix B; Stuttgart, sounding E). Levels in the lowest 500 m with temperature gradients exceeding 12 K km^{-1} (e.g., appendix B; Erzurum, sounding B) were corrected to 12 K km^{-1} by adjusting temperature values (as this issue may significantly affect most-unstable parcel calculations). Profiles with unrealistic changes in boundary layer moisture profile, such as spikes in dewpoints (e.g., appendix B; Tucson, sounding C) or reaching unreliable surface θ_e values relative to local climatology (e.g., appendix B; Brindisi and Munich, soundings D and G), were excluded as well. Soundings with a mean dewpoint depression (considering the entire profile) lower than 2°C were also removed (e.g., appendix B; Edwards, sounding I). For selected temperature, moisture, and wind parameters, we removed all soundings that had values higher than the 0.99999th or lower than the 0.00001th percentile of the dataset distribution. Although we are aware that this technique may remove some portion of tail distribution of good-quality profiles, a random manual check of several dozen examples indicated that the majority of these profiles had errors that were not detected in the preliminary quality-control phase. As it is not possible to manually investigate each profile in the dataset, we are aware that no perfect computational solution will provide a database free from erroneous soundings. However, the approach used here ensures that the data are of high quality, and the small fraction of poor-quality soundings should have a negligible influence on the results.

In the initial database, a number of soundings had a CAPE exceeding $10\,000 \text{ J kg}^{-1}$, mostly due to major errors in the profile such as spikes in low-level temperature and/or moisture (e.g., appendix B; Edwards, sounding I). After the quality-control phase, the largest CAPE sounding for the United States was 9413 J kg^{-1} at 0000 UTC 3 July 1999 in North Platte, whereas for Europe it was 6216 J kg^{-1} at 0000 UTC 13 September 2008 in Trapani (Fig. 2). Both of these soundings do not feature

any major errors, which partially confirm our filters to be successful in removing poor-quality profiles. The ERA5-equivalent grid for the North Platte sounding indicated CAPE of 5639 J kg^{-1} and for MERRA-2 5226 J kg^{-1} , while for Trapani CAPE was 4698 J kg^{-1} for ERA5 and 3610 J kg^{-1} for MERRA-2. This comparison further highlights the poor ability of reanalyses to represent the localized rare and extreme convective environments, especially those characterized by high CAPE and/or vertical wind shear.

In total, 3.7 million cases, including 1.6 million for the United States and 2.1 million for Europe, met quality-control assumptions. Measurements from 0000 and 1200 UTC were the most frequent (92% of all profiles); 0600 and 1800 UTC had a share of 7.5%. Soundings performed at 0300, 0900, 1500, and 2100 UTC were sporadic and accounted only for 0.5% of the overall sample (Table 1).

b. ERA5 and MERRA-2 reanalyses

ERA5 is the fifth-generation ECMWF global atmospheric reanalysis (Hersbach et al. 2020) openly available through the Copernicus Climate Change Service (2017). It has a $0.25^\circ \times 0.25^\circ$ horizontal grid spacing with hourly temporal steps and 137 terrain-following hybrid model vertical levels, contrasting many prior reanalyses with lower spatial, temporal, and vertical resolution. One of the significant advantages of ERA5 is the 28 levels through the 0–2-km layer at hourly steps, allowing the possibility to more accurately sample boundary layer conditions and study diurnal cycles of convective environments in ways not possible before (e.g., compared to 6-h steps in ERA-Interim, NCEP–NCAR, and 20CR, and 3-h steps in NARR and MERRA-2; Mesinger et al. 2006; Saha et al. 2010; Dee et al. 2011; Compo et al. 2011; Gelaro et al. 2017). The ERA5 analysis is computed using 4D-VAR data assimilation in CY41R2 of the Integrated Forecast System. Further details on the parameterizations are available in the Climate Data Store (<https://cds.climate.copernicus.eu/>).

MERRA-2 is the global reanalysis developed by NASA's Global Modeling and Assimilation Office based on a period of regular conventional and satellite observations era starting in 1980. This reanalysis has a $0.5^\circ \times 0.625^\circ$ horizontal grid spacing with 3-h steps and 72 terrain-following hybrid model levels including 14 layers in the 0–2-km layer. The analysis is computed using a 3D-VAR algorithm based on the GSI under 6-h update cycle and first guess at the appropriate time (FGAT) procedure. Further details on this reanalysis are available in Gelaro et al. (2017), Randles et al. (2017), and Bosilovich et al. (2017).

For each sounding used in this study, we collocated corresponding ERA5 and MERRA-2 atmospheric profiles by temporal and spatial proximity (the same time step in the closest reanalysis grid to the sounding site). To ensure soundings and reanalysis profiles were comparable, 10 m was used as the first level for each dataset.

c. Parameters evaluated

The choice of parameters evaluated in this study was based on indices commonly applied in studies focusing on severe thunderstorm environments and their corresponding climatologies

Soundings with the highest MU and ML CAPE across the United States and Europe

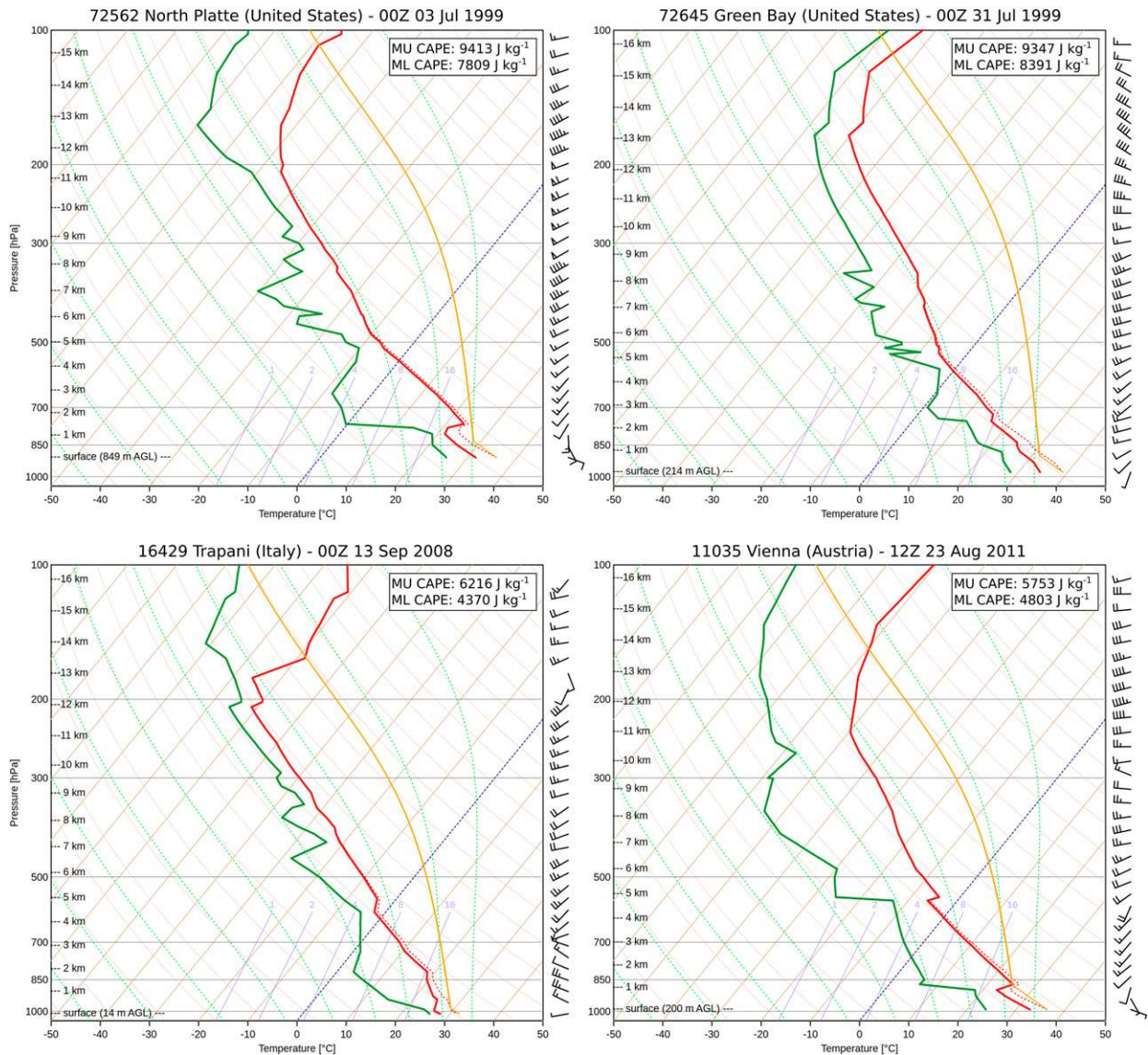


FIG. 2. Soundings with the highest measured (left) most-unstable and (right) mixed-layer CAPE for (top) the United States and (bottom) Europe that passed a quality-control phase in our study (1980–2018). Most-unstable parcel curves are indicated by orange lines, temperatures by red lines, and dewpoints by green lines. Generated with thundeR R language package.

over both Europe and the United States (Rasmussen and Blanchard 1998; Thompson et al. 2003; Craven and Brooks 2004; Groenemeijer and van Delden 2007; Kaltenböck et al. 2009; Gensini and Ashley 2011; Thompson et al. 2012, 2013; Allen et al. 2015; Tang et al. 2019; Liu et al. 2020; Taszarek et al. 2020, 2021). For each sounding and reanalysis-located profile, we used temperature, humidity, pressure, height, and u and v wind. Parcel parameter calculations were performed for two versions. The first used a mean mixed-layer parcel from the 0–500-m layer (ML) to account for shallower moisture layers in Europe (Taszarek et al. 2020,2021), while

the second was a most-unstable parcel with the highest θ_e over the 0–3000-m layer (MU). For both approaches, a virtual temperature correction (Doswell and Rasmussen 1994) was applied. For computations of 0–1- and 0–3-km SRH we used the right-moving storm motion following Bunkers et al. (2000). Wind shear parameters were computed by taking into account a difference between wind vectors at 10 m and the wind profile interpolated to the desired height (1, 3, or 6 km). To compute SCP and STP, we used the most recent formulas from Gropp and Davenport (2018) and Coffey et al. (2019). For computing the significant hail parameter (SHIP),

TABLE 1. Number of soundings (after quality-control phase) used in the study.

Domain	Number of soundings for specific hour								Total
	0000 UTC	0300 UTC	0600 UTC	0900 UTC	1200 UTC	1500 UTC	1800 UTC	2100 UTC	
North America	756 218	1278	9549	3156	791 577	5487	25 602	3928	1 596 795
Europe	989 331	617	142 332	1591	938 300	704	104 966	4114	2 181 955

the formula used was as described by the NOAA Storm Prediction Center (<https://www.spc.noaa.gov>), also available in appendix A of Taszarek et al. (2020). In addition to severe weather parameters, we also evaluated temperature, moisture, and wind speed at 10 m and at standard pressure levels throughout the profile (500, 700, and 850 hPa) as these are often used in the operational forecasting and climatological studies (Grams et al. 2012). It is also important to highlight that exactly the same computational script was used to derive convective parameters from sounding, ERA5, and MERRA-2 vertical profiles to ensure this did not cause differences in the calculated parameters. The aforementioned script has been under development over the last 4 years and allows us to compute over 130 variables in less than 0.02 s per profile. This kind of efficiency is crucial in processing large reanalysis datasets with a reasonable amount of time [the same script was also applied in ERA5 climatologies in Taszarek et al. (2020, 2021)]. The authors plan to make this script publicly available in 2021 as an R language CRAN library called “thunder.”

d. Limitations

The most significant limitation of the approach used herein is that some of the soundings that passed quality-control check may still contain issues with the vertical profile of temperature, humidity, or wind. We also note that the quality of sounding data has changed throughout time, which may influence results; however, similar issues are likely to impact the reanalysis. In such cases it is uncertain whether soundings evaluate quality of the reanalysis or rather reanalysis is testing the credibility of sounding measurements. To address this potential limitation, we focus part of our analysis on showing how the mean absolute error (MAE) and correlation coefficient have changed over the years. The increased vertical resolution of soundings is also an important factor, with the number of available sounding levels in the lower troposphere and midtroposphere constantly increasing over time, with major changes in 1996 for Europe and 2006 for the United States (Fig. 1b).

Profiles may also have issues due to the grid size of the reanalysis ($0.25^\circ \times 0.25^\circ$ in ERA5 and $0.5^\circ \times 0.625^\circ$ in MERRA-2), which average the vertical profile of the atmosphere over a certain area. This can produce differences as compared to sounding profiles that are not a reflection of reanalysis performance but are rather due to the grid size poorly representing regional processes. This may be especially evident over areas with complex orography and strong horizontal gradients such as coastal zones, islands, mountains, or valleys; thus, we also explore whether reanalysis performance varies for stations located within these sharp gradients.

Caveats regarding numerical formulation of reanalysis may also influence the results. The horizontal grid spacing of ERA5 and MERRA-2 requires convective parameterization schemes that may lead to errors in the vertical profile of temperature and moisture, potentially creating biases in thermodynamic parameters like CAPE or CIN. Similarly, low-level conditions may be impacted by the presence of boundary layer parameterization schemes. The spatial location of profile extraction is also of great importance (e.g., before or after a cold front) and can lead to dramatic changes in sensible conditions over small spatiotemporal scales. Differences in the reanalysis timing of such boundaries may be responsible for large differences in thermodynamic or kinematic variables when compared to observations. The temporal interval of the ERA5 and MERRA-2 as well as data assimilation windows may lead to slower or faster development of convection in the model and subsequent contamination of the profile (Allen and Karoly 2014; King and Kennedy 2019). Reanalysis parameterization schemes and data assimilation issues have been discussed as likely candidates responsible for biases in the low-level thermodynamic fields (Gensini et al. 2014; Allen and Karoly 2014; Tippett et al. 2014; Taszarek et al. 2018; King and Kennedy 2019). Although it is beyond the scope of this paper to diagnose the source of the errors, the aforementioned issues provide a caveat on the presented results.

3. Results

a. Parcel parameters

Thermodynamic parcel-related parameters are better represented by ERA5, with higher correlations and lower MAE compared to MERRA-2 (Fig. 3, Table 2). As CAPE is highly spatially variable, this partly accounts for the discrepancy between soundings and reanalysis (Figs. 3 and 4). However, this discrepancy is generally lower than in previous studies (Allen and Karoly 2014; Gensini et al. 2014; Taszarek et al. 2018). Among parcel types, ML CAPE has slightly better performance metrics compared to MU CAPE (Table 2). MAE is lower for ML CAPE (283 and 331 J kg^{-1} for ERA5 and MERRA-2), compared to the MU version (378 and 414 J kg^{-1}), which can be explained by climatologically higher values of the latter. Both reanalyses tend to underestimate CAPE with respect to observed profiles as reflected by the negative mean error (ME) and distribution of percentiles, especially when considering extreme values (Fig. 4).

Overall parcel parameters had some of the lower correlations for all parameters, with ML parcels generally providing higher correlations but little improvement to MAE. The lifted index (LI) has a generally high correlation but, similar to CAPE, both reanalyses underestimate parcel buoyancy with

Comparison of selected thermodynamic variables with collocated profiles from ERA5 and MERRA2

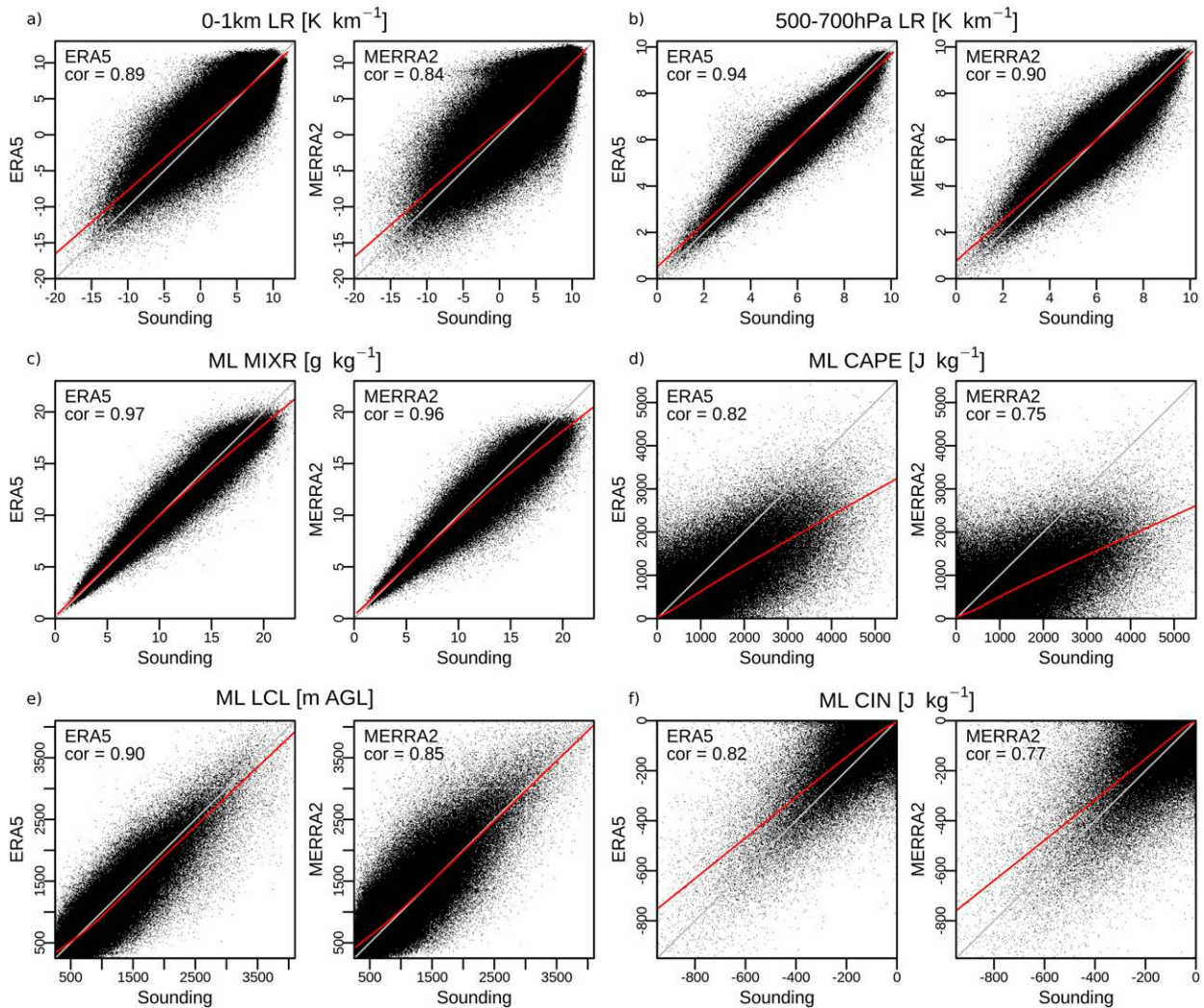


FIG. 3. Comparison of sounding observations with collocated ERA5 and MERRA-2 proximity profiles for (a) 0–1-km lapse rate, (b) 500–700-hPa lapse rate, (c) mixed-layer mixing ratio, (d) mixed-layer CAPE, (e) mixed-layer lifted condensation level, and (f) mixed-layer convective inhibition. The gray line denotes a one-to-one ratio. The red line denotes locally estimated scatterplot smoothing (LOESS). Value in the top-left corner of each plot denotes Pearson correlation coefficient.

MAE of around 1.5°C (Table 2). CIN correlations for ML parcels are generally higher (0.82 for ERA5 and 0.77 for MERRA-2) compared to MU (0.71 and 0.63), while the MAE for both parcel types is around 30 J kg^{-1} . Better overall performance of ML compared to MU parcels is also observed within the lifted condensation level (LCL), level of free convection (LFC), and equilibrium level (EL), where in all of them ERA5 is superior to MERRA-2 in terms of correlations and MAE (Fig. 3, Table 2). From all parameters in this group, LFC and LCL for the MU parcel are characterized by the lowest correlations. Although both reanalyses tend to overestimate LFC height, differences are found for LCL, which is slightly underestimated in ERA5 (ME is -24 and -30 m for MU and ML parcels) and overestimated in MERRA-2 (86 and 78 m for

MU and ML). A similar pattern can also be seen in the distribution of percentiles (Fig. 4).

b. Moisture parameters

Atmospheric moisture is also better represented in the reanalysis by ML compared to MU parcels, although for both parcel types correlations are high. The ML mixing ratio (MIXR) has a correlation of 0.97 and 0.96, respectively, for ERA5 and MERRA-2, while for MU MIXR the correlations are smaller (0.94 and 0.91; Table 2). Quantifying MAE for MU MIXR also shows larger errors as compared to ML MIXR (0.74 and 0.94 g kg^{-1} for ERA5 and MERRA-2, respectively). Both reanalyses perform equally for less extreme values of ML MIXR, but ERA5 has better accuracy for larger values as

TABLE 2. Selected statistical indices comparing sounding-derived parameters with collocated ERA5 and MERRA-2 profiles for parcel, moisture, and temperature parameters. ML-based parameters are calculated only for ML CAPE $> 0 \text{ J kg}^{-1}$, and MU-based parameters are calculated only for MU CAPE $> 0 \text{ J kg}^{-1}$. Numbers in italics indicate parameters where the Pearson correlation coefficient in ERA5 is higher compared to MERRA-2 by at least 0.05. The units for mean error, mean absolute error, and root-mean-square error are the same as for the parameters. An asterisk indicates that only unstable environments are considered in the computations.

	Pearson correlation coefficient (R)		Spearman correlation coefficient (R_s)		Mean error (ME)		Mean absolute error (MAE)		Root-mean-square error (RMSE)	
	ERA5	MERRA-2	ERA5	MERRA-2	ERA5	MERRA-2	ERA5	MERRA-2	ERA5	MERRA-2
Parcel parameters										
MU CAPE (J kg^{-1})*	0.82	0.77	0.78	0.72	-246.57	-231.09	378.16	414.16	650.49	705.76
ML CAPE (J kg^{-1})*	0.82	0.75	0.80	0.76	-149.63	-203.71	283.22	331.03	519.63	606.71
MU LI ($^{\circ}\text{C}$)*	0.91	0.87	0.89	0.85	0.65	0.42	1.34	1.51	1.75	1.96
ML LI ($^{\circ}\text{C}$)*	0.96	0.94	0.93	0.90	0.28	0.38	1.17	1.39	1.53	1.82
MU CIN (J kg^{-1})*	0.71	0.63	0.58	0.51	-2.61	0.61	28.93	32.54	59.23	68.18
ML CIN (J kg^{-1})*	0.82	0.77	0.70	0.67	5.39	4.56	26.97	31.26	51.65	61.20
MU LCL (m AGL)*	0.69	0.62	0.68	0.61	-24.64	86.56	375.33	451.42	572.11	664.46
ML LCL (m AGL)*	0.90	0.85	0.86	0.81	-30.92	78.86	170.43	224.17	233.59	306.19
MU LFC (m AGL)*	0.68	0.63	0.68	0.64	118.02	90.77	561.30	607.99	861.26	933.29
ML LFC (m AGL)*	0.69	0.66	0.69	0.68	21.83	14.51	530.67	565.60	812.50	881.73
MU EL (m AGL)*	0.89	0.85	0.88	0.85	-699.45	-439.79	1321.67	1445.41	1985.33	2156.08
ML EL (m AGL)*	0.91	0.88	0.92	0.89	-385.99	-503.28	1179.80	1414.45	1837.33	2151.37
Moisture parameters										
MU MIXR (g kg^{-1})*	0.94	0.91	0.93	0.91	-0.37	-0.51	1.16	1.36	1.57	1.85
ML MIXR (g kg^{-1})*	0.97	0.96	0.98	0.97	-0.18	-0.46	0.74	0.94	1.02	1.32
PWATER (mm)	0.98	0.97	0.98	0.97	-0.36	0.08	1.40	1.83	2.13	2.68
10-m dewpoint ($^{\circ}\text{C}$)	0.97	0.96	0.97	0.96	-0.09	-0.04	1.51	1.86	2.10	2.56
850-hPa dewpoint ($^{\circ}\text{C}$)	0.94	0.91	0.96	0.93	-0.06	0.38	2.23	2.75	3.57	4.28
700-hPa dewpoint ($^{\circ}\text{C}$)	0.91	0.87	0.93	0.90	-0.07	0.28	3.17	3.90	4.96	5.87
500-hPa dewpoint ($^{\circ}\text{C}$)	0.86	0.81	0.87	0.82	0.34	0.46	3.70	4.55	5.51	6.56
Temperature parameters										
0-1-km LR (K km^{-1})	0.89	0.85	0.89	0.84	0.33	0.17	1.41	1.80	1.93	2.40
0-3-km LR (K km^{-1})	0.94	0.91	0.94	0.90	0.09	0.03	0.47	0.61	0.63	0.81
3-6-km LR (K km^{-1})	0.94	0.90	0.94	0.89	0.01	-0.02	0.22	0.29	0.30	0.38
500-700-hPa LR (K km^{-1})	0.94	0.90	0.93	0.89	-0.01	-0.02	0.24	0.32	0.33	0.43
10-m temperature ($^{\circ}\text{C}$)	0.98	0.98	0.98	0.98	-0.03	-0.20	1.36	1.74	1.85	2.30
850-hPa temperature ($^{\circ}\text{C}$)	0.99	0.99	0.99	0.99	-0.09	-0.07	0.61	0.78	0.85	1.07
700-hPa temperature ($^{\circ}\text{C}$)	0.99	0.99	0.99	0.99	-0.06	0.01	0.53	0.66	0.72	0.89
500-hPa temperature ($^{\circ}\text{C}$)	0.99	0.99	0.99	0.99	-0.03	0.04	0.47	0.61	0.67	0.83

Distribution of selected thermodynamic variables from soundings and collocated ERA5 and MERRA2 profiles

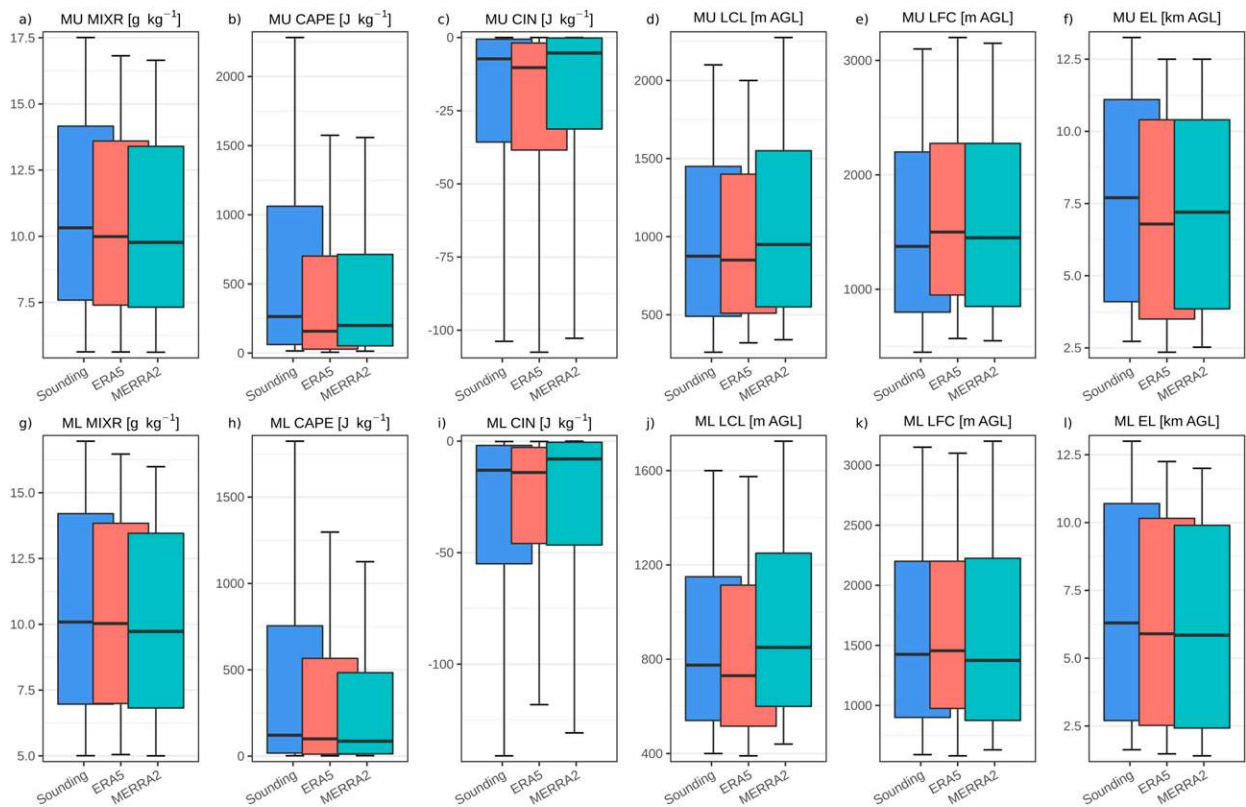


FIG. 4. Box-and-whisker plots for soundings (blue), ERA5-collocated profiles (red), and MERRA-2-collocated profiles (turquoise) for (a) most-unstable mixing ratio, (b) most-unstable CAPE, (c) most-unstable convective inhibition, (d) most-unstable lifted condensation level, (e) most-unstable level of free convection, (f) most-unstable equilibrium level, (g) mixed-layer mixing ratio, (h) mixed-layer CAPE, (i) mixed-layer convective inhibition, (j) mixed-layer lifted condensation level, (k) mixed-layer level of free convection, and (l) mixed-layer equilibrium level. The median is represented as a horizontal line inside the box and the edges of the box represent the 25th and 75th percentiles, whereas whiskers represent the 10th and 90th percentiles.

indicated by the locally estimated scatterplot smoothing line [locally estimated scatterplot smoothing (LOESS) with a span of 0.75; Wickham 2016] in Fig. 3c. Box-and-whisker plots show similar distributions of 10th and 25th percentiles for reanalysis and observations, both for MU and ML MIXR (Fig. 4). Differences between analyzed datasets are seen for the high percentiles.

Precipitable water (PWATER) shows high correlations for ERA5 (0.98) and MERRA-2 (0.97). However, ERA5 has a smaller MAE of 1.4 mm compared to 1.8 mm in MERRA-2 (Table 2). Evaluating dewpoint at 10 m and pressure levels (850, 700, and 500 hPa), we note decreasing correlation and increasing MAE with height in both reanalyses. At 10-m dewpoint correlations are 0.97 and 0.96 (ERA5 and MERRA-2), whereas by 500 hPa the correlations drop to 0.86 and 0.81. For 10-m dewpoint, MAE is below 2°C (for both reanalyses), whereas for 500 hPa it increases to as much as 3.7°C for ERA5 and 4.5°C for MERRA-2 (Table 2).

c. Temperature parameters

Overall, temperature-related parameters are better performers than moisture and wind in the reanalyses (Figs. 3 and 5,

Table 2). Air temperatures at fixed levels (10 m; 850, 700, and 500 hPa) are well represented by both reanalyses with slightly better accuracy for ERA5 as evidenced by correlations (Table 2). Contrary to dewpoints, MAE for temperature is the highest near the surface (1.36°C for ERA5 and 1.74°C for MERRA-2) and decreases with height to 0.47°C for ERA5 and 0.61°C for MERRA-2 at 500 hPa. ME is close to zero for temperature variables. This suggests that one of the primary factors driving errors in instability is low-level moisture, which has higher uncertainty in reanalyses compared to temperatures.

In contrast with the increasing errors with height in moisture, the low-level (0–1 km) lapse rate is characterized by the smallest correlations (0.89 for ERA5 and 0.85 for MERRA-2) and MAE of 1.4 K km⁻¹ (ERA5) and 1.8 K km⁻¹ (MERRA-2). Correlations are consistent in height for lapse rates within 0–3 km, 3–6 km, and 500–700 hPa with 0.94 for ERA5 and 0.90 for MERRA-2. Mean errors are around zero while MAE ranges from 0.22 K km⁻¹ for the 3–6-km layer in ERA5 up to 0.61 K km⁻¹ for the 0–3-km layer in MERRA-2 (Table 2). As evidenced by LOESS in Fig. 3, both reanalyses have lower accuracy in sampling temperature inversions (negative values).

Distribution of selected kinematic variables from soundings and collocated ERA5 and MERRA2 profiles

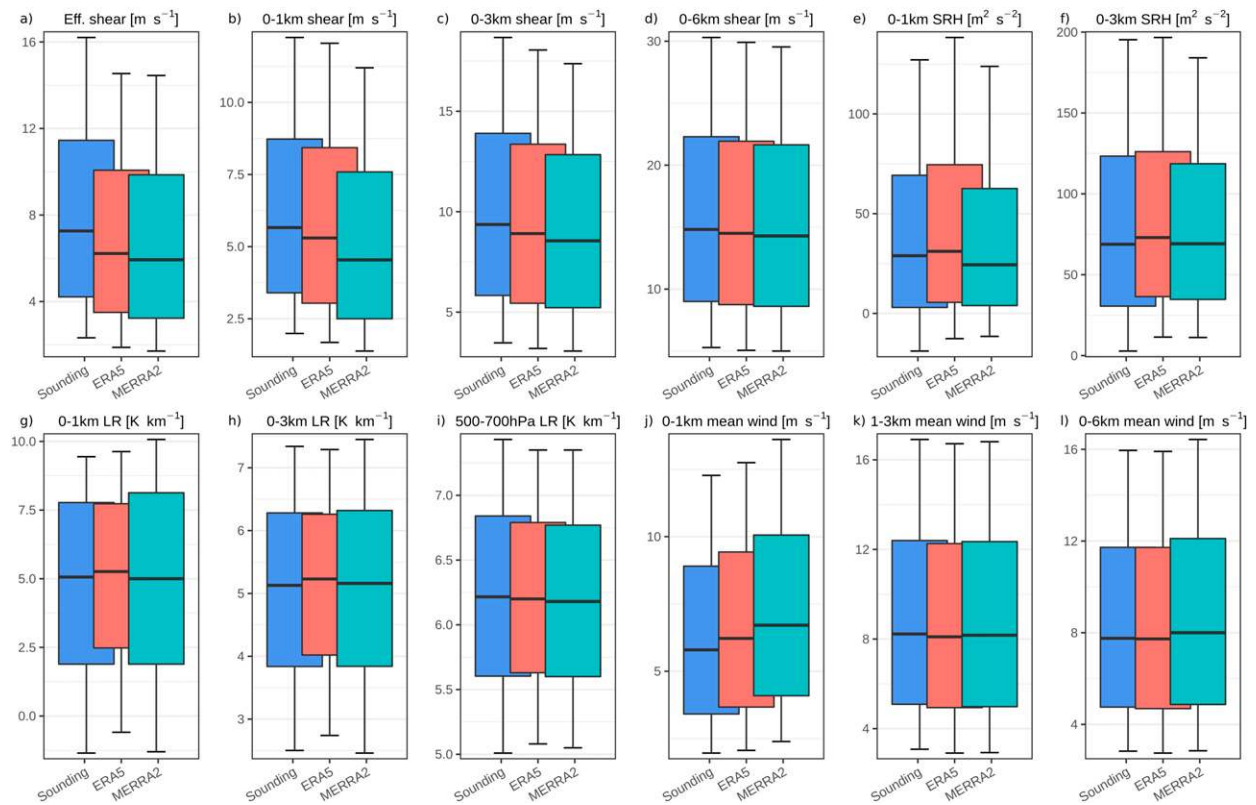


FIG. 5. As in Fig. 4. but for (a) effective shear, (b) 0–1-km wind shear, (c) 0–3-km wind shear, (d) 0–6-km wind shear, (e) 0–1-km storm-relative helicity, (f) 0–3-km storm-relative helicity, (g) 0–1-km lapse rate, (h) 0–3-km lapse rate, (i) 500–700-hPa lapse rate, (j) 0–1-km mean wind speed, (k) 1–3-km mean wind speed, and (l) 0–6-km mean wind speed.

Similar to other thermodynamic parameters MERRA-2 provides a greater spread of values for lapse rates compared to ERA5 (Fig. 3). Relatively good agreement is observed on the interquartile spread among all lapse rate variables (Fig. 5).

d. Wind parameters

Statistical analysis of wind parameters reveals that there is greater sensitivity to how layers and quantities are defined, with bulk measures being closest to those observed. The smallest correlation is for effective shear (0.74 for ERA5 and 0.70 MERRA-2; Table 3), reflecting the dependence of this parameter on the vertical structure of instability (Thompson et al. 2007), which can vary considerably between reanalysis and sounding, as also noted by King and Kennedy (2019). Underestimation of CAPE is related to a lower EL height and a shallower depth for the effective shear calculation (Table 2), which tend to reduce values of this parameter in a baroclinic environment with winds increasing with height. Effective shear, as well as wind shear of 0–1-, 0–3-, and 0–6 km, are underestimated by both reanalyses (Table 3, Figs. 5 and 6). Effective shear has MAE of 2.93 m s^{-1} for ERA5 and 3.25 m s^{-1} for MERRA-2 while fixed layer wind shear parameters have MAE of around 2 m s^{-1} for ERA5 and 2.5 m s^{-1} for MERRA-2. However, correlation increases along with shear layer height with

0–1 km having correlations of 0.81 and 0.76, and 0–6 km having correlations of 0.96 and 0.94 for ERA5 and MERRA-2, respectively. As indicated in the scatterplots in Fig. 6, low-level wind variables such as 0–1-km wind shear and 10-m wind exhibit greater statistical spread of values. This result is broadly consistent with other reanalyses that were compared with soundings (Allen and Karoly 2014; Gensini et al. 2014; Taszarek et al. 2018). ERA5 performance is once again superior to that of MERRA-2, which tends to have a slightly bigger magnitude of wind shear underestimation (Table 3, Figs. 5 and 6).

SRH is an important parameter for characterizing the environments associated with both supercells and tornadoes, particularly for 0–1- and 0–3-km layers (Thompson et al. 2003, 2012; Anderson-Frey et al. 2016; Coffey et al. 2019, 2020; Coniglio and Parker 2020). ERA5 tends to overestimate SRH with ME of around $5 \text{ m}^2 \text{ s}^{-2}$, whereas for MERRA-2 ME is close to zero, while MAE for both reanalyses is around $30 \text{ m}^2 \text{ s}^{-2}$ (Table 3, Fig. 5). Reanalysis performance for mean wind speed over 0–1-, 1–3-, and 0–6-km layers is similar to the results for wind shear, where a greater accuracy is observed for the midtroposphere. However, contrary to underestimation of 0–1-km wind shear, both reanalyses overestimate 0–1-km mean wind (ME of 0.35 and 0.89 m s^{-1} for ERA5 and MERRA-2). High correlations are observed for 0–6-km mean wind (0.98 for

TABLE 3. As in Table 2, but for wind and composite parameters. Here, an asterisk indicates that only cases exceeding 0 are considered in computations.

	Pearson correlation coefficient (R)		Spearman correlation coefficient (R_s)		Mean error (ME)		Mean absolute error (MAE)		Root-mean-square error (RMSE)	
	ERA5	MERRA-2	ERA5	MERRA-2	ERA5	MERRA-2	ERA5	MERRA-2	ERA5	MERRA-2
Wind parameters										
Effective shear (m s^{-1})*	0.74	0.70	0.68	0.64	-1.06	-1.27	2.93	3.25	4.05	4.40
0–1-km shear (m s^{-1})	0.81	0.76	0.76	0.71	-0.32	-0.98	1.95	2.30	2.62	3.02
0–3-km shear (m s^{-1})	0.91	0.86	0.89	0.71	-0.46	-0.85	1.94	2.41	2.60	3.18
0–6-km shear (m s^{-1})	0.96	0.94	0.95	0.93	-0.31	-0.53	2.07	2.61	2.83	3.50
0–1-km SRH ($\text{m}^2 \text{s}^{-2}$)	0.83	0.77	0.80	0.74	5.21	-1.44	28.02	30.96	42.21	46.37
0–3-km SRH ($\text{m}^2 \text{s}^{-2}$)	0.85	0.79	0.82	0.76	4.82	-0.61	32.99	38.16	49.70	56.52
0–1-km mean wind (m s^{-1})	0.90	0.86	0.88	0.84	0.35	0.89	1.38	1.77	1.93	2.44
1–3-km mean wind (m s^{-1})	0.97	0.95	0.96	0.94	-0.15	-0.09	1.00	1.29	1.42	1.77
0–6-km mean wind (m s^{-1})	0.98	0.96	0.97	0.96	-0.05	0.25	0.80	1.04	1.12	1.43
10-m wind speed (m s^{-1})	0.71	0.67	0.67	0.60	-0.19	0.29	1.40	1.62	1.90	2.12
850-hPa wind (m s^{-1})	0.94	0.91	0.92	0.60	-0.28	-0.14	1.37	1.71	1.91	2.31
700-hPa wind (m s^{-1})	0.96	0.93	0.95	0.92	-0.31	-0.27	1.31	1.72	1.83	2.30
500-hPa wind (m s^{-1})	0.98	0.96	0.98	0.96	-0.28	-0.20	1.39	1.92	1.99	2.62
Composite parameters										
STP*	0.63	0.61	0.57	0.52	-0.07	-0.04	0.16	0.17	0.35	0.37
SCP*	0.67	0.56	0.64	0.53	-0.79	-0.62	1.25	1.39	2.40	2.60
SHIP*	0.71	0.59	0.61	0.49	-0.11	-0.09	0.16	0.18	0.26	0.29
ML WMAXSHEAR*	0.80	0.73	0.73	0.68	-47.39	-52.57	118.65	134.02	181.53	207.99
MU WMAXSHEAR*	0.79	0.71	0.73	0.66	-85.34	-59.86	151.04	165.68	220.94	244.21

ERA5 and 0.96 for MERRA-2) with ME of -0.05 and 0.26 m s^{-1} for ERA5 and MERRA-2, respectively.

Analysis of wind speeds at fixed levels suggests that the reanalyses handle wind fields typically well except for the near-surface level (Table 3, Fig. 6). MAE for ERA5 over all layers is in the range between 1.31 and 1.40 m s^{-1} and slightly higher for MERRA-2 at 1.62 – 1.92 m s^{-1} . Mean errors reflect small underestimates for all wind speed variables (around -0.3 m s^{-1}), with the exception of 10-m wind of MERRA-2 where ME indicates overestimation of 0.29 m s^{-1} . Wind speed at 10 m also has the lowest correlation (0.71 for ERA5 and 0.67 for MERRA-2), reflecting the limitations of boundary layer and topographical representation in the reanalysis, also found in prior studies (Thompson et al. 2003; Allen and Karoly 2014; Gensini et al. 2014; Taszarek et al. 2018; King and Kennedy 2019).

e. Composite parameters

Composite parameters, by their design, seek overlapping quantification of the likelihood for severe hazards, meaning that any limitation of the reanalysis components can be exacerbated or minimized by the combination. This leads to a correlation no greater than 0.8 considering all composite parameters (Table 3). Performance for parameters that involve SRH and effective shear (SCP and STP) is worse (Table 3, Fig. 7) than a simple combination of CAPE and fixed vertical wind shear (WMAXSHEAR; Taszarek et al. 2017, 2020;

Rodríguez and Bech 2021). However, even in this case the lower correlations of CAPE lead to a decrease in the performance of the covariate quantity. In comparison to the individual parameters, differences between the reanalyses are also more stark. For STP both reanalyses have similar correlations (0.63 and 0.61), whereas for SCP and SHIP there are noticeably higher correlations for ERA5 (0.67 and 0.71) compared to MERRA-2 (0.56 and 0.59) (Table 3). However, despite these differences, MAE are similar between reanalyses (Table 3). In contrast, for WMAXSHEAR there is generally a slight advantage to ERA5, although the difference is relatively small (Fig. 7).

The parameter phase space of ML CAPE and 0–6-km wind shear indicates that both reanalyses tend to underestimate high values of instability ($>3000 \text{ J kg}^{-1}$; Fig. 8). However, it can be also noted that MERRA-2 has overall better performance in sampling high-shear, low-CAPE environments (HSLC; Sherburn and Parker 2014; Anderson-Frey et al. 2019; Gatzen et al. 2020) than ERA5. Compared to observations, both reanalyses have more frequent situations with marginal both instability and wind shear. The best overall agreement is obtained for CAPE 500 – 1000 J kg^{-1} and 0–6-km wind shear of 10 – 20 m s^{-1} , but these cases are also the most frequent and thus make it less sensitive to changes in fractions. Focusing on low-level moisture and wind shear yields that especially extreme values (ML MIXR $> 18 \text{ g kg}^{-1}$ and 0–1-km wind shear $> 25 \text{ m s}^{-1}$) are not well represented by both reanalyses. Conversely, very low values of 0–1-km

Comparison of selected kinematic variables with collocated profiles from ERA5 and MERRA2

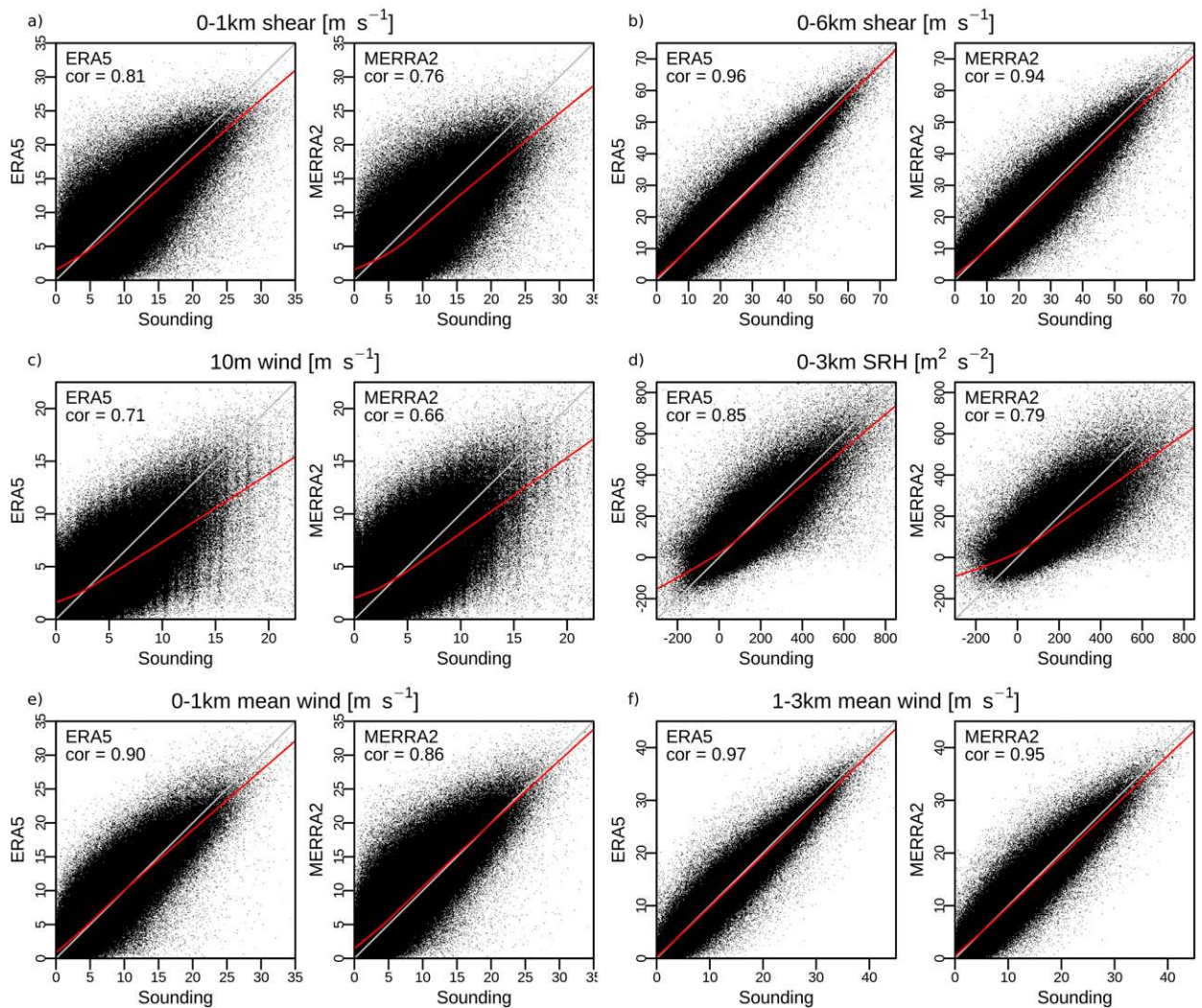


FIG. 6. As in Fig. 3 but for (a) 0–1-km wind shear, (b) 0–6-km wind shear, (c) 10-m wind speed, (d) 0–3-km storm-relative helicity, (e) 0–1-km mean wind speed, and (f) 1–3-km mean wind speed.

wind shear are also overestimated, especially for MERRA-2 (Fig. 8).

f. Long-term changes

For a series of representative parameters, temporal changes over the past 39 years associated with more frequent observations being available for the reanalyses have been reflected in decreasing MAE and increasing correlations for both reanalyses. Among wind parameters, the upper-level wind speed and direction had smaller decreases to MAE, while near the surface there were more significant changes to wind accuracy (Fig. 9). For the 0–1-km mean wind, MAE has dropped considerably, particularly after 1991. Curiously, although MAE of low-level wind shear has decreased over time for ERA5, an increase was observed for MERRA-2, even though correlations for both reanalyses have been consistently increasing (Fig. 9).

Changes in MERRA-2 wind data over time may be related to an increasing number of assimilated observations, especially those that started to rapidly increase in the 1990s, such as remotely sensed upper-air measurements, aircraft data, and geostationary/polar satellite observations (Koster et al. 2016). Similar increases in quantity and quality of assimilated data are observed also within ERA5 (Hersbach et al. 2020). This indicates that specific variables in the reanalysis may be vulnerable for inhomogeneities in time and space, and thus the use of such data in studying long-term climatological trends should be performed with caution.

Near-surface temperature and lapse rates have high correlations for the whole dataset (Table 2), and this is matched by smaller improvements in MAE and correlations over time as compared to wind parameters (Fig. 10). Only a very slight change was observed for 10-m temperature with a decrease in MAE of 0.1°C comparing the 1980s with the 2010s. Only slight

Comparison of selected composite indices with collocated profiles from ERA5 and MERRA2

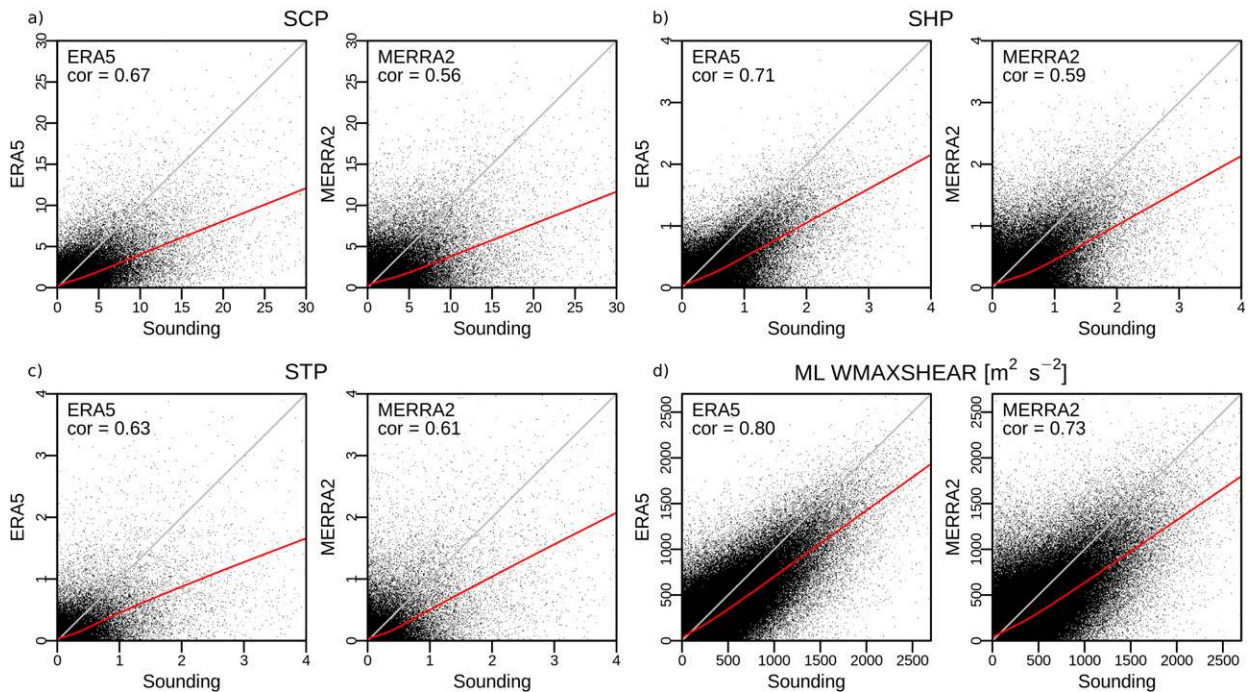


FIG. 7. As in Fig. 3 but for (a) supercell composite parameter, (b) significant hail parameter, (c) significant tornado parameter, and (d) mixed-layer WMAXSHEAR.

improvements were observed for mid- and lower-tropospheric lapse rates, and the rate of change in the latter was smaller for MERRA-2 compared to ERA5 (Fig. 10). The greatest improvements with respect to instability (using only environments with positive CAPE) relate to moisture parameters and parcel properties. MAE for ML MIXR has decreased from a mean of 0.75 and 1.05 g kg^{-1} in the 1980s to 0.65 and 0.85 g kg^{-1} after 2010 for ERA5 and MERRA-2, respectively (Fig. 11). For ML CAPE, MAE has decreased from 350 J kg^{-1} for MERRA-2 and 300 J kg^{-1} for ERA5 prior to 2004, and then to around 275 J kg^{-1} for MERRA-2 and 225 J kg^{-1} for ERA5 over the past 5 years. However, the differences between the reanalyses are small despite large differences in correlation. MAE for ML CIN has remained relatively steady over time (around 32 J kg^{-1} for MERRA-2 and 22 J kg^{-1} for ERA5); however, there is a larger increase in correlation over the record, with notable difference between the two reanalyses, consistent with the benefits of higher vertical resolution (Fig. 11).

Finally, we assess changes in the midtropospheric wind, temperature, and dewpoint, which also show long-term decreases in MAE (Fig. 12). While relative changes in 500-hPa temperature and dewpoint are equivalent for both reanalyses (drop of MAE by 0.2°C for temperature and 0.5°C for dewpoint between the 1980s and 2010s), the change in MAE for 500-hPa wind speed is observed mainly for MERRA-2 (Fig. 12). The highest slope of change of MAE for 500-hPa temperature predominantly occurred prior to 2000, and since that time only minor increases in accuracy have been observed. Despite the

larger changes in MAE, correlations for 500-hPa wind speed and temperature remained stable. In contrast, correlations of 500-hPa dewpoint have seen increases on the order of 15% for both MERRA-2 and ERA5.

While significant quality control has been performed for the sounding observations, it is uncertain whether observed long-term changes are related to changes in the sounding quality (e.g., the use of different sensors), the number of available levels (Fig. 1b), or rather an increase in the accuracy of reanalysis (Koster et al. 2016; Hersbach et al. 2020). Variables such as 500-hPa wind speed, which change in MAE only in MERRA-2, or 0–1-km wind shear, where MAE increases with time in MERRA-2, suggest either greater sensitivity to assimilated observations over time or a limitation in the underlying model to these changes in observations. However, it is likely that changes in the quality of sounding data over time may also be an important factor in assessing long-term changes in the performance metrics of reanalyses.

g. Annual variability

Correlation and MAE also vary through the year as a function of the annual cycle. Wind parameters are characterized by the highest MAE and correlation during the winter months and conversely the lowest MAE and lower correlations during the summer (Fig. 9). Similar annual cycles in MAE are also observed for temperature and lapse rates, but the amplitude is the largest near the surface and decreases with height (Figs. 10 and 12). Differences between the reanalyses can be

Fraction of reanalysis profiles among soundings in a given parameter space

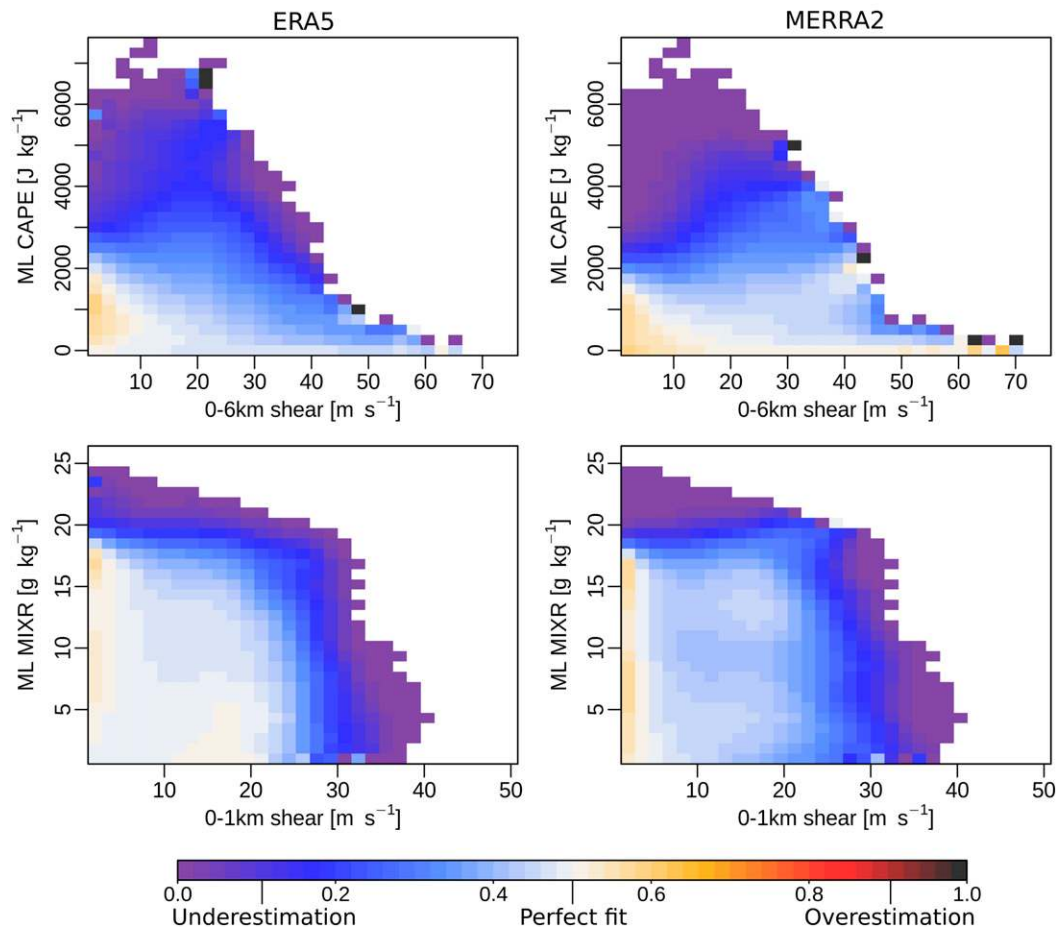


FIG. 8. Fraction of reanalysis cases (ERA5 at left, MERRA-2 at right) among soundings in a given parameter space of (top) mixed-layer CAPE and 0–6-km wind shear and (bottom) mixed-layer mixing ratio and 0–1-km wind shear. A value of 0.5 indicates an equal ratio between soundings and reanalysis; <0.5 indicates underestimation while >0.5 indicates overestimation of reanalysis environments. Please note that a slight smoothing (3×3 grid focal median) has been applied for better readability of results.

found in the annual cycle of MAE for 10-m temperature where MERRA-2 has a clear difference between summer (1.6°C) and winter (1.9°C) while ERA5 errors are relatively stable. Correlations for MERRA-2 0–1-km lapse rates are the one exception, where a considerable decrease in MAE is accompanied by large increases in correlation from the spring through fall (Fig. 10). This contrasts with the very consistent 0–1-km lapse rates that are almost unchanged throughout the year for ERA5.

The largest intra-annual differences are found for moisture and instability parameters, where ML MIXR, CAPE, and CIN peak in MAE during summer and are considerably lower in the cool season, reflecting the annual cycle of these parameters (Fig. 11). Correlations for ML CAPE peak in the spring and fall for both reanalyses, and are similar in both summer and winter, whereas ML CIN tends to have higher correlations for the early part of the year through April (Fig. 11).

h. Spatial variability

In this section we explore the spatial distribution of convective parameters in reanalysis relative to climatological distributions of the respective sounding stations to identify local biases. We focus only on the 0000 and 1200 UTC time steps due to consistent availability of sounding observations in these hours. Generally, ML MIXR (Fig. 13a) shows a strong correlation between the reanalyses and observations over North America with a clear peak over the Gulf Coast (95th percentile $> 18 \text{ g kg}^{-1}$) and the corridor of poleward moist advection with enhanced values over the Southeast, Midwest, and Great Plains ($> 14 \text{ g kg}^{-1}$). Despite this general agreement, there are a few small differences between MERRA-2 and both ERA5 and the soundings. Moisture values near the Gulf are generally lower, while over the Great Plains the gradient is displaced eastward. Both reanalyses underestimate the northward extent of moisture

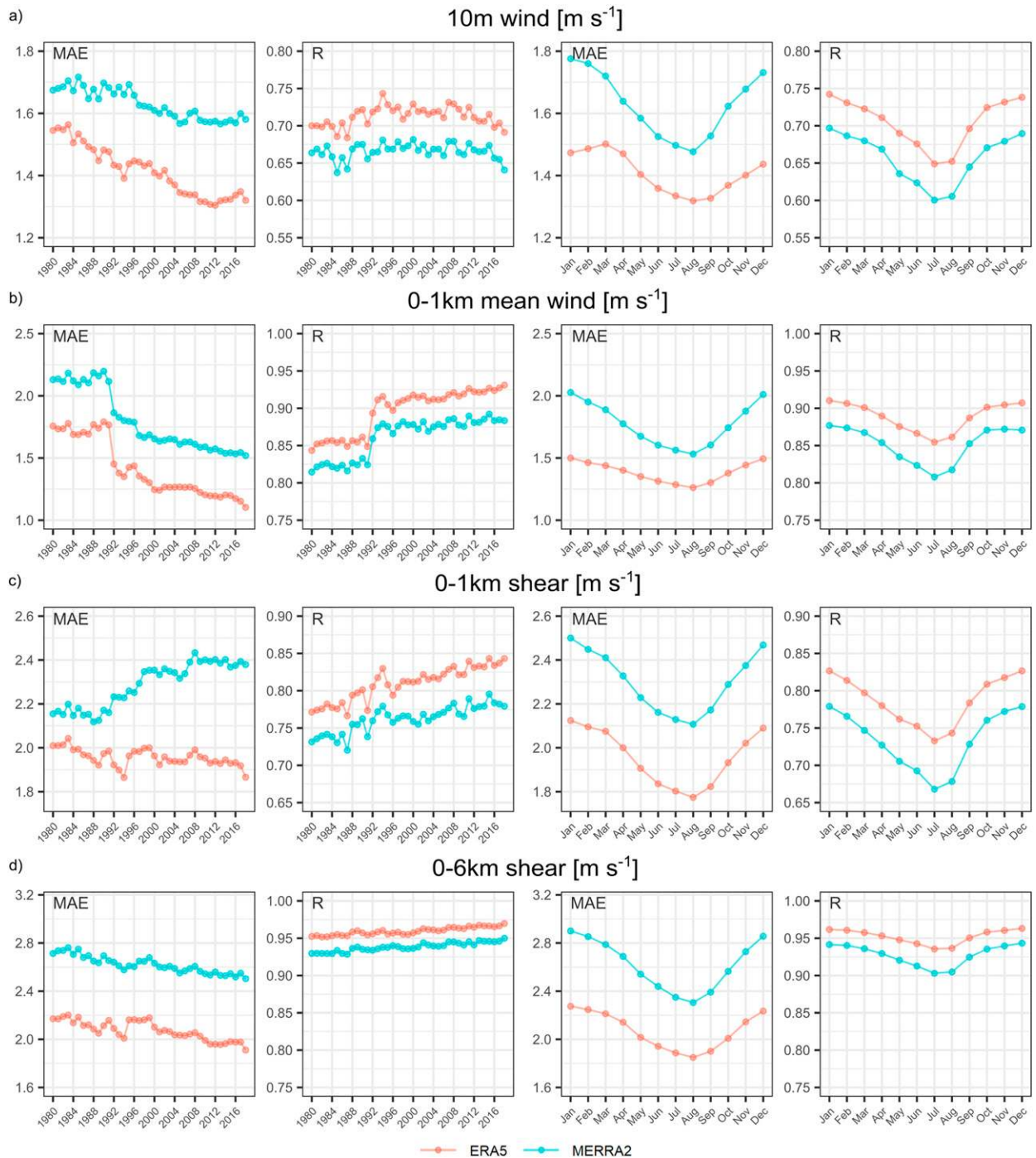


FIG. 9. (first and second columns) Long-term variability and (third and fourth columns) annual cycle of mean absolute error (MAE) and Pearson correlation coefficient (R) for (a) 10-m wind speed, (b) 0–1-km mean wind, (c) 0–1-km wind shear, and (d) 0–6-km wind shear for ERA5 (red) and MERRA-2 (turquoise) in comparison with soundings.

over the East Coast close to the Gulf Stream. Over Europe MERRA-2 has notably lower moisture over central Europe and higher values over western Russia compared to ERA5 (Fig. 13a).

There is relatively good agreement between all datasets for the midtropospheric lapse rates (Fig. 13b) with the highest

values observed over the Rocky Mountains over the United States (95th percentile $> 9 \text{ K km}^{-1}$) with smaller peaks over the northern Atlantic and northwestern Africa ($> 8 \text{ K km}^{-1}$). Peak values are slightly higher in both reanalyses; however, there are no sounding stations near the lapse rate maximum

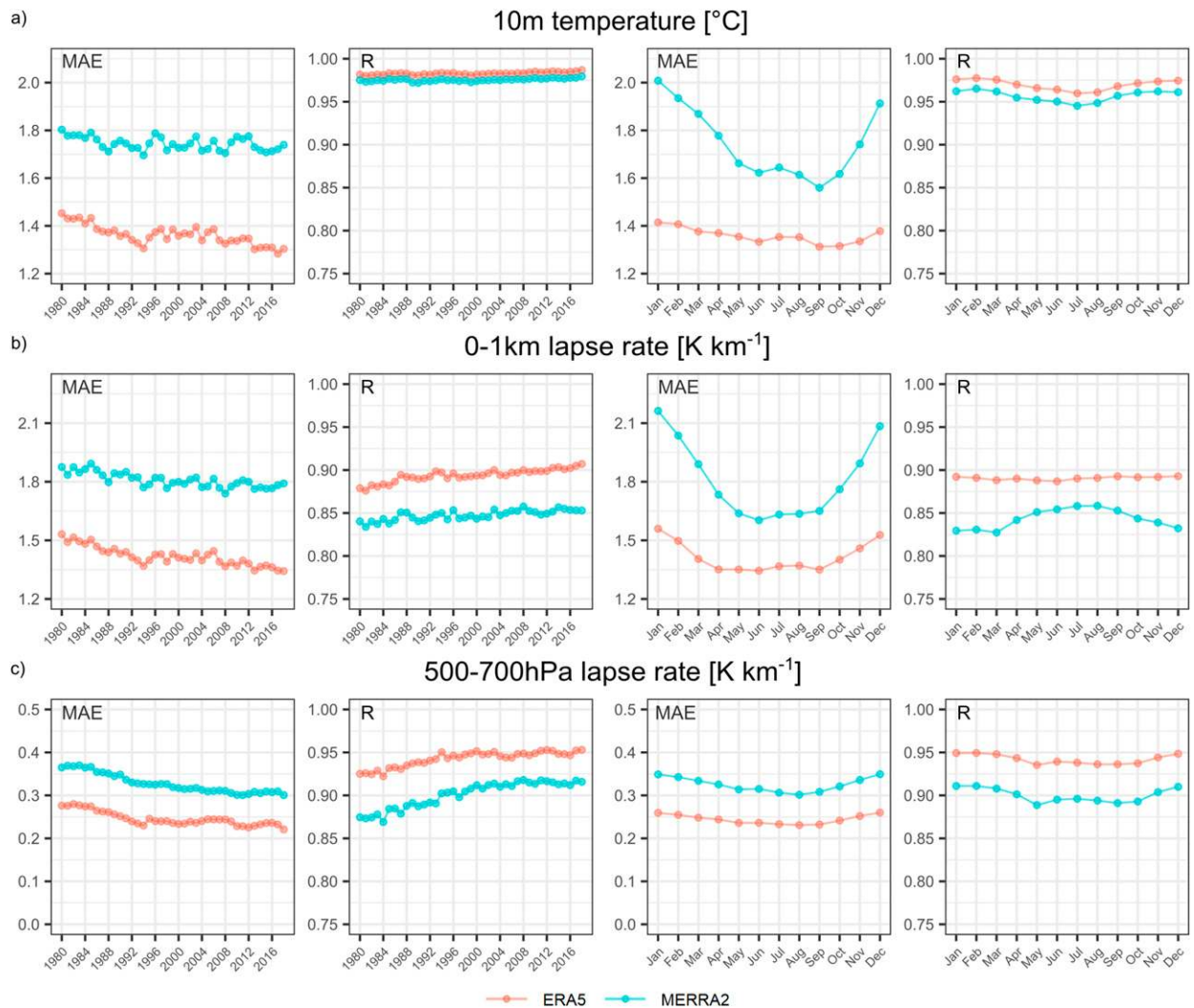


FIG. 10. As in Fig. 9, but for (a) 10-m temperature, (b) 0–1-km lapse rate, and (c) 500–700-hPa lapse rate.

across central Colorado. For ML CAPE there are larger differences, with lower 95th percentiles for both ERA5 and MERRA-2 particularly over the Great Plains and southeastern United States (Fig. 14a). Over Europe these differences are generally lower outside of stations that are close to coastal or mountainous zones. Between the reanalyses, ERA5 overall has better accuracy than MERRA-2. There are a few substantive biases in MERRA-2, including underestimation over the Gulf Coast, North Dakota, western and central Europe, the Iberian Peninsula, and the Balkans. Conversely, over western Russia and the Atlantic, MERRA-2 indicates higher instability. Biases at locations for ML CIN are considerably lower compared to ML CAPE across North America; however, magnitudes are still underestimated despite higher resolution of ERA5 (Fig. 14b). Over Europe CIN underestimation can be observed in both reanalyses across the entire continent.

Similar to results presented in the previous sections, 0–6-km wind shear is represented well by both reanalyses (Fig. 15a). From a climatological perspective, wind shear is considerably

stronger across northern parts of the United States (50th percentile $> 18 \text{ m s}^{-1}$), compared to $> 15 \text{ m s}^{-1}$ over northwestern Europe. However, combining the 0–6-km wind shear with ML CAPE into a bivariate severe-thunderstorm proxy with ML WMAXSHEAR illustrates that covariates can produce much larger differences, and are primarily related to biases in ML CAPE. Both reanalyses notably underestimate ML WMAXSHEAR (Fig. 15b) over the central and southeastern United States. Across Europe, ERA5 has much better agreement with soundings for ML WMAXSHEAR, while MERRA-2 has a notable underestimation over central Europe and the Balkans, and an overestimation over the coastal zones of the Atlantic.

Finally, we compare a mean correlation for all analyzed 45 variables for each sounding site (Fig. 16). The highest values are typically observed for stations away from complex orography (i.e., the central and eastern United States, and western, central, and eastern Europe). Consistent with results discussed in prior sections, ERA5 has clearly better accuracy over both

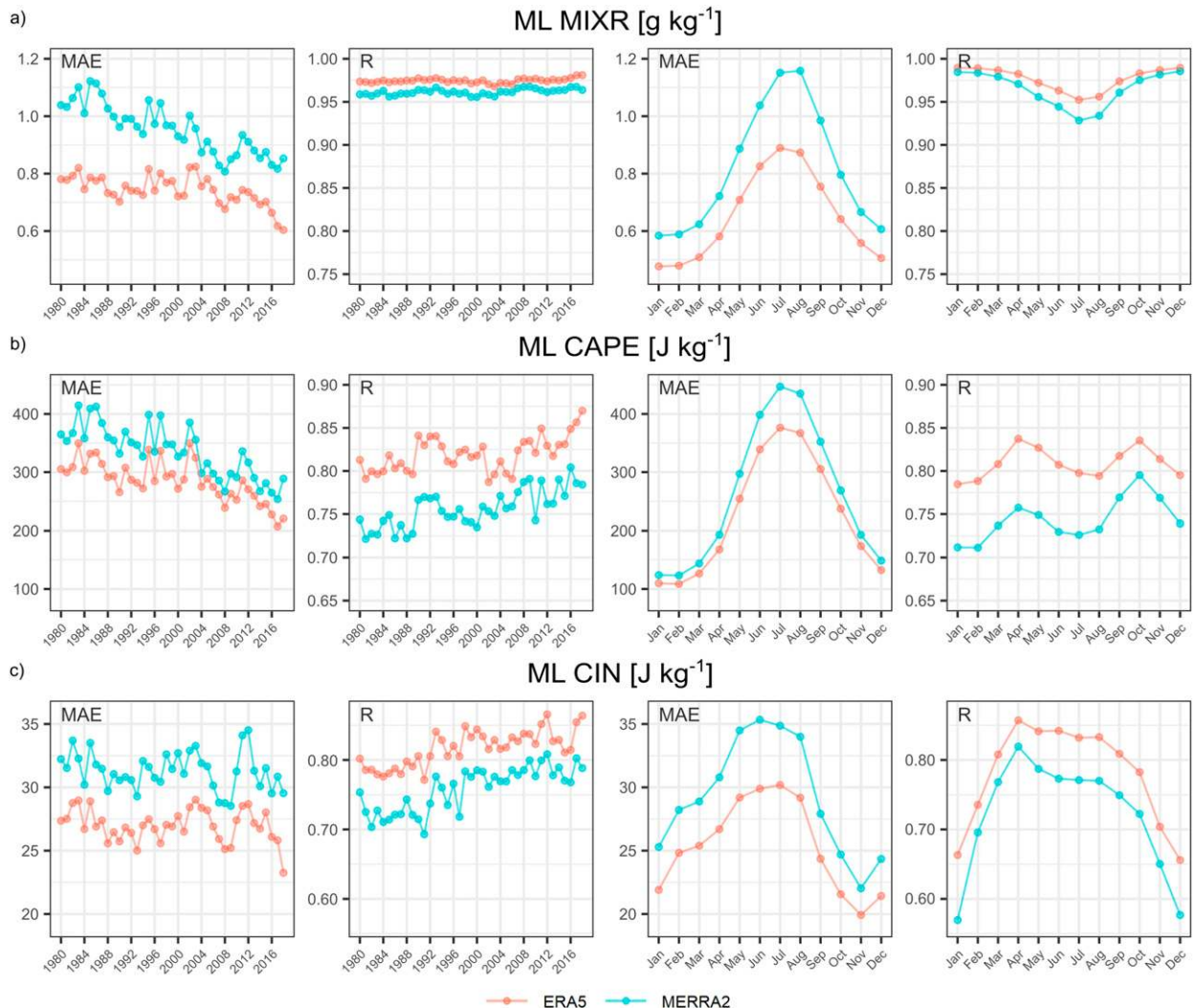


FIG. 11. As in Fig. 9, but for (a) mixed-layer mixing ratio, (b) mixed-layer CAPE, and (c) mixed-layer convective inhibition.

continents in representing atmospheric parameters. This performance is generally higher everywhere except over the mountainous western United States where topographic gradients play a significant role in the underlying meteorology. In contrast, MERRA-2 shows more spatial biases and lower correlations over a wider area.

One of the more intriguing results is obtained for mean correlations computed between both reanalyses (Fig. 16c). The best agreement between ERA5 and MERRA-2 can be found over a broad corridor from the British Isles to Russia (mean correlation from 45 variables > 0.9) with lower values over southern Europe. Over North America, spatial patterns in this parameter can be linked to orography, with the highest mean correlation values exceeding 0.9 over central and eastern United States and the lowest over the mountainous west and Mexico. This result likely highlights that the decrease in grid size to 0.25° for ERA5 can make a substantive difference in performance for many areas, which may be a reason for the spatial patterns in reanalysis performance. It may also suggest

that the lower density of sounding data entering data assimilation in both reanalyses can be resolved in different ways between the reanalyses and lead to contrasting performance.

4. Summary and concluding remarks

In this study we compared 45 observed sounding parameters from 232 stations over Europe and North America (3.7 million measurements) with collocated ERA5 and MERRA-2 profiles to examine how well reanalysis depicts observed and derived variables, specifically focusing on those relevant to convection. The frequency of sounding measurements and number of available levels has generally increased over time. The analysis of temperature, moisture, wind, MU/ML parcel parameters, and composite indices highlighted several findings of interest, the most important of which are listed below.

- The largest differences between soundings and reanalysis are found for low-level wind parameters (10-m wind speed,

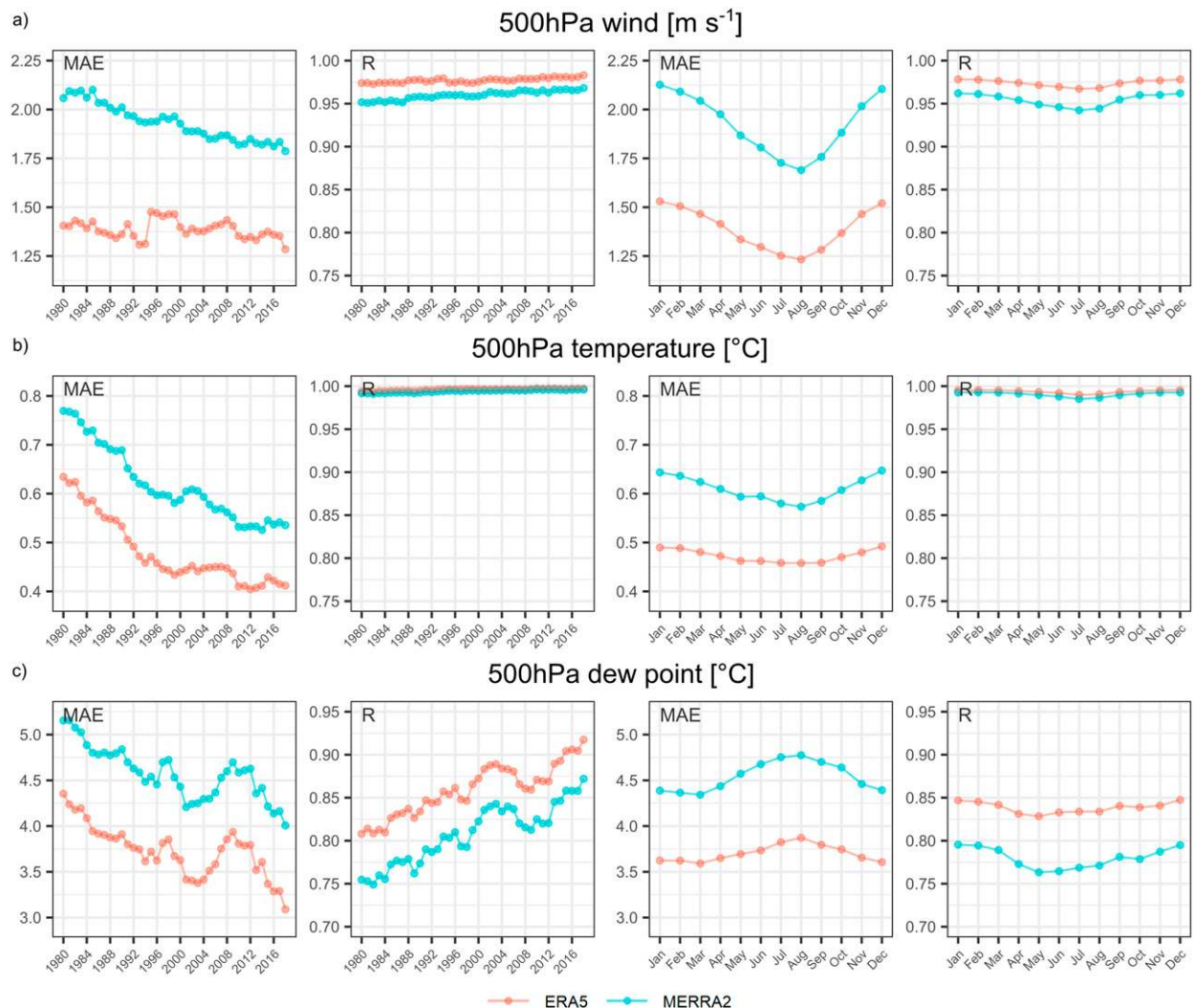
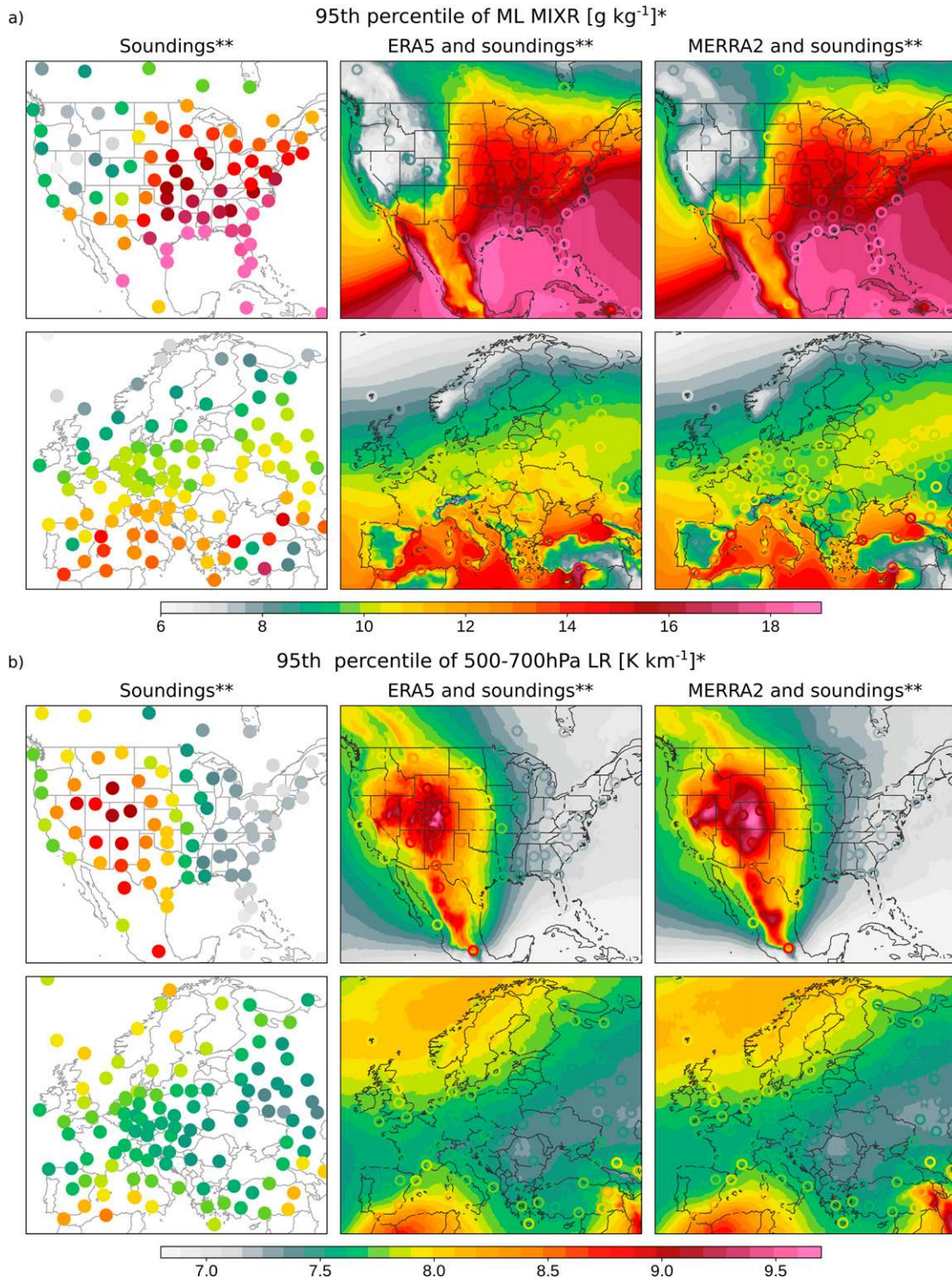


FIG. 12. As in Fig. 9, but for (a) 500-hPa wind speed, (b) 500-hPa temperature, and (c) 500-hPa dewpoint.

- 0–1-km wind shear, mean wind, SRH), MU and ML parcel parameters (CAPE, CIN, LCL, LFC), low-level lapse rates (0–1 km), and composite indices (STP, SCP, SHP, WMAXSHEAR). In contrast, reanalyses do well in representing temperature and moisture variables (at 10 m, 850, 700, and 500 hPa, and PWATER), midtropospheric lapse rates (3–6 km, 500–700 hPa), and mean wind over deeper layers such as 1–3 and 0–6 km.
- Both reanalyses tend to underestimate CAPE, EL, low-level moisture, wind shear, wind speed, and all composite parameters. Slight overestimation is observed for low-level lapse rates, midtropospheric moisture, and LFC. ERA5 underestimates LCL and 10-m wind, while slightly overestimating SRH. Both reanalyses have problems in characterizing extreme environments, especially considering situations with CAPE exceeding 3000 J kg^{-1} and boundary layer-related parameters (e.g., ML MIXR $> 18 \text{ g kg}^{-1}$, 0–1-km shear $> 25 \text{ m s}^{-1}$).
 - Among parcel parameters ML parcels are more accurately represented in the reanalyses than MU parcels with

lower mean absolute errors and higher correlations with observations.

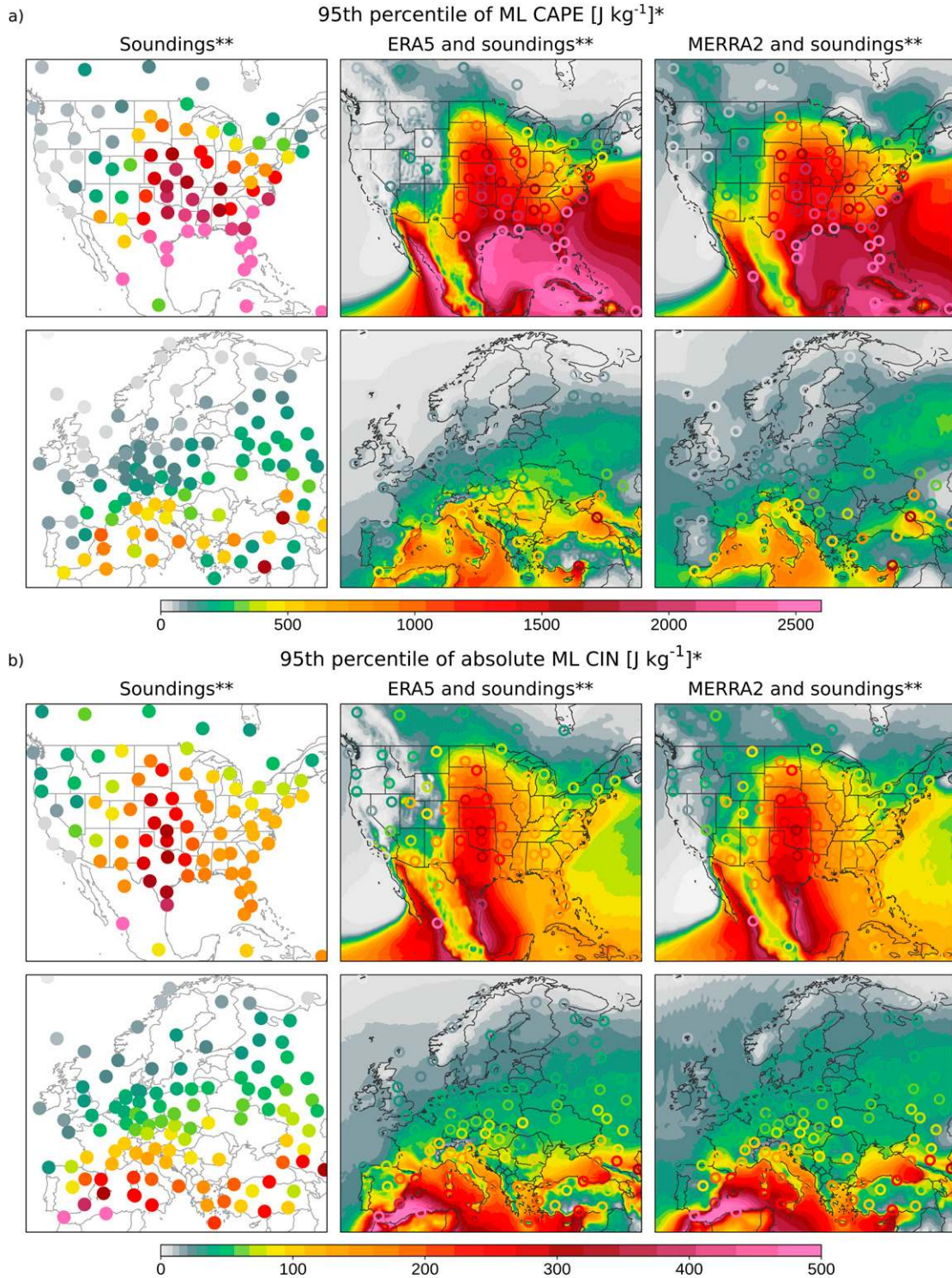
- In every analyzed variable, ERA5 had higher correlation and lower mean absolute errors than MERRA-2.
- Over time, mean absolute errors between soundings and reanalyses have decreased and correlations have increased for almost every variable. An exception to this was 0–1-km wind shear in MERRA-2 for which mean absolute error has increased. The largest increases in accuracy were for wind parameters, and the lowest for temperature variables. Notable temporal discontinuities were found in the 0–1-km mean wind and dewpoint at 500 hPa between 1991 and 1992.
- Mean absolute error and correlation have well-defined annual cycles. Wind and temperature parameters typically have their lowest mean absolute errors and correlation during summer, and their highest values during winter. Moisture and parcel (CAPE, CIN) variables typically have the highest mean absolute errors during summer. The



* - considering only 00 and 12 UTC timesteps over a period of 1980-2018

** - considering only stations with at least 2000 measurements from 00 and 12 UTC, and with 00 and 12 UTC sample size ratio in the range from 0.75 and 1.25

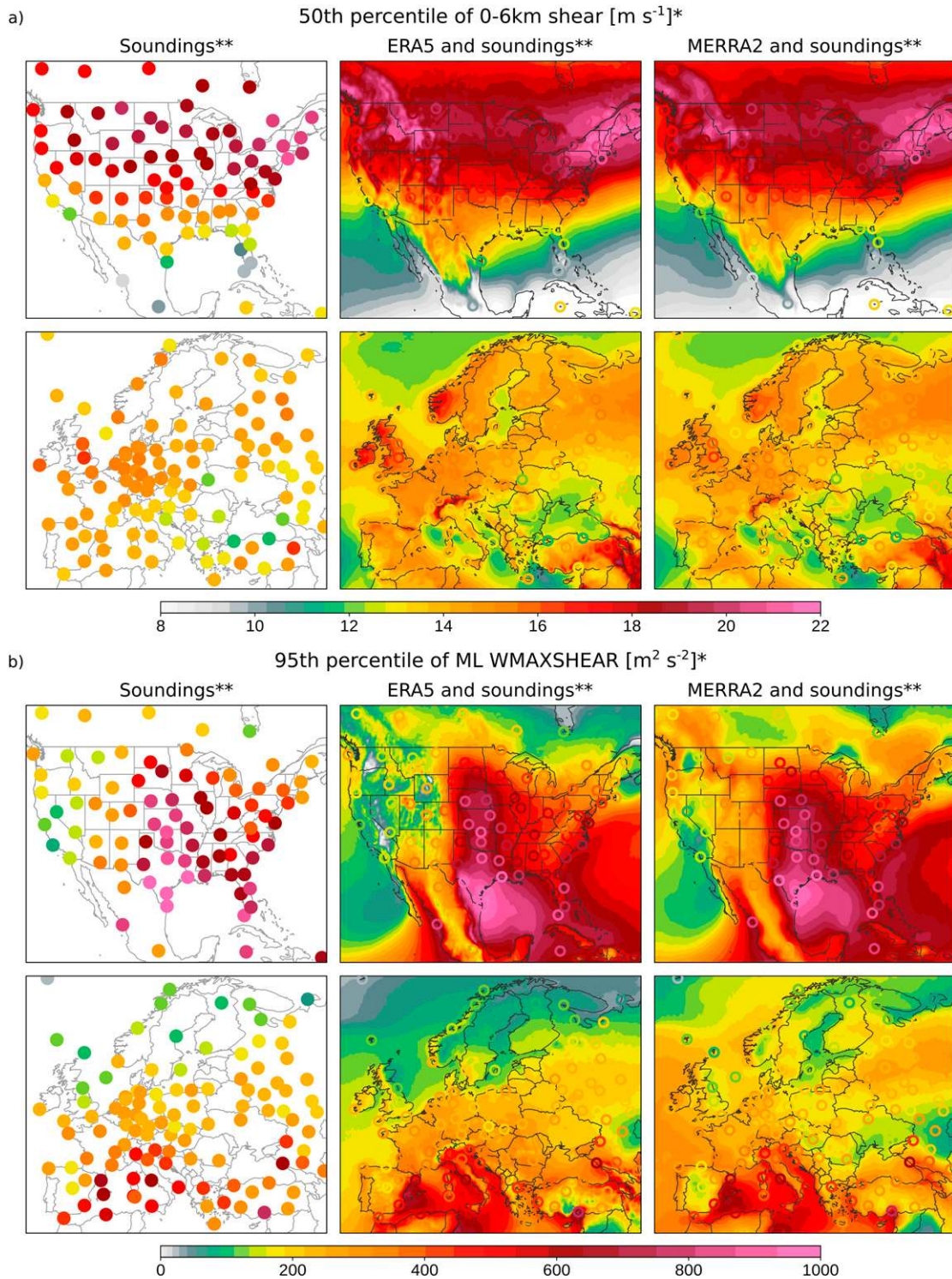
FIG. 13. Climatological distribution of 95th percentile of (a) mixed-layer mixing ratio and (b) 500–700-hPa lapse rate for sounding, ERA5, and MERRA-2. Only 0000 and 1200 UTC time steps over a period of 1980–2018 are considered. Only sounding stations with at least 2000 measurements for both 0000 and 1200 UTC with a ratio between these two ranging from 0.75 and 1.25 are considered.



* - considering only 00 and 12 UTC timesteps over a period of 1980-2018

** - considering only stations with at least 2000 measurements from 00 and 12 UTC, and with 00 and 12 UTC sample size ratio in the range from 0.75 and 1.25

FIG. 14. As in Fig. 13 but for (a) mixed-layer CAPE, and (b) mixed-layer convective inhibition.



* - considering only 00 and 12 UTC timesteps over a period of 1980-2018

** - considering only stations with at least 2000 measurements from 00 and 12 UTC, and with 00 and 12 UTC sample size ratio in the range from 0.75 and 1.25

FIG. 15. As in Fig. 13 but for (a) 50th percentile of 0-6-km wind shear and (b) 95th percentile of mixed-layer WMAXSHEAR.

Mean correlation coefficient for all analyzed parameters among individual locations

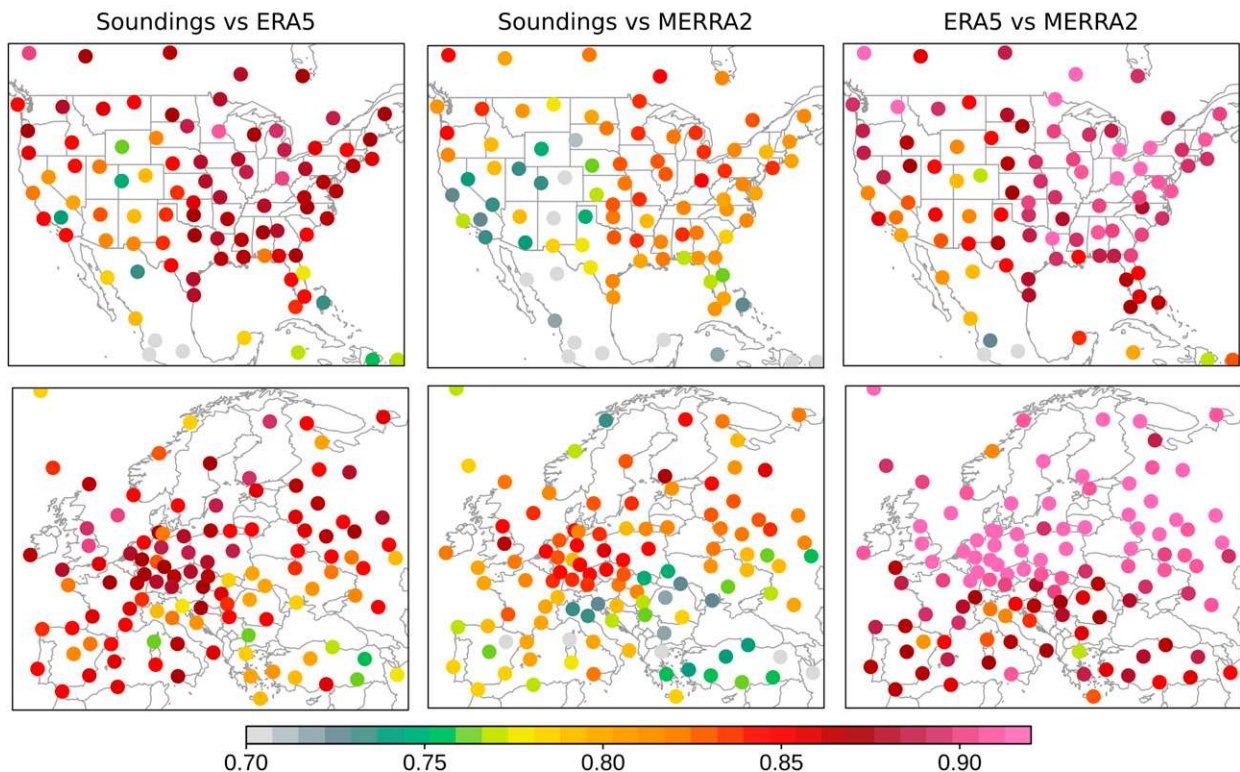


FIG. 16. Pearson correlation coefficient as a mean from all 45 parameters listed in Table 2 and 3 (considering only stations with at least 5000 measurements), for (left) soundings and ERA5, (center) soundings and MERRA-2, and (right) ERA5 and MERRA-2. Higher values indicate a better fit between datasets. Please note, that this concerns mainly convective parameters and unstable profiles (most-unstable CAPE $> 0 \text{ J kg}^{-1}$).

correlation for CAPE is the highest during transitional seasons, whereas CIN is maximized mainly during spring.

- Spatial analysis indicates that both reanalyses are generally able to capture the pattern of convective parameters very well, with MERRA-2 typically a bit drier and less unstable (lower MIXR, CAPE, and WMAXSHEAR) over central Europe and the Balkans with the opposite pattern over western Russia. Both reanalyses notably underestimate CIN across Europe, and CAPE and CIN over the United States.
- Mean correlations by station for the two reanalyses illustrate that the best representation is over flat or lower elevations. In contrast, performance is poor near sharp topographic boundaries (coastal zones and mountain ranges). The correlation between reanalyses is the highest for stations located over northern Europe and decreases toward the south. Over North America the best agreement is for stations located over the east and decreases toward the west.

Atmospheric reanalyses provide a best guess to the state of the past atmosphere, and unlike sparse sounding data with limited spatiotemporal resolution, they are an invaluable tool for construction of severe thunderstorm climatologies (Brooks et al. 2003; Gensini and Ashley 2011; Tippett et al. 2012; Thompson et al. 2013; Allen and Karoly 2014; Allen et al. 2015; Tazarek

et al. 2018; Chen et al. 2020; Li et al. 2020; Tazarek et al. 2020), assessment of long-term historical trends (Mohr and Kunz 2013; Gensini and Brooks 2018; Rädler et al. 2018; Tang et al. 2019; Tazarek et al. 2021), and investigation of environments associated with specific convective hazards leading to their better prediction (Thompson et al. 2003, 2007, 2012; Gropp and Davenport 2018; Coffey et al. 2020; Ingrosso et al. 2020; Tazarek et al. 2020; Rodríguez and Bech 2021). The new ERA5 reanalysis provides higher spatial, vertical, and temporal resolution that represents convective environments with a resolution that was not available in prior global reanalysis, opening new research possibilities.

While reanalyses do provide an invaluable source of data, as highlighted in this study, there are areas of possible future improvement, especially related to boundary layer representation. As shown in our results and those of prior studies, one of the biggest limitations of the reanalyses is their inability to represent rare and extreme convective environments, especially those that are characterized by high CAPE and/or vertical wind shear. Reanalyses are also inconsistent through time for parameters relevant to severe convection, and hence trend analyses require careful consideration to ensure shifts are not solely the response to changes in data quality. Here, we demonstrated that convective profiles derived from ERA5 and MERRA-2 hybrid-sigma model levels are overall better

Fraction of measurements that passed quality-control phase

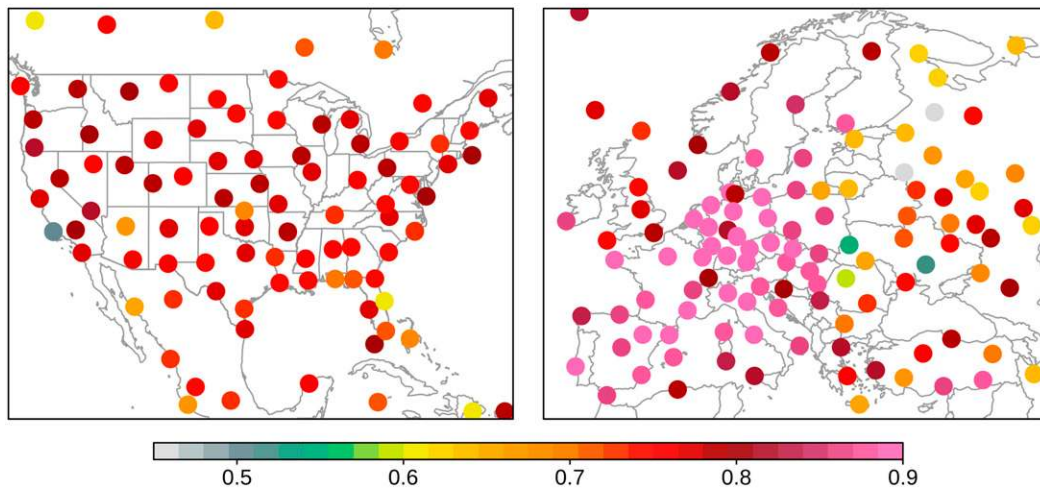


FIG. A1. Fraction of measurements that passed preliminary quality-control assumptions for individual location over North America and Europe.

sampled compared to previously evaluated profiles from both ERA-Interim (Allen and Karoly 2014; Taszarek et al. 2018) and NARR (Gensini et al. 2014). Correlations in our study were also higher than in the work of King and Kennedy (2019), who derived convective variables from six different reanalyses to compare with the RUC2 collocated profiles. A potential explanation for this difference is that in contrast to King and Kennedy (2019), who only considered profiles proximal to known supercell cases, we analyze all available time steps including situations not contaminated by convection.

We would also suggest that in a view of best practice, it is recommended that computation of convective variables, if available, should be performed with the native coordinates at the full resolution the reanalysis was produced. As an example, hybrid-sigma model levels follow orography allowing for consistent vertical resolution and can more accurately sample boundary layer conditions (e.g., 28 levels at 0–2 km in ERA5) and depict important features such as capping inversions. Conversely, commonly used pressure levels are interpolated from model levels, and are subsampled to reduce data volume (typically 37 levels for the entire profile) while also not following orography. This also leads to a reduction in the number of boundary layer data available in high elevation areas. We also conclude that researchers using reanalysis and model datasets more broadly for climatological purposes should ensure that they are aware of their respective parameter biases and (where possible) cross-validate these results with observations. Finally, based on the results from our work and prior studies we suggest that ERA5 is likely one of the most reliable available reanalyses for exploration of convective environments, mainly because of its improved resolution.

Acknowledgments. This research was supported by grants from the Polish National Science Centre (2017/27/B/ST10/

00297 and 2019/33/N/ST10/00403) and the Polish National Agency for Academic Exchange–The Bekker Programme (PPN/BEK/2018/1/00199). J. T. Allen acknowledges support from the National Science Foundation under Grant AGS-1945286. The reanalysis computations were performed in Poznan Supercomputing and Networking Center (project number 448).

Data availability statement. All datasets used in this study are openly available. ERA5 data were downloaded from the European Centre for Medium-Range Weather Forecasts (ECMWF), Copernicus Climate Change Service (C3S) available at <https://cds.climate.copernicus.eu/>. MERRA-2 data were downloaded from the National Aeronautics and Space Administration (NASA), Goddard Earth Sciences Data and Information Services Center (GES DISC) available at <https://disc.gsfc.nasa.gov/>. Rawinsonde data were downloaded from the University of Wyoming upper air database available at <http://weather.uwyo.edu/upperair/>.

APPENDIX A

Fraction of Measurements that Passed Quality-Control Phase

Figure A1 shows the fraction of measurements that passed preliminary quality control assumptions.

APPENDIX B

Example of Errors Detected in Quality-Control Phase

Figure B1 shows an example of erroneous sounding measurements detected during the quality-control phase.

Example of erroneous sounding measurements detected during quality-control phase

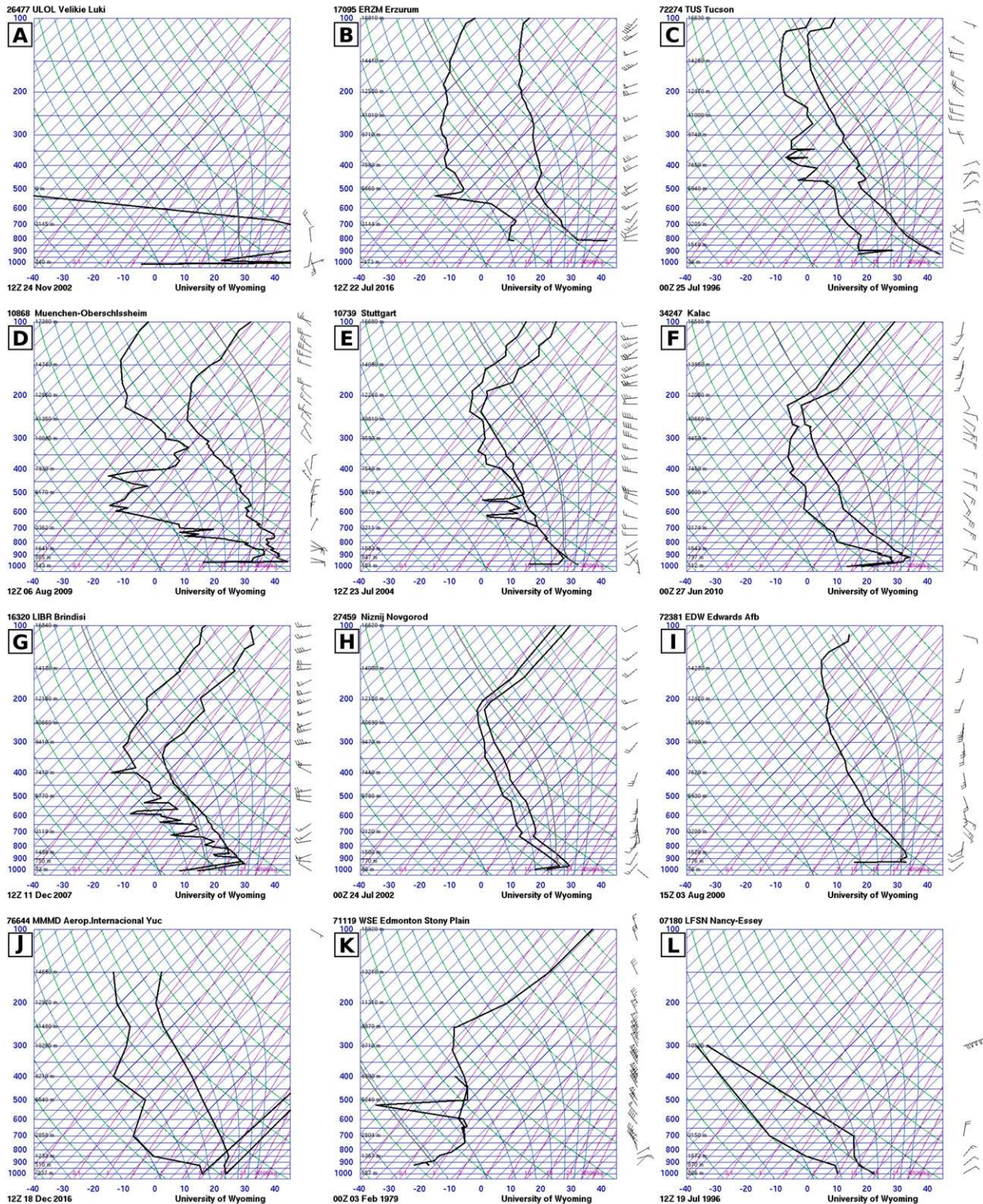


FIG. B1. Example of erroneous sounding measurements detected during the quality-control phase. Plots derived from University of Wyoming sounding database <http://weather.uwyo.edu/upperair/>.

REFERENCES

- 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, <https://doi.org/10.1002/joc.3667>.
- , —, and G. A. Mills, 2011: A severe thunderstorm climatology for Australia and associated thunderstorm environments. *Aust. Meteor. Oceanogr. J.*, **61**, 143–158, <https://doi.org/10.22499/2.6103.001>.
- , M. K. Tippett, and A. H. Sobel, 2015: An empirical model relating U.S. monthly hail occurrence to large-scale meteorological environment. *J. Adv. Model. Earth Syst.*, **7**, 226–243, <https://doi.org/10.1002/2014MS000397>.
- Anderson-Frey, A. K., Y. P. Richardson, A. R. Dean, R. L. Thompson, and B. T. Smith, 2016: Investigation of near-storm environments for tornado events and warnings. *Wea. Forecasting*, **31**, 1771–1790, <https://doi.org/10.1175/WAF-D-16-0046.1>.
- , —, —, —, and —, 2019: Characteristics of tornado events and warnings in the southeastern United States. *Wea. Forecasting*, **34**, 1017–1034, <https://doi.org/10.1175/WAF-D-18-0211.1>.
- Balsamo, G., E. Dutra, C. Albergel, S. Munier, J. C. Calvet, J. Munoz-Sabater, and P. de Rosnay, 2018: ERA-5 and ERA-Interim driven ISBA land surface model simulations: Which one performs better? *Hydrol. Earth Syst. Sci.*, **22**, 3515–3532, <https://doi.org/10.5194/hess-22-3515-2018>.
- Benjamin, S. G., and Coauthors, 2004: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518, [https://doi.org/10.1175/1520-0493\(2004\)132<0495:AHACTR>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0495:AHACTR>2.0.CO;2).
- Bosilovich, M. G., F. R. Robertson, L. Takacs, A. Molod, and D. Mocko, 2017: Atmospheric water balance and variability in the MERRA-2 reanalysis. *J. Climate*, **30**, 1177–1196, <https://doi.org/10.1175/JCLI-D-16-0338.1>.
- Brooks, H. E., J. W. Lee, and J. P. Craven, 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.*, **67–68**, 73–94, [https://doi.org/10.1016/S0169-8095\(03\)00045-0](https://doi.org/10.1016/S0169-8095(03)00045-0).
- , A. R. Anderson, K. Riemann, I. Ebberts, and H. Flachs, 2007: Climatological aspects of convective parameters from the NCAR/NCEP reanalysis. *Atmos. Res.*, **83**, 294–305, <https://doi.org/10.1016/j.atmosres.2005.08.005>.
- Bunkers, M. J., B. A. Klimowski, J. W. Zeitler, R. L. Thompson, and M. L. Weisman, 2000: Predicting supercell motion using a new hodograph technique. *Wea. Forecasting*, **15**, 61–79, [https://doi.org/10.1175/1520-0434\(2000\)015<0061:PSMUAN>2.0.CO;2](https://doi.org/10.1175/1520-0434(2000)015<0061:PSMUAN>2.0.CO;2).
- , J. R. Wetenkamp Jr., J. J. Schild, and A. Fischer, 2010: Observations of the relationship between 700-mb temperatures and severe weather reports across the contiguous United States. *Wea. Forecasting*, **25**, 799–814, <https://doi.org/10.1175/2009WAF2222333.1>.
- Chen, J., A. Dai, Y. Zhang, and K. L. Rasmussen, 2020: Changes in convective available potential energy and convective inhibition under global warming. *J. Climate*, **33**, 2025–2050, <https://doi.org/10.1175/JCLI-D-19-0461.1>.
- Coffer, B. E., M. D. Parker, R. L. Thompson, B. T. Smith, and R. E. Jewell, 2019: Using near-ground storm relative helicity in supercell tornado forecasting. *Wea. Forecasting*, **34**, 1417–1435, <https://doi.org/10.1175/WAF-D-19-0115.1>.
- , M. Taszarek, and M. D. Parker, 2020: Near-ground wind profiles of tornadic and nontornadic environments in the United States and Europe from ERA5 reanalyses. *Wea. Forecasting*, **35**, 2621–2638, <https://doi.org/10.1175/WAF-D-20-0153.1>.
- Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, <https://doi.org/10.1002/qj.776>.
- Coniglio, M. C., and M. D. Parker, 2020: Insights into supercells and their environments from three decades of targeted radiosonde observations. *Mon. Wea. Rev.*, **148**, 4893–4915, <https://doi.org/10.1175/MWR-D-20-0105.1>.
- Copernicus Climate Change Service, 2017: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), accessed 1 January 2020, <https://cds.climate.copernicus.eu/#/search?text=ERA5&type=dataset>.
- Craven, J. P., and H. E. Brooks, 2004: Baseline climatology of sounding derived parameters associated with deep moist convection. *Natl. Wea. Dig.*, **28**, 13–24.
- Davies, J. M., 1993: Hourly helicity, instability, and EHI in forecasting supercell tornadoes. Preprints, *17th Conf. on Severe Local Storms*, Kansas City, MO, Amer. Meteor. Soc., 107–111.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Doswell, C. A., III, and E. N. Rasmussen, 1994: The effect of neglecting the virtual temperature correction on CAPE calculations. *Wea. Forecasting*, **9**, 625–629, [https://doi.org/10.1175/1520-0434\(1994\)009<0625:TEONTV>2.0.CO;2](https://doi.org/10.1175/1520-0434(1994)009<0625:TEONTV>2.0.CO;2).
- , and D. M. Schultz, 2006: On the use of indices and parameters in forecasting severe storms. *Electron. J. Severe Storms Meteor.*, **1** (3), <http://www.ejssm.org/ojs/index.php/ejssm/article/viewArticle/11/12>.
- Gatzen, C., A. H. Fink, D. M. Schultz, and J. G. Pinto, 2020: An 18-year climatology of derechos in Germany. *Nat. Hazards Earth Syst. Sci.*, **20**, 1335–1351, <https://doi.org/10.5194/nhess-2019-234>.
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Gensini, V. A., and W. S. Ashley, 2011: Climatology of potentially severe convective environments from North American Regional Reanalysis. *Electron. J. Severe Storms Meteor.*, **6** (8), <http://www.ejssm.org/ojs/index.php/ejssm/article/viewArticle/85>.
- , and T. L. Mote, 2014: Estimations of hazardous convective weather in the United States using dynamical downscaling. *J. Climate*, **27**, 6581–6589, <https://doi.org/10.1175/JCLI-D-13-00777.1>.
- , and H. E. Brooks, 2018: Spatial trends in United States tornado frequency. *npj Climate Atmos. Sci.*, **1**, 38, <https://doi.org/10.1038/s41612-018-0048-2>.
- , T. L. Mote, and H. E. Brooks, 2014: Severe-thunderstorm reanalysis environments and collocated radiosonde observations. *J. Appl. Meteor. Climatol.*, **53**, 742–751, <https://doi.org/10.1175/JAMC-D-13-0263.1>.
- Grams, J. S., R. L. Thompson, D. V. Snively, J. A. Prentice, G. M. Hodges, and L. J. Reames, 2012: A climatology and comparison of parameters for significant tornado events in the United States. *Wea. Forecasting*, **27**, 106–123, <https://doi.org/10.1175/WAF-D-11-00008.1>.
- Groenemeijer, P. H., and A. van Delden, 2007: Sounding-derived parameters associated with large hail and tornadoes in the Netherlands. *Atmos. Res.*, **83**, 473–487, <https://doi.org/10.1016/j.atmosres.2005.08.006>.

- Gropp, M. E., and C. E. Davenport, 2018: The impact of the nocturnal transition on the lifetime and evolution of supercell thunderstorms in the Great Plains. *Wea. Forecasting*, **33**, 1045–1061, <https://doi.org/10.1175/WAF-D-17-0150.1>.
- Hart, J. A., and W. Korotky, 1991: The SHARP workstation v1.50 users guide. National Weather Service NOAA, U.S. Department of Commerce, 30 pp.
- Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Hoogewind, K. A., M. E. Baldwin, and R. J. Trapp, 2017: The impact of climate change on hazardous convective weather in the United States: Insight from high-resolution dynamical downscaling. *J. Climate*, **30**, 10 081–10 100, <https://doi.org/10.1175/JCLI-D-16-0885.1>.
- Ingrasso, R., P. Lionello, M. M. Miglietta, and G. Salvadori, 2020: A statistical investigation of mesoscale precursors of significant tornadoes: The Italian case study. *Atmosphere*, **11**, 301, <https://doi.org/10.3390/atmos11030301>.
- Johns, R. H., and C. A. Doswell III, 1992: Severe local storms forecasting. *Wea. Forecasting*, **7**, 588–612, [https://doi.org/10.1175/1520-0434\(1992\)007<0588:SLSF>2.0.CO;2](https://doi.org/10.1175/1520-0434(1992)007<0588:SLSF>2.0.CO;2).
- Kaltenböck, R., G. Diendorfer, and N. Dotzek, 2009: Evaluation of thunderstorm indices from ECMWF analyses, lightning data and severe storm reports. *Atmos. Res.*, **93**, 381–396, <https://doi.org/10.1016/j.atmosres.2008.11.005>.
- King, A. T., and A. D. Kennedy, 2019: North American supercell environments in atmospheric reanalysis and RUC-2. *J. Appl. Meteor. Climatol.*, **58**, 71–92, <https://doi.org/10.1175/JAMC-D-18-0015.1>.
- Koster, R. D., and Coauthors, 2016: MERRA-2 input observations: Summary and assessment. NASA/TM-2016-104606/Vol. 46, 51 pp., <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20160014544.pdf>.
- Lee, J. W., 2002: Tornado proximity soundings from the NCEP/NCAR reanalysis data. M.S. thesis, Dept. of Meteorology, University of Oklahoma, 61 pp.
- Lei, Y., H. Letu, H. Shang, and J. Shi, 2020: Cloud cover over the Tibetan Plateau and eastern China: A comparison of ERA5 and ERA-Interim with satellite observations. *Climate Dyn.*, **54**, 2941–2957, <https://doi.org/10.1007/s00382-020-05149-x>.
- Li, F., D. R. Chavas, K. A. Reed, and D. T. Dawson II, 2020: Climatology of severe local storm environments and synoptical-scale features over North America in ERA5 reanalysis and CAM6 simulation. *J. Climate*, **33**, 8339–8365, <https://doi.org/10.1175/JCLI-D-19-0986.1>.
- Liu, N., C. Liu, B. Chen, and E. Zipser, 2020: What are the favorable large-scale environments for the highest-flash-rate thunderstorms on Earth? *J. Atmos. Sci.*, **77**, 1583–1612, <https://doi.org/10.1175/JAS-D-19-0235.1>.
- Markowski, P. M., J. M. Straka, E. N. Rasmussen, and D. O. Blanchard, 1998: Variability of storm-relative helicity during VORTEX. *Mon. Wea. Rev.*, **126**, 2959–2971, [https://doi.org/10.1175/1520-0493\(1998\)126<2959:VOSRHD>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<2959:VOSRHD>2.0.CO;2).
- Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, <https://doi.org/10.1175/BAMS-87-3-343>.
- Mohr, S., and M. Kunz, 2013: Recent trends and variabilities of convective parameters relevant for hail events in Germany and Europe. *Atmos. Res.*, **123**, 211–228, <https://doi.org/10.1016/j.atmosres.2012.05.016>.
- Olauson, J., 2018: ERA5: The new champion of wind power modelling? *Renewable Energy*, **126**, 322–331, <https://doi.org/10.1016/j.renene.2018.03.056>.
- Potvin, C. K., K. L. Elmore, and S. J. Weiss, 2010: Assessing the impacts of proximity sounding criteria on the climatology of significant tornado environments. *Wea. Forecasting*, **25**, 921–930, <https://doi.org/10.1175/2010WAF2222368.1>.
- Pučík, T., P. Groenemeijer, D. Rýva, and M. Kolář, 2015: Proximity soundings of severe and nonsevere thunderstorms in central Europe. *Mon. Wea. Rev.*, **143**, 4805–4821, <https://doi.org/10.1175/MWR-D-15-0104.1>.
- Rädler, A. T., P. Groenemeijer, E. Faust, and R. Sausen, 2018: Detecting severe weather trends using an Additive Regressive Convective Hazard Model (AR-CHaMo). *J. Appl. Meteor. Climatol.*, **57**, 569–587, <https://doi.org/10.1175/JAMC-D-17-0132.1>.
- Randles, C. A., and Coauthors, 2017: The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. *J. Climate*, **30**, 6823–6850, <https://doi.org/10.1175/JCLI-D-16-0609.1>.
- Rasmussen, E. N., 2003: Refined supercell and tornado forecast parameters. *Wea. Forecasting*, **18**, 530–535, [https://doi.org/10.1175/1520-0434\(2003\)18<530:RSATFP>2.0.CO;2](https://doi.org/10.1175/1520-0434(2003)18<530:RSATFP>2.0.CO;2).
- , and D. O. Blanchard, 1998: A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Wea. Forecasting*, **13**, 1148–1164, [https://doi.org/10.1175/1520-0434\(1998\)013<1148:ABCOSD>2.0.CO;2](https://doi.org/10.1175/1520-0434(1998)013<1148:ABCOSD>2.0.CO;2).
- Robinson, E. D., R. J. Trapp, and M. E. Baldwin, 2013: The geospatial and temporal distributions of severe thunderstorms from high-resolution dynamical downscaling. *J. Appl. Meteor. Climatol.*, **52**, 2147–2161, <https://doi.org/10.1175/JAMC-D-12-0131.1>.
- Rodríguez, O., and J. Bech, 2021: Tornadic environments in the Iberian Peninsula and the Balearic Islands based on ERA5 reanalysis. *Int. J. Climatol.*, <https://doi.org/10.1002/joc.6825>, in press.
- Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1058, <https://doi.org/10.1175/2010BAMS3001.1>.
- Sherburn, K. D., and M. D. Parker, 2014: Climatology and ingredients of significant severe convection in high-shear, low-CAPE environments. *Wea. Forecasting*, **29**, 854–877, <https://doi.org/10.1175/WAF-D-13-00041.1>.
- Tang, B. H., V. A. Gensini, and C. R. Homeyer, 2019: Trends in United States large hail environments and observations. *npj Climate Atmos. Sci.*, **2**, 45, <https://doi.org/10.1038/s41612-019-0103-7>.
- Tarek, M., F. P. Brissette, and R. Arsenault, 2020: Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrol. Earth Syst. Sci.*, **24**, 2527–2544, <https://doi.org/10.5194/hess-24-2527-2020>.
- Taszarek, M., H. E. Brooks, and B. Czernecki, 2017: Sounding-derived parameters associated with convective hazards in Europe. *Mon. Wea. Rev.*, **145**, 1511–1528, <https://doi.org/10.1175/MWR-D-16-0384.1>.
- , —, —, P. Szuster, and K. Fortuniak, 2018: Climatological aspects of convective parameters over Europe: A comparison of ERA-Interim and sounding data. *J. Climate*, **31**, 4281–4308, <https://doi.org/10.1175/JCLI-D-17-0596.1>.
- , J. T. Allen, T. Púčik, K. Hoogewind, and H. E. Brooks, 2020: Severe convective storms across Europe and the United States. Part II: ERA5 environments associated with lightning, large hail, severe wind and tornadoes. *J. Climate*, **33**, 10 263–10 286, <https://doi.org/10.1175/JCLI-D-20-0346.1>.
- , —, H. E. Brooks, N. Pilguy, and B. Czernecki, 2021: Differing trends in United states and European severe thunderstorm environments in a warming climate. *Bull. Amer. Meteor. Soc.*, <https://doi.org/10.1175/BAMS-D-20-0004.1>, in press.

- Thompson, R. L., R. Edwards, J. A. Hart, K. L. Elmore, and P. Markowski, 2003: Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Wea. Forecasting*, **18**, 1243–1261, [https://doi.org/10.1175/1520-0434\(2003\)018<1243:CPSWSE>2.0.CO;2](https://doi.org/10.1175/1520-0434(2003)018<1243:CPSWSE>2.0.CO;2).
- , C. M. Mead, and R. Edwards, 2007: Effective storm-relative helicity and bulk shear in supercell thunderstorm environments. *Wea. Forecasting*, **22**, 102–115, <https://doi.org/10.1175/WAF969.1>.
- , B. T. Smith, J. S. Grams, A. R. Dean, and C. Broyles, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part II: Supercell and QLCS tornado environments. *Wea. Forecasting*, **27**, 1136–1154, <https://doi.org/10.1175/WAF-D-11-00116.1>.
- , —, A. Dean, and P. Marsh, 2013: Spatial distributions of tornadic near-storm environments by convective mode. *Electron. J. Severe Storms Meteor.*, **8** (5), <http://www.ejssm.org/ojs/index.php/ejssm/article/viewArticle/125>.
- Tippett, M. K., A. H. Sobel, and S. J. Camargo, 2012: Association of U.S. tornado occurrence with monthly environmental parameters. *Geophys. Res. Lett.*, **39**, L02801, <https://doi.org/10.1029/2011GL050368>.
- , —, —, and J. T. Allen, 2014: An empirical relation between U.S. tornado activity and monthly environmental parameters. *J. Climate*, **27**, 2983–2999, <https://doi.org/10.1175/JCLI-D-13-00345.1>.
- Trapp, R. J., B. A. Halvorson, and N. S. Diffenbaugh, 2007: Telescoping, multimodel approaches to evaluate extreme convective weather under future climates. *J. Geophys. Res.*, **112**, D20109, <https://doi.org/10.1029/2006JD008345>.
- , E. D. Robinson, M. E. Baldwin, N. S. Diffenbaugh, and B. R. J. Schwedler, 2011: Regional climate of hazardous convective weather through high-resolution dynamical downscaling. *Climate Dyn.*, **37**, 677–688, <https://doi.org/10.1007/s00382-010-0826-y>.
- Westermayer, A. T., P. Groenemeijer, G. Pistotnik, R. Sausen, and E. Faust, 2017: Identification of favorable environments for thunderstorms in reanalysis data. *Meteor. Z.*, **26**, 59–70, <https://doi.org/10.1127/metz/2016/0754>.
- Wickham, H., 2016: *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, 213 pp.