The DOI for this manuscript is doi: 10.1175/BAMS-D-15-00256.1

On the climatological use of radar data mosaics: Possibilities and challenges

Frédéric Fabry[1], Véronique Meunier[1], Bernat Puigdomènech Treserras[1], Alexandra Cournoyer[1], and Brian Nelson[2],

[1]McGill University, Montreal, QC, Canada

[2]NOAA

Bulletin of the American Meteorological Society

Corresponding author address:

Frédéric Fabry
Department of Atmospheric and Oceanic Sciences,
McGill University
805 Sherbrooke Street West
Montreal, Quebec, Canada H3A 0B9
Tel.: +1 514-398-3652
Fax: +1 514-398-6115
frederic.fabry@mcgill.ca

ABSTRACT

Continental mosaics of radar data have now been generated for more than twenty years. These offer information on precipitation climatology that is simply not available or archived elsewhere: How often does it rain at any particular location? At what time? And with what intensity distribution? What are the geographical and temporal patterns of precipitation occurrence, formation, and decay? What is the climatology of severe weather? Answers to these questions have value on their own and also invariably trigger more questions about the processes causing these patterns as well as suggest some answers. They also have considerable pedagogical value to illustrate in the classroom the impacts on precipitation of different processes such as sea-land breezes, topography, and seasons.

In this work, U.S. mosaics of radar data from 1996 to 2015 are used to demonstrate the possibilities offered by such a data set. Three topics are touched: a) climatologies and daily cycles of precipitation and convection, and what they can teach us about precipitation mechanisms; b) the spatial and temporal distribution of the appearance and occurrence of convection, and what it reveals on the importance of surface terrain properties for these events; and c) the power and challenges of looking for a small signal in even such a large dataset using the influence of weekly activity cycles and of cities on precipitation as an illustration.

CAPSULE

Where the radar climatology of weather echoes is used to reveal how surface properties shape precipitation occurrence and to explore the ease or difficulty to unambiguously detect these effects.

1. Precipitation climatology and radars

Radar has historically transformed the way we study storms thanks to its ability to take frequent and regular 3-D measurements even through clouds and precipitation. As a result, it is commonly used both operationally for weather surveillance as well as for research to help understand the dynamics and microphysical processes of atmospheric phenomena (Atlas et al. 1990; Wakimoto and Srivastava 2003; Fabry 2015).

The first national Doppler radar network in the world was deployed in the United States in the mid-1990s. More importantly, a framework and process for monitoring and maintaining radar data quality was implemented and adhered to since. From late 1995 onwards, the reflectivity data from all these radars have been made into national mosaics by a variety of actors, including private companies, research institutes, and the National Weather Service itself. A unique dataset now exists to study radar echoes collected by the same radars over a period of more than 20 years (and counting) over the contiguous United States.

Though country-wide climatological information on precipitation exists, for example from the US Climate Normals (Applequist et al. 2012, Arguez et al. 2012), the information available is not as rich as it could be. As an illustration, we challenge all readers to find the answer to a simple climatology question: what fraction of the time does it rain or snow at your location (by opposition to how many days per year)? Historically, the data required to answer such a basic question have not been available primarily because even though the existing technology could have been used, the detailed information required to derive a statistic like this was not archived. Radar data offer information that is simply not available or archived elsewhere: how often does it rain at any particular location? At what time? And with what intensity distribution? What are the geographical and temporal patterns of precipitation occurrence, formation, and decay? What is the climatology of severe weather? Answers to these questions invariably trigger more questions about the processes causing these patterns as well as suggest some answers. These tend to be of a different nature than those arising from individual case studies because the specificity of atmospheric conditions leading to one storm instead of another loses its significance. What is left are the persistent features that often or always influence precipitation occurrence, which, in the end, are the most important to get right both in the context of process studies and of numerical modeling. Many of those are the result of variations in terrain type and orography. We have also found several of these results to be extremely powerful illustrations of the effects of atmospheric phenomena taught in classes such as sea-land and mountain-valley breezes, lake-effect snow, diurnal cycles and spatial patterns of convection climatology, among others.

While radar climatologies have been attempted early on in radar meteorology (Riggs and Truppi 1957) and on and off since (e.g., references in Arnold 2005; Wilson 1977), it is only thanks to the work of Richard Carbone and colleagues that it has achieved a timid rebirth in the United States (Carbone et al. 2002; Carbone and Tuttle 2008), followed by a few efforts here and elsewhere (e.g., Parker and Knievel 2005; Overeem et al. 2009; Mohee and Miller 2010; Weckwerth et al. 2011; Fairman et al. 2015, 2016; Lock and Houston 2015), the focus being primarily on precipitation mapping and convection studies, the natural strengths of radar.

Of course radar data processing and interpretation are fraught with complications. Are all radars properly calibrated? Have the data been properly cleaned of ground echoes, of insects, of birds? Is radar coverage sufficient everywhere? Are there range or topography dependent biases? These questions both complicate the interpretation of a radar echo climatology and can also be partially answered by it (see the sidebar on *Data, Processing, and Quality Issues* for some details). In parallel, radar has unique strengths, in particular for measuring the coverage and instantaneous intensity of precipitation, more so than for quantifying precipitation accumulation.

Given the strengths and expected limitations of the available dataset, we strove to use the radar data to provide otherwise unavailable climatological information as opposed to try to displace existing good quality products such as those derived from dense gauge networks. We first focused our attention on data quality and echo coverage issues, as they determine what can and cannot be achieved with radar mosaic maps. Next we studied phenomena and processes as well as used approaches for which these complications would be minimized, such as convection-related topics that are less sensitive to data coverage issues or contamination by weak echoes, and diurnal cycles that naturally cancel time-invariant biases.

2. Building a radar climatology

For reasons of simplicity, and because we did not have easy access to the raw radar data for the whole U.S. over such a long period, we have chosen to build the radar echo climatology from existing mosaics. But the capabilities of radars collecting the data have changed, and so has the process of cleaning radar data and making them into a national mosaic. We must hence contend with radar mosaic maps generated in real-time whose recipe has changed over the years (Table 1). This changing process with time made us shy away from studying trends over the 20-year period.

In the end, mosaics from two sources were combined for this climatology. The first (October 1995 to August 2007) is the Weather Services International (WSI) NowRAD MASTER15 mosaic with a latitude-dependent spatial resolution of approximately 2 km and a reflectivity resolution of 5 dBZ until 2001 (Zhang et al. 2015) and 1 dB afterwards. The second (September 2007 to December 2015) was made by the Warning Decision Support System–Integrated Information (WDSS-II; Lakshmanan et al. 2006, 2007) and has a resolution of spatial approximately 1 km. Both data sets attempt to characterize the echo strength and coverage in the lower troposphere, preferably free of non-meteorological echoes. Mosaics were analyzed at 15 min resolution, to "limit" the analysis to just under 700,000 radar maps. Because of changes in data and its processing over time, we first need to examine how realistic the derived statistics look like.

To get a first feel for the overall quality of the radar mosaic data, a 20-year precipitation accumulation was computed from them and compared with an analysis derived from gauges over the same period (Fig. 1, see also Fig. SB1). Here, gauge-based accumulations were available over land areas (Fig. 1b) and simplistically extrapolated over water using a 1/distance weighting. What it confirms is that in the eastern two-thirds of the conterminous United States, with the exception of the Appalachian area, radar-based precipitation R_{radar} and gauge-based precipitation R_{gauge} are comparable enough (equivalent to reflectivity biases of less than 2.5 dB) that meaningful intensity statistics can be derived there. In mountainous area, a combination of radar beam blockage and measurements far away from the ground surface limit the usefulness of radar data climatology.

3. Occurrence and intensity of precipitation

A first illustration of the kind of information retrievable by years of radar data is a set of maps of the likelihood of observing surface precipitation with different reflectivities (Fig. 2). Because such statistics are likely to be wrong if systematic biases caused by beam blockage and frequent under- or over-estimation aloft affect the data, we masked areas where radar-estimated precipitation differ too much from gauge-estimated precipitation. We arbitrarily chose to stripe in gray areas that did not meet the criteria $2/3\, R_{gauge} < R_{radar} < 3/2\, R_{gauge}$, as we believe we could not trust derived statistics outside of that interval. Precipitation exceeding the reflectivity of light snow and moderate drizzle ($Z \geq 5$ dBZ, corresponding to about 0.1 mm/h, Fig. 2a) is most frequent in mid-latitude regions to the north, especially near the oceans or the Great Lakes area. It is observed on average 3 hrs a day immediately east of each Great Lakes and 4 hrs a day just east of Seattle on the foothills of the Cascades, but 30 mins a day in Los Angeles and 1.25 hrs in Miami. Note that the "bullseyes" patterns around each radar in the Great Lakes area primarily reflect the difficulty of the mosaics to correctly account for weak snow and drizzle at far ranges. As we increase the reflectivity threshold, the area of higher occurrence shifts southward. Significant convective rainfall (≥ 45 dBZ, corresponding to about 20 mm/h) is rarely observed on the West Coast, detected 3.5 hours per year in Buffalo, but 16 hours per year in Miami. If we further increase the threshold to 60 dBZ (Fig. 2c), a reflectivity that can only be associated with hail (Fabry 2015), the peak of occurrence shifts towards the west of the Central Great Plains where it averages 10 mins per year. Interestingly, the map compares well with that of severe hail occurrence made by the Storm Prediction Center (SPC) from 48 years of significant hail reports (available at the time of this writing at http://www.srh.noaa.gov/images/oun/spotter/sighail.jpg), except that it shifts the hail capital away from central Oklahoma and is more aligned with the much shorter hail climatology of Cintineo et al. (2012).

Precipitation occurrence has a strong annual cycle, and this is well documented in precipitation climatology maps. The frequency at which convection occurs also follows an annual cycle, but different areas see a peak in convection at different times of the year (Fig. 3). For example, we know that convection peaks in late spring in the Central Plains when upper-level support is still important, later elsewhere for which strong upper level support is less critical to the occurrence of convection. Thanks to images like Fig. 3, the results of all these processes can be nicely illustrated.

While no truly surprising results came out of this exercise, this section illustrates the value of using long-term statistics derived from radar mosaics for meteorological teaching purposes. We will now shift our attention towards convection occurrence.

4. Convection occurrence and diurnal cycle

Convective rain has a strong diurnal cycle. The diurnal cycle of summer convection in the continental United States (Fig. 4, electronic supplement) has become the classic result of radar-based climatology since Carbone et al. (2002, 2008). Figure 4 illustrates how convection forms at various locations during daytime, in particular over the Rockies, and later on the Great Plains, and then travels eastward during the night. Particularly striking for basic meteorology teaching is the effect of sea- and land-breezes on the timing of convection from the Gulf coast of Texas to the Carolinas, as well as the local hotspots forming immediately east of peaks of the Rocky Mountains where convection starts first from 18:00 UTC fed by valley breezes. The electronic supplement showing an animation of the diurnal cycle of convection in the warm season is particularly telling, and has a richness that is difficult to describe; if a picture is worth a thousand words, that particular animation could feed a small textbook.

Among the remarkable results from the diurnal cycle of convection is the rapid morning decimation of nighttime convection, especially in the Midwestern United States. On average, during the night, convection is tracking eastward with only a very gradual decay as can be seen from the limited change in echo patterns between Fig. 4f and Fig. 4a. This is likely associated with the maintenance of convective instability over long periods thanks to low-level jets (e.g., Uccellini and Johnson 1979; Kumjian et al. 2006; Coniglio et al. 2007). The local maximum in convection occurrence moving from the Great Plains however decays very rapidly in the morning in the Midwest, showing that the added solar energy destroyed the support for nighttime convection well before support for daytime convection can be re-established. Of particular interest is that convection occurrence seems to diminish particularly along some specific corridors such as the Mississippi, lower Missouri, and Ohio river valleys, perhaps because the descent branches of the solenoid circulations associated with these valleys either suppress the advecting storms or prevent the replacement of older naturally decaying storms by fresh new ones.

The various processes affecting the diurnal cycle of convection also shape the time at which convection is most likely to be observed (Fig. 5a): Morning over the warm waters of the south, early afternoon just east of major peaks in the Rockies and on the southern coasts, late afternoon in the east, in the night in the Central Plains and over the Great Lakes, with two weak maxima being observed in the Midwest. In addition of being of meteorological interest, this information could have practical importance, such as for hazard preparedness purposes: for example, if flash flooding is more likely to occur at night in some areas, additional training may be needed for the nighttime flood management crews who are the most likely to face a difficult situation.

Patterns of time of peak convection occurrence such as Fig. 5a arise from the blend of two somewhat different phenomena: “daytime” convection where surface heating plays a critical role, and “nighttime” convection where atmospheric destabilization is dominated by processes occurring aloft. If we want to focus on only one phenomenon, say daytime convection, we found that looking for temporal maxima during the day gives a misleading picture. For such a purpose, focusing on the time of fastest intensification of convection occurrence proved to be a better alternative, though it is a considerably noisier quantity to estimate. By computing the rate of increase in occurrence of convection over 4-hr windows, we were able to obtain Fig. 5b that illustrates how daytime convection starts earlier in some areas compared to others. In particular, over the Great Plains and east of the Appalachians, convection generally starts in late afternoon instead of in early afternoon in other regions away from strong orography and the influence of water bodies. There are also many other smaller-scale patterns whose statistical and physical significance remains uncertain.

What we also found remarkable is that it does not require a large topographic feature to affect the occurrence of convection. Changes in the timing and frequency of occurrence of convective events (Fig. 2b) occur associated with lakes and topographic features that are not very large: for example, Lake Pontchartrain (southern Louisiana) reduces afternoon thunderstorm occurrence (Fig. 5a) while the Cumberland Plateau (eastern Tennessee) experiences an earlier onset of daytime thunderstorms than neighboring areas (Fig. 5b). Other man-made reservoirs may also make such changes (Haberlie et al. 2016). One of the largest unexpected signature found on such mosaic maps was a local minimum in mid-summer afternoon convection associated with the Mississippi valley where the combination of a) weaker initial static stability as the surrounding elevated terrain protruded above the nocturnal inversion and b) the mechanical lifting of impinging flow over the terrain stimulated convection around the valley and created a local minimum within it (Kirshbaum et al. 2016). Many more possibly significant local signatures can be seen on this map that are not clearly associated with definite topographic features and may deserve to be studied.

5. The challenge of finding meaningful signals

The above being said, the search for meaningful signal from long-term radar data is fraught with challenges. Some are related to the measurement of rainfall from radar mosaics: The location of radar with respect to the features of interest, data availability, terrain blockage, ground clutter, and the vertical profile of reflectivity all introduce biases and other artifacts in the radar data, many of which can be seen on the maps in Figs. 1 and 2. To help control for these artifacts, the use of complementary data that have different measurement problems such as lightning maps helps. Other challenges are due to properties of atmospheric patterns themselves.

An intrinsic property of atmospheric and geophysical fields is that they are correlated in space. For example, if it has been anomalously wet in New York City, it has very likely been the case in Newark 15 km to the west, and probably also in Philadelphia 130 km to the southwest. These fields are also correlated in time: if it rains now, there is a much higher chance than climatology that it will rain an hour from now, or even a day from now. The extent of the correlation of precipitation patterns in time can be illustrated using power spectra of radar-derived precipitation such as the one in Fig. 6: As long as power spectra have a non-zero slope, the patterns observed at one time scale are correlated with those at other scales. The correlations in space and time are clearly linked: it is generally the same propagating weather systems, or the same instabilities and forced disturbances, that will alternatively dictate precipitation patterns in Philadelphia and Newark before they affect New York City.

A consequence of the spatiotemporal correlations in atmospheric fields is that it creates many convincing-looking patterns by chance that make the detection of meaningful signals more challenging. Classical statistical approaches rely on two assumptions generally violated in atmospheric fields, independence and stationarity. Independence between samples at two locations or over two periods implies that events affecting one location do not affect the other; with weather systems extending up to continental scales and oscillations such as the El Niño Southern Oscillation lasting years, this is clearly not the case. Stationarity implies that statistical properties such as mean and variance do not vary over time, while they clearly vary over the course of seasons and years. Independence of samples and stationarity of standard deviations are the bedrock on which are based statistical tests such as the computation of p-values (the probability that two samples could occur by chance from one process having a unique mean and variance) as well as analysis approaches such as the resampling of data sets to generate new possible samples using bootstrapping or permutation methods (Efron and Tibshirani 1992; Manly 2006). The net result is that if these tests are not run appropriately, it is very easy to wrongly find that two samples are unlikely to come from one process when in fact they may (Daniel et al. 2012).

The search for a significant weekly cycle in precipitation using remote-sensed data (e.g., Bell et al. 2008; Tuttle and Carbone 2011) provides such an example. If you were to look for the difference between weekday and weekend precipitation occurrence using the 20-yr period between 1996 and 2015, you would get a map like Fig. 7a: In the northeast, near the area of peak deposition of sulfates and nitrates associated with combustion (e.g., Zhang et al. 2012), precipitation occurrence is 10-15% more frequent on Tuesdays to Fridays than on Saturdays to Mondays. If we focus on severe convection occurrence (Fig. 7b), or on rainfall accumulations associated with convection that show similar patterns, there is an overall tendency for greater occurrence of 50 dBZ echoes in the Gulf Coast on weekdays than on weekends. Having gotten very excited ourselves by such a finding (Fabry et al. 2013) and its possible link with the weekly cycle of particulate emissions, we felt the need to investigate it further.

The uncertainty and the spatial variability of patterns of precipitation are difficult to study quantitatively because of the episodic yet spatiotemporally correlated nature of precipitation as well as its non-Gaussian statistics. Given such a beast, one of the best and most common technique to study the significance of a signal uses the bootstrap method: At each location, the available sample of data, here the 20 years of reflectivity data, is resampled in two or more categories multiple times to generate a large number of plausible time series. In this example, plausible datasets of two categories, weekdays (Tuesdays to Fridays) and weekends (Saturdays to Mondays), are created by randomly resampling available data on those days. These new plausible datasets of similar length to the original one are then used to evaluate the likelihood with which the two categories can have similar or different values and be statistically different with a certain probability. What Fig. 6 reveals is that cycles of seven days are not quite on the flat section of the power spectra of precipitation, which implies that there is some correlation left between successive seven day cycles. In other words, given the weather on a particular Sunday, the weather on the next Wednesday will be a subset of the weather expected for all Wednesdays, even given the same climatology. Hence, if for a given Sunday any Wednesday is chosen as part of a data resampling process, the difference between Wednesdays and Sundays will be overestimated compared to what it can be in reality. Therefore, when resampling the dataset, it is essential to do it in blocks that are longer than a week to ensure that a plausible Wednesday follows the chosen Sunday. For the resampling to have any value though, there must be enough useful data blocks to get some useful randomization. What our experience and that of others (Daniel et al. 2012) show is that two-week blocks are a good compromise. In parallel, it is also essential to sample all months proportionally in order to respect the climatology of annual cycles. When these two factors are properly taken care of, it is found that the minimum in weekend precipitation in the North East is a 2-σ event (p=0.05), hardly unexpected to occur by chance given the number of mostly independent local maxima and minima one can observe on this map. But the fact that this signature occurred at a physically plausible location pushed us to continue looking for clues. Following additional investigations, several other factors reinforced the likelihood that this pattern is an accidental signature: a) Signatures of similar strengths can be obtained when looking for meaningless 6-day cycles (Figs. 7b and 7d), even if the rainfall patterns are slightly more correlated over six day than over seven day periods; b) the power spectra of precipitation occurrence and amount show no peak for 7-day cycles (inset of Fig. 6); and, c) an analysis of gauge data going back further in time shows that the difference observed in the past 20 years in the North East has been an anomaly even if aerosols were as much or more prevalent 30 years ago. Noting again well after Thomas Henry Huxley (1822-1895) that "*the* *great tragedy of Science [is] the slaying of a beautiful hypothesis by an ugly fact*", we finally accepted that this particularly enticing signature was probably a fluke.

Trying to learn from that experience, we wondered how strong a locally-forced precipitation pattern has to be to be detected amidst patterns caused by natural variability. To answer this question, we must first determine the magnitude of that natural variability, that we define as the one associated with the random passage of weather events, separate from the one associated with the spatial or temporal variability of climate. Climate-related variability is expected to have long time scales (from subseasonal periods to years), and it has left its mark on the power spectra of Fig. 6 through a strong annual peak (with its second and third harmonics) together with gradually increasing variability with increasing years. Weather systems have at most continental scales and affect a given area for a few days at most. They are characterized by smaller-scale structures embedded within larger ones as illustrated by the sloped power spectra for periods shorter than a few days. In between the "weather" and the "climate" regime, the temporal variability of precipitation is dictated by the mostly random uncorrelated sequence of weather events as illustrated by the constant power spectrum. This peculiar split of frequency between weather and climate variability can be taken advantage of to estimate the magnitude of the variability in precipitation associated with weather.

Let us assume that the precipitation sampled at each location over 20 years is on average a standard deviation σ away from the true climatology that would have been measured with an infinitely long dataset in an unchanging climate. Let us split the available sample into two halves, alternatively binning one week of data in one category and the next week of data in the other, as if we were trying to look for a 14-day cycle. What this process does is to separate equally among the two halves the variability in time associated with climate processes and with changes in data processing over the years, while randomly assigning the spatio-temporal variability of weather events among the two halves. Since the weather component, responsible for the average to be a standard deviation σ away from the true climatology in the original dataset, is now split in two independent half-samples, the average of each of those half samples will now be on average $\sqrt{2}\sigma$ away from the true climatology, the variance on the average of these half-sized datasets being doubled. If we then subtract those two half-sample averages, the result will have a standard deviation of 2σ spatially because the variance on the result of the subtraction is doubled again. The resulting field will have numerous maxima and minima than can be used to estimate the correlation matrix on the "noise" induced by weather on the climatology while the local average-squared value can be used to estimate $(2\sigma)^2$ and then 2σ. This can be repeated for cycle periods slightly greater or smaller than 2 weeks to get additional semi-independent estimates of 2σ.

We applied this procedure to evaluate the significance of the effect of cities on the occurrence of severe convection (Lowry 1998; Sheppard 2005). We first looked for all US cities above 1 million inhabitants that were away from oceans and lakes whose breeze would confuse the precipitation analysis, and also away from mountains and other obstacles causing beam blockage and significant clutter. The resulting 13 cities selected are hence mostly concentrated on the eastern half of the continent and away from the coasts. The summer data (May to August) for the 13 selected cities (Atlanta, Birmingham, Cincinnati, Columbus, Dallas, Denver, Indianapolis, Kansas City, Memphis, Minneapolis, Nashville, Oklahoma City, and Pittsburgh) were then averaged with the city center in the middle. The resulting map, shown in Fig. 8a, illustrates that over and immediately east of cities, severe convection occurs 10% more frequently than in surrounding areas (0.057 % of the time over and east of cities compared to 0.052% of the time around the cities). If we perform the half samples differencing test, we find that the 2σ uncertainty on this pattern is about 0.0023%, making the strength of the city signal for each pixel four standard deviations above the expected uncertainty. With precipitation instead of severe convection, the signature of cities is five standard deviations above the expected uncertainty due to the random passage of weather events. What this suggests is that the signature we observe above cities and east of cities is significant in the dataset, but the cause of the signature could have many origins. To confirm that this signature is due to the effect of cities, we computed similar mosaics for nighttime (0:00-8:00 solar) and afternoon (12:00-20:00), and found that the signature was strongest in the afternoon when the air affected by the city can feed storms, smallest at night when inversions tend to isolate storms from surface influences (Figs. 9b and 9c). Finally, to ensure that measurement biases are not fooling us, radars tending to gravitate not far from cities after all, we performed the same mosaics with 23 years of lightning data (1990-2012, NCDC 2012) and obtained similar results (Figs. 8d-8f). The convection and precipitation enhancement associated with cities appears to be real.

An interesting observation we can make on Fig. 8a is that even though the 2σ-uncertainty due to weather is only 0.0023%, spatial patterns of larger amplitudes can be observed on the mosaic maps. Those patterns are caused by all other confounding effects from varying radar coverage to variations in precipitation climatology caused by other processes that we tried to control for by selecting cities and averaging them. The 4σ-signature observed only appeared because we combined the data from 13 cities and had an effective 250-yr dataset to analyze; with only 20 years of data even at such high resolution, it would be difficult for the influence of individual cities to exceed the expected variability caused by the random passage of storms. This result partly explains the lack of consistency of findings obtained on the influence of cities on precipitation.

6. Future

The derivation of precipitation and convection statistics done above is only a small sample of what is possible to do with many years of radar data over large areas. Recently, a reanalysis of radar data combined with other data sources (Ortega et al. 2015) has become available and adds Doppler information, while other efforts seek to better combine the instantaneous estimates of radar with the stability of gauges (e.g., Nelson et al. 2010). These represent our most complete information on severe storms and their evolution, and possibilities are limitless for people with the imagination and drive to mine such a dataset. What will you do with it?

Acknowledgements

Sincere thanks to everyone who participated in the creation of this dataset and its collection at McGill, including Weather Services International, the Warning Decision Support System -- Integrated Information group, and Weather Decision Technologies for their great work creating and relaying these mosaics, and Urs Germann and Marc Berenguer for helping with the dataset acquisition at McGill University. This project was undertaken with the financial support of the Government of Canada provided through the Department of the Environment and Climate Change as well as through the Natural Science and Engineering Research Council of Canada.

References

Applequist, S., A. Arguez, I. Durre, M.F. Squires, R.S. Vose, and X. Yin, 2012: 1981–2010 U.S. hourly normals. *Bulletin of the American Meteorological Society*, **93**, 1637–1640, doi:10.1175/BAMS-D-11-00173.1.

Arguez, A., I. Durre, S. Applequist, R.S. Vose, M.F. Squires, X. Yin, R.R. Heim, Jr., and T.W. Owen, 2012: NOAA's 1981–2010 U.S. climate normals: An overview. *Bulletin of the American Meteorological Society*, **93**, 1687–1697, doi:10.1175/BAMS-D-11-00197.1.

Arnold, D.L., 2005: Radar, climatic applications. In the *Encyclopedia of World Climatology*, Oliver, J.E. (Ed.), Springer, doi:10.1007/1-4020-3266-8_166.

Atlas, D. (Ed.), 1990: *Radar in Meteorology*. American Meteorological Society, Boston, 806 pp.

Bell, T.L., D. Rosenfeld, K.-M. Kim, J.-M. Yoo, M.-I. Lee, and M. Hahnenberger, 2008: Midweek increase in U.S. summer rain and storm heights suggests air pollution invigorates rainstorms. *J. Geophys. Res.*, **113**, D02209, doi: 10.1175/1520-0469(2002)059<2033:IOPAWW>2.0.CO;2.

Carbone, R.E., J.D. Tuttle, D. Ahijevych, and S.B. Trier, 2002: Inferences of predictability associated with warm season precipitation episodes. *J. Atmospheric Sci.*, **59**, 2033–2056, doi:10.1175/1520-0496(2002)059<2033.

Carbone, R.E., and J.D. Tuttle, 2008: Rainfall occurrence in the U.S. warm season: the diurnal cycle. *Journal of Climate*, **21**, 4132–4146, doi:10.1175/2008JCLI2275.1.

Cintineo, J. L., T. M. Smith, V. Lakshmanan, H. E. Brooks, and K. L. Ortega, 2012: An objective high-resolution hail climatology of the contiguous United States. *Wea Forecasting*, **27**, 1235–1248, doi:10.1175/WAF-D-11-00151.1.

Clayton, H. H., 1894: Six and seven day weather periods. *Amer. J. Science*, 3rd ser., **47**, 223–231.

Coniglio, M. C., H. E. Brooks, S. J. Weiss, and S. F. Corfidi, 2007: Forecasting the maintenance of quasi-linear mesoscale convective systems. *Wea. Forecasting*, **22**, 556–570, doi: 10.1175/WAF1006.1.

Daniel, J., R. Portmann, S., Solomon, and D. Murphy, 2012: Identifying weekly cycles in meteorological variables: The importance of an appropriate statistical analysis. *J. Geophys. Res.*, **117**, D13203, doi:10.1029/2012JD017574.

Efron, B., and R. J. Tibshirani, 1992: *An Introduction to the Bootstrap*. Chapman and Hall/CRC, 480 pp.

Fabry, F., 2015: *Radar Meteorology – Principles and Practice.* Cambridge University Press, 272 pp.

Fabry, F., Q. Cazenave, and R. Basivi, 2013: Echo climatology, impact of cities, and initial convection studies: New horizons opened using 17 years of conterminous U.S. radar mosaics. *Proceedings*, 36th Conf. on Radar in Meteorology, Breckenridge CO, 16-20 September 2013, paper 10.1, https://ams.confex.com/ams/36Radar/webprogram/Manuscript/Paper228783/EchoClimatoEtcRad36.pdf.

Fairman, J. G. Jr., D. M. Schultz, D. J. Kirshbaum, S. L. Gray, and A. I. Barrett, 2015: A radar-based rainfall climatology of Great Britain and Ireland. *Weather*, **70**(5), 153–158, doi:10.1002/wea.2486.

Fairman, J. G. Jr., D. M. Schultz, D. J. Kirshbaum, S. L. Gray, and A. I. Barrett, 2016: Climatology of banded precipitation over the contiguous United States. *Mon. Wea. Rev.*, **144**, 4553–4568, doi:10.1175/MWR-D-16-0015.1.

Haberlie, A. M., W. S. Ashley, A. J. Fultz, and S. M. Eagan, 2016: The effect of reservoirs on the climatology of warm-season thunderstorms in Southeast Texas, USA. *Int. J. Climatol.*, **36**, 1808–1820. doi:10.1002/joc.4461.

Horton, R.E., 1941: An approach toward a physical interpretation of infiltration-capacity. *Soil Science Society of America Journal*, **5**(C), 399–417.

Joss, J., and A. Waldvogel, 1970: A method to improve the accuracy of radar-measured amounts of precipitation. *Preprints*, 14th Radar Meteorology Conf., Tucson, AZ, Amer. Meteor. Soc., 237–238.

Kirshbaum, D., F. Fabry, and Q. Cazenave, 2016. The Mississippi Valley convection minimum on summer afternoons: Observations and simulations. *Mon. Wea. Rev.*, **144**, 263–272, doi:10.1175/MWR-D-15-0238.1.

Kumjian, M., J. Evans, and J. Guyer 2006: The relationship of the Great Plains low level jet to nocturnal MCS development. 23rd Conf. on Severe Local Storms, AMS, St. Louis, MO.

Lakshmanan V., T. Smith, K. Hondl, G.J Stumpf, and A. Witt, 2006: A real-time, three dimensional, rapidly updating, heterogeneous radar merger technique for reflectivity, velocity and derived products. *Wea. Forecasting*, **21**, 802–823, doi: 10.1175/WAF942.1.

Lakshmanan, V., A. Fritz, T. Smith, K. Hondl, and G. Stumpf, 2007: An automated technique to quality control radar reflectivity data. *J. Applied Meteorol. Climatol.*, **46**, 288–305, doi:10.1175/JAM2460.1.

Lock, N.A., and A.L. Houston, 2015: Spatiotemporal distribution of thunderstorm initiation in the US Great Plains from 2005 to 2007. *Int. J. Climatol.*, **35**, 4047–4056, doi: 10.1002/JOC.4261.

Lowry, W.P., 1998: Urban effects on precipitation amount, *Prog. Phys. Geogr.*, **22**, 477–520, doi:10.1177/030913339802200403.

Manly, B. F. J., 2006: *Randomization, Bootstrap and Monte Carlo Methods in Biology, Third Edition*. Chapman and Hall/CRC, 480 pp.

Mohee, F.M., and C. Miller, 2010: Climatology of thunderstorms for North Dakota, 2002–06. *J. Applied Meteor. Clim.*, **49**, 1881–1890, doi:10.1175/2010JAMC2400.1.

National Climate Data Center, 2012: Lightning Products and Services – Gridded Summaries. Accessed July 2013. [Available online at: http://www.ncdc.noaa.gov/data-access/severe-weather/lightning-products-and-services]

Nelson, B.R., D-J. Seo, and D. Kim, 2009: Multisensor precipitation reanalysis. *J. Hydrometeor.*, **11**, 666–682, doi:10.1175/2010JHM1210.1.

Ortega, K. L., T. M. Smith, S. E. Stevens, S. S. Williams, D. M. Kingfield, and R. A. Lagerquist, 2015: The multi-year reanalysis of remotely sensed storms (MYRORSS): Data processing and severe weather projects. Presentation at the 37th Conf. on Radar Meteorology, Norman OK, 14-18 September 2015, https://ams.confex.com/ams/37RADAR/webprogram/Handout/Paper275486/205_ortega_et_al_myrorss.pdf.

Overeem, A., I. Holleman, and A. Buishand, 2009: Derivation of a 10-year radar-based climatology of rainfall. *J. Applied Meteor. Clim.*, **48**, 1448–1463, doi: 10.1175/2009JAMC1954.1.

Parker, M. D., and J. C. Knievel, 2005: Do meteorologists suppress thunderstorms?: Radar-derived statistics and the behavior of moist convection. *Bull. Amer. Meteor. Soc.*, **86**, 341–358, doi:10.1175/BAMS-86-3-341.

Parker, M. D., and D. A. Ahijevych, 2007: Convective episodes in the East-Central United States. *Mon. Wea. Rev.*, **135**, 3707–3727, doi:10.1175/2007MWR2098.1.

PRISM, 2016: Annual- and monthly-averaged precipitation dataset AN81m at 4 km resolution over the conterminous United States. Subset used: 1996–2015, http://www.prism.oregonstate.edu/recent/, accessed 30 May 2016.

Riggs, L.P, and L.E. Truppi, 1957: A survey of radar climatology. Proceedings of the Sixth Weather Radar Conference, Cambridge, MA, March 26-28 1957, 227-232.

Sanchez-Lorenzo, A., P. Laux, H.-J. Hendricks-Franssen, J. Calbo, S. Vogl, A. K. Georgoulias, and J. Quaas, 2012: Assessing large-scale weekly cycles in meteorological variables: A review. *Atmos. Chem. Phys. Discuss.*, **12**, 1451–1491, doi:10.5194/acpd-12-1451-2012.

Shepherd, J. M., 2005: A review of current investigations of urban-induced rainfall and recommendations for the future. *Earth Interact.*, **9**, 1–12, doi:10.1175/EI156.1.

Tuttle, J. D., and R. E. Carbone, 2011: Inferences of weekly cycles in summertime rainfall. J. Geophys. Res., **116**, D20213, doi:10.1029/2011JD015819.

Uccellini, L. W., and D. R. Johnson, 1979: The coupling of upper and lower tropospheric jet streaks and implications for the development of severe convective storms. *Mon. Wea. Rev.*, **107**, 682-703, doi:10.1175/1520-0493(1979)107<0682:TCOUAL>2.0.CO;2.

Wakimoto, R.M., and R. Srivastava (Eds.), 2003: *Radar and Atmospheric Science – A Collection of Essays in Honor of David Atlas*. American Meteorological Society, Boston, 270 pp.

Weckwerth, T.M., J.W. Wilson, M. Hagen, T.J. Emerson, J.O. Pinto, D.L. Rife, and L. Grebe, 2011: Radar climatology of the COPS region. *Quart. J. Royal Meteor. Soc.*, **137**, 31–41, doi:10.1002/QJ.747.

Wilson, J.W., 1977: Effect of Lake Ontario on precipitation. *Monthly Weather Review*, **105**, 207–214, doi:10.1175/1520-0493(1977)105<0207.

Zhang, L., D. J. Jacob, E. M. Knipping, N. Kumar, J. W. Munger, C. C. Carouge, A. van Donkelaar, Y. X. Wang, and D. Chen, 2012: Nitrogen deposition to the United States:

Distribution, sources, and processes. *Atmos. Chem. Phys.* **12**, 4539–4554, doi: 10.5194/acp-12-4539-2012.

Zhang, Y., D. Kitzmiller, D. Seo, D. Kim, and R. Cifelli, 2015: Creation of multisensor precipitation products from WSI NOWrad reflectivity data. *J. Hydrol. Eng.*, E4015001, doi: 10.1061/(ASCE)HE.1943-5584.0001216.

Sidebar (Appendix)

Data, Processing, and Quality Issues

Radar measures the echo strength, or the equivalent radar reflectivity factor (often simply called “reflectivity”) from all the targets large enough to be detected. The actual reflectivity of a target depends on its nature (rain, snow, insects, birds…) and its vertical structure (affected by precipitation growth, the presence of melting particles, etc.). Our ability to measure that reflectivity is affected by, among others, radar sensitivity, calibration, and scanning strategy, blockage by obstacles, and how chirurgically ground clutter can be removed without affecting the echo strength from other targets. As a result, raw reflectivity radar images and statistics derived from them can be “dirty”. Even if we never expected rainfall accumulations derived from gauges and from radar mosaics alone to match perfectly, a comparison between these two (Fig. SB1) can help reveal which problems likely affect more the final statistics. The effect of blockage by topography and uneven radar coverage stand out as expected, and so do a few pixels of persistent clutter; a couple of abnormally “hot” (read “overestimating”) radars can be spotted, such as in northwestern Texas; and if one knows the location of individual radars (see Fig. 1a), one may start to notice some systematic range-dependent behavior that are more visible in Fig. SB2.

The cleaning of reflectivity maps at the radar data processor site and in the process of making radar mosaics has been an evolving endeavor: For example, at the time of this writing, most radars are transitioning to the 17th major revision to the radar data processing system since the beginning of the WSR-88D program. The massive size of the current radar dataset (we evaluated that it would take two years non-stop just to download the data on our university network) makes the reprocessing and regeneration of mosaics possible only by large organizations. For radar climatology work, we must hence largely rely on mosaic maps that were generated in real-time with the approaches used at the time. Finally, mosaic products are often put together with a given goal in mind, e.g., obtaining reflectivity at a given height or at the surface (like the one made by WSI, top of Fig. SB2) versus obtaining reflectivity at the lowest possible level (like the one made by WDSSII, bottom of Fig. SB2), and that goal also affects the climatology obtained as the average estimated rainfall differs by 11% between the two. In our case, availability of mosaics dictated the use of two different datasets over two different periods (see Table 1). The only "reprocessing" of the nearly 700,000 mosaics maps used in this study was the suppression of maps badly affected by blunders (e.g., incorrect remapping, or incorrect reflectivities): an automatic algorithm first flagged times of suspiciously rapid changes in echo statistics; then we manually looked at those time periods to determine what caused these anomalies, and removed clearly damaged mosaic maps.

The net result is that any climatological analysis of radar data from ready-made mosaics will be imperfect and we should accept those imperfections. These will determine what useful results can be obtained as well as how to interpret them. Hence, except for the computation of frequency of occurrence of different echo intensities (Fig. 2), we focused our analysis on processes less likely to be affected by data quality issues, primarily relative changes in annual and daily cycles for which many biases get canceled out, and focusing on convection not affected by weak non-weather echoes. Also, data in areas where the long-term accumulation of precipitation differs significantly from that observed with gauges are extremely doubtful and have been masked in most figures.

Table

TABLE 1: Mosaic radar maps used in this study (0.0181°of latitude = 2 km).

Period	Source	Resolution	Processing	Stated goal
10/1995-12/2001	Weather Services International (WSI)	5 dB(Z); 0.0181° lat. * 0.0191° lon.; 15 min	Zhang et al. (2015)	Estimate surface reflectivity
02/2002-08/2007	Weather Services International (WSI)	1 dB(Z); 0.0181° lat. * 0.0191° lon.; 15 min		Estimate surface reflectivity
09/2007-03/2011	NSSL / WDSSII	<.5 dB(Z); 0.01° lat. * 0.01° lon.; 5 min	Lakshmanan et al. (2006, 2007) US low altitude	Mosaic the lowest-available reflectivity
04/2011-12/2015	NSSL / WDSSII via Weather Decision Technologies	.33 dB(Z); 0.009° lat. * 0.0116° lon.; 5 min	Lakshmanan et al. (2006, 2007) US low altitude	Mosaic the lowest-available reflectivity

Figure captions

Fig. 1. a) WSR-88D radar coverage over the conterminous United States (original image courtesy of NOAA); b) Computed annual precipitation from radar mosaics between 1996 and 2015 using the Joss and Waldvogel (1970) reflectivity (Z) to rainfall rate (R) relationship $Z = 300R^{1.5}$, limiting the peak rainfall to 100 mm hr^{-1}; c) Gauge-derived annual precipitation over the same period as derived from the data of the PRISM Climate Group of the Oregon State University (PRISM 2016).

Fig. 2: Frequency of observation of echoes of a) at least 5 dBZ, b) at least 45 dBZ, and c) at least 60 dBZ. Areas stripped in gray did not meet the criteria $2/3\ R_{gauge} < R_{radar} < 3/2\ R_{gauge}$. Artifact-wise, the fingerprints of individual radars are more obvious at low reflectivity than at high reflectivity. Meteorology-wise, precipitation is more frequent in the mid-latitudes (West Coast & north east). Convective rain occurrence is highest on the Gulf Coast and southern Atlantic Coast where sea breezes often play a major role in convection initiation, and lowest on the West Coast bathed by cold ocean water. Hail echoes are most frequent in the Great Plains. Note how the three images show very different patterns. For reference, a frequency of 4% corresponds to 1 hr day^{-1}, 0.1% is 9 hrs yr^{-1}, and 0.001% is 5 min yr^{-1}.

Fig. 3: Contrast between the frequency of echoes exceeding 45 dBZ in a) late spring (May and June) and b) middle of the summer (July and August). Changes in patterns of convection between the two seasons reflect the changes in the larger-scale processes driving them.

Fig. 4: Diurnal cycle of the frequency of occurrence of echoes exceeding 40 dBZ between the months of April and September starting from the late night on the upper left (2:00-5:45 CST in the middle of the continent) and ending on the middle of the night on the lower right.

Fig. 5: Solar time of a) the preferred occurrence of echoes exceeding 40 dBZ in the warm season, and of b) the fastest daytime growth in the occurrence of such echoes. In both plots, a two-dimensional color scale is used to characterize the timing of events: The hue or frequency of the color used shows the average time or the time of the fastest occurrence increase (e.g., reds indicating peak of occurrence or fastest increase in the afternoon); the saturation and brightness of the color illustrates whether the diurnal cycle of convection or the rate of convection increase is strong and unimodal (saturated bright colors) or weak or multimodal (unsaturated dark colors). Black pixels indicate areas too contaminated by clutter or without enough data to make a proper peak time determination.

Fig. 6: Power spectra of 20-yr long time series of radar-derived precipitation rate (blue curve) and fractional area of precipitation occurrence (≥ 5 dBZ, red curve). Each curve is an average of spectra for 554 small areas 0.25° longitude by 0.25° latitude wide (approximately 24-by-28 km in size) centered on every 1° in longitude and latitude in the eastern two-thirds of the conterminous United States where radar coverage is expected to be good ($2/3\ R_{gauge} < R_{radar} < 3/2\ R_{gauge}$). For time scales under a week, sloping spectra characteristic of

precipitation structures embedded within smaller/shorter precipitation structures can be observed. Superposed on these, the signature of diurnal and annual cycles and some of their harmonics (half and third of a day and a year) can be detected. In inset, a zoom of the curves around the one-week period has been added.

Fig. 7: Patterns of relative difference in the occurrence of echoes exceeding 5 dBZ (left column) and 50 dBZ (right column) observed when separating the 20-year dataset in two groups A and B using two different strategies. a) and b) Difference in echo occurrence between week-ends (Saturdays to Mondays, group A) and week-days (Tuesdays to Fridays, group B). In the north-east, precipitation is notably less frequent on week-ends while in southern Texas, week-ends tend to be wetter. c) and d) Difference between Days 1-3 of an arbitrary 6-day cycle starting 1 January 1996 (group A) and Days 4-6 of the same cycle (group B). Early in the six-day cycle, precipitation occurrence is noticeably lower in the Midwest and higher in Louisiana, and conversely late in that cycle. This obviously accidental pattern is stronger and more statistically significant than any weekday-weekend patterns.

Fig. 8: Occurrence of echoes stronger than 50 dBZ (top row) and of lightning (bottom row) around major cities between May and August for the whole day (left column), the late night (middle column) and the afternoon (right column). The lightning and radar data around 13 cities with over 1 million inhabitants away from both major topographic features (oceans, Great Lakes, significant orography) and areas of poor radar data quality (due to clutter and beam blockage)

were combined to make this figure. On average, an enhancement of afternoon convection and especially lightning occurrence can be observed immediately over and east of these cities.

Fig. SB1: Ratio of the radar-derived precipitation accumulation between 1996 and 2015 shown in Fig. 1b and of the gauge-derived precipitation accumulation over the same period shown in Fig. 1c.

Fig. SB2: Radar-derived mean annual precipitation derived from two different mosaics and for two different periods: a) Precipitation derived from WSI mosaics (1996-2006); b) Precipitation derived from WDSSII mosaics (2008-2015). Key differences to notice are not as much the overall difference in derived precipitation, as those do change with time, as how the patterns of precipitation accumulation around individual radars changed between the two mosaics, concentric patterns being more visible in b) than in a) in the eastern half of the United States.

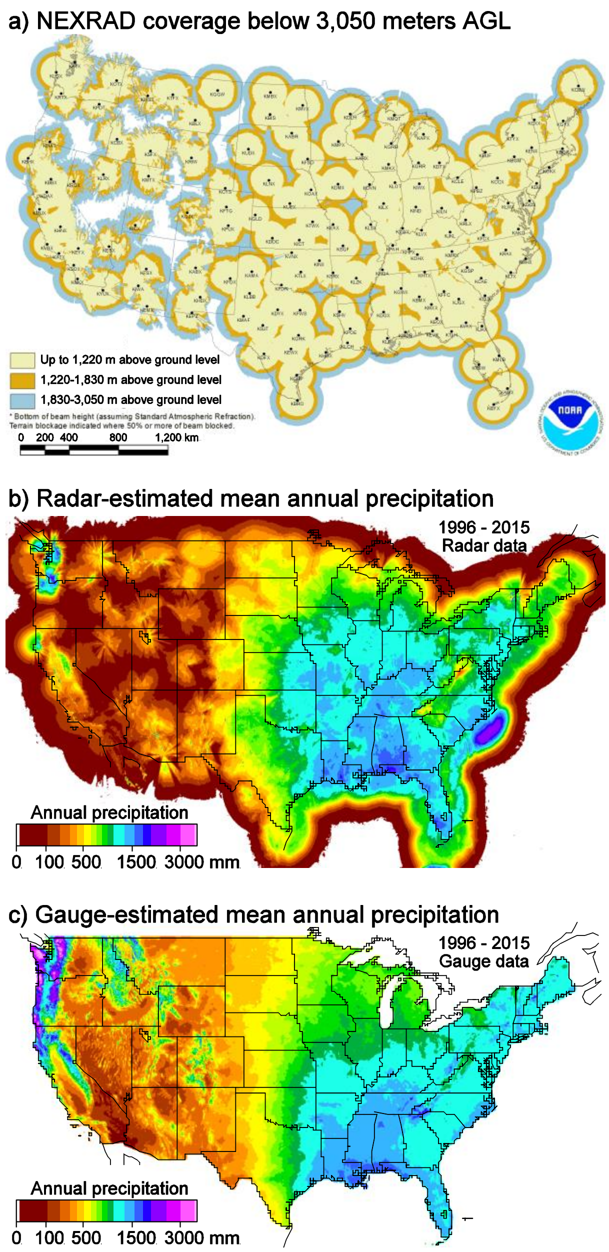


Fig. 1. a) WSR-88D radar coverage over the conterminous United States (original image courtesy of NOAA); b) Computed annual precipitation from radar mosaics between 1996 and 2015 using the Joss and Waldvogel (1970) reflectivity (Z) to rainfall rate (R) relationship $Z = 300R^{1.5}$, limiting the peak rainfall to 100 mm hr^{-1}; c) Gauge-derived annual precipitation over the same period as derived from the data of the PRISM Climate Group of the Oregon State University (PRISM 2016).

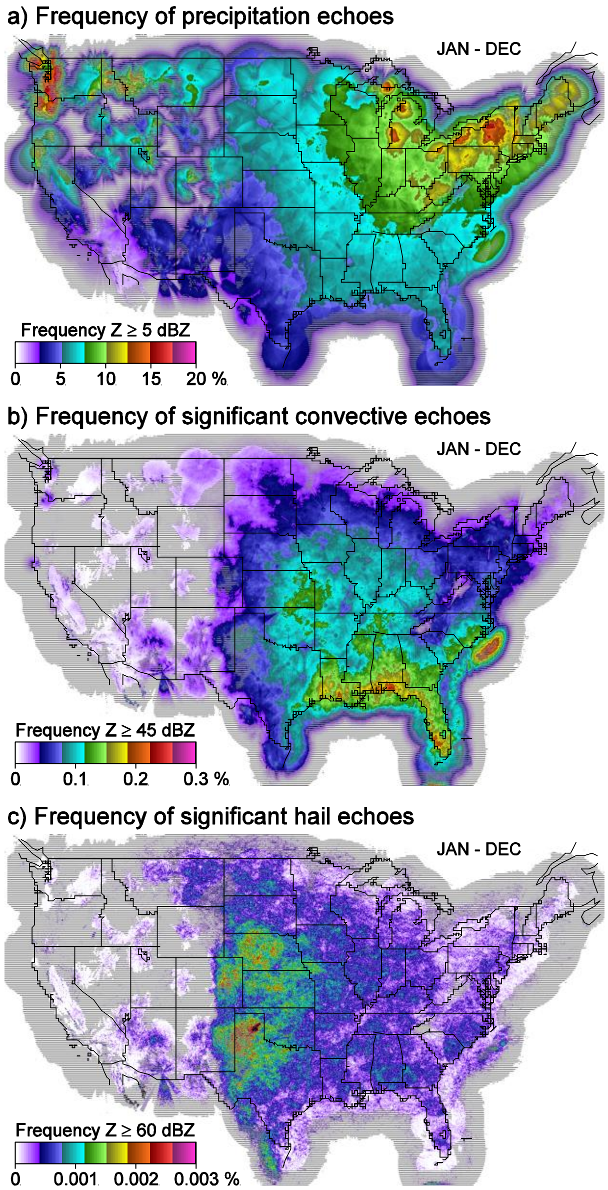


Fig. 2: Frequency of observation of echoes of a) at least 5 dBZ, b) at least 45 dBZ, and c) at least 60 dBZ. Areas stripped in gray did not meet the criteria $2/3\, R_{gauge} < R_{radar} < 3/2\, R_{gauge}$. Artifact-wise, the fingerprints of individual radars are more obvious at low reflectivity than at high reflectivity. Meteorology-wise, precipitation is more frequent in the mid-latitudes (West Coast & north east). Convective rain occurrence is highest on the Gulf Coast and southern Atlantic Coast where sea breezes often play a major role in convection initiation, and lowest on the West Coast bathed by cold ocean water. Hail echoes are most frequent in the Great Plains. Note how the three images show very different patterns. For reference, a frequency of 4% corresponds to 1 hr day^{-1}, 0.1% is 9 hrs yr^{-1}, and 0.001% is 5 min yr^{-1}.

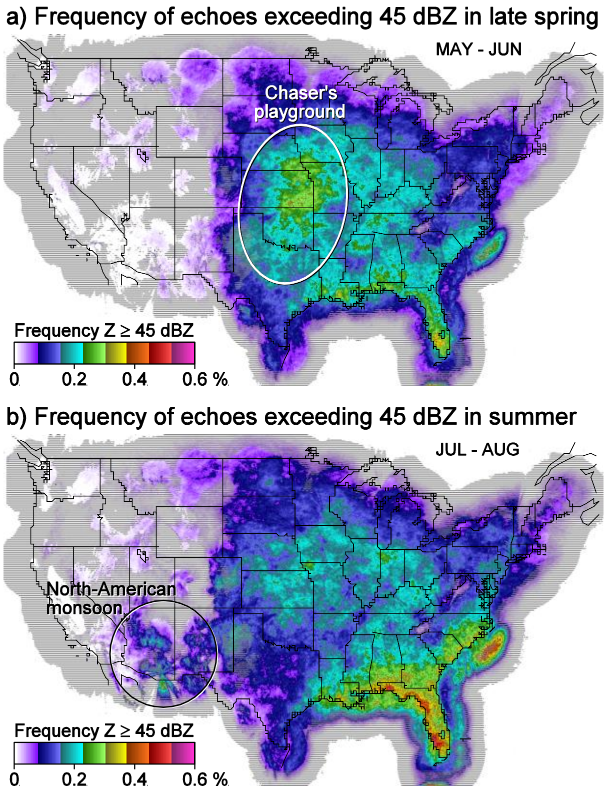


Fig. 3: Contrast between the frequency of echoes exceeding 45 dBZ in a) late spring (May and June) and b) middle of the summer (July and August). Changes in patterns of convection between the two seasons reflect the changes in the larger-scale processes driving them.

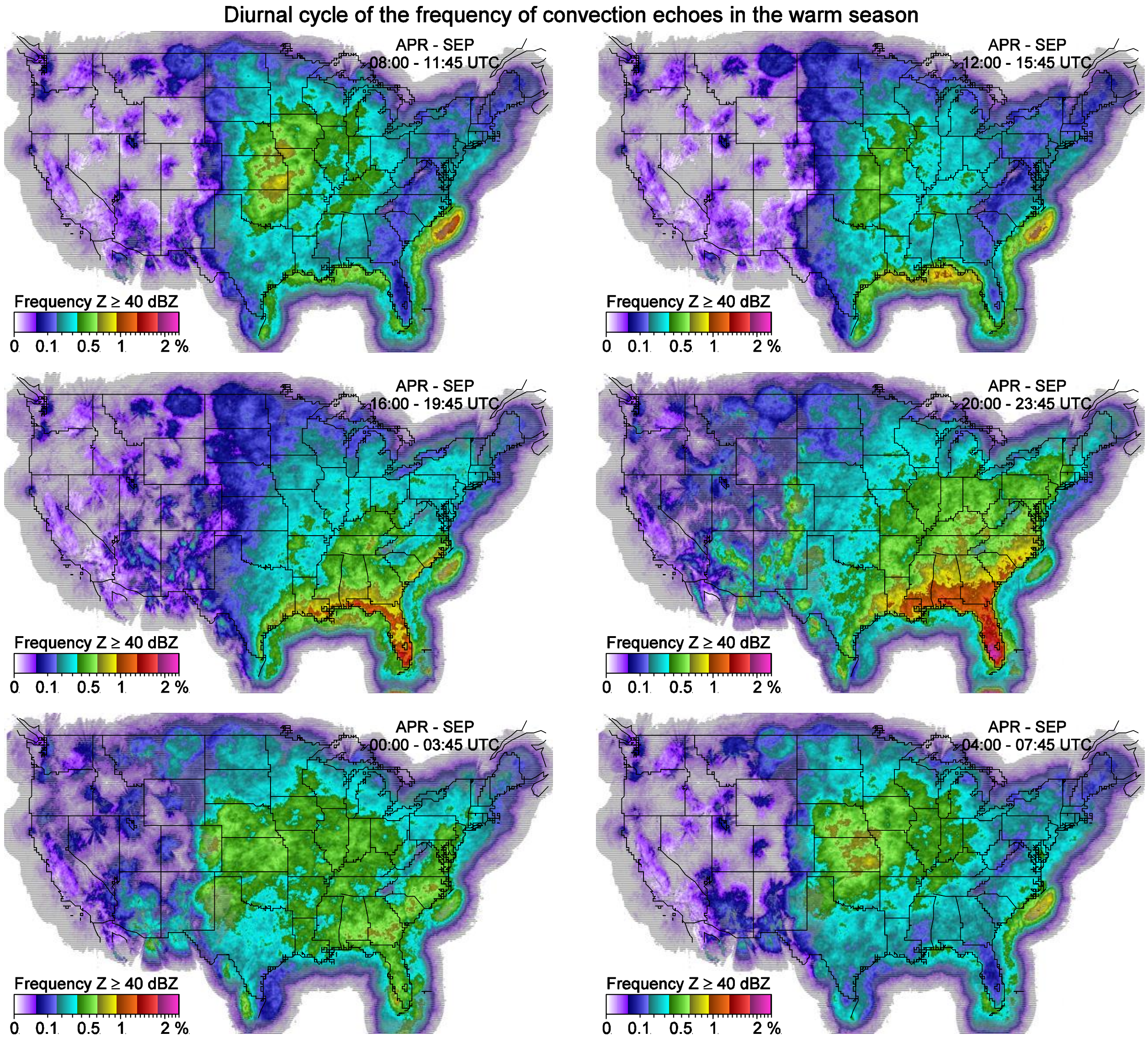
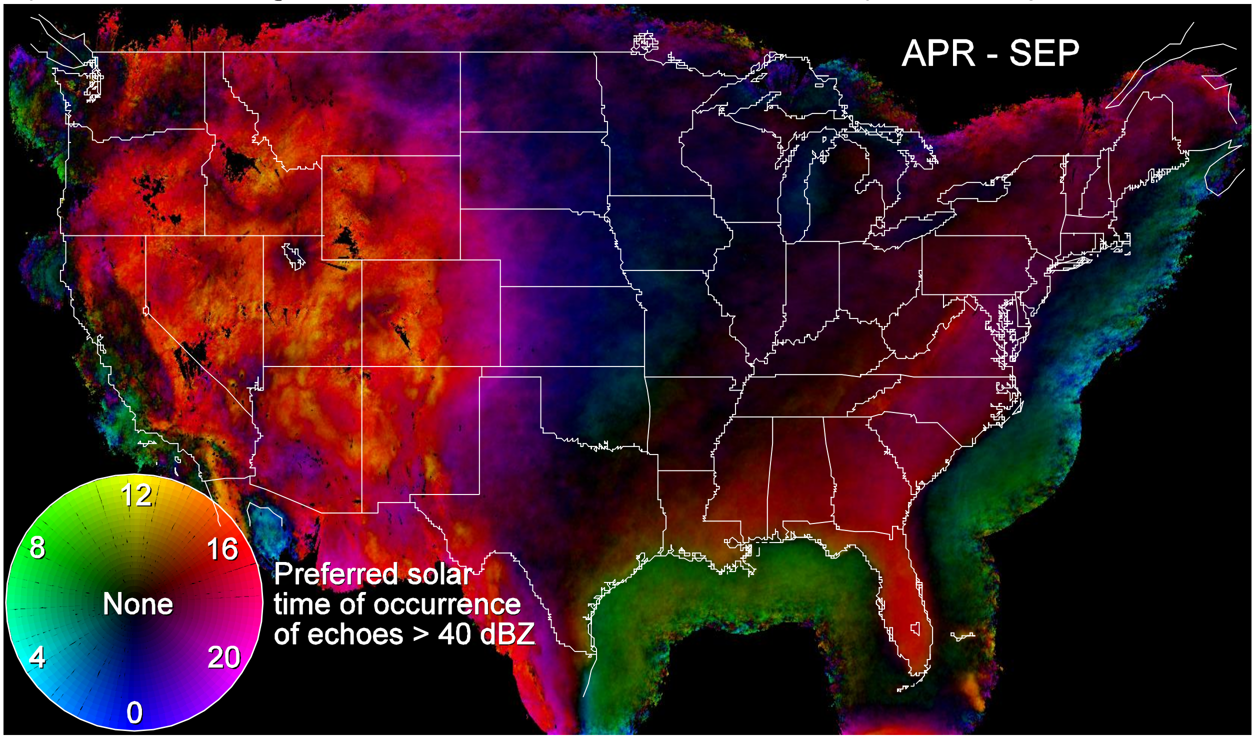


Fig. 4: Diurnal cycle of the frequency of occurrence of echoes exceeding 40 dBZ between the months of April and September starting from the late night on the upper left (2:00-5:45 CST in the middle of the continent) and ending on the middle of the night on the lower right.

a) Preferred/average solar time at which convective echoes (Z > 40 dBZ) are observed



b) Solar time of the fastest increase in daytime convection (Z > 40 dBZ) occurrence

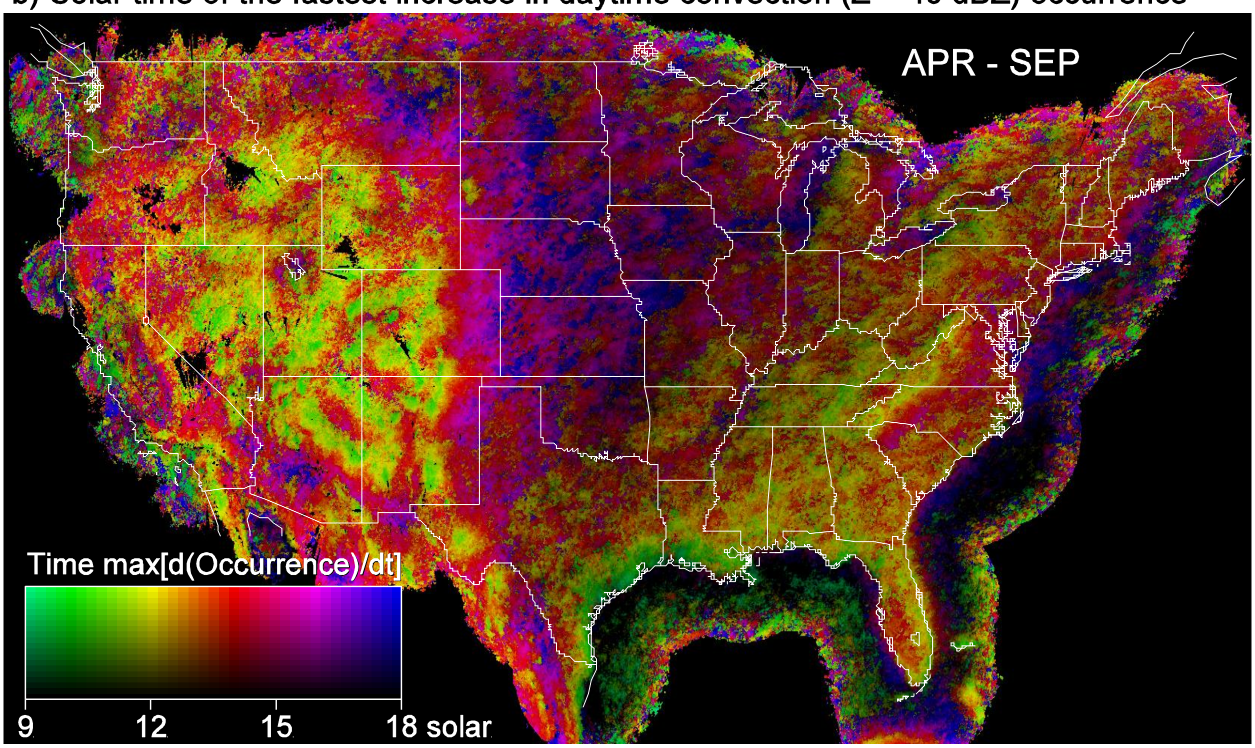


Fig. 5: Solar time of a) the preferred occurrence of echoes exceeding 40 dBZ in the warm season, and of b) the fastest daytime growth in the occurrence of such echoes. In both plots, a two-dimensional color scale is used to characterize the timing of events: The hue or frequency of the color used shows the average time or the time of the fastest occurrence increase (e.g., reds indicating peak of occurrence or fastest increase in the afternoon); the saturation and brightness of the color illustrates whether the diurnal cycle of convection or the rate of convection increase is strong and unimodal (saturated bright colors) or weak or multimodal (unsaturated dark colors). Black pixels indicate areas too contaminated by clutter or without enough data to make a proper peak time determination.

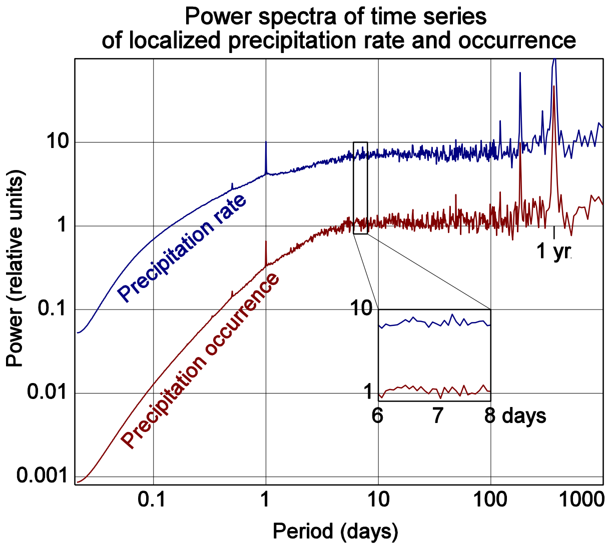


Fig. 6: Power spectra of 20-yr long time series of radar-derived precipitation rate (blue curve) and fractional area of precipitation occurrence (≥ 5 dBZ, red curve). Each curve is an average of spectra for 554 small areas 0.25° longitude by 0.25° latitude wide (approximately 24-by-28 km in size) centered on every 1° in longitude and latitude in the eastern two-thirds of the conterminous United States where radar coverage is expected to be good ($2/3\, R_{gauge} < R_{radar} < 3/2\, R_{gauge}$). For time scales under a week, sloping spectra characteristic of precipitation structures embedded within smaller/shorter precipitation structures can be observed. Superposed on these, the signature of diurnal and annual cycles and some of their harmonics (half and third of a day and a year) can be detected. In inset, a zoom of the curves around the one-week period has been added.

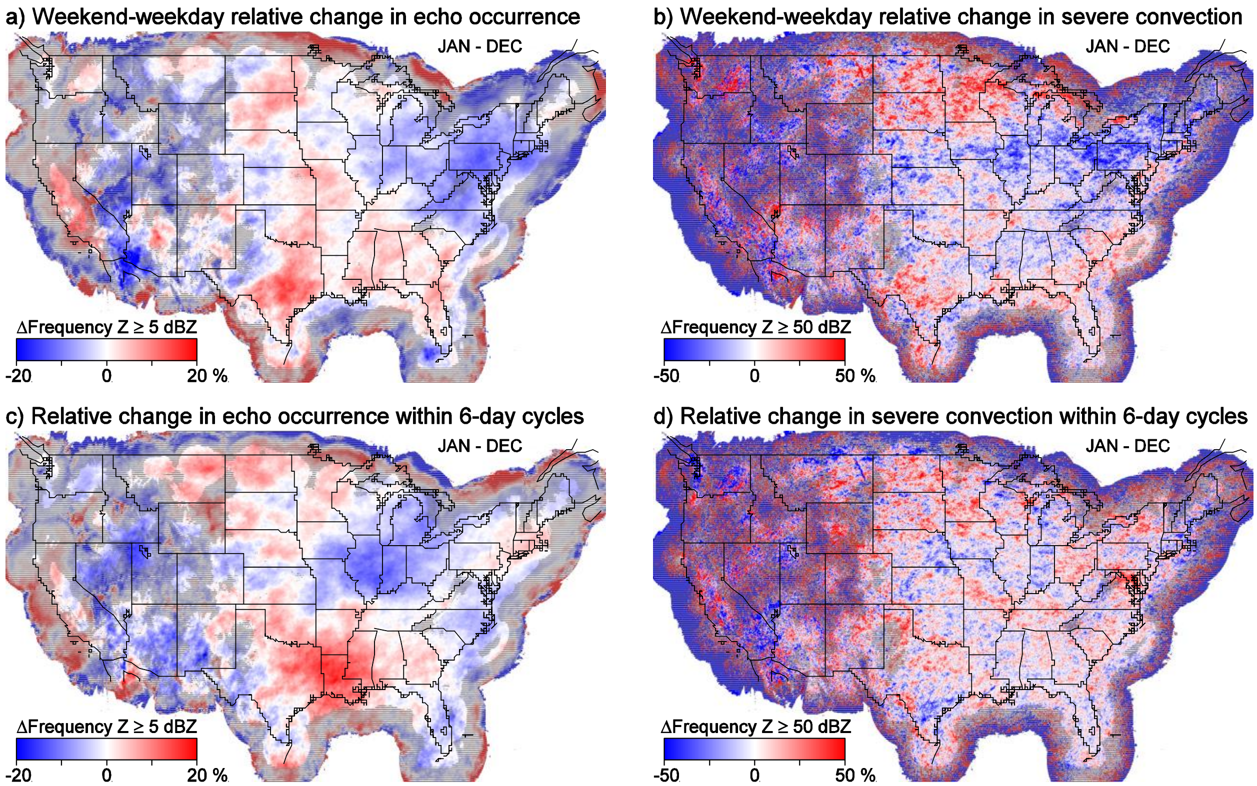


Fig. 7: Patterns of relative difference in the occurrence of echoes exceeding 5 dBZ (left column) and 50 dBZ (right column) observed when separating the 20-year dataset in two groups A and B using two different strategies. a) and b) Difference in echo occurrence between week-ends (Saturdays to Mondays, group A) and week-days (Tuesdays to Fridays, group B). In the north-east, precipitation is notably less frequent on week-ends while in southern Texas, week-ends tend to be wetter. c) and d) Difference between Days 1-3 of an arbitrary 6-day cycle starting 1 January 1996 (group A) and Days 4-6 of the same cycle (group B). Early in the six-day cycle, precipitation occurrence is noticeably lower in the Midwest and higher in Louisiana, and conversely late in that cycle. This obviously accidental pattern is stronger and more statistically significant than any weekday-weekend patterns.

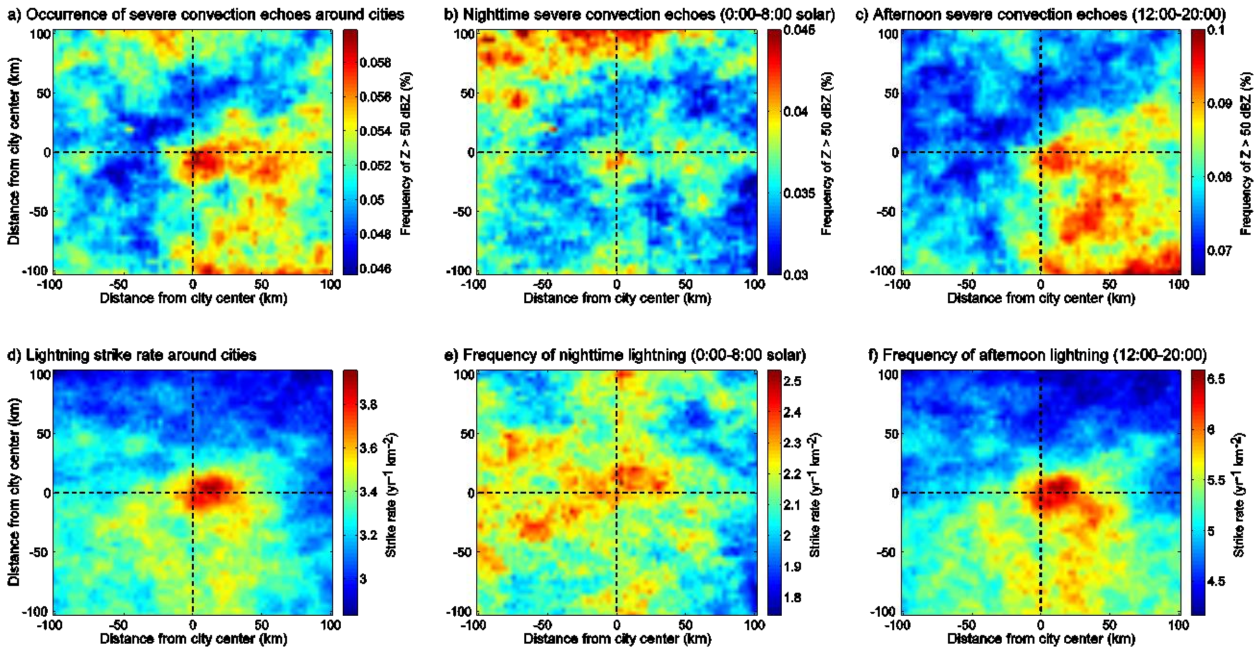


Fig. 8: Occurrence of echoes stronger than 50 dBZ (top row) and of lightning (bottom row) around major cities between May and August for the whole day (left column), the late night (middle column) and the afternoon (right column). The lightning and radar data around 13 cities with over 1 million inhabitants away from both major topographic features (oceans, Great Lakes, significant orography) and areas of poor radar data quality (due to clutter and beam blockage) were combined to make this figure. On average, an enhancement of afternoon convection and especially lightning occurrence can be observed immediately over and east of these cities.

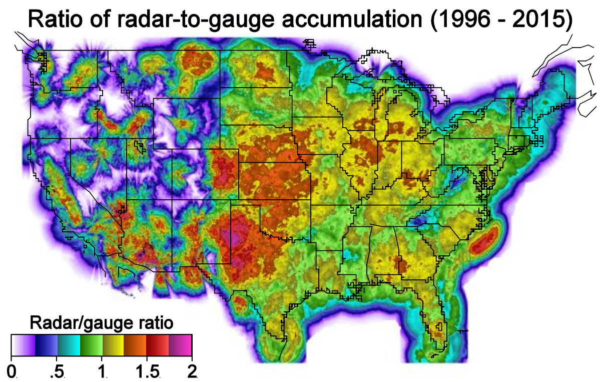


Fig. SB1: Ratio of the radar-derived precipitation accumulation between 1996 and 2015 shown in Fig. 1b and of the gauge-derived precipitation accumulation over the same period shown in Fig. 1c.

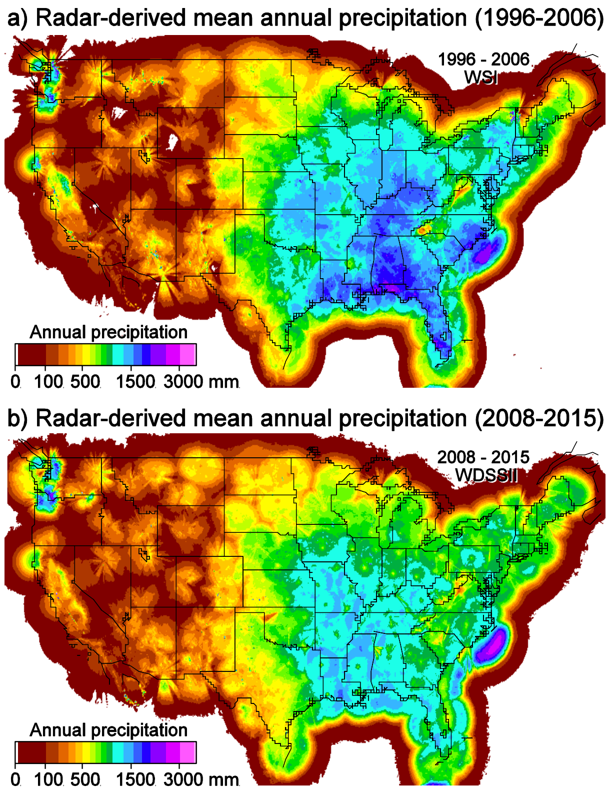


Fig. SB2: Radar-derived mean annual precipitation derived from two different mosaics and for two different periods: a) Precipitation derived from WSI mosaics (1996-2006); b) Precipitation derived from WDSSII mosaics (2008-2015). Key differences to notice are not as much the overall difference in derived precipitation, as those do change with time, as how the patterns of precipitation accumulation around individual radars changed between the two mosaics, concentric patterns being more visible in b) than in a) in the eastern half of the United States.