AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2017GL074781

Key Points:
- Probability forecasts of U.S. March-May severe thunderstorm activity show skill over about a quarter of the country
- Forecasts are based on either the observed or predicted ENSO state during the preceding December-February
- Hail event numbers are predicted more skillfully than tornado counts

Supporting Information:
- Supporting Information S1

Correspondence to:
C. Lepore,
clepore@ldeo.columbia.edu

Citation:
Lepore, C., M. K. Tippett, and J. T. Allen (2017), ENSO-based probabilistic forecasts of March–May U.S. tornado and hail activity, *Geophys. Res. Lett.*, *44*, 9093–9101, doi:10.1002/2017GL074781.

Received 1 JUL 2017
Accepted 22 AUG 2017
Accepted article online 25 AUG 2017
Published online 12 SEP 2017

ENSO-based probabilistic forecasts of March–May U.S. tornado and hail activity

Chiara Lepore[1], **Michael K. Tippett**[2,3], **and John T. Allen**[4]

[1]Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York, USA, [2]Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York, USA, [3]Center of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz University, Jeddah, Saudi Arabia, [4]Department of Earth and Atmospheric Sciences, Central Michigan University, Mount Pleasant, Michigan, USA

Abstract Extended logistic regression is used to predict March–May severe convective storm (SCS) activity based on the preceding December–February (DJF) El Niño–Southern Oscillation (ENSO) state. The spatially resolved probabilistic forecasts are verified against U.S. tornado counts, hail events, and two environmental indices for severe convection. The cross-validated skill is positive for roughly a quarter of the U.S. Overall, indices are predicted with more skill than are storm reports, and hail events are predicted with more skill than tornado counts. Skill is higher in the cool phase of ENSO (La Niña like) when overall SCS activity is higher. SCS forecasts based on the predicted DJF ENSO state from coupled dynamical models initialized in October of the previous year extend the lead time with only a modest reduction in skill compared to forecasts based on the observed DJF ENSO state.

1. Introduction

Severe convective storms (SCSs; thunderstorms with tornadoes, large hail, or damaging straight-line winds) occur around the world but are especially frequent in the U.S. Annual average insured losses from thunderstorms in the U.S. during the period 2003–2015 were $11.23 billion (in 2016 USD), comparable to the losses from hurricanes during the same period [*Gunturi and Tippett*, 2017].

In the U.S., official forecasts and outlooks provide information about expected SCS activity up to a week in advance. However, prospects for extended-range prediction are supported by evidence that U.S. SCS activity is related to predictable climate signals such as the Madden-Julian Oscillation (MJO) [*Barrett and Gensini*, 2013; *Thompson and Roundy*, 2013], the El Niño–Southern Oscillation (ENSO) [*Cook and Schaefer*, 2008; *Lee et al.*, 2013; *Allen et al.*, 2015a; *Lee et al.*, 2016; *Cook et al.*, 2017], and Gulf of Mexico sea surface temperature (SST) [*Jung and Kirtman*, 2016; *Molina et al.*, 2016]. Similar relations of U.S. precipitation and near-surface temperature with ENSO, combined with the ability to predict ENSO, were the basis of the first seasonal climate forecasts for the U.S. [*Barnston et al.*, 1994]. More recently, improved model representation and prediction of the MJO and its teleconnections have been the basis for skillful subseasonal prediction of U.S. precipitation and near-surface temperature [*Vitart*, 2014; *DelSole et al.*, 2017].

A particular challenge of extended-range prediction of SCS activity compared to the prediction of temperature or precipitation (or even hurricanes) is that physics-based dynamical models do not directly resolve thunderstorms at that lead time [*Tippett et al.*, 2015]. This limitation means that a completely physics-based extended-range SCS forecast is not presently practical, and statistical methods are needed to bridge this gap. One approach is to use numerical weather prediction (NWP) models to predict environmental factors (indices) that measure the favorability of conditions for tornado and hail [*Tippett et al.*, 2012; *Carbin et al.*, 2016]. An alternative forecast method is to construct a statistical model that relates antecedent conditions with subsequent SCS activity, much in the way that early seasonal climate forecasts were constructed [*Barnett and Preisendorfer*, 1987]. *Elsner and Widen* [2013], for example, developed a statistical model based on trends, year-to-year autocorrelation, and February SST in the western Caribbean and Gulf of Alaska to predict April–June tornado counts for a region in the central Great Plains. *Allen et al.* [2015a], instead, used an ENSO-based regression to form categorical probabilities for the level of March–April tornado and hail environmental indices in a region including northern Texas and Oklahoma.

Here we construct and assess spatially resolved March–May (MAM) probabilistic forecasts for SCS activity over the contiguous U.S. (CONUS). The December–February (DJF) ENSO state is the predictor for the probability of below-normal, normal, and above-normal SCS activity. SCS activity is defined by pooling tornado reports, hail events, and environmental indices. By pooling the SCS data, we take the view that observed variations in the ENSO response across the report and index data reflect unpredictable weather noise rather than robust signal differences. The practical consequence of pooling the SCS activity data is that the resulting SCS forecasts provide no differentiation between the expected levels of tornado, hail, or index activity. The SCS activity forecasts are verified against tornado reports, hail events, and environmental indices separately. Observed values of the DJF ENSO state are available early in March. However, the DJF ENSO state can be predicted several months in advance with high skill, and we demonstrate that forecasts of MAM SCS activity based on predicted ENSO state available in early October of the previous year have skill that is comparable to those based on the observed DJF ENSO state.

The paper is organized as follows. Data and methods are described in section 2. Results are presented in section 3, and a summary and conclusions in section 4.

2. Data and Methods

In this section we describe the data (storm reports, severe weather indices, and observed and forecast DJF ENSO state) and methods used to construct (extended logistic regression) and verify (Brier skill score) cross-validated probabilistic forecasts of MAM SCS activity.

2.1. Data

Tornado and hail reports are taken from the NOAA Storm Prediction Center (SPC) Severe Weather Database and aggregated on a seasonal basis (MAM) to a $1^\circ \times 1^\circ$ grid for the 37 years 1979–2015. The gridded tornado data correspond to the number of MAM tornado reports within each grid box. The gridded hail data are the number of MAM hail *events*, which are defined as 3-hourly periods with one or more reports of hail greater than 25.4 mm in diameter [*Allen et al.*, 2015b]. Tornado and hail reports correspond to observed events but have limitations related to the reporting process [*Verbout et al.*, 2006; *Allen and Tippett*, 2015].

To offset report limitations, we use two SCS activity indices: the *Tornado Environment Index* (TEI) *Tippett et al.* [2012, 2014] and the *Hail Environment Index* (HEI) [*Allen et al.*, 2015b], which show good associations with corresponding storm reports but are spatially complete, unrelated to the storm reporting process, and less noisy. TEI and HEI provide spatially complete descriptions from the environmental viewpoint and reflect storm potential rather than occurrence. TEI and HEI are functions of the monthly averages of convective precipitation, 0–3 km storm relative helicity, and in the case of HEI, the 180 hPa mixed-layer convective available potential energy. Low-level moisture is not included in the HEI used here. Both indices are here calculated using data from the North American Regional Reanalysis (NARR) [*Mesinger et al.*, 2006] for the same time period as the storm reports and averaged on the same $1^\circ \times 1^\circ$ grid.

The 37 antecedent DJF values of the seasonal Oceanic Nino Index (ONI; the NINO 3.4 index computed using the ERSST version 4 data set) are obtained from the NOAA Climate Prediction Center. Forecast DJF averages of the NINO 3.4 index initialized in 1982–2014 (33 forecasts) come from the North American Multimodel Ensemble (NMME) project [*Kirtman et al.*, 2014]. The details of the models used to form multimodel ensemble mean and its deterministic and probabilistic skill for monthly averages are found in *Barnston et al.* [2017] and *Tippett et al.* [2017]. We consider NMME forecasts of DJF NINO 3.4 with lead times ranging from 5 months (starts in July) to 0 months (starts in December); real-time NMME forecast data are available by the eighth of the start month.

2.2. Extended Logistic Regression

The forecast format here consists of probabilities for the occurrence of below-normal, normal, and above-normal SCS activity. These three categories are defined to be equally likely and have climatological frequencies of 33%. The likelihood of the categories conditional on the DJF ONI value is estimated using extended logistic regression (ELR), which has been previously applied to model output in medium-range, subseasonal and seasonal prediction [*Wilks and Hamill*, 2007; *Wilks*, 2009; *Allen et al.*, 2015a; *Vigaud et al.*, 2017]. In ELR, the probability p of not exceeding a threshold q conditional on a single predictor variable (here the DJF ONI value) is

$$\log\left(\frac{p}{1-p}\right) = b_0 + b_1 \mathrm{ONI} + g(q)\,, \tag{1}$$

where b_0 and b_1 are the intercept and regression coefficient, respectively, and we take $g(q) = b_2 q$. Here q is the climatological percentile of SCS activity, which we convert to ranks. The terciles of the standard logistic distribution are − log 2 and log 2, and the probabilities of the occurrence of below-normal, above-normal, and normal activity are computed from, respectively,

$$P_B = \frac{\exp\left(b_0 + b_1 \mathrm{ONI} - b_2 \log 2\right)}{1 + \exp\left(b_0 + b_1 \mathrm{ONI} - b_2 \log 2\right)},$$
$$P_A = 1 - \frac{\exp\left(b_0 + b_1 \mathrm{ONI} + b_2 \log 2\right)}{1 + \exp\left(b_0 + b_1 \mathrm{ONI} + b_2 \log 2\right)}, \quad (2)$$
$$P_N = 1 - \left(P_A + P_B\right).$$

The left-hand side of (1) is the log-odds ratio, which is an increasing function of the probability p. Consequently, a positive value of b_1 indicates that the probability of the below normal category is increased (reduced SCS activity) for positive ONI (El Niño-like) values and decreased (increased SCS activity) for negative ONI (La Niña-like) values.

2.3. Estimation Procedure

Values for the ELR parameters b_0, b_1, and b_2 are estimated at each grid point using historical values of the DJF ONI and SCS activity as measured by the pooled tornado, hail, TEI, and HEI data. By pooling the SCS data, we do not estimate separate regression parameters (or produce separate forecasts) for tornado, hail, and index data but instead estimate parameters for forecasts of "generic" SCS activity. Since each data set has different units and ranges, each is converted to percentile ranks before pooling. For forecasts of tercile probabilities, the ELR fitting procedure requires classifying the predictand data as falling into one of the three equally likely categories of below-normal, above-normal, and normal activity. This classification fails when more than 13 years of the predictand data are zero, since in this case the division between the below-normal and normal categories (the 33rd percentile) is ambiguous. This problem of locations with fewer than 24 nonzero values is an issue only for the report data, since the indices are always nonzero. This filtering of the data means that at a few grid points west of the Rockies (see Figure S1 in the supporting information), 1 or 2 of the report data sets are not used in the fitting. TEI and HEI values are nonzero and used in the fitting at all grid points. Statistical significance of the estimated parameters at the 5% level is determined by likelihood ratio test.

Varying levels of smoothing are applied to the four SCS activity data sets, and the performance of resulting forecasts is used to choose the appropriate level of smoothing. Spatial smoothing is applied prior to conversion to percentile ranks and reduces the frequency of grids and years with no reports. In particular, the training data are spatially smoothed using 3×3, 7×7, and 9×9 kernel smoothers with weights that follow a bivariate Gaussian function with standard deviation = 1° and are normalized to sum to 1. Coastal area weights are normalized accordingly to the number of land grid points included in the smoothing. The terciles levels for the four data sets with 7×7 smoothing are shown in supporting information Figure S1.

In the skill assessment, verification data are also spatially smoothed with the same kernel smoothers, and forecasts are made using leave-one-out cross validation in which regression parameters are estimated withholding all data from the year being forecast. Tercile definitions and percentile rankings are also cross validated.

The observed DJF ONI values in the ELR are replaced by the forecast values of NINO 3.4 for the SCS activity forecasts based on predicted values of the NINO 3.4 index, and the SCS activity forecasts are verified for the 33 years 1983–2015.

2.4. Validation

Forecasts are validated at each grid point using the Brier skill score (BSS) [*Brier*, 1950], which for a particular category is

$$\mathrm{BSS} = 1 - \frac{\sum_{i=1}^{N} \left(P_i - O_i\right)^2}{\sum_{i=1}^{N} \left(R_i - O_i\right)^2} \quad (3)$$

where P_i is the forecast probability in year i; O_i is 1 if the category occurs and 0 otherwise; R_i is a reference forecast, here the climatological frequency; and N is the number of forecasts. Here we show the BSS for forecasts of the occurrence of the below and above categories separately and note that their average corresponds to the ranked probability skill score [*Bradley and Schwartz*, 2011]. Forecasts with BSS > 0 are more accurate

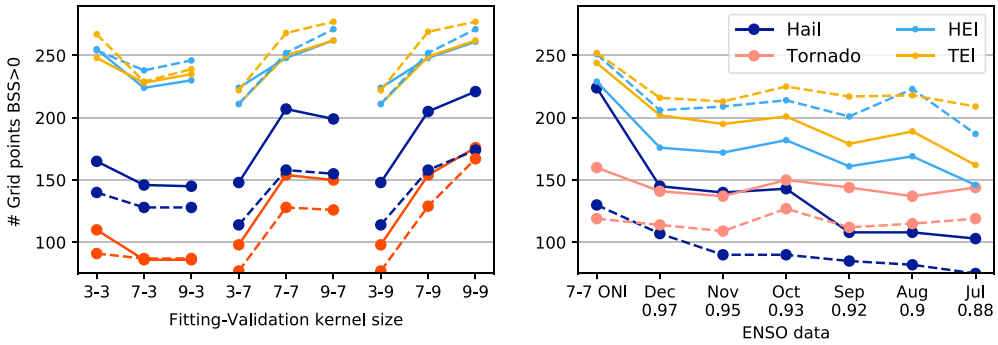

Figure 1. Number of grid points whose BSS is greater than 0.0, (left) for different fitting-validation kernel size combinations and (right) for different NMME lead times. Solid lines are for below category (P_B) forecast; dashed lines for above category (P_A) forecast. *X* axis labels in Figure 1 (right) indicates correlation of the forecast with ONI. The total number of grid points of the domain is 867.

than the climatological reference forecast, and we denote them here as skillful. The numerical values of the BSS tend to be small compared to more familiar measures like correlation. For instance, probability forecasts for joint Gaussian variables with correlation r have BSS $\approx 1 - \sqrt{1 - r^2}$, so that BSS = 0.1 (a typical value for seasonal predictions of precipitation outside of the tropics) corresponds to a correlation of approximately 0.44 [*Tippett et al.*, 2010, 2017].

3. Results

Our physical understanding of the mechanisms (shifts of the jet stream, changes in moisture availability, etc.) by which ENSO modulates SCS activity means that a ENSO-based SCS seasonal activity forecast should be relatively large scale and locally homogeneous. On the other hand, estimates of the ENSO signal from the report data include considerable small-scale variability due to the small spatial scale and relative rarity of severe thunderstorms. Moreover, there is considerable SCS variability that is not explained by ENSO, especially at the local level. Therefore, we smooth the training data so that the resulting forecast is relatively large scale and locally smooth. Likewise, we smooth the verification data and acknowledge that the ENSO signal only provides information about SCS activity on fairly large spatial scales [*Gong et al.*, 2003; *Allen et al.*, 2015a]. To select reasonable levels of smoothing, we use the number of grid points with positive BSS as a summary measure of overall skill. Figure 1 (left) shows the number of grid points with positive BSS for the below-normal and above-normal categories, different combinations of smoothing of the training and verification data, and verification against different measures of SCS activity (reports and indices). When the 3×3 smoother is applied to the verification data, the choice of smoothing for the training data has little impact, with some indication that more smoothing of the training data reduces skill (Figure 1, left). There is little difference in the number of grid points with positive BSS for the 7×7 and 9×9 smoothing of the validation data, and thus, to avoid excessive smoothing, we choose the 7×7 smoother for the validation data. For the 7×7 smoothed validation data, the number of grid points with positive BSS is roughly optimal with 7×7 smoothing of the training data. We use this smoothing combination (7×7 smoothing of the training and validation data, referred to as 7-7) in the results that follow.

We also explore how the number of grid points with positive BSS varies as the SCS forecast is verified against different data sets (Figure 1, left). Verification against indices produces substantially more grid points with positive BSS than does verification against report data. This difference is partially a reflection of the greater predictability of the indices but is also due in part to the fact that forecasts for the indices can be verified at all grid points, while verification against report data is only possible at grid points with fewer than 13 zero values. In particular, forecasts verified against indices count as skillful even in regions were few if any severe thunderstorms occur (as indicated by both low values of the indices and lack of reports; see Figure S1) as long as there is an ENSO signal in the environmental factors. Zero values in the report data may reflect actual nonoccurence or simply a lack of reports. Verification against hail events gives more points with positive BSS than does verification against tornado reports, likely due to the greater variability and smaller spatial footprint of tornadoes. In this case, the tornado reports and hail events have sufficient nonzero values at roughly the same number of grid points, and, in fact, the tornado grid tends to have more points.

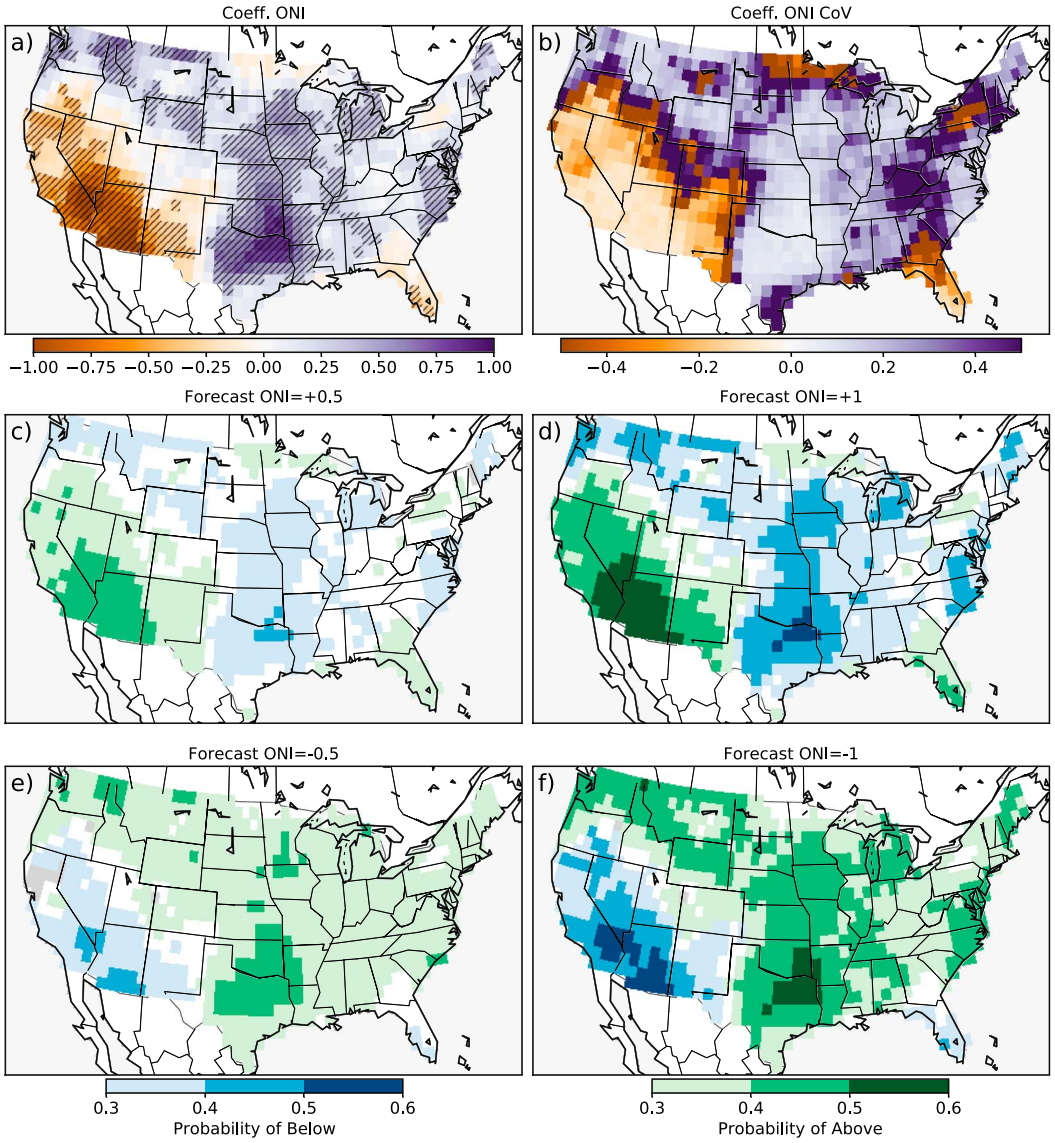

Figure 2. (a) Regression coefficient for ONI (b_1) and (b) its coefficient of variation in the leave-one-out cross validation; hatched areas in Figure 2a are for grid points where b_1 is significant (p value < 0.05). Model forecast for (c) weak (0.5) and (d) moderate (1.0) El Niño, and for (e) weak (−0.5) and (f) moderate (−1.0) La Niña. Only the dominant (≥ 35%) probability is shown, and white indicates areas where no category has 35% or more probability.

To assess the potential for increasing the lead of spring SCS activity forecasts, Figure 1 (right) compares the number of grid points with positive BSS for forecasts using observed DJF ONI with those in which NMME forecast values of the NINO 3.4 index are used. The NMME forecast initializations range from July (longest lead) to December (shortest lead) of the proceeding year, and the correlations with the observed ONI values range from 0.97 to 0.88 (Figure 1, right). The number of grid points with positive BSS drops when forecast NINO 3.4 values are used, but the drop does not increase substantially as lead time increases until September starts, suggesting that SCS forecasts based on NINO 3.4 predicted initialized in October are still skillful.

The regression coefficient (b_1) that multiplies ONI (Figure 2a; computed using all years of data) determines, for the most part, the spatial structure of the probability forecasts shown for four ENSO states (Figures 2c–2f). The other two regression coefficients (not shown) are relatively uniform spatially. Positive values of the regression coefficient indicate enhanced SCS activity during La Niña conditions and are found for most of the U.S., including the south and central regions where severe thunderstorm activity is most frequent during this season. Areas where the values of b_1 are statistically significantly different from zero at the 5% level are

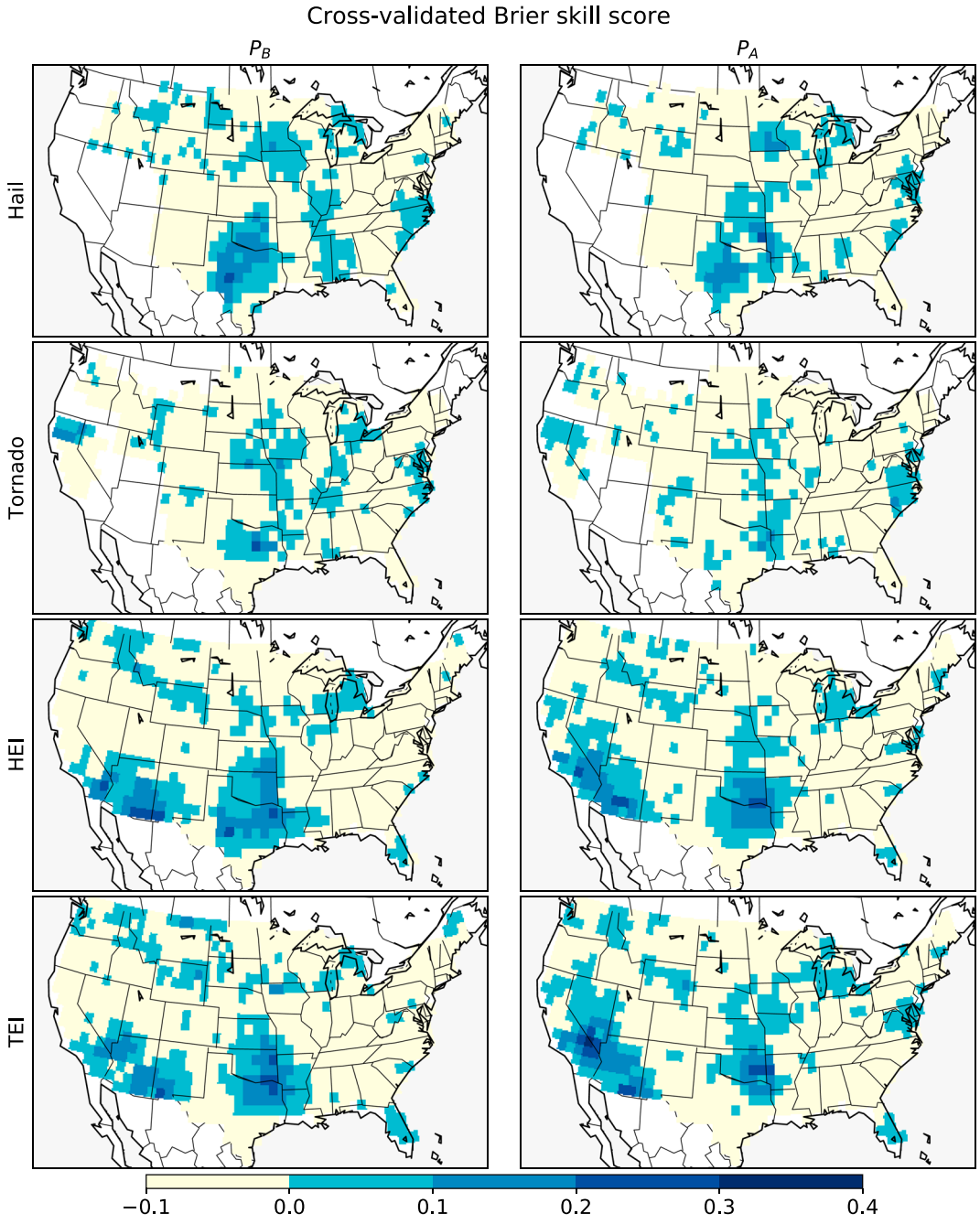

Figure 3. Brier skill score (BSS) maps for the SCS activity forecasts verified against hail events, tornado counts, HEI, and TEI. Areas with insufficient (see text) nonzero data values are indicated in white.

indicated with hatching and match well the areas where the coefficient of variation (CoV; mean/standard deviation) of b_1 obtained from the leave-one-out cross validation are small. Contrastingly, Florida and the western U.S. including the southwest, California, and parts of Oregon show the opposite relation between SCS activity and ENSO. West of the Rockies, the indices are the primary source of information in the fitting of the regression parameters, SCS activity is low, and shifts of the indices to the above category still mean that the expected SCS activity in absolute terms is low (see definition of above normal in Figure S1).

Figures 2c–2f show the probabilistic forecasts of MAM tornado and hail activity with only the dominant category probability indicated, for representative ONI values: weak (0.5) and moderate (1.0) El Niño (EN), and weak (−0.5) and moderate (−1.0) La Niña (LN). As mentioned before, the spatial structure of the probability

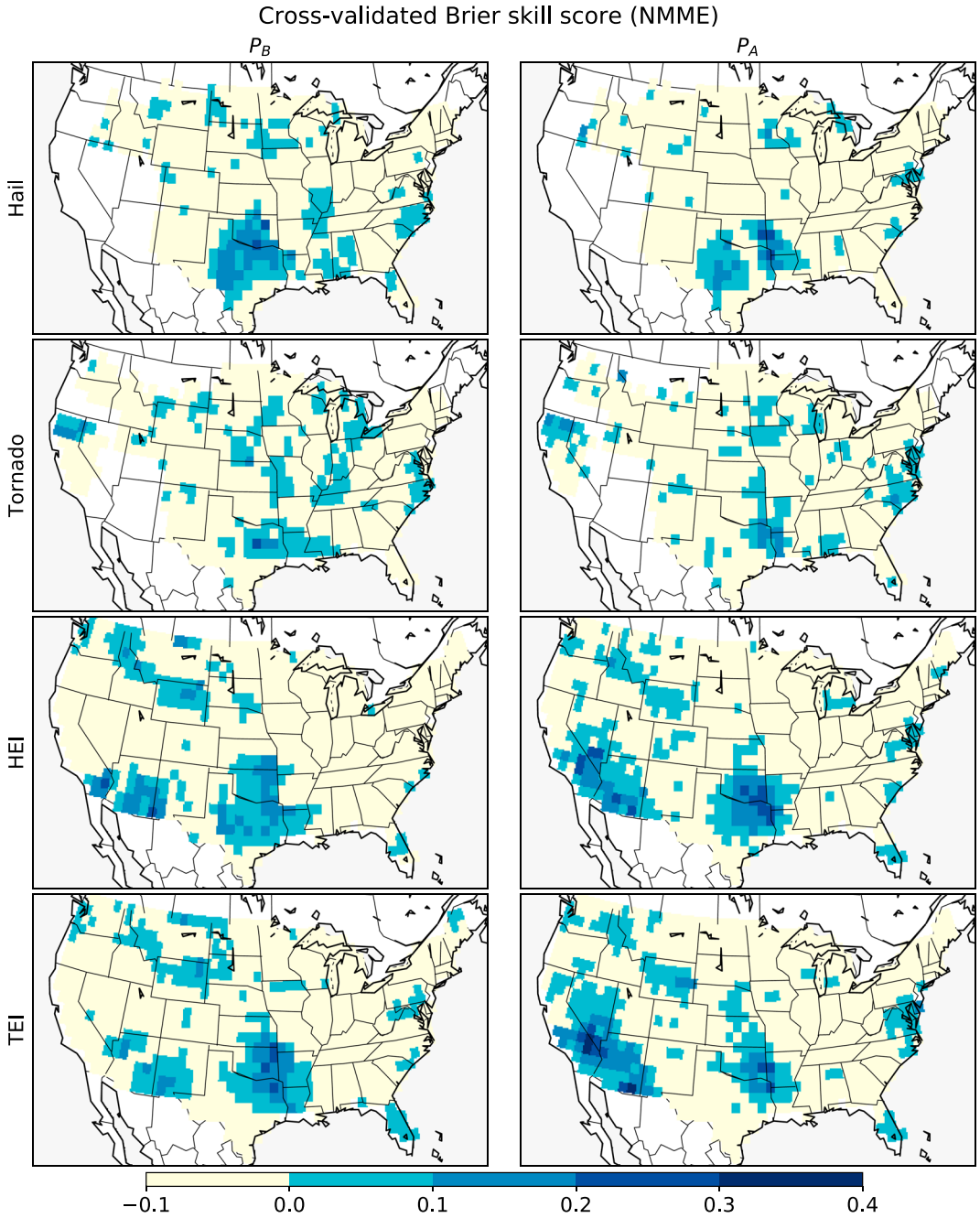

Figure 4. As in Figure 3 except using NINO 3.4 values from NMME forecasts initialized in early October 1982–2014.

forecast maps matches that of the regression coefficient b_1. Weak EN gives a shift of a few percent away from climatological values for the below category throughout the Great Plains and Midwest and somewhat stronger shifts in the above category probabilities over Florida and the southwest of the continent. The probability shifts are amplified for moderate EN and exceed 10% in some regions, most notably over the southern to central Great Plains for the below-normal category, while Florida and the desert Southwest exhibit particularly strong responses in above-normal category. In contrast, for moderate LN, the probability shifts are in the opposite direction and overall reflect somewhat stronger probabilistic likelihoods. The probability shifts are not symmetric with respect to ONI because of the intercept term (b_0) which is negative. This asymmetry may be a reason for the larger number of grid points with positive BSS for the below category seen in the report verifications and to a lesser degree in the index verifications (Figure 1, left).

Maps of BSS values in Figure 3 show the skill of the SCS activity forecasts verified against the four data sets. Grid points with more than 13 zero values are not included in the verification; these are limited to areas west of the Rockies for hail events, and to AZ, MT, and part of NV and ID for the tornado counts. The area with the strongest dependence on ENSO (TX, OK, LA, and AR) has the highest overall skill for both the below and above categories. The skill for hail over this region is westwardly displaced relative to that for tornadoes, reflecting the difference in spatial climatology between the phenomena. Florida, which has an opposite dependence on ENSO, as the rest of the eastern U.S. is skillfully predicted only for the indices, reflecting large variability in the tornado frequency. The skill for TEI is also higher over a wider area relative to HEI, reflecting the lower instability available over Florida. CA, OR, and the southwest have positive BSS values for the above category only, in line with the results shown in Figure 2. This signal is consistent with the substantial reductions in available moisture that occur in this region during LN years, and contrasting higher values during EN. This change in available moisture has a consequent influence in available instability, thus leading to substantial skill in both patterns, though the relatively sparse population precludes sufficient report data to validate the environments of the indices. Another region with appreciable skill is the Midwest, in particular over Michigan.

Both categories have greater skill for ONI < -0.1 than for ONI > 0.1 (Figure S2). This behavior is consistent with the fact that SCS activity shows a reduced shift from normal level during warm (El Niño-like) ENSO conditions than during cool (La Niña-like) ENSO conditions [*Allen et al.*, 2015a]. This differing level of skill is also consistent with the asymmetry of the probability forecasts with respect to the ONI. The BSS measures skill relative to a climatological forecast and rewards more forecasts with larger excursions from the climatological probabilities.

The spatial structure of the skill shows little change when forecast values of the DJF NINO 3.4 index are used in place of observed DJF ONI. Figure 4 shows the skill for the four data sets when using NINO 3.4 values from NMME forecasts initialized in early October. The high skill areas seen in Figure 2 maintain similar values of BSS when predicted values of NINO 3.4 are used but are somewhat more limited in extent. These results shows that MAM SCS activity can be predicted with some skill 5–6 months ahead.

4. Summary and Conclusions

The modulation of U.S. climate by the El Niño–Southern Oscillation (ENSO) provides a basis for seasonal prediction of precipitation and near-surface temperature anomalies. Recently, the evidence that ENSO also has an impact on U.S. severe convective storm (SCS; tornadoes and hail) activity has become clearer [*Cook and Schaefer*, 2008; *Allen et al.*, 2015a]. The ENSO signal in SCS activity is difficult to detect due to the large variability (noise) in storm occurrence. Environmental indices that measure how conducive conditions are for storm occurrence are less susceptible to the variability posed by incomplete storm observations and provide a more coherent picture of the ENSO signal in SCS activity. Since ENSO is fairly persistent from winter to spring, the winter ENSO state, here the December–February (DJF) NINO 3.4 value, is a plausible predictor of springtime, March–May (MAM) SCS activity. Utilizing this predictability, we have constructed cross-validated and spatially resolved probability forecasts of SCS activity based on the DJF ONI value and verified them against MAM tornado reports, hail events, a Hail Environment Index, and a Tornado Environment Index.

Overall, the forecast skill is most robust in the southern U.S. including parts of Texas, Oklahoma, Kansas, Louisiana, and Arkansas. The environmental indices can be predicted more skillfully and in more regions than the storm report data. The indices show ENSO signals west of the Rockies where their climatological values are sufficiently low (see Figure S1) that above-normal values still indicate little favorability of the environment for SCS activity. Hail events are forecast skillfully in more locations than are tornado counts, likely as a result of their higher frequency and larger spatial footprint. The forecasts show more skill on average for cool ENSO conditions (La Niña like) when SCS activity is enhanced overall than for warm ENSO (El Niño like). The forecast probability shifts away from normal activity are smaller during warm conditions than during cool conditions, consistent with *Allen et al.* [2015a].

From the perspective of an application to seasonal forecasting of spring SCS activity, the DJF ONI value is predictable several months in advance. In light of this, we have explored longer-lead SCS forecasts of MAM based on NMME predictions of the NINO 3.4 index. SCS predictions based on NMME forecasts initialized in October show only modest decreases in skill compared to those based on observed DJF values. These findings suggest that probabilistic guidance for hail and tornado occurrence based on ENSO state is feasible at leads up to 5 months. However, as is the case with other seasonal climate forecasts, there is substantial variability in

SCS activity that is unexplained by ENSO and may or may not be predictable. In any case, a seasonal forecast of SCS activity is of limited value to the general public, though it provides a good opportunity to encourage preparedness. On the other hand, organizations with broad exposure to SCS risk and the ability to incorporate uncertain information into their decision making might be able to profit from a seasonal SCS forecast.

Acknowledgments
This work was partially supported by a Columbia University Research Initiatives for Science and Engineering (RISE) award, the Office of Naval Research awards N00014-12-1-0911 and N00014-16-1-2073, NOAA's Climate Program Office's Modeling, Analysis, Predictions and Projections program award NA14OAR4310185, and the Willis Research Network. Severe Weather Database tornado and hail reports are available from the Storm Events Database at http://www.spc.noaa.gov/wcm/#data. NARR data are downloaded from the Research Data Archive (RDA) at the National Center for Atmospheric Research (NCAR), Computational and Information Systems Laboratory at http://rda.ucar.edu/datasets/ds608.0/. The Oceanic Nino Index (ONI) from the NOAA Climate Prediction Center is available at http://www.cpc.ncep.noaa.gov/data/indices/oni.ascii.txt. The NMME project and data dissemination is supported by NOAA, NSF, NASA, and DOE. We acknowledge the help of CPC, IRI, and NCAR personnel in creating, updating, and maintaining the NMME archive. NMME forecasts are available for download from the IRI Data Library at http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/.

References

Allen, J. T., and M. K. Tippett (2015), The characteristics of United States hail reports: 1955–2014, *Electron. J. Severe Storms Meteor.*, *10*, 1–31.
Allen, J. T., M. K. Tippett, and A. H. Sobel (2015a), Influence of the El Niño/Southern Oscillation on tornado and hail frequency in the United States, *Nat. Geosci.*, *8*, 278–283, doi:10.1038/ngeo2385.
Allen, J. T., M. K. Tippett, and A. H. Sobel (2015b), An empirical model relating U.S. monthly hail occurrence to large-scale meteorological environment, *J. Adv. Model. Earth Syst.*, *7*, 226–243, doi:10.1002/2014MS000397.
Barnett, T. P., and R. Preisendorfer (1987), Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis, *Mon. Weather Rev.*, *115*, 1825–1850.
Barnston, A. G., et al. (1994), Long-lead seasonal forecasts—Where do we stand?, *Bull. Am. Meteorol. Soc.*, *75*, 2097–2114, doi:10.1175/1520-0477(1994)075<2097:LLSFDW>2.0.CO;2.
Barnston, A. G., M. K. Tippett, M. Ranganathan, and M. L. L'Heureux (2017), Deterministic skill of ENSO predictions from the North American Multimodel Ensemble, *Clim. Dyn.*,1–20, doi:10.1007/s00382-017-3603-3.
Barrett, B. S., and V. A. Gensini (2013), Variability of central United States April–May tornado day likelihood by phase of the Madden-Julian Oscillation, *Geophys. Res. Lett.*, *40*, 2790–2795, doi:10.1002/grl.50522.
Bradley, A. A., and S. S. Schwartz (2011), Summary verification measures and their interpretation for ensemble forecasts, *Mon. Weather Rev.*, *139*, 3075–3089, doi:10.1175/2010MWR3305.1.
Brier, G. W. (1950), Verification of forecasts expressed in terms of probability, *Mon. Weather Rev.*, *78*, 1–3, doi:10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2.
Carbin, G. W., M. K. Tippett, S. P. Lillo, and H. E. Brooks (2016), Visualizing long-range severe thunderstorm environment guidance from CFSv2, *Bull. Am. Meteorol. Soc.*, *97*, 1021–1031, doi:10.1175/BAMS-D-14-00136.1.
Cook, A. R., and J. T. Schaefer (2008), The relation of El Niño–Southern Oscillation (ENSO) to winter tornado outbreaks, *Mon. Weather Rev.*, *136*, 3121–3137, doi:10.1175/2007MWR2171.1.
Cook, A. R., L. M. Leslie, D. B. Parsons, and J. T. Schaefer (2017), The impact of the El Niño Southern Oscillation (ENSO) on winter and early spring U.S. tornado outbreaks, *J. Appl. Meteorol. Climatol.*, *56*, 2455–2478, doi:10.1175/JAMC-D-16-0249.1.
DelSole, T., L. Trenary, M. K. Tippett, and K. Pegion (2017), Predictability of week 3–4 average temperature and precipitation over the contiguous United States, *J. Clim.*, *30*, 3499–3512, doi:10.1175/JCLI-D-16-0567.1.
Elsner, J. B., and H. M. Widen (2013), Predicting spring tornado activity in the Central Great Plains by 1 March, *Mon. Weather Rev.*, *142*, 259–267, doi:10.1175/MWR-D-13-00014.1.
Gong, X., A. G. Barnston, and M. N. Ward (2003), The effect of spatial aggregation on the skill of seasonal precipitation forecasts, *J. Clim.*, *16*(18), 3059–3071.
Gunturi, P., and M. K. Tippett (2017), Managing severe thunderstorm risk: Impact of ENSO on U.S. tornado and hail frequencies, *Tech. Rep.*, WillisRe, Minneapolis, Minn.
Jung, E., and B. P. Kirtman (2016), Can we predict seasonal changes in high impact weather in the United States?, *Environ. Res. Lett.*, *11*, 074018.
Kirtman, B., et al. (2014), The North American Multi-Model Ensemble (NMME): Phase-1 seasonal to interannual prediction; Phase-2 toward developing intra-seasonal prediction, *Bull. Am. Meteorol. Soc.*, *95*, 585–601, doi:10.1175/BAMS-D-12-00050.1.
Lee, S.-K., R. Atlas, D. Enfield, C. Wang, and H. Liu (2013), Is there an optimal ENSO pattern that enhances large-scale atmospheric processes conducive to tornado outbreaks in the United States?, *J. Clim.*, *26*(5), 1626–1642.
Lee, S.-K., A. T. Wittenberg, D. B. Enfield, S. J. Weaver, C. Wang, and R. Atlas (2016), US regional tornado outbreaks and their links to spring ENSO phases and North Atlantic SST variability, *Environ. Res. Lett.*, *11*, 044008.
Mesinger, F., et al. (2006), North American Regional Reanalysis, *Bull. Am. Meteorol. Soc.*, *87*, 343–360.
Molina, M. J., R. P. Timmer, and J. T. Allen (2016), Importance of the Gulf of Mexico as a climate driver for U.S. severe thunderstorm activity, *Geophys. Res. Lett.*, *43*, 12,295–12,304, doi:10.1002/2016GL071603.
Thompson, D. B., and P. E. Roundy (2013), The relationship between the Madden-Julian oscillation and U.S. violent tornado outbreaks in the spring, *Mon. Weather Rev.*, *141*, 2087–2095, doi:10.1175/MWR-D-12-00173.1.
Tippett, M. K., A. G. Barnston, and T. Delsole (2010), Comment on "Finite samples and uncertainty estimates for skill measures for seasonal prediction", *Mon. Weather Rev.*, *138*, 1487–1493.
Tippett, M. K., A. H. Sobel, and S. J. Camargo (2012), Association of U.S. tornado occurrence with monthly environmental parameters, *Geophys. Res. Lett.*, *39*, L02801, doi:10.1029/2011GL050368.
Tippett, M. K., A. H. Sobel, S. J. Camargo, and J. T. Allen (2014), An empirical relation between U.S. tornado activity and monthly environmental parameters, *J. Clim.*, *27*, 2983–2999, doi:10.1175/JCLI-D-13-00345.1.
Tippett, M. K., J. T. Allen, V. A. Gensini, and H. E. Brooks (2015), Climate and hazardous convective weather, *Curr. Clim. Change Rep.*, *1*, 60–73, doi:10.1007/s40641-015-0006-6.
Tippett, M. K., M. Ranganathan, M. L'Heureux, A. G. Barnston, and T. DelSole (2017), Assessing probabilistic predictions of ENSO phase and intensity from the North American Multimodel Ensemble, *Clim. Dyn.*, 1–22, doi:10.1007/s00382-017-3721-y.
Verbout, S. M., H. E. Brooks, L. M. Leslie, and D. M. Schultz (2006), Evolution of the U.S. tornado database: 1954–2003, *Weather Forecasting*, *21*, 86–93.
Vigaud, N., A. W. Robertson, and M. K. Tippett (2017), Multi-model ensembling of subseasonal precipitation forecasts over North America, *Mon. Weather Rev.*, doi:10.1175/MWR-D-17-0092.1.
Vitart, F. (2014), Evolution of ECMWF sub-seasonal forecast skill scores, *Q. J. R. Meteorol. Soc.*, *140*(683), 1889–1899, doi:10.1002/qj.2256.
Wilks, D. S. (2009), Extending logistic regression to provide full-probability-distribution MOS forecasts, *Meteorol. Appl.*, *16*, 361–368.
Wilks, D. S., and T. M. Hamill (2007), Comparison of ensemble-MOS methods using GFS reforecasts, *Mon. Weather Rev.*, *135*, 2379–2390, doi:10.1175/MWR3402.1.