

Diagnosing an Artificial Trend in NLDAS-2 Afternoon Precipitation[Ⓞ]

CRAIG R. FERGUSON

*Atmospheric Sciences Research Center, and Department of Atmospheric and Environmental Sciences,
University at Albany, State University of New York, Albany, New York*

DAVID M. MOCKO

*Hydrological Sciences Laboratory, NASA GSFC, Greenbelt, Maryland, and
Science Applications International Corporation, McLean, Virginia*

(Manuscript received 26 October 2016, in final form 30 December 2016)

ABSTRACT

While investigating linkages between afternoon peak rainfall amount and land–atmosphere coupling strength, a statistically significant trend in phase 2 of the North American Land Data Assimilation System (NLDAS-2) warm season (April–September) afternoon (1700–2259 UTC) precipitation was noted for a large fraction of the conterminous United States, namely, two-thirds of the area east of the Mississippi River, during the period from 1979 to 2015. To verify and better characterize this trend, a thorough statistical analysis is undertaken. The analysis focuses on three aspects of precipitation: amount, frequency, and intensity at 6-hourly time scale and for each calendar month separately. At the NLDAS-2 native resolution of $0.125^\circ \times 0.125^\circ$, Kendall's tau and Sen's slope estimators are used to detect and estimate trends and the Pettitt test is used to detect breakpoints. Parallel analyses are conducted on both NARR and Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2), subdaily precipitation estimates. Widespread breakpoints of field significance at the $\alpha = 0.05$ level are detected in the NLDAS-2 frequency and intensity series for all months and 6-h periods that are absent from the analogous NARR and MERRA-2 datasets. These breakpoints are shown to correspond with a July 1996 NLDAS-2 transition away from hourly $2^\circ \times 2.5^\circ$ NOAA/CPC precipitation estimates to hourly 4-km stage II Doppler radar precipitation estimates in the temporal disaggregation of CPC daily gauge analyses. While NLDAS-2 may provide the most realistic diurnal precipitation cycle overall, users should be aware of this discontinuity and its direct effect on long-term trends in subdaily precipitation and indirect effects on trends in modeled soil moisture, surface temperature, surface precipitation and water fluxes, snow cover, snow water equivalent, and runoff/streamflow.

1. Introduction

The $0.125^\circ \times 0.125^\circ$ 1-hourly meteorological forcing dataset from phase 2 of the North American Land Data Assimilation System (NLDAS-2; Xia et al. 2012a,b) is a very widely used forcing dataset for offline hydrological simulations conducted over the conterminous United States (CONUS) for the period from 1979 to present. The NLDAS has a rich history dating back to the initiation of its first phase in 2000 (NLDAS-1; Mitchell et al.

2004). In design, NLDAS closely resembles the Project for the Intercomparison of Land-Surface Parameterization Schemes (PILPS; e.g., Henderson-Sellers et al. 1995; Pitman and Henderson-Sellers 1998; Pitman et al. 1999; Wood et al. 1998) conducted under the auspices of the World Climate Research Programme's Global Energy and Water Exchanges project (GEWEX). Specifically, the NLDAS best-estimate forcing dataset enables intercomparisons among participating models to isolate parameterization-related differences. However, a valid criticism of both PILPS and NLDAS is that through focusing attention on model physics they have effectively shifted attention away from important data forcing uncertainties.

Uncertainties in precipitation forcing have the largest impact on offline and coupled hydrological simulations (e.g., Ferguson et al. 2010; Gochis et al. 2016; Pan et al. 2010).

[Ⓞ] Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JHM-D-16-0251.s1>.

Corresponding author e-mail: Craig R. Ferguson, crferguson@albany.edu

DOI: 10.1175/JHM-D-16-0251.1

© 2017 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSRuseLicenses).

The diurnal cycle of precipitation, including precipitation phase (i.e., solid or liquid) and aspect (amount, frequency, and intensity), helps determine surface runoff partitioning and thereby flood frequency and severity, groundwater recharge, evapotranspiration, and vegetative water stress. Currently, resolving the summertime diurnal rainfall cycle over the central United States is a particular weakness of weather and climate models (e.g., Lee et al. 2007; Liang et al. 2004; Yamada et al. 2012) and one that is closely related to their systematic warm and dry biases over the region (e.g., Klein et al. 2006).

Land–atmosphere coupling [for a review, see Seneviratne et al. (2010)] serves as one mechanism through which wetting and drying trends over land can translate into concomitant changes in the diurnal precipitation cycle. Under conditions of strong land–atmosphere coupling not uncommon to the central United States and other mid-continental regions (e.g., Guo et al. 2006; Koster et al. 2006), soil moisture anomalies can nudge the diurnal precipitation cycle through their effect on the surface heat flux partitioning (e.g., Findell and Eltahir 2003a,b). Song et al. (2016) showed that the central United States' normal nocturnal (0300 local time) precipitation peak climatology (e.g., Wallace 1975) was interrupted on strong land–atmosphere coupling days with afternoon peak precipitation. Because of the extent that precipitation is dependent on precursor cloud formation, any changes in subdaily precipitation frequency would likely be accompanied by shifts in subdaily cloudiness, which could have wide-ranging impacts on the surface radiation budget (Dai et al. 1999). For example, a shift toward enhanced nighttime (precipitation) cloudiness would suggest increased surface net radiation and warming. Conversely, a shift toward enhanced daytime (precipitation) cloudiness would suggest decreased surface net radiation and cooling.

High-resolution (7 km or finer) mesoscale models run with explicit (as opposed to parameterized) convection and precipitation processes have been shown to better capture the diurnal precipitation cycle (Dirmeyer et al. 2012; Erlingis and Barros 2014; Prein et al. 2015; Sato et al. 2009; Tao et al. 2013). However, more extensive verification of these convection-resolving models over a larger range of conditions with high-temporal-resolution precipitation datasets (1-hourly or less), like that of NLDAS, is needed. An evaluation of the long-term diurnal precipitation cycle for nonstationarity is also needed because its detection would signal a moving target in model development. Thus, verification products are required to be highly accurate and free from data discontinuities.

In this study, we address the following question: Are there detectable trends in NLDAS-2 subdaily precipitation characteristics (amount, frequency, and intensity) over

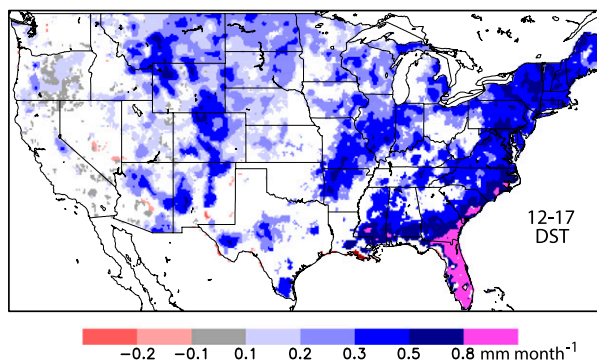


FIG. 1. Sen's slopes in NLDAS-2 12–17 DST precipitation P amount between 1979 and 2015, computed from the 6-month (April–September) warm season (mm month^{-1} ; i.e., $0.8 \text{ mm month}^{-1} \times 6 \text{ months} = 4.8 \text{ mm yr}^{-1}$). All plotted slopes are locally significant at the 0.05 level. These results exclude FDR correction and subsequent testing of field significance.

the period of record (1979–2015)? A climatology over a 10-yr period (1998–2007) using consistent input data sources (detailed below in section 2a) of the diurnal cycle of NLDAS-2 precipitation was examined in Matsui et al. (2010). However, as far as we could determine, the interannual variability of diurnal precipitation cycle over CONUS has only previously been conducted at relatively coarse (i.e., $2.5^\circ \times 2.0^\circ$) horizontal resolution and for an earlier period from 1963 to 1993 (Dai 1999; Dai et al. 1999). Preliminary analyses revealed statistically significant trends in afternoon [1700–2259 UTC or 1200–1759 daylight saving time (DST)] precipitation amount from the start of the NLDAS-2 record to the current most recently completed year (1979–2015), especially in regions east of the Mississippi River (Fig. 1). The ensuing analyses, undertaken to further diagnose these trends in the context of the full diurnal cycle and all aspects of precipitation, are detailed here. By design, this paper serves both as a critical assessment of NLDAS-2 hourly precipitation forcing and as a tutorial on a thorough statistical methodology appropriate for long-term trend detection on gridded datasets. When dealing with long-term datasets, it is important to test for natural, as well as artificial, data-driven discontinuities. The methodology presented fully considers breakpoints as well as field significance and is appropriate whether or not there is a priori knowledge of dataset inhomogeneities.

2. Data and methods

a. NLDAS-2

NLDAS-2 is a multimodel land modeling and assimilation system run in an uncoupled mode on a common $0.125^\circ \times 0.125^\circ$ grid covering CONUS, the southern part of Canada, and the northern portion of Mexico

(25°–53°N, 125°–67°W). NLDAS is a collaboration project among several groups: NOAA/NCEP/Environmental Modeling Center (EMC), NASA Goddard Space Flight Center (GSFC), Princeton University, the University of Washington (UW), the NWS Office of Hydrologic Development (OHD), and the NCEP/Climate Prediction Center (CPC). NLDAS is supported by the NOAA Climate Program Office's Modeling, Analysis, Predictions, and Projections (MAPP) program. NLDAS-2 may be considered as the follow-on to NLDAS-1 (Mitchell et al. 2004). Like NLDAS-1, it provides standard near-surface hourly gridded meteorological forcing (i.e., forcing file A; 2-m air temperature and specific humidity, 10-m zonal u and meridional v wind speed, surface pressure, downward shortwave and longwave radiation, and precipitation) as well as collections of hourly model outputs from several different LSMs driven by the same forcing, including soil moisture/temperature, land surface temperature, surface energy and water fluxes, snow cover, snow water equivalent, and runoff/streamflow (Xia et al. 2009, 2012a,b). NLDAS-2 provides a secondary forcing dataset, including model-generated precipitation, based on the NCEP North American Regional Reanalysis (NARR; Mesinger et al. 2006) lowest prognostic model level (forcing file B). Major differences between NLDAS-1 and NLDAS-2 include record length and choice of atmospheric reanalysis used for nonprecipitation forcing. NLDAS-1 covers the 11-yr period from 1 August 1996 to 31 December 2007, whereas NLDAS-2 covers the complete modern era from 1979 to present. NLDAS-1 used the 40-km, 3-hourly NCEP Eta Data Assimilation System (EDAS; Rogers et al. 1996) analysis fields (Cosgrove et al. 2003) whereas NLDAS-2 used analysis fields interpolated from the 32-km ($\sim 0.3^\circ \times 0.3^\circ$), 3-hourly NARR.

In this study, we are only concerned with the NLDAS-2 precipitation dataset as distributed in the forcing file A (Xia et al. 2009). Discontinuities in the underlying inputs are of direct relevance to our assessment of NLDAS-2 precipitation trends. In the following paragraphs, we will provide an overview of the time-varying precipitation inputs over CONUS, the focus region of our study. We note that the NLDAS-2 record comprises three modes of production: a 30-yr (from 1 January 1979 to 31 December 2008) retrospective implementation completed in 2009 (Xia et al. 2010a, 2014); a subsequent near-real-time implementation (from 1 January 2009 to 4 August 2014); and most recently, a transition to NCEP Operational Implementation (from 5 August 2014 to present). It is important to stress that, for our purposes, knowing the date of a shift from retrospective to real-time production is not all telling. It is instead knowing the date of any shifts from "retrospective" to "real time" versions of the contributing gauge-based analysis that is paramount.

The NLDAS-2 precipitation analysis is composed of two components: a daily gauge-based precipitation analysis and a subdaily gauge-, radar-, satellite-, or model-based precipitation analysis used to perform temporal disaggregation to hourly time scale. Over CONUS, NLDAS-2 used an improved CPC/OHD daily gauge-based analysis (Fan et al. 2006; Shafran et al. 2004) from 1979 to 2011. This product differs from the more widely used (and documented) CPC Unified Raingauge Dataset (URD; Higgins et al. 2000) in three aspects. First, CPC/OHD applied a Parameter-Elevation Regressions on Independent Slopes Model (PRISM; Daly et al. 1994) orographic adjustment based on the 1961–90 climate normal versus no terrain adjustment in the URD. Second, CPC/OHD applied an inverse-square distance-weighting interpolation (modified Shepard's method; Renka 1988) with a 50-km search radius versus a modified Cressman (1959) scheme (Charba et al. 1992; Glahn et al. 1985) with a 200-km search radius used by URD. Third, CPC/OHD was produced on the $0.125^\circ \times 0.125^\circ$ U.S. NLDAS grid versus the URD 0.25° grid. The CPC/OHD processing algorithm is widely accredited to John Schaake (e.g., Luo et al. 2005; Mesinger et al. 2006). Starting on 1 January 2002, the CPC/OHD product went "real time"; the number of contributing stations ($\sim 12\,000$) dropped by nearly one-third (Mo et al. 2012, their Fig. 2) but the methodology remained the same. Since 1 January 2012, NLDAS-2 has used the global real-time $0.125^\circ \times 0.125^\circ$ CPC Unified daily gauge-based precipitation analysis (Chen et al. 2008).

Which subdaily precipitation estimates are used to temporally disaggregate the underlying daily gauge-based analysis depends on data availability, and varies in space and time. It is important to note that the temporal disaggregation to hourly only determines the time of day when the precipitation falls; the daily amounts remain the same from the gauge-based analysis. Over CONUS, during the period from July 1996 to present, available 4-km hourly Doppler radar stage II precipitation estimates (Lin and Mitchell 2005) are used to temporally disaggregate the CPC/OHD daily gauge-based analysis into hourly precipitation. From December 2002 to present, if stage II data are unavailable, then 8-km, 30-min CPC morphing technique (CMORPH; Joyce et al. 2004) precipitation estimates are used. If neither stage II nor CMORPH data are available (especially prior to July 1996), then $2.0^\circ \times 2.5^\circ$ hourly CPC Hourly Precipitation Dataset (HPD; Higgins et al. 1996) estimates are used for temporal disaggregation. Outside of HPD's period of coverage (from January 1979 to September 2002), NARR 32-km ($\sim 0.3^\circ \times 0.3^\circ$), 3-hourly model-simulated precipitation (available from January 1979 to present) serves as the ultimate fallback dataset.

Unfortunately, no log files were generated in the production of NLDAS-2 to document which precipitation sources were used for which days and grid cells.

b. NARR

The NARR record is composed of a retrospective simulation from 1979 to 2002 and a real-time continuation of NARR from 2003 to present, using the Regional Climate Data Assimilation System (R-CDAS). NARR assimilates gauge-based precipitation inputs by altering the model's vertical profiles of latent heating, water vapor, and cloud water during the hourly assimilation intervals (Lin et al. 2001). Thus, there should be a very good correspondence between precipitation amount and diurnal cycle and observations (e.g., Ruane 2010).

Over CONUS, during NARR's retrospective period from 1979 to 2002, the NARR hourly precipitation analysis is obtained by temporally disaggregating the 24-h CPC/OHD analysis (discussed above) using the CPC HPD (Higgins et al. 1996). Between January 2003 and mid-2013, both NLDAS-2 and real-time NARR (R-CDAS) shared the same precipitation analysis (Mo et al. 2012). In mid-2013, NCEP ported the NARR code from its IBM P6 Advanced Interactive eXecutive (AIX) Central Computing System (CC) to a new IBM iDataPlex Intel Linux Weather and Climate Operational Supercomputing System (WCOSS). NARR runs conducted on the WCOSS were inadvertently set to ingest the North American Mesoscale Forecast System (NAM; WRF-NMM; Janjić 2003) predicted precipitation instead of the CPC Unified gauge-based analysis. Because NCEP was running NARR on both CCS and WCOSS for a period of time, the exact date of the WCOSS transition is difficult to pinpoint, but it falls between March and July 2013. Officially, NCEP Central Operations (NCO) NARR WCOSS data transmissions began on 1 July 2013 (W. Ebisuzaki 2016, personal communication). By mid-January 2015, NCO NARR had returned to using the CPC Unified product, consistent with NLDAS-2. A separate, experimental NARR run led by Wesley Ebisuzaki contains the input precipitation fix from September 2014 (W. Ebisuzaki 2016, personal communication).

Because NARR and NLDAS-2 use a consistent daily gauge-based analysis for most of the study period (1979–2015), their precipitation fields should be comparable, especially at the daily time scale. Major differences between NLDAS-2 and NARR subdaily precipitation characteristics should be attributable to the use of different datasets in their respective hourly temporal disaggregation. Namely, NLDAS-2 has employed stage II radar data since July 1996 and NARR/R-CDAS has not.

In this study, we analyzed NARR hourly precipitation obtained from the NLDAS-2 forcing file B dataset. The

NARR dataset was interpolated from its native $0.3^\circ \times 0.3^\circ$ resolution to $0.125^\circ \times 0.125^\circ$ resolution following the budget bilinear method, which uses 25 neighboring points instead of the typical 4 neighboring points used in basic bilinear interpolation (NASA 2016).

c. Other datasets

In parallel and in an identical manner, we also assessed long-term trends in both NASA's Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2; Molod et al. 2015), precipitation and in Livneh et al. (2015, hereafter L15) precipitation. MERRA-2 covers the modern satellite era (from 1980 to present) using a new version of the Goddard Earth Observing System Model, version 5 (GEOS-5). We analyzed both MERRA-2 forecast precipitation (prectot) and the MERRA-2 bias-corrected precipitation (prectotcorr) obtained from the 0.625° (longitude) \times 0.5° (latitude) hourly surface flux diagnostics collection (tavgl_2d_flux_Nx; GMAO 2015). Both fields were bilinearly interpolated to the NLDAS $0.125^\circ \times 0.125^\circ$ grid. The prectotcorr dataset is used in place of prectot to force MERRA-2's land surface scheme, as similarly performed in MERRA-Land (Reichle et al. 2011). From 1980 to 2005 and 2006 to 2015, prectotcorr over CONUS is based on the global CPC Unified (Chen et al. 2008) retrospective and real-time product, respectively. In the case of MERRA-2, the CPC Unified daily gauge-based analysis is temporally disaggregated using MERRA model-simulated precipitation. Because of the sparseness of contributing gauge stations to CPC Unified at high latitudes, prectotcorr tapers back to prectot northward of 42.5°N and is entirely prectot northward of 62.5°N . As an aside, trend and breakpoint results are nearly identical when MERRA-2 is instead interpolated using conservative box averaging.

The $0.0625^\circ \times 0.0625^\circ$ daily gauge-based L15 dataset covering Mexico, CONUS, and southern Canada for the period 1950–2013 was developed to remedy a paucity of gauge data in Mexico, the lack of orographic precipitation adjustment over Mexico and Canada, and significant discontinuities at Mexico–U.S.–Canada boundaries present in earlier UW heritage gridded precipitation datasets (Livneh et al. 2013; Maurer et al. 2002; Wood and Lettenmaier 2006). The UW heritage precipitation datasets are distinct from the NLDAS-2 precipitation dataset (described above) in that they use only a subset of National Climatic Data Center's (NCDC) Cooperative Observer Program gauge stations selected on the basis of data quality and stability (Andreadis et al. 2005). For example, L15 uses only 8844 gauge stations over CONUS that satisfy a 20-yr minimum record length. In other regards, the L15 and NLDAS-2 precipitation approaches are relatively consistent. For example, L15 applied the

Synagraphic Mapping System (SYMAP) interpolation algorithm (Maurer et al. 2002; Shepard 1984), which is very similar to the modified Shepard's interpolation applied in NLDAS-2. And both L15 and NLDAS-2 applied PRISM to account for topographical effects, although L15 used a more modern climate normal period (i.e., 1981–2010) that coincides with the monitoring era (from 1979 to present) of the high-elevation USDA Natural Resources Conservation Service (NRCS) Snowpack Telemetry (SNOTEL) network (Crook 1977; Serreze et al. 1999). Because of L15's greater attention to network stability and topographic correction, it represents an important comparison dataset at the daily time scale. For our analyses, L15 was resampled to the $0.125^\circ \times 0.125^\circ$ NLDAS-2 grid using box averaging.

d. Methods

All UTC hourly datasets (i.e., NLDAS-2, NARR, and MERRA-2) were first postprocessed into DST hourly datasets and subsequently used to generate 6-hourly DST diurnal monthly mean time series of precipitation amount, frequency, and intensity. The 6-h time windows are defined as 0–5, 6–11, 12–17, and 18–23 DST, where 0–5 DST corresponds with 0000–0559 DST, for example, and DST is computed by solving for the local standard time (LST) within each time zone (using its UTC offset) and then adding 1 h. Precipitation amount is calculated as the hourly mean accumulation over the specified 6-h time window. Precipitation frequency is calculated as the number of hours within the specified 6-h time window for which $0.125^\circ \times 0.125^\circ$ precipitation rate exceeded 0.0004 mm h^{-1} . Precipitation intensity is calculated as the quotient of precipitation accumulation and frequency (i.e., number of wet hours) over the specified 6-h time window. Over CONUS, the minimum rate filter screens out 0.5% of NLDAS-2 hourly precipitation events. An equivalent precipitation rate threshold (i.e., 0.1 mm h^{-1} at 2° upscaled from 0.125°) was applied by Dai et al. (1999).

The ensuing statistical analysis was conducted on a $0.125^\circ \times 0.125^\circ$ grid cell basis for all of CONUS ($m = 52412$ land grids) using the 6-hourly DST diurnal monthly mean time series of precipitation amount, frequency, and intensity. The analysis consisted of simultaneous slope estimation and detection of breakpoints, followed by an assessment of their field significance (e.g., Livezey and Chen 1983). The phrase “field significance” will denote field significance on the CONUS domain at the two-tailed, $\alpha = 0.05$ level, unless otherwise specified. The subdaily analysis time scale served our objective of understanding the nature of the 1979–2015 wetting trend in 12–17 DST precipitation amount (Fig. 1).

As a starting point, we checked for a trend and positive lag-1 serial correlation—both at the $\alpha = 0.05$ local

significance level—using the nonparametric rank-based Mann–Kendall test (MK) (Kendall 1975; Mann 1945) and the autocorrelation function (ACF; Venables and Ripley 2002), respectively. If no significant trend or serial correlation was detected, then the Pettitt (1979) breakpoint test is applied to the original time series. Otherwise, a modified trend-free prewhitening (TFPW; Yue et al. 2002) procedure is implemented as follows:

- 1) A Theil–Sen (Sen 1968; Theil 1950) nonparametric slope is estimated from the original data and is used to compute a residual or detrended series. Theil–Sen has previously been shown to be less sensitive to outliers than linear regression (Hirsch et al. 1982; Lettenmaier et al. 1994).
- 2) ACF is used to compute the positive lag-1 serial correlation of the detrended series. If serial correlation is not found to be significant, then both the MK and Pettitt tests are applied to the original series (end here). If the serial correlation is significant at the $\alpha = 0.05$ level, then the detrended series is prewhitened with an autoregressive lag-1 [AR(1)] model.
- 3) The Theil–Sen trend (step 1) is added back to the TFPW series of step 2 to produce a blended TFPW series that includes the trend of the original series (if existent) without autocorrelation.
- 4) The MK trend significance and Pettitt breakpoint tests are applied to the blended TFPW series. If a significant breakpoint is detected, then the same procedure described above, as pertains to the MK test, is applied to each homogenous subseries segment (i.e., prior = 1:break date; posterior = break date + 1:n; where n is the record length).

The nonparametric Pettitt test is based on the Mann–Whitney test for evaluating whether two independent samples come from the same population [for a review, see Mallakpour and Villarini (2016)]. It is designed to detect a single abrupt change in the mean of the distribution of the variable of interest at an unknown point in time. It has been widely used with hydroclimatological data (e.g., Bárdossy and Caspary 1990; Busuioc and von Storch 1996; Ferguson and Villarini 2012; Liu et al. 2012; Villarini et al. 2009; Wijngaard et al. 2003) and is most sensitive to breaks that occur in the middle of the series.

Accounting for trends and breakpoints in each other's presence is essential. The presence of a monotonic trend could lead to false detection of a breakpoint by the Pettitt test (e.g., Beaulieu et al. 2012). Likewise, the presence of a breakpoint could lead to a false MK trend detection. For our analysis, in cases when both a significant monotonic trend and a breakpoint are detected, the MK test is applied separately to each homogenous subseries segment (see above). If a significant trend ($p \leq 0.05$) of the

same sign is detected for each segment, then the breakpoint is disregarded as an artifact of the trend. Otherwise, the breakpoint is accepted and the trend is disregarded. The confidence level used to evaluate significance of the ACF was computed from the 97.5th percentile of the standard normal distribution. Our results were obtained using R 3.2.3 (R Core Development Team 2011) with the trend v0.0.1 package (Pohlert 2015).

After trends and breakpoints of local significance at the $\alpha = 0.05$ level are found, they are subsequently tested for field significance on the CONUS domain following the false discovery rate (FDR) approach as described by Benjamini and Hochberg (1995). Field significance testing allows for the determination of the percentage of simultaneous tests that are expected to show a trend, at a given local significance level, purely by chance (i.e., type I errors). The family-wise type I error rate (FWER), as it is called, is based on both the individual error rate and the number of comparisons (e.g., domain size). For a single (grid cell) comparison, the type I family error rate is equal to the local significance level ($\alpha = 0.05$). However, each additional (grid cell) comparison causes the FWER to increase in a cumulative manner. Only by accounting for FWER may the significance of local (grid cell) statistics be supported with appropriate confidence.

The FDR procedure controls this multiplicity (selection) effect by requiring local p values [computed per Eq. 2.12 in Pettitt (1979)] be less than or equal to a rescaled global significance level α_{global} :

$$p_i \leq \frac{i}{m} \alpha_{\text{global}}, \quad (1)$$

where $p_1 \leq p_2 \leq \dots \leq p_m$ are the ordered local p values, m is the number of tests (cells; $m = 52414$), and α_{global} is set in our case to equal 0.05. Only those points where p_i satisfies (\leq) Eq. (1) are deemed field significant at the global 5% significance level. Field significance was determined in this way for each analysis case (i.e., month/6-h time window/precipitation aspect) for trends and breakpoints. Importantly, the full distribution of local p values over CONUS is included in the ordered p set, not just the subset satisfying local significance at the 0.05 level (Tramblay et al. 2013). If none of the local p values satisfy the bracketed conditions in Eq. (1), then none of them are deemed to be field significant, and consequently the global null hypothesis is not rejected (Wilks 2006). Field significance depends on the choice of domain boundaries. In this case, the CONUS domain was selected because it encompasses the area impacted by stage II radar data integration.

Prior studies have proven the FDR procedure of Benjamini and Hochberg (1995) robust to positive cross correlations (including spatial correlations) and suitable for work with any statistical test for which one can

generate a p value (Khaliq et al. 2009; Renard et al. 2008; Ventura et al. 2004; Wilks 2006). Relative to the conventional counting and Walker's tests for field significance, Wilks (2006) found FDR to be preferable in most instances. A convenience of the FDR is that its ceiling (and proportion of false rejections) may be controlled in the context of the global test [Eq. (1)]. Further explanation of the FDR method is provided by Miller et al. (2001).

3. Results

a. CONUS breakpoint and trend detection

An initial step in our methodology (see section 2d) was to screen for lag-1 autocorrelation in the data. When data are significantly autocorrelated, any test statistics derived from those data will be based on an effective sample size that is less than the actual sample size. The average percent of CONUS cells affected by autocorrelation is 1.8%, 8.4%, and 5.0% across all months and 6-h time windows for NLDAS-2 precipitation amount, frequency, and intensity, respectively. The equivalent statistics for NARR are 1.9%, 2.6%, and 2.2%. For MERRA-2 prectot and prectotcorr, the statistics are comparably less: $\sim 1.3\%$, $\sim 1.4\%$, and $\sim 1.1\%$. For the monthly precipitation amount analyses (discussed below), the average percent of CONUS cells affected by autocorrelation is 0.8%–1.6% across all months and data products (i.e., NLDAS-2, NARR, MERRA-2 prectot and prectotcorr, and L15). Accordingly, while correcting for autocorrelation is important, it was probably not essential to ensuring validity of our statistical analysis in this case.

Another aspect of our approach, field significance testing, should also be discussed at this point. To satisfy field significance, local (grid cell) p values needed to be substantially less than 0.05. For example, NLDAS-2 local p values less than or equal to 0.015, 0.042, and 0.029 were required to satisfy field significance at $\alpha = 0.05$, respectively, for 6-h precipitation amount, frequency, and intensity. In the case of NARR, those values were 0.009, 0.008, and 0.011, respectively. For MERRA-2 prectot and prectotcorr, all local p values had to satisfy $p \leq 4.5 \times 10^{-4}$ for precipitation amount, frequency, and intensity. For trends and breakpoints in daily precipitation amount, Eq. (1) is rarely satisfied. In summary, our results satisfy a much stricter significance criterion than studies testing local significance only.

To first rule out the possibility of spurious shifts in the NLDAS-2 daily mean precipitation record, we performed trend and breakpoint analyses on monthly mean precipitation amount from NLDAS-2 (1979–2015) and four comparison datasets at $0.125^\circ \times 0.125^\circ$ resolution: NARR (1979–2015), MERRA-2 prectot (1980–2015),

MERRA-2 prectotcorr (1980–2015), and L15 (1979–2013). Over CONUS, no field significant trends or breakpoints were detected in NLDAS-2, MERRA-2 (prectot and prectotcorr), or L15—for any month (i.e., January, February, . . . , December). Only a negligible number (less than 30) were found in the case of NARR (not shown). Having verified consistency among the five products at the daily time scale in this sense, we proceeded with our planned analyses at subdaily time scale.

Analysis of subdaily precipitation amount yielded noteworthy results for the 0–5 DST time window. More than 10% of CONUS was found to have field significant breakpoints for the months of January–March, May, and November (Figs. 2a, 3). The coverage of field significant breakpoints in NARR 0–5 DST precipitation amount was generally below 10% (Fig. 2b). Both NLDAS-2 and NARR 0–5 DST precipitation amount time series yielded field significant trends in March over 7% of CONUS, largely organized in the central United States (Fig. S1 in the supplemental material). Over the 37-yr period (1979–2015), the detected trends of $0.1 \text{ mm h}^{-1} \text{ yr}^{-1}$ amount to a total reduction of 2.2 cm in 0–5 DST precipitation amount (Fig. S1 in the supplemental material). The CONUS field significance coverage for trends in all other months and 6-h periods (i.e., 6–11, 12–17, and 18–23 DST) was less than 2.7% and 1% for NLDAS-2 and NARR, respectively (Figs. S2–S4, S11–S14 in the supplemental material).

The trends in NLDAS-2 April–September 12–17 DST precipitation amounts that originally motivated this study (Fig. 1) were determined to be artifacts of breakpoints. Of the statistically significant trends detected, most—93% east of the Mississippi River—can be attributed to abrupt, statistically significant breakpoints. Figure 4 underscores the pervasiveness of breakpoints in frequency and intensity, especially. It also shows the role of time-averaging period on detection of field significant breaks in NLDAS-2 precipitation amount, frequency, and intensity. Relative to precipitation frequency and intensity, the detectability of breakpoints in amount is much more sensitive to the averaging period. Specifically, detectability undergoes a steady decline from 6 to 3 months, and a precipitous drop from 3 to 1 month. This result is likely related to differences in the natural variability of precipitation amount versus frequency and intensity; a higher natural variability in precipitation amount could explain lower breakpoint detectability.

Field significant breakpoints in NLDAS-2 precipitation frequency cover 24%–87% of CONUS for any given month and 6-h time window with median coverages of 63%, 49%, 39%, and 60% for 0–5, 6–11, 12–17, and 18–23 DST, respectively (Figs. 5a–d, 6; Figs. S5–S7 in the supplemental material). The coverage of field significant

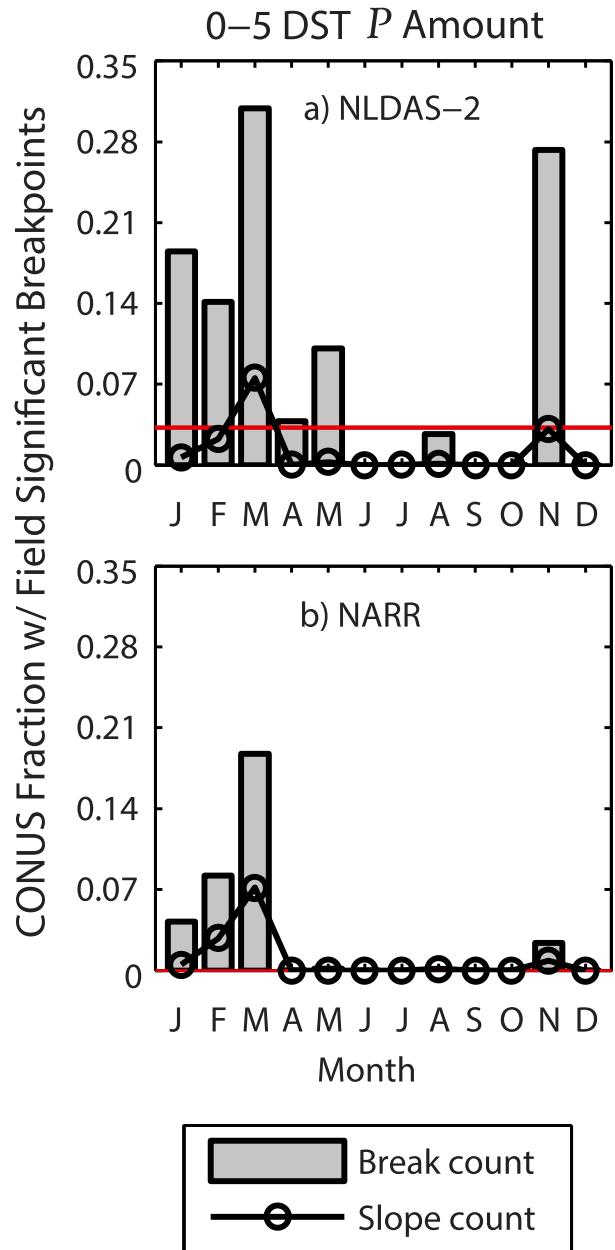


FIG. 2. The fraction (0–1) of CONUS $0.125^\circ \times 0.125^\circ$ grid cells affected by either a Pettitt breakpoint (gray bars) or Sen's slope (line with circles) of field significance at the 0.05 level for (a) NLDAS-2 and (b) NARR 0–5 DST *P* amount during the period from 1979 to 2015. Horizontal red lines denote monthly median values for breakpoint coverage.

breakpoints in NARR precipitation frequency ranges from only 0% to 16% of CONUS for all months and 6-h time windows (Figs. 5i–l; Figs. S15–S18 in the supplemental material). Less than 2% of CONUS is found to exhibit a field significant trend in NLDAS-2 or NARR precipitation frequency for all months and 6-h time windows (not shown).

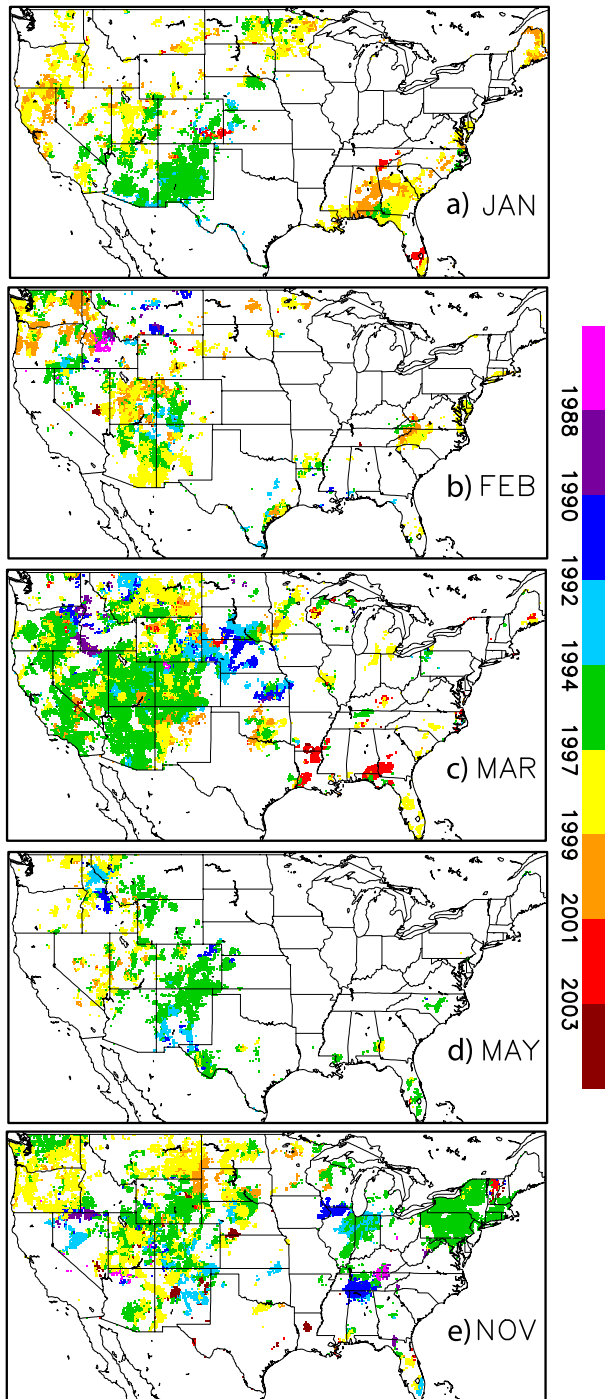


FIG. 3. Maps of Pettitt field significant break dates at the 0.05 level for 0–5 DST NLDAS-2 P amount, for months with greater than 10% field significant coverage (Fig. 2a): (a) January, (b) February, (c) March, (d) May, and (e) November.

Field significant breakpoints in NLDAS-2 precipitation intensity were generally less widespread than those detected in frequency, but still remarkable, with a median CONUS coverage of 2.7%, 24%, 40%, and 15% for

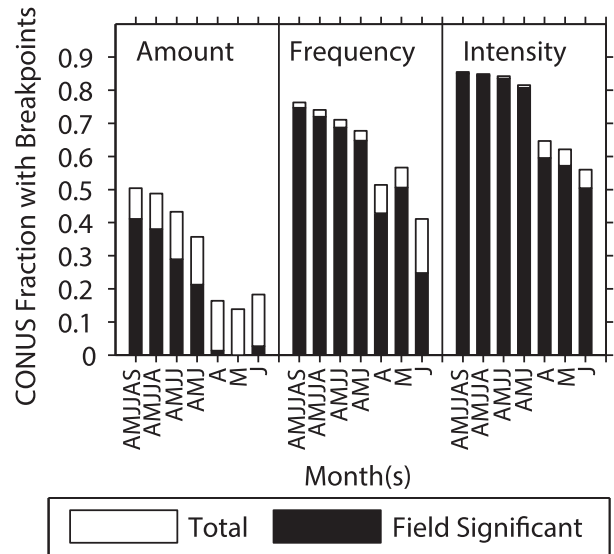


FIG. 4. For 12–17 DST P amount, frequency, and intensity, the CONUS fraction with detected Pettitt breakpoints of local significance at the 0.05 level (hollow bar) as well as of field significance at the 0.05 level (filled bar). Statistics are provided for a cascade of diminishing time periods, including: April–September (AMJJAS), April–August (AMJJA), April–July (AMJJ), April–June (AMJ), April (A), May (M), and June (J).

0–5, 6–11, 12–17, and 18–23 DST (Figs. 5e–h, 7; Figs. S8–S10 in the supplemental material). The median CONUS breakpoint count in NARR intensity is approximately 7% for 0–5 DST and less than 0.3% for all other 6-h time windows (Figs. 5m–p). For the 0–5 DST window, it is interesting to note that whereas NLDAS-2 experienced a large number of breakpoints in frequency and a negligible number of breakpoints in intensity, the opposite occurs with NARR (Figs. 5a,e,i,m). Accordingly, the breakpoints in NARR 0–5 DST precipitation amount (Fig. 2b) are more likely to be explained by breakpoints in intensity than frequency.

Overall, there is no discernible seasonality to the breakpoint count results (Figs. 2, 5). Spatially, however, a few seasonal tendencies can be noted: 1) breakpoints in NLDAS-2 0–5 DST precipitation amount tend to occur in the west (Fig. 3); 2) breakpoints in NLDAS-2 12–17 DST precipitation frequency occur in the south and Southeast in December–February and expand north and northeast over the course of the year (Fig. 6); and 3) breakpoints in NLDAS-2 12–17 DST precipitation intensity are focused over the eastern United States in July and August, are additionally detected in the Northwest from November–March and in the high plains from April to June, and generally are not detected over the south and Southwest (Fig. 7). Only the breakpoint count maps corresponding to NLDAS-2 12–17 DST frequency and intensity are included here in the interest of space and keeping with the

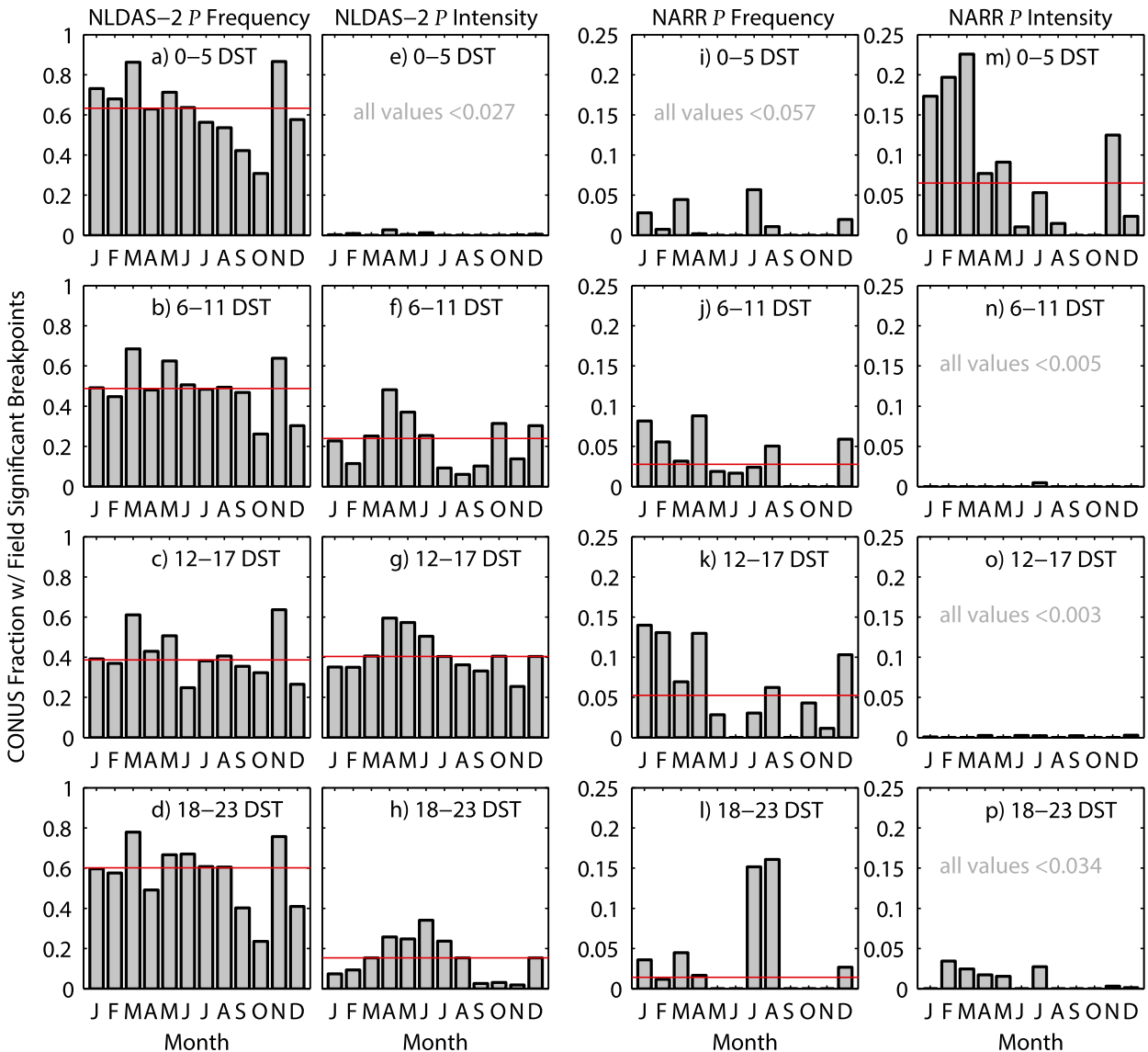


FIG. 5. The fraction (0–1) of CONUS $0.125^\circ \times 0.125^\circ$ grid cells affected by a Pettitt breakpoint (gray bars) of field significance at the 0.05 level during the period from 1979 to 2015 computed for (a)–(h) NLDAS-2 and (i)–(p) NARR 6-hourly P frequency and intensity. Note the y-axis scales for NLDAS-2 and NARR differ. Horizontal red lines denote monthly median values where nonnegligible.

underlying land–atmosphere coupling theme (Song et al. 2016). However, breakpoint count maps for both NLDAS-2 and NARR, for all other months, 6-hourly time periods, and precipitation aspects are included as supplemental material (Figs. S2–S22).

Figure 8 provides the break-date histograms corresponding to the breakpoint count bar plots provided in Figs. 2 and 5. Figure 8a shows the break-date histograms for NLDAS-2 (bar) and NARR (line) 0–5 DST amount. Figures 8b and 8c and Figs. 8d and 8e show the break-date histograms for NLDAS-2 and NARR frequency and intensity, respectively. Note that in Figs. 8b–e, the bar graph corresponds to the full sample of all 6-h time window

results, whereas each grayscale line corresponds with a different 6-h time period. Recall that a break date marks the last year in a homogenous segment, or the year prior to the shift. The median break date of the total sample (bar histograms) is marked by a vertical red line, which is shown to fall between 1996 and 1999. Using the Bai–Perron breakpoint test (Bai and Perron 2003; Ferguson and Villarini 2014), we tested the same time series (i.e., NLDAS-2 and NARR estimates, for all months, 6-h time windows, and precipitation aspects) for significant breakpoints in 1996 and 1997 (i.e., using a segment length of $h = 18$ yr). All instances of a Pettitt field significant breakpoint were found to satisfy Bai–Perron

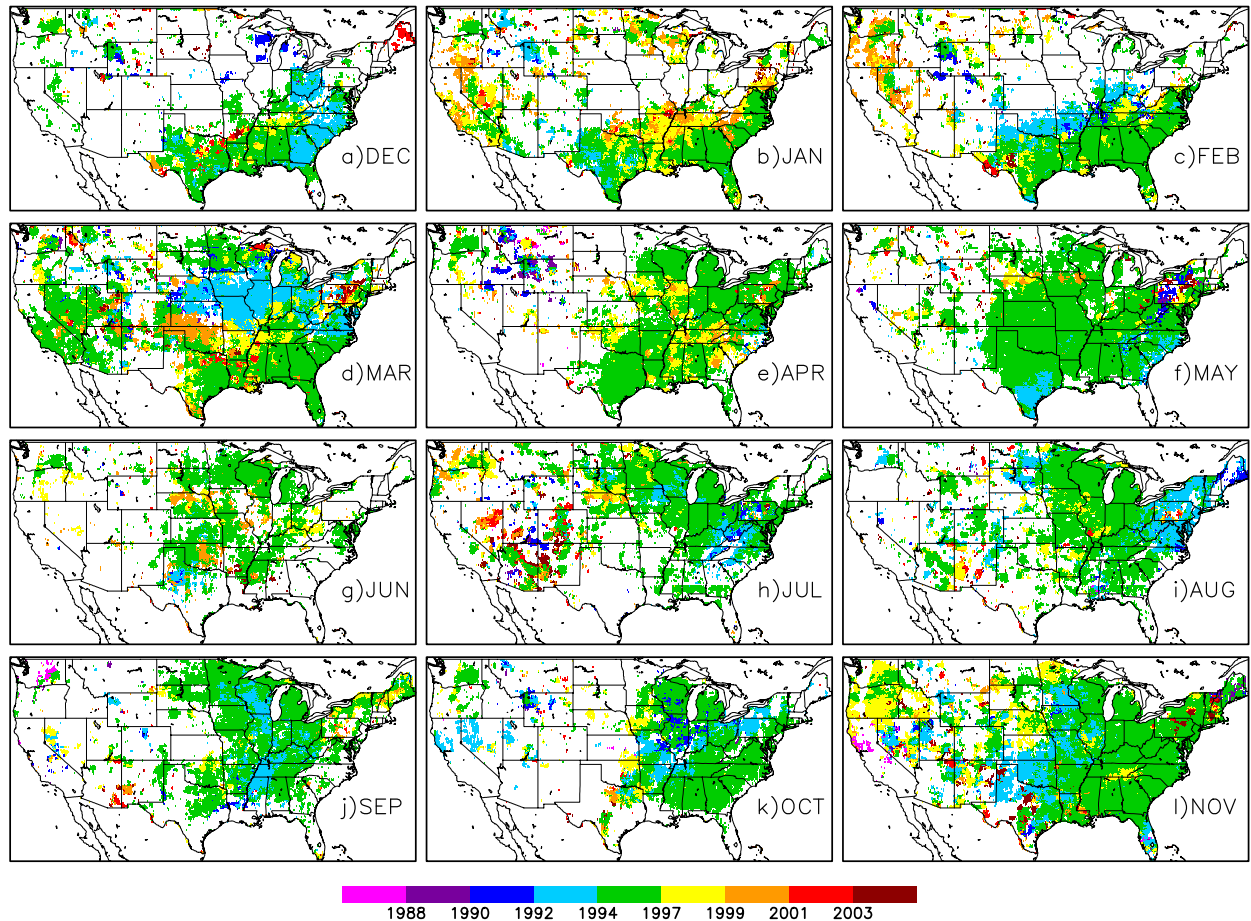


FIG. 6. Maps of Pettitt field significant break dates at the 0.05 level for 12–17 DST NLDAS-2 P frequency. Maps for other 6-h windows (i.e., 0–5, 6–11, and 18–23 DST) are provided as supplemental material.

breakpoint requirements at the 0.05 local significance level. Conversely, for all instances (i.e., NLDAS-2 and NARR estimates, for all months, 6-h time windows, and precipitation aspects) of locally significant Bai–Perron breakpoints fewer than six were uncorroborated by a Pettitt field significant breakpoint.

The fact that the median break date for NLDAS-2 corresponds with the known July 1996 transition to stage II hourly radar data for temporal disaggregation is strongly suggestive of a causal link. However, the fact that the limited number of NARR cells with breakpoints also tend to experience breaks around that time, especially in 0–5 DST precipitation intensity (Fig. 8e), suggests the potential of either a shared bias in the underlying CPC/OHD data, and/or the effect of natural variability. Indeed, 1997–98 is well known for an extremely strong El Niño event followed by a strong La Niña event (Cai et al. 2014; Changnon 2000; Compo and Sardeshmukh 2010; McPhaden 1999). However, a natural variability hypothesis is not supported by the MERRA-2 prectot and prectotcorr results; less than

0.3% of all CONUS grids were determined to have field significant breakpoints or trends in the case of all months, 6-h time windows, and precipitation aspects (not shown).

In Fig. 9, the results for four grid points selected for significant increasing trends in NLDAS-2 12–17 DST precipitation amount adequately demonstrate the methodology in its application. Readily apparent is the influence of trends (red lines) and breakpoints (gray columns) on the detection of one another. Significant trends in precipitation amount at three grids are disregarded due to the presence of breakpoints (Figs. 9a–c). Further, the issue of local versus field significance is illustrated. While the breakpoints in Figs. 9a and 9b appear substantial when taken in isolation, only breakpoints in Figs. 9c, 9e, 9h, and 9j–l are field significant at the global 0.05 level. As is typical, none of the breakpoint-generated series segments shown have a statistically significant slope, even at the local 0.05 level. In Fig. 9d, NLDAS 12–17 DST P is shown to be statistically homogenous at the 0.05 level with a significant slope. If the series is instead tested at the

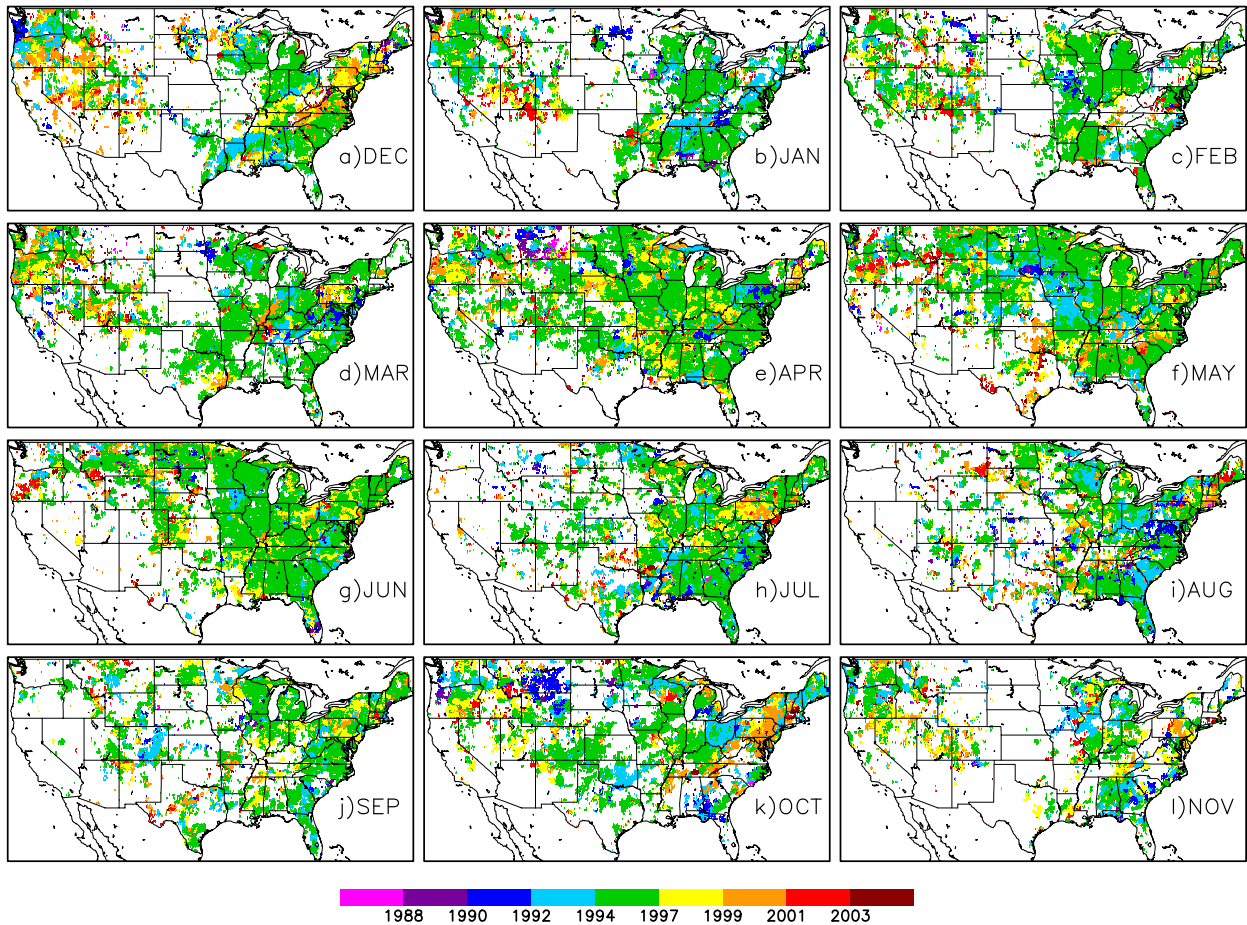


FIG. 7. Maps of Pettitt field significant break dates at the 0.05 level for 12–17 DST NLDAS-2 *P* intensity. Maps for other 6-h windows (i.e., 0–5, 6–11, and 18–23 DST) are provided as supplemental material.

0.11 level, then one would find a significant 1996 breakpoint (not shown). In summary, Fig. 9 highlights the sensitivity of results to the choice of significance threshold and the prioritization of slope versus breakpoint detection.

b. Regional impacts

We have summarized the impact of breakpoints on NLDAS-2 time-averaged precipitation amount, frequency, and intensity for six subdomains of CONUS: Northeast, Southeast, Midwest, high plains, south, and west (see Fig. 10 for delineations). Specifically, Fig. 11 illustrates the difference in pre- and post-July 1996 area-averaged means for each month, 6-h time window, and precipitation aspect. Note that all land grids were included in the calculation, whether or not a field significant breakpoint was detected at their location. Filled circles denote a difference at the 0.01 significance level.

The sharpest declines in precipitation amount between pre- and post-July 1996 periods across all regions and months occur for 0–5 DST. The largest increases

occur in the 12–17 DST time window over the Northeast and Southeast. For the west, precipitation amount increased for 6–11 and 12–17 DST and decreased for 18–23 and 0–5 DST. In the south, a similar subdaily statement holds, except that 18–23 DST precipitation declined for May–September and changed very little for October–April. In the Southeast, precipitation declined for all 6-h windows, except 12–17 DST. In the Northeast, precipitation generally declined for 0–5 and 6–11 DST and increased for 12–17 and 18–23 DST (Figs. 11a–f).

The Northeast and Southeast experienced 5%–20% and 10%–35% declines in precipitation frequency, respectively, with the largest declines occurring between May and August for the 18–23 DST period. In fact, the 18–23 DST declines in the Southeast are the largest of declines for all regions, months, and 6-h time windows. The Midwest and high plains experienced 10%–25% and 0%–15% declines, respectively, while the Midwest results displayed perhaps the most well-defined seasonality of all regions. Overall, the smallest precipitation frequency

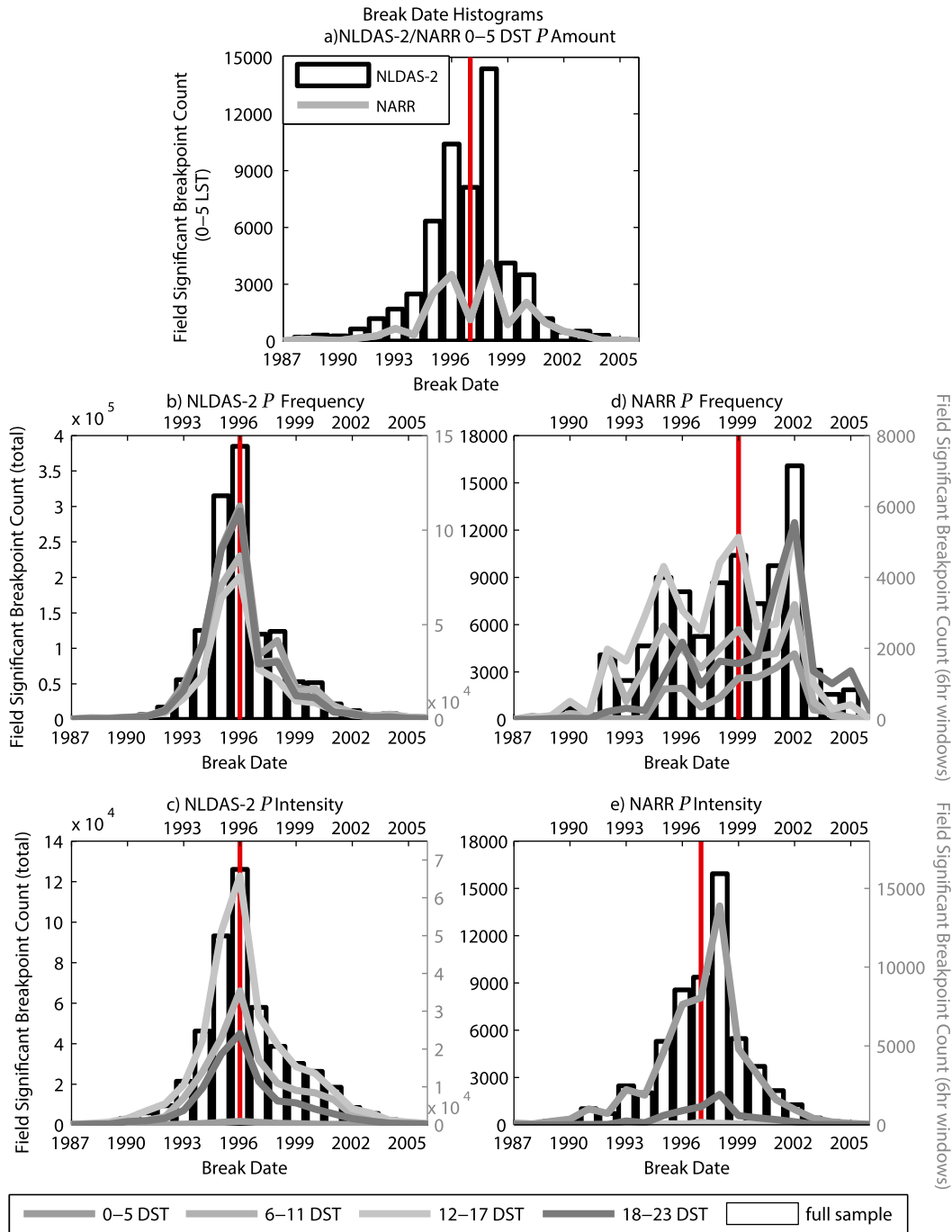


FIG. 8. Histograms of Pettitt field significant break dates at the 0.05 level corresponding to Figs. 2 and 5: (a) 0-5 DST NLDAS-2 and NARR *P* amount, NLDAS-2 *P* (b) frequency and (c) intensity, and NARR *P* (d) frequency and (e) intensity. The hollow black bars denote the histogram of the full sample, whereas the grayscale lines denote the respective histogram for each 6-h time window (i.e., 0-5, 6-11, 12-17, and 18-23 DST), except in (a), where the gray line denotes the NARR histogram. Vertical red lines denote the median break dates for each full sample histogram.

declines occurred for the December-February period over the Northeast, Midwest, and high plains. The south experienced 10%-20% declines on average with slightly more substantial declines during May-July for

the 18-23 DST window. The west experienced 0%-20% declines, with the exception of a slight increase in July for the 12-17 DST period. The largest declines over the west occurred between November and March (Figs. 11g-l).

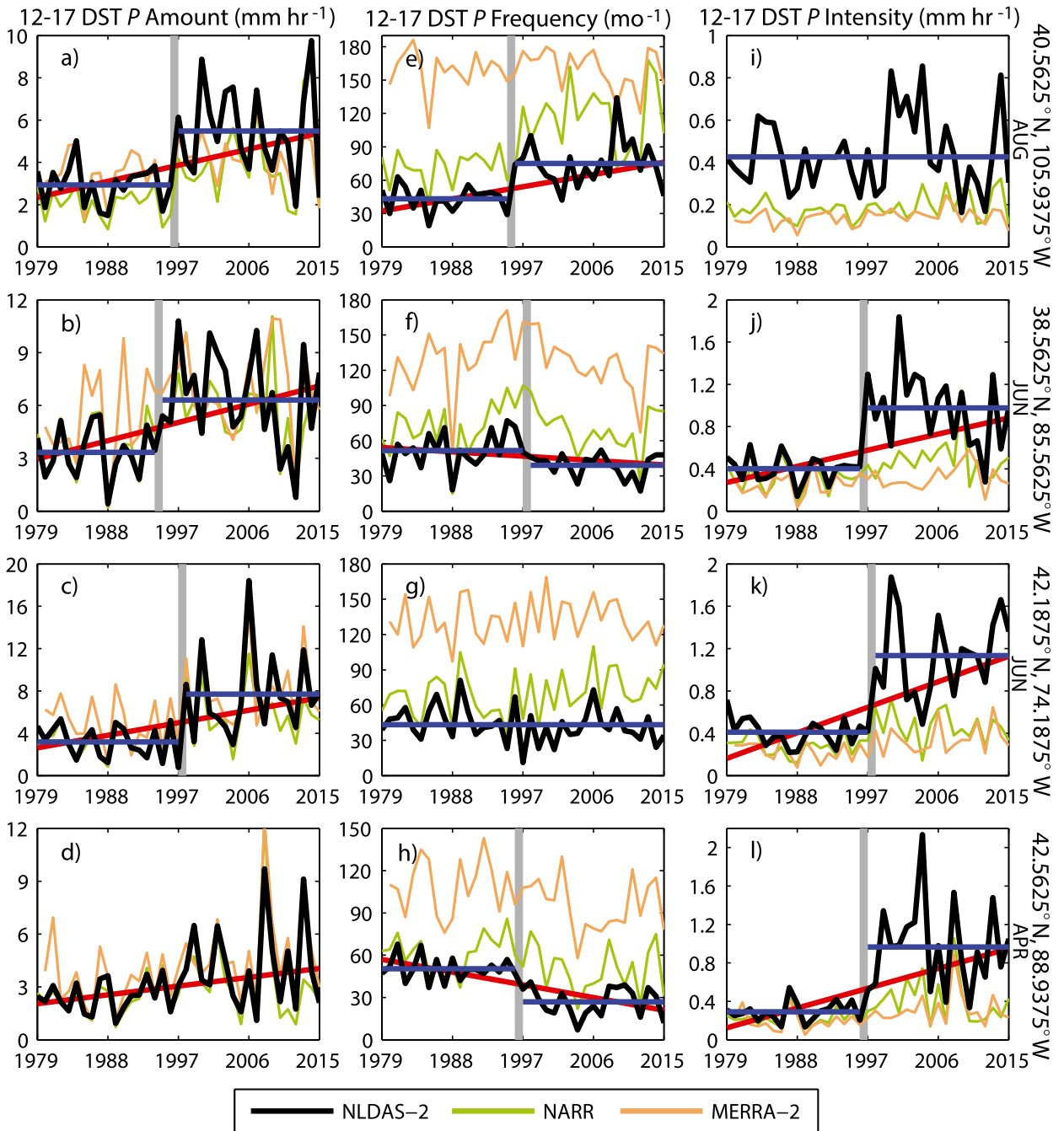


FIG. 9. Theil-Sen slope (red) and Pettitt breakpoint (gray) results for the NLDAS-2 12-17 DST P (a)–(d) amount (mm h^{-1}), (e)–(h) frequency (h month^{-1}), and (i)–(l) intensity (mm h^{-1}) time series at four grids (see Fig. 10). From top to bottom, the time series correspond to the following months: August, June, June, and April. Slope and breakpoint results are shown only when they meet local significance at the 0.05 level. Blue lines denote the mean of statistically homogenous series (g) and (i), as well as series segments [all other panels, except (d)]. Significant lag correlation was detected for (k), so analysis was conducted on a blended TFPW series (see section 2d). For context, the NARR P and MERRA-2 prectotcorr time series are included.

All regions experienced an increase in precipitation intensity in the range of 0–0.6 mm h^{-1} , with the exception of January and March in the high plains and May–September in the west for which declines were found in 0–5 DST

precipitation intensity. Notably, 0–5 DST saw the least increases of all 6-h time windows in all regions and months. The largest increase in precipitation intensity occurred in September over the Northeast during the 12–17 DST time

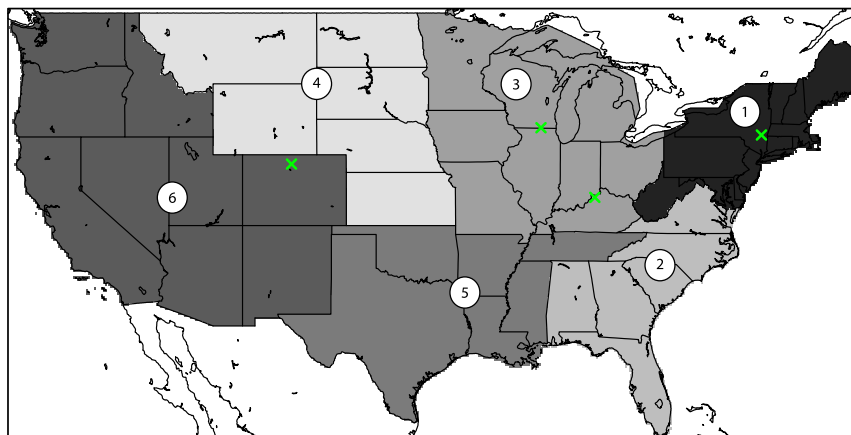


FIG. 10. Delineated in grayscale, the six CONUS subdomains selected to quantify broad area-averaged impacts of stage II hourly Doppler radar P estimates on NLDAS-2 subdaily P characteristics since their July 1996 introduction. The domains are 1) Northeast, 2) Southeast, 3) Midwest, 4) high plains, 5) south, and 6) west. They are consistent with the climate regions of the U.S. Drought Monitor with slight variation in the high plains and west. The green crosses correspond to grid locations called out in Fig. 9.

period. The Midwest's response portrays a strong seasonality, with a June–July maximum. Compared to other regions, the Northeast and Southeast demonstrate a relatively larger variation in subdaily response (i.e., among 6-h time periods; Figs. 11m–r). The fact that subdaily variation (i.e., dispersion) in mean area-averaged response is least for frequency and substantially greater for amount and intensity is supportive of the 1996 radar hypothesis. This result is also consistent with the greater prominence of field significant breaks in NLDAS-2 frequency as compared to amount and intensity (Figs. 2, 5).

A corresponding plot (to Fig. 11) for NARR is provided as supplemental material (Fig. S23). In the case of precipitation amount, comparable shifts in magnitude and seasonality are observed for all region–month–time windows, although we do not see the same pronounced peak impact in the Southeast at 12–17 DST. With regard to precipitation frequency, declines are much less pronounced in NARR with frequency increases even found for some regions. For example, whereas NLDAS-2 frequency has diminished over the high plains and west, NARR frequency has increased. Shifts in precipitation intensity are relatively smaller in NARR—ranging from -0.1 to 0.2 —relative to -0.2 to 0.6 for NLDAS-2. The NARR intensity shifts exhibit relatively less pronounced seasonality over the Midwest and high plains relative to NLDAS-2. Otherwise, for other regions the NARR and NLDAS-2 seasonality is similar.

4. Summary and conclusions

We have presented a rigorous trend assessment of NLDAS-2 subdaily precipitation amount, frequency, and

intensity over CONUS during the period from 1979 to 2015. Two critical aspects of the methodology, the simultaneous consideration of breakpoints and the evaluation of field significance, were essential to reaching our conclusion that the July 1996 switch to stage II radar data for use in temporal disaggregation of daily gauge-based inputs introduced a significant discontinuity in NLDAS-2 subdaily precipitation frequency and intensity. Unaccounted for, this discontinuity, or breakpoint, leads to the erroneous detection of significant trends in precipitation. Breakpoint-induced trends are particularly evident in warm season (April–September) afternoon (12–17 DST) precipitation records east of the Mississippi River (Figs. 1, 4). When breakpoints are accounted for, less than 2% of CONUS is found to exhibit field significant trends in any aspect of subdaily precipitation at the monthly scale, with the exception of 0–5 DST precipitation amount over the central United States (Fig. S1 in the supplemental material). Field significant testing was undertaken to account for FWER, which is especially relevant considering spatial correlation lengths of precipitation, and enabled us to assign appropriate confidence to the significance of local (grid cell) statistics. Unfortunately, the issue of field significance is commonly overlooked in weather and climate studies.

The attribution of detected breakpoints is always a difficult task. Breakpoint detectability is a function of several factors, including the spatial and temporal scales at which the analysis is conducted (e.g., Fig. 4), the background natural variability of the target variable (e.g., precipitation) over the study domain and period (e.g., season), the record length, the number of (natural and artificial) inhomogeneities, as well as the sensitivity (or

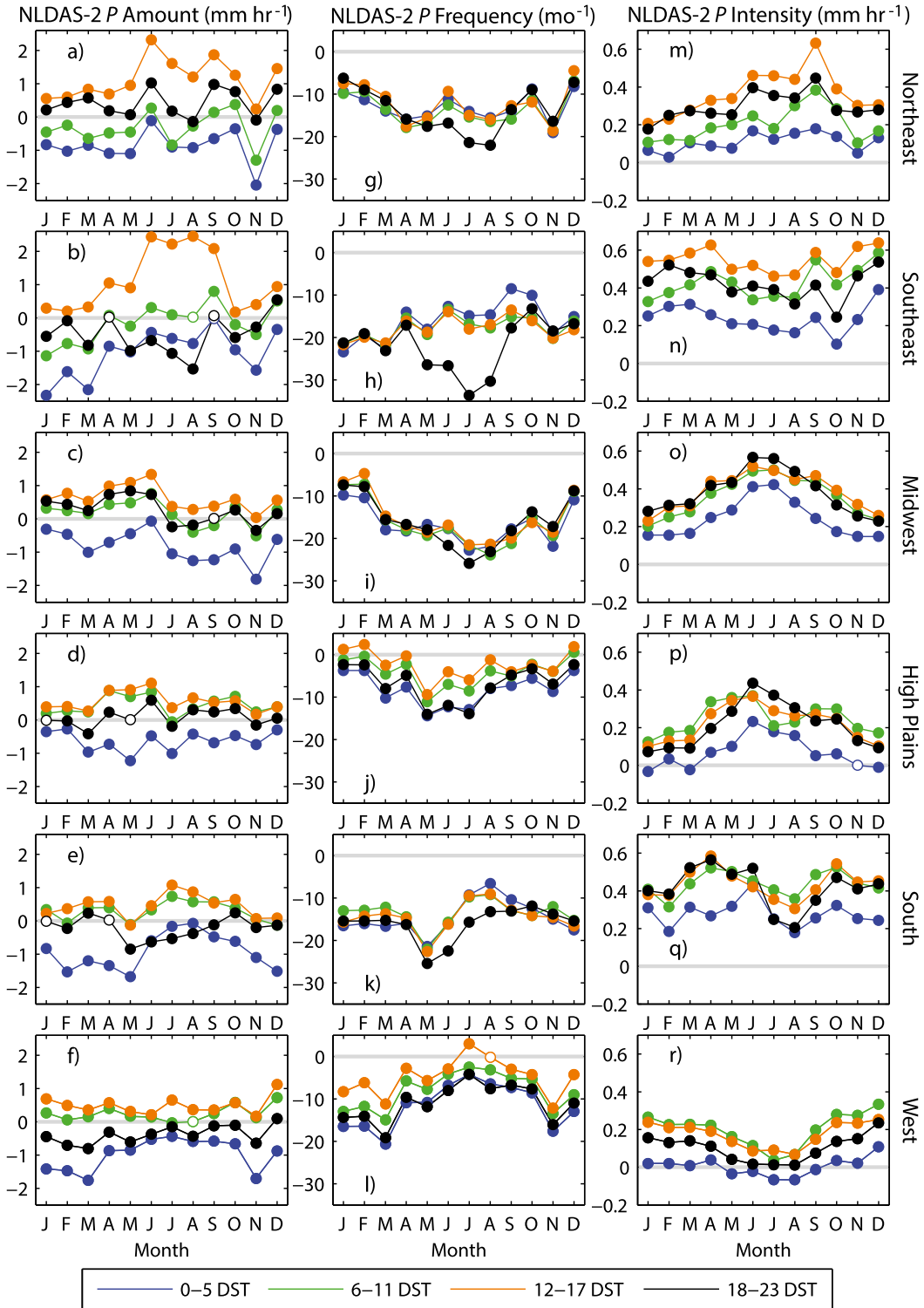


FIG. 11. The shift in 0.125° subdaily NLDAS-2 P (a)–(f) amount (mm h^{-1}), (g)–(l) frequency (h month^{-1}), and (m)–(r) intensity (mm h^{-1}) between preradar (from January 1979 to July 1996) and postradar (from August 1996 to December 2015) period means (i.e., postradar minus preradar) for the six CONUS subregions shown in Fig. 9. Lines are colored according to the corresponding 6-h time window. Filled circles signify that the difference of the period means is statistically significant at the 0.01 level according to a two-tailed Student’s t test; only 13 (1.5%) differences failed to achieve significance at this level. A gray horizontal line is drawn to highlight $y = 0$ (no change in mean).

limitations) of the detection algorithm itself (e.g., [Ferguson and Villarini 2012](#); [Mallakpour and Villarini 2016](#)). Because of the issues of detectability, discontinuities—even those occurring over large spatial scales—are not uniformly detectable across affected grids. In our application, near coincidence of a natural breakpoint, the very strong 1997/98 El Niño, and the July 1996 median break date ([Fig. 8](#)) could confound breakpoint attribution. Ideally, supporting metadata are available in these instances to enable the distinction between natural and artificial breakpoints (i.e., test the observational network hypothesis). Indeed, the NLDAS-2 July 1996 system-wide introduction of stage II radar data is well documented.

Other important aspects of the NLDAS-2 precipitation analysis were not documented, however, such as the precise precipitation product used in hourly temporal disaggregation at the daily, grid-specific scale. Forcing data discontinuities of this kind also serve to confound breakpoint attribution. Currently, the lessons learned in this and similar studies are being applied toward making the third phase of NLDAS more transparent. Specifically, fields will be included in each hourly forcing file that detail how many (and which) stations were used in the precipitation analysis, which dataset was used for temporal disaggregation at each grid point for each hour, data quality indicators, etc. Other data providers have already made progress along similar lines. Namely, observation feedback archives that document assimilated observations at each analysis time were recently released for the NOAA Cooperative Institute for Research in Environmental Sciences (CIRES) Twentieth Century Reanalysis (20CR, version 2c; [Compo et al. 2011](#)) and the 111-yr European Centre for Medium-Range Weather Forecasts (ECMWF) pilot reanalysis of the twentieth century (ERA-20C; [Poli et al. 2013](#)).

In spite of the aforementioned detectability issues and the likely existence of additional natural and artificial discontinuities in the NLDAS-2 precipitation record, our results indicate that the single most remarkable discontinuity occurs circa 1996 and is linked to the July 1996 radar transition. The strongest evidence in support of the radar hypothesis is the detection of large numbers of breakpoints in subdaily precipitation frequency and intensity, as compared to a negligible number in precipitation amount ([Figs. 2, 5](#)). Furthermore, breakpoint counts were largely undetected in NARR, MERRA-2 prectot, and MERRA-2 prectotcorr, across all aspects of subdaily precipitation. All of these products use a similar daily gauge-based analysis. So, any NLDAS-2 differences should be linked to its hourly temporal disaggregation. Coincidentally, the median break date (1996–99) matches irrefutably well with the July 1996 transition to stage II

radar for temporal disaggregation ([Fig. 8](#)). The likelihood that circa 1996 breakpoints, including those detected in 0–5 DST NARR precipitation amount, are linked to natural variability is low because they are not corroborated by MERRA-2.

The extent of the 1996 discontinuity's impact on regional mean subdaily precipitation amount, frequency, and intensity was quantified for six subdomains of CONUS. The severity of the discontinuity's knock-on effect in land surface modeling applications is likely to vary widely according to surface scheme (e.g., physics, soil, and vegetative parameterizations), study domain, and simulation period. Issues with the precipitation will affect the soil moisture, land surface temperature, surface energy and water fluxes, snow cover, snow water equivalent, and runoff/streamflow. [Nearing et al. \(2016\)](#) benchmarked the NLDAS-2 participating models using observations of soil moisture and evapotranspiration to separate uncertainty contributions of the forcing, the model physics, and the parameters. They found that parameters dominated uncertainty in the simulated soil moisture while forcing dominated uncertainty in the simulated evapotranspiration. The largest effects will likely occur in the densely vegetated Northeast and Southeast due to canopy interception and evapotranspiration processes (e.g., [Reichle et al. 2011](#)). In any case, NLDAS-2 subdaily precipitation is unfit for long-term trend assessment starting from before mid-1996. Concerning the most recent 20-yr record (1996–2015), no field significant Pettitt breakpoints were detected for any 6-h period, month, or precipitation aspect over CONUS.

As a standard practice, we recommend that users perform a statistical analysis on all key forcing datasets at the spatial–temporal resolution of their intended application prior to conducting model simulations. Separately, offline simulations may be used to assess the precise sensitivity of a given land scheme's hydrologic representativeness to shifts in the forcing (diurnal precipitation cycle). Importantly, our analysis did not discriminate between precipitation phase (solid vs liquid) and character (convective vs stratiform), which may be important considerations depending on the application. We note that the ongoing work of the GEWEX Hydroclimatology Panel (GHP) crosscut project Intelligent Use of Climate Models for Adaptation to Nonstationary Hydrological Extremes (INTENSE) stands to enhance the accessibility and quality of subdaily gauge-based precipitation records (e.g., [Blenkinsop et al. 2017](#)).

Acknowledgments. This research was funded through the NYSUNY 2020 Challenge Grant Program, the

Climate, Environment, and Energy: Impacts, Linkages, and Synergies; NASA Grant NNX16AM13G; and NOAA Climate Program Office Grant GC14-194A. The NLDAS-2 and MERRA-2 products were obtained through from the NASA GESDISC Data Archive accessed via the NASA Earth Observing System Data and Information System (EOSDIS). The L15 data were obtained directly from Ben Livneh. Yun Fan and Youlong Xia were instrumental in corroborating the timing and nature of changes to the NLDAS precipitation forcing production algorithm over the period of record.

REFERENCES

- Andreadis, K. M., E. A. Clark, A. W. Wood, A. F. Hamlet, and D. P. Lettenmaier, 2005: Twentieth-century drought in the conterminous United States. *J. Hydrometeorol.*, **6**, 985–1001, doi:10.1175/JHM450.1.
- Bai, J., and P. Perron, 2003: Computation and analysis of multiple structural change models. *J. Appl. Econ.*, **18**, 1–22, doi:10.1002/jae.659.
- Bárdossy, A., and H. J. Caspary, 1990: Detection of climate change in Europe by analyzing European atmospheric circulation patterns from 1881 to 1989. *Theor. Appl. Climatol.*, **42**, 155–167, doi:10.1007/BF00866871.
- Beaulieu, C., J. Chen, and J. L. Sarmiento, 2012: Change-point analysis as a tool to detect abrupt climate variations. *Philos. Trans. Roy. Soc. London*, **370A**, 1228–1249, doi:10.1098/rsta.2011.0383.
- Benjamini, Y., and Y. Hochberg, 1995: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.*, **57B**, 289–300.
- Blenkinsop, S., E. Lewis, S. C. Chan, and H. J. Fowler, 2017: Quality-control of an hourly rainfall dataset and climatology of extremes for the UK. *Int. J. Climatol.*, **37**, 722–740, doi:10.1002/joc.4735.
- Busuioc, A., and H. von Storch, 1996: Changes in the winter precipitation in Romania and its relation to the large-scale circulation. *Tellus*, **48A**, 538–552, doi:10.1034/j.1600-0870.1996.t01-3-00004.x.
- Cai, W. J., and Coauthors, 2014: Increasing frequency of extreme El Niño events due to greenhouse warming. *Nat. Climate Change*, **4**, 111–116, doi:10.1038/nclimate2100.
- Changnon, S. A., 2000: *El Niño 1997–1998: The Climate Event of the Century*. Oxford University Press, 232 pp.
- Charba, J. P., A. W. Harrell III, and A. C. Lackner III, 1992: A monthly precipitation amount climatology derived from published atlas maps: Development of a digital database. TDL Office Note 92-7, NOAA/NWS, 20 pp. [Available online at <http://www.nws.noaa.gov/mdl/pubs/Documents/OfficeNotes/OfficeNote1992-07.pdf>.]
- Chen, M. Y., W. Shi, P. P. Xie, V. B. S. Silva, V. E. Kousky, R. W. Higgins, and J. E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res.*, **113**, D04110, doi:10.1029/2007JD009132.
- Compo, G. P., and P. D. Sardeshmukh, 2010: Removing ENSO-related variations from the climate record. *J. Climate*, **23**, 1957–1978, doi:10.1175/2009JCLI2735.1.
- , and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, doi:10.1002/qj.776.
- Cosgrove, B. A., and Coauthors, 2003: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *J. Geophys. Res.*, **108**, 8842, doi:10.1029/2002JD003118.
- Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374, doi:10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.
- Crook, A. G., 1977: SNOTEL: Monitoring climatic factors to predict water supplies. *J. Soil Water Conserv.*, **32**, 294–295.
- Dai, A., 1999: Recent changes in the diurnal cycle of precipitation over the United States. *Geophys. Res. Lett.*, **26**, 341–344, doi:10.1029/1998GL900318.
- , F. Giorgi, and K. E. Trenberth, 1999: Observed and model-simulated diurnal cycles of precipitation over the contiguous United States. *J. Geophys. Res.*, **104**, 6377–6402, doi:10.1029/98JD02720.
- Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteorol.*, **33**, 140–158, doi:10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2.
- Dirmeyer, P. A., and Coauthors, 2012: Simulating the diurnal cycle of rainfall in global climate models: Resolution versus parameterization. *Climate Dyn.*, **39**, 399–418, doi:10.1007/s00382-011-1127-9.
- Erlingis, J. M., and A. P. Barros, 2014: A study of the role of daytime land–atmosphere interactions on nocturnal convective activity in the Southern Great Plains during CLASIC. *J. Hydrometeorol.*, **15**, 1932–1953, doi:10.1175/JHM-D-14-0016.1.
- Fan, Y., H. M. Van den Dool, D. Lohmann, and K. Mitchell, 2006: 1948–98 U.S. hydrological reanalysis by the Noah land data assimilation system. *J. Climate*, **19**, 1214–1237, doi:10.1175/JCLI3681.1.
- Ferguson, C. R., and G. Villarini, 2012: Detecting inhomogeneities in the Twentieth Century Reanalysis over the central United States. *J. Geophys. Res.*, **117**, D05123, doi:10.1029/2011JD016988.
- , and —, 2014: An evaluation of the statistical homogeneity of the Twentieth Century Reanalysis. *Climate Dyn.*, **42**, 2841–2866, doi:10.1007/s00382-013-1996-1.
- , J. Sheffield, E. F. Wood, and H. L. Gao, 2010: Quantifying uncertainty in a remote sensing–based estimate of evapotranspiration over continental USA. *Int. J. Remote Sens.*, **31**, 3821–3865, doi:10.1080/01431161.2010.483490.
- Findell, K. L., and E. A. B. Eltahir, 2003a: Atmospheric controls on soil moisture–boundary layer interactions. Part I: Framework development. *J. Hydrometeorol.*, **4**, 552–569, doi:10.1175/1525-7541(2003)004<0552:ACOSML>2.0.CO;2.
- , and —, 2003b: Atmospheric controls on soil moisture–boundary layer interactions. Part II: Feedbacks within the continental United States. *J. Hydrometeorol.*, **4**, 570–583, doi:10.1175/1525-7541(2003)004<0570:ACOSML>2.0.CO;2.
- Glahn, H. R., T. L. Chambers, W. S. Richardson, and H. P. Perrotti, 1985: Objective map analysis for the local AFOS MOS Program. NOAA Tech. Memo. NWS TDL 75, 34 pp. [Available online at <http://www.nws.noaa.gov/mdl/pubs/Documents/TechMemo/TechMemo75.pdf>.]
- GMAO, 2015: MERRA-2 tavg1_2d_flux_Nx: 2d,1-hourly,time-averaged,single-level,assimilation,surface flux diagnostics V5.12.4. Goddard Earth Sciences Data and Information Services Center, accessed 29 July 2016, doi:10.5067/7MCPBJ41Y0K6.
- Gochis, D. J., and Coauthors, 2016: Operational Hydrometeorological Predictions in Mountain Watersheds using the New National Water Model Implementation of the WRF-Hydro Modeling System. *17th Conf. on Mountain Meteorology*, Burlington, VA, Amer. Meteor. Soc., 15.6. [Available online at <https://ams.confex.com/ams/17Mountain/webprogram/Paper296619.html>.]

- Guo, Z. C., and Coauthors, 2006: GLACE: The Global Land–Atmosphere Coupling Experiment. Part II: Analysis. *J. Hydrometeorol.*, **7**, 611–625, doi:10.1175/JHM511.1.
- Henderson-Sellers, A., A. J. Pitman, P. K. Love, P. Irannejad, and T. H. Chen, 1995: The Project for Intercomparison of Land Surface Parameterization Schemes (PILPS): Phases 2 and 3. *Bull. Amer. Meteor. Soc.*, **76**, 489–503, doi:10.1175/1520-0477(1995)076<0489:TPFIOL>2.0.CO;2.
- Higgins, R. W., J. Janowiak, and J. P. Yao, 1996: A gridded hourly precipitation data base for the United States (1963–1993). NCEP/Climate Prediction Center Atlas 1, NOAA, 47 pp. [Available online at http://www.cpc.ncep.noaa.gov/research_papers/ncep_cpc_atlas/1/cover.html.]
- , W. Shi, E. S. Yarosh, and R. Joyce, 2000: Improved United States precipitation quality control system and analysis. NCEP/Climate Prediction Center Atlas 7, NOAA, 40 pp. [Available online at http://www.cpc.ncep.noaa.gov/research_papers/ncep_cpc_atlas/7/toc.html.]
- Hirsch, R. M., J. R. Slack, and R. A. Smith, 1982: Techniques of trend analysis for monthly water quality data. *Water Resour. Res.*, **18**, 107–121, doi:10.1029/WR018i001p00107.
- Janjić, Z. I., 2003: A nonhydrostatic model based on a new approach. *Meteor. Atmos. Phys.*, **82**, 271–285, doi:10.1007/s00703-001-0587-6.
- Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.*, **5**, 487–503, doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.
- Kendall, M. G., 1975: *Rank Correlation Methods*. 4th ed. Griffin, 202 pp.
- Khalik, M. N., T. B. M. J. Ouarda, P. Gachon, L. Sushama, and A. St-Hilaire, 2009: Identification of hydrological trends in the presence of serial and cross correlations: A review of selected methods and their application to annual flow regimes of Canadian rivers. *J. Hydrol.*, **368**, 117–130, doi:10.1016/j.jhydrol.2009.01.035.
- Klein, S. A., X. N. Jiang, J. Boyle, S. Malyshev, and S. C. Xie, 2006: Diagnosis of the summertime warm and dry bias over the U.S. Southern Great Plains in the GFDL climate model using a weather forecasting approach. *Geophys. Res. Lett.*, **33**, L18805, doi:10.1029/2006GL027567.
- Koster, R. D., and Coauthors, 2006: GLACE: The Global Land–Atmosphere Coupling Experiment. Part I: Overview. *J. Hydrometeorol.*, **7**, 590–610, doi:10.1175/JHM510.1.
- Lee, M. I., and Coauthors, 2007: An analysis of the warm-season diurnal cycle over the continental United States and northern Mexico in general circulation models. *J. Hydrometeorol.*, **8**, 344–366, doi:10.1175/JHM581.1.
- Lettenmaier, D. P., E. F. Wood, and J. R. Wallis, 1994: Hydroclimatological trends in the continental United States, 1948–88. *J. Climate*, **7**, 586–607, doi:10.1175/1520-0442(1994)007<0586:HCTITC>2.0.CO;2.
- Liang, X. Z., L. Li, A. Dai, and K. E. Kunkel, 2004: Regional climate model simulation of summer precipitation diurnal cycle over the United States. *Geophys. Res. Lett.*, **31**, L24208, doi:10.1029/2004GL021054.
- Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 1.2. [Available online at https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.]
- , M. E. Baldwin, K. E. Mitchell, E. Rogers, and G. DiMego, 2001: Spring 2001 changes to the NCEP Eta Analysis and Forecast System: Assimilation of observed precipitation. *18th Conf. Weather Analysis Forecasting/14th Conf. on Numerical Weather Prediction*, Fort Lauderdale, FL, Amer. Meteor. Soc., J2.7. [Available online at https://ams.confex.com/ams/WAF-NWP-MESO/techprogram/paper_23306.htm.]
- Liu, L., Z. X. Xu, and J. X. Huang, 2012: Spatio-temporal variation and abrupt changes for major climate variables in the Taihu basin, China. *Stochastic Environ. Res. Risk Assess.*, **26**, 777–791, doi:10.1007/s00477-011-0547-8.
- Livezey, R. E., and W. Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.*, **111**, 46–59, doi:10.1175/1520-0493(1983)111<0046:SFSRID>2.0.CO;2.
- Livneh, B., E. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. Andreadis, E. Maurer, and D. Lettenmaier, 2013: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *J. Climate*, **26**, 9384–9392, doi:10.1175/JCLI-D-12-00508.1.
- , T. J. Bohn, D. W. Pierce, F. Munoz-Arriola, B. Nijssen, R. Vose, D. R. Cayan, and L. Brekke, 2015: A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and southern Canada 1950–2013. *Sci. Data*, **2**, 150042, doi:10.1038/sdata.2015.42.
- Luo, Y., E. H. Berbery, and K. E. Mitchell, 2005: The operational Eta model precipitation and surface hydrologic cycle of the Columbia and Colorado basins. *J. Hydrometeorol.*, **6**, 341–370, doi:10.1175/JHM435.1.
- Mallakpour, I., and G. Villarini, 2016: A simulation study to examine the sensitivity of the Pettitt test to detect abrupt changes in mean. *Hydrol. Sci. J.*, **61**, 245–254, doi:10.1080/02626667.2015.1008482.
- Mann, H. B., 1945: Nonparametric tests against trend. *Econometrica*, **13**, 245–259, doi:10.2307/1907187.
- Matsui, T., D. Mocko, M. I. Lee, W. K. Tao, M. J. Suarez, and R. A. Pielke, 2010: Ten-year climatology of summertime diurnal rainfall rate over the conterminous U.S. *Geophys. Res. Lett.*, **37**, L13807, doi:10.1029/2010GL044139.
- Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *J. Climate*, **15**, 3237–3251, doi:10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2.
- McPhaden, M. J., 1999: El Niño: The child prodigy of 1997–98. *Nature*, **398**, 559–562, doi:10.1038/19193.
- Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, doi:10.1175/BAMS-87-3-343.
- Miller, C. J., and Coauthors, 2001: Controlling the false-discovery rate in astrophysical data analysis. *Astron. J.*, **122**, 3492–3505, doi:10.1086/324109.
- Mitchell, K. E., and Coauthors, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.*, **109**, D07S90, doi:10.1029/2003JD003823.
- Mo, K. C., L. C. Chen, S. Shukla, T. J. Bohn, and D. P. Lettenmaier, 2012: Uncertainties in North American Land Data Assimilation Systems over the contiguous United States. *J. Hydrometeorol.*, **13**, 996–1009, doi:10.1175/JHM-D-11-0132.1.
- Molod, A., L. Takacs, M. Suarez, and J. Bacmeister, 2015: Development of the GEOS-5 atmospheric general circulation

- model: Evolution from MERRA to MERRA2. *Geosci. Model Dev.*, **8**, 1339–1356, doi:10.5194/gmd-8-1339-2015.
- NASA, 2016: Land Data Toolkit: LDT 7.1 users' guide. NASA GSFC Doc., 129 pp. [Available online at https://modelingguru.nasa.gov/servlet/JiveServlet/previewBody/2555-102-3-6450/LDT_usersguide.pdf.]
- Nearing, G. S., D. M. Mocko, C. D. Peters-Lidard, S. V. Kumar, and Y. L. Xia, 2016: Benchmarking NLDAS-2 soil moisture and evapotranspiration to separate uncertainty contributions. *J. Hydrometeorol.*, **17**, 745–759, doi:10.1175/JHM-D-15-0063.1.
- Pan, M., H. B. Li, and E. Wood, 2010: Assessing the skill of satellite-based precipitation estimates in hydrologic applications. *Water Resour. Res.*, **46**, W09535, doi:10.1029/2009WR008290.
- Pettitt, A. N., 1979: A non-parametric approach to the change-point problem. *Appl. Stat.*, **28**, 126–135, doi:10.2307/2346729.
- Pitman, A. J., and A. Henderson-Sellers, 1998: Recent progress and results from the Project for the Intercomparison of Landsurface Parameterization Schemes. *J. Hydrol.*, **212–213**, 128–135, doi:10.1016/S0022-1694(98)00206-6.
- , and Coauthors, 1999: Key results and implications from phase 1(c) of the Project for Intercomparison of Land-Surface Parameterization Schemes. *Climate Dyn.*, **15**, 673–684, doi:10.1007/s003820050309.
- Pohlert, T., 2015: Trend: Non-parametric trend tests and change-point detection, version 0.0.1. R package. [Available online at <https://www.rdocumentation.org/packages/trend/versions/0.0.1>.]
- Poli, P., and Coauthors, 2013: The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20c). ERA Rep. Series 14, ECMWF, 62 pp. [Available online at <http://www.ecmwf.int/en/elibrary/11699-data-assimilation-system-and-initial-performance-evaluation-ecmwf-pilot-reanalysis>.]
- Prein, A. F., and Coauthors, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53**, 323–361, doi:10.1002/2014RG000475.
- R Core Development Team, 2011: R: A language and environment for statistical computing. R Foundation for Statistical Computing. [Available online at <http://www.R-project.org/>.]
- Reichle, R. H., R. D. Koster, G. J. M. de Lannoy, B. A. Forman, Q. Liu, S. P. P. Mahanama, and A. Touré, 2011: Assessment and enhancement of MERRA land surface hydrology estimates. *J. Climate*, **24**, 6322–6338, doi:10.1175/JCLI-D-10-05033.1.
- Renard, B., and Coauthors, 2008: Regional methods for trend detection: Assessing field significance and regional consistency. *Water Resour. Res.*, **44**, W08419, doi:10.1029/2007WR006268.
- Renka, R. J., 1988: Multivariate interpolation of large sets of scattered data. *ACM Trans. Math. Software*, **14**, 139–148, doi:10.1145/45054.45055.
- Rogers, E., T. Black, D. Deaven, G. DiMego, Q. Zhao, M. Baldwin, N. Junker, and Y. Lin, 1996: Changes to the operational “early” Eta analysis/forecast system at the National Centers for Environmental Prediction. *Wea. Forecasting*, **11**, 391–413, doi:10.1175/1520-0434(1996)011<0391:CTTOEA>2.0.CO;2.
- Ruane, A. C., 2010: NARR's atmospheric water cycle components. Part II: Summertime mean and diurnal interactions. *J. Hydrometeorol.*, **11**, 1220–1233, doi:10.1175/2010JHM1279.1.
- Sato, T., H. Miura, M. Satoh, Y. N. Takayabu, and Y. Q. Wang, 2009: Diurnal cycle of precipitation in the tropics simulated in a global cloud-resolving model. *J. Climate*, **22**, 4809–4826, doi:10.1175/2009JCLI2890.1.
- Sen, P. K., 1968: Estimates of the regression coefficient based on Kendall's tau. *J. Amer. Stat. Assoc.*, **63**, 1379–1389, doi:10.1080/01621459.1968.10480934.
- Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Sci. Rev.*, **99**, 125–161, doi:10.1016/j.earscirev.2010.02.004.
- Serreze, M. C., M. P. Clark, R. L. Armstrong, D. A. McGinnis, and R. S. Pulwarty, 1999: Characteristics of the western United States snowpack from Snowpack Telemetry (SNOTEL) data. *Water Resour. Res.*, **35**, 2145–2160, doi:10.1029/1999WR900090.
- Shafraan, P. C., J. Woolen, W. Ebisuzaki, W. Shi, Y. Fan, R. Grumbine, and M. Fennessy, 2004: Observational data used for assimilation in the NCEP North American Regional Reanalysis. *14th Conf. on Applied Climatology*, San Diego, CA, Amer. Meteor. Soc., 1.4. [Available online at <http://ams.confex.com/ams/pdfpapers/71689.pdf>.]
- Shepard, D. S., 1984: Computer mapping: The SYMAP interpolation algorithm. *Spatial Statistics and Models*, G. L. Gaile and C. J. Willmott, Eds., D. Reidel, 133–145.
- Song, H.-J., C. R. Ferguson, and J. K. Roundy, 2016: Land-atmosphere coupling at the Southern Great Plains Atmospheric Radiation Measurement (ARM) field site and its role in anomalous afternoon peak precipitation. *J. Hydrometeorol.*, **17**, 541–556, doi:10.1175/JHM-D-15-0045.1.
- Tao, W. K., and Coauthors, 2013: Precipitation intensity and variation during MC3E: A numerical modeling study. *J. Geophys. Res. Atmos.*, **118**, 7199–7218, doi:10.1002/jgrd.50410.
- Theil, H., 1950: A rank-invariant method of linear and polynomial regression analysis. *Proc. K. Ned. Akad. Wet.*, **53**, 386–392, 512–525, 1397–1412.
- Tramblay, Y., S. El Adlouni, and E. Servat, 2013: Trends and variability in extreme precipitation indices over Maghreb countries. *Nat Hazard Earth Sys*, **13**, 3235–3248, doi:10.5194/nhess-13-3235-2013.
- Venables, W. N., and B. D. Ripley, 2002: *Modern Applied Statistics with S*. 4th ed. Springer, 498 pp.
- Ventura, V., C. J. Paciorek, and J. S. Risbey, 2004: Controlling the proportion of falsely rejected hypotheses when conducting multiple tests with climatological data. *J. Climate*, **17**, 4343–4356, doi:10.1175/3199.1.
- Villarini, G., F. Serinaldi, J. A. Smith, and W. F. Krajewski, 2009: On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resour. Res.*, **45**, W08417, doi:10.1029/2008WR007645.
- Wallace, J. M., 1975: Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States. *Mon. Wea. Rev.*, **103**, 406–419, doi:10.1175/1520-0493(1975)103<0406:DVIPAT>2.0.CO;2.
- Wijngaard, J. B., A. M. G. K. Tank, and G. P. Konnen, 2003: Homogeneity of 20th century European daily temperature and precipitation series. *Int. J. Climatol.*, **23**, 679–692, doi:10.1002/joc.906.
- Wilks, D. S., 2006: On “field significance” and the false discovery rate. *J. Appl. Meteor. Climatol.*, **45**, 1181–1189, doi:10.1175/JAM2404.1.
- Wood, A. W., and D. P. Lettenmaier, 2006: A test bed for new seasonal hydrologic forecasting approaches in the western United States. *Bull. Amer. Meteor. Soc.*, **87**, 1699–1712, doi:10.1175/BAMS-87-12-1699.

- Wood, E. F., and Coauthors, 1998: The Project for Intercomparison of Land-Surface Parameterization Schemes (PILPS) phase 2 (c) Red–Arkansas River basin experiment: 1. Experiment description and summary intercomparisons. *Global Planet. Change*, **19**, 115–135, doi:[10.1016/S0921-8181\(98\)00044-7](https://doi.org/10.1016/S0921-8181(98)00044-7).
- Xia, Y. L., and Coauthors, 2009: NLDAS primary forcing data L4 hourly 0.125×0.125 degree V002. Goddard Earth Sciences Data and Information Services Center, accessed 3 August 2016, doi:[10.5067/6J5LHHOHZHN4](https://doi.org/10.5067/6J5LHHOHZHN4).
- , and Coauthors, 2012a: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res.*, **117**, D03109, doi:[10.1029/2011JD016048](https://doi.org/10.1029/2011JD016048).
- , and Coauthors, 2012b: Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow. *J. Geophys. Res.*, **117**, D03110, doi:[10.1029/2011JD016051](https://doi.org/10.1029/2011JD016051).
- , J. Sheffield, M. B. Ek, J. Dong, N. Chaney, H. Wei, J. Meng, and E. F. Wood, 2014: Evaluation of multi-model simulated soil moisture in NLDAS-2. *J. Hydrol.*, **512**, 107–125, doi:[10.1016/j.jhydrol.2014.02.027](https://doi.org/10.1016/j.jhydrol.2014.02.027).
- Yamada, T. J., M. I. Lee, M. Kanamitsu, and H. Kanamaru, 2012: Diurnal characteristics of rainfall over the contiguous United States and northern Mexico in the dynamically downscaled reanalysis dataset (US10). *J. Hydrometeor.*, **13**, 1142–1148, doi:[10.1175/JHM-D-11-0121.1](https://doi.org/10.1175/JHM-D-11-0121.1).
- Yue, S., P. Pilon, B. Phinney, and G. Cavadias, 2002: The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrol. Processes*, **16**, 1807–1829, doi:[10.1002/hyp.1095](https://doi.org/10.1002/hyp.1095).