A New Mode of Subseasonal Predictability Over the US: Boreal Summer

Paul Buchmann,[a] , and Timothy DelSole[a,b]

[a] *George Mason University, Fairfax, VA 22030, USA*

[b] *Center for Ocean-Land-Atmospheric Studies, George Mason University, Fairfax, VA 22030, USA*

Corresponding author: Timothy DelSole, tdelsole@gmu.edu

ABSTRACT: This study identifies the most predictable modes of subseasonal temperature over the United States during boreal summer for weeks 1-2 and, separately, for weeks 3-4. Surprisingly, Granger Causality tests reveal that these modes are unrelated to standard indices of subseasonal predictability, such as El Niño or the Madden-Julian Oscillation. Lagged regression analysis indicates that the leading week 1-2 mode is driven by western Pacific precipitation and exhibits enhanced persistence due to interactions with soil moisture. Similarly, the leading week 3-4 mode is linked to western Pacific precipitation. While these modes share features with the Boreal Summer Intraseasonal Oscillation (BSISO), the resemblance is not exact, and the chain of mechanisms leading to predictability over the U.S., particularly involving soil moisture, appears to be new. NOAA's Climate Forecast System v2 (CFSv2) successfully captures the leading week 1-2 mode but fails to represent the leading week 3-4 mode. The lagged relationships identified here may provide insights into model adjustments that could enhance subseasonal predictability. These modes were identified using Canonical Correlation Analysis (CCA), which is capable of uncovering predictability without prior assumptions about its source. While CCA is a well-established statistical method, its application to climate data has been limited due to challenges in significance testing and feature selection. This study addresses these limitations by employing a recently developed Mutual Information Criterion (MIC) to optimize feature selection, using Monte Carlo techniques to establish rigorous significance tests for small samples, and formulating a comprehensive procedure for validating predictability in independent datasets.

SIGNIFICANCE STATEMENT: Accurate subseasonal forecasts, covering the 2- to 8-week time frame, would provide significant societal benefits in areas such as public health, agriculture, water resource management, energy, utilities, and early warnings for extreme events. This paper integrates rigorous statistical procedures into a framework with the potential to uncover new sources of predictability. The underlying idea is that if weather and climate are predictable on subseasonal time scales, there ought to be some correlation between events separated in time. Such correlations can be identified with no pre-conceived notion as to their source. After identifying these correlations, known sources of predictability may be removed statistically one-by-one. Any predictability that remains after this sifting process indicates a new source of predictability.

1. Introduction

Predictions of temperature and precipitation on subseasonal time scales have been made for at least a decade. These predictions typically forecast one- or two-week means up to six weeks in advance (Johnson et al. 2014). Subseasonal predictions are generally less skillful than either weather or seasonal forecasts because the lead time is long enough for information from atmospheric initial conditions to degrade, yet the averaging period is short enough that not all weather noise is smoothed out. Despite these challenges, accurate subseasonal forecasts could provide significant benefits to water management, agriculture, disaster preparedness, and health (White et al. 2017). Case studies have shown that subseasonal forecasts can support decision-making in areas such as public health, agriculture, water resource management, energy, and utilities, as well as early warnings for extreme events (White et al. 2021; Domeisen et al. 2022). The importance of subseasonal forecasting led to subseasonal forecast competitions, with substantial cash prizes, by the U.S. Bureau of Reclamation in 2017 (Hwang et al. 2019) and 2019 (Nowak et al. 2020), and by the United Nations' World Meteorological Organization in 2021 (Vitart et al. 2022).

Several known phenomena contribute to subseasonal predictability in the United States, including the El Niño Southern Oscillation (ENSO), the Madden-Julian Oscillation (MJO), the Pacific-North American teleconnection pattern (PNA), the North Atlantic Oscillation (NAO), Sudden Stratospheric Warming events (SSWs), and land-atmosphere coupling (Robertson and Vitart 2019; National Research Council 2010; National Academies of Sciences Engineering and Medicine 2016). A more recently identified source, which plays a significant role in this study, is the Boreal Summer Intraseasonal Oscillation (BSISO). The BSISO is a summer mode of the MJO characterized by northward-propagating precipitation anomalies extending from India to the western Pacific. Most studies on the BSISO focus on its influence on the Asian summer monsoons or on tropical cyclones. Few studies examine its effects on the United States' 2m temperature. One such study by Krishnamurthy et al. (2021) describes an oscillation in tropical winds over the eastern Pacific related to the BSISO, with the response of 2m temperature over the U.S. quantified through regression maps. Jenney et al. (2019) also assessed the seasonal impact of the MJO and BSISO on surface temperatures in the U.S and concluded that its impact on summer predictability is small compared to the MJO's influence on winter predictability.

All of the above phenomena influence U.S. temperatures on subseasonal timescales. However, in each case, the phenomenon was identified first, and its impact on temperature was determined afterward. This raises the question of whether there might be other mechanisms driving subseasonal predictability that have gone unnoticed simply because they have not yet been identified. If an unknown source of predictability exists, how would we discover it? Our goal is to identify subseasonal predictability without requiring prior knowledge or a hypothesis about the source.

A mechanism that produces predictability on subseasonal time scales should produce a temporal dependence between a pattern at time t and a (potentially different) pattern at time $t+\tau$. Accordingly, we employ methods that identify temporal correlations in multivariate time series. Several approaches exist for this, including multichannel singular spectrum analysis (MSSA, Ghil et al. 2002; Krishnamurthy et al. 2021), coherence spectrum analysis (Madden and Julian 1971), lead-lag regression between leading EOFs, Canonical Correlation Analysis (CCA, Barnett and Preisendorfer 1987; Barnston and Smith 1996; Huth 2002; DelSole and Tippett 2022), and machine learning approaches (McGovern et al. 2014; Hwang et al. 2019; He et al. 2021; Trenary and DelSole 2023). However, each method has limitations in its current form. MSSA does not explicitly maximize a measure of predictability. Coherence spectrum analysis is univariate and thus does not capture multivariate dependencies. Lead-lag regression assumes that individual EOF patterns represent the full response to a mechanism, a highly restrictive assumption. Machine learning approaches, while powerful, require large datasets for training and validation, which poses a challenge for subseasonal prediction due to the relatively small sample sizes involved.

Among the available statistical methods, we adopt Canonical Correlation Analysis (CCA) for this study. While CCA has its own limitations–such as the need to select the number of EOFs for analysis and the reliance on asymptotic significance tests–these issues are addressable. First, a relatively new selection criterion, the Mutual Information Criterion (MIC), has been developed specifically for CCA (DelSole and Tippett 2021). Second, Monte Carlo techniques can be employed to derive small-sample significance tests (DelSole and Tippett 2022). The goal of this work is to leverage these tools to develop CCA into a rigorous and objective procedure for identifying predictability in multivariate time series. Once predictability is detected, lagged regression maps can be used to describe its temporal evolution and its relationship to other physically relevant variables. To ensure that the predictability identified by CCA is not the result of overfitting, we verify the findings using independent datasets. This is done by projecting the predictable components onto an independent dataset and comparing the correlation in that data to the correlation determined by CCA. A key contribution of this work is the development of a comprehensive and rigorous validation procedure in independent datasets. Interestingly, this process may yield uncertainty ranges that do not contain either the in-sample or out-of-sample correlations, which may surprise some readers.

Since CCA does not require a prior hypothesis about the underlying mechanism, it has the potential to uncover previously unknown forms of predictability. Given our goal of identifying new sources of predictability, we focus primarily on analyzing observational data. There is no doubt that ENSO contributes to subseasonal predictability; any reasonable method will detect ENSO as a dominant influence. To explore additional sources of predictability beyond ENSO, this study removes the seasonal ENSO influence by subtracting the seasonal mean from the temperature data before performing the predictability analysis. Because the mechanisms driving temperature predictability vary by season, we will analyze each season separately. This paper presents the results for boreal summer, while results for other seasons are discussed in Buchmann (2024).

For ease of communication, a pair of temperature patterns along with their time series will be referred to as a 'mode.' The leading mode identified by CCA will have the highest possible correlation. The second mode will have the highest possible correlation that is uncorrelated with the leading mode, and so on. After confirming the detection of predictability, we investigate whether it is linked to a known source by regressing that phenomenon out of the data and recalculating CCA. We then compare the resulting modes and correlations to the original ones. If the correlation of a mode becomes insignificant or if the mode disappears entirely, we can conclude that the phenomenon we removed is responsible for that mode. Finally, we assess whether subseasonal dynamical models capture the modes identified by CCA. This is done by projecting the predictable components onto subseasonal temperature forecasts and comparing the resulting correlations with those derived from CCA. If a mode is not well-represented by a model, it may offer a target for model developers to improve the representation of specific features in their models.

This paper is organized as follows. The next section reviews our data and methods, particularly CCA and associated selection criteria, significance tests, and connections to Granger Causality Analysis. Section 3 describes a new, comprehensive, and rigorous procedure for validating canonical components in independent data. Section 4 describes our results of applying CCA to

June-July-August (JJA) 2m temperature over CONUS. Section 5 discusses results of assessing whether a state-of-the-art climate forecast model captures the predictability identified here from observational data sets. This paper ends with a summary and discussion of our results. This work is a partial summary of a PhD thesis by Paul Buchmann. More comprehensive discussion of these results and other results for other seasons can be found in the thesis Buchmann (2024).

2. Data and Methods

a. ERA40 Reanalysis

The main dataset used is the daily ERA40 reanalysis. This reanalysis covers September 1957 to August 2002, making it one of the longest reanalysis data sets. The variables are on a 1.25 by 1.25 uniform longitude-latitude grid. The following daily variables are used from this reanalysis: 2m temperature, total precipitation, the Nino 3.4 index of SST, and soil moisture in the top layer (0-7cm underground). BSISO indices used are based on Kikuchi (2020); the EOFs of outgoing longwave radiation (OLR) were obtained from the International Pacific Research Center, and intraseasonal OLR from ERA40 was projected onto the EOFs to obtain the time series of the BSISO indices. The 2m temperature data is used to investigate intraseasonal predictability. The other data is used to explore the source of the predictability in 2m temperature.

b. Observed Data

To verify the correlations found from the ERA40 reanalysis, we use observed daily 2m temperature over CONUS from the NOAA ESRL. We utilize only the time period September 1, 2002 to February 28, 2022, which does not intersect with the ERA40 dataset. The data was interpolated to the ERA40 grid.

To investigate the sources of predictability in the ERA40 2m temperature data, we use the NAO and PNA indices provided by NOAA's Climate Prediction Center (CPC). These indices overlap with the ERA40 data range and are used for convenience rather than being recalculated directly from the ERA40 dataset.

c. *Data Preprocessing*

Data in ERA40 is output in 6 hour increments. Except for precipitation, the data is converted to daily averages. For precipitation, the data is converted to daily total. Then anomalies are calculated at the gridpoint level by removing a trend and 3 annual harmonics. To convert to EOFs, the data is separated into seasons (DJF, MAM, JJA, SON) and averaged into two-week means.

An important step is that the mean from each season is removed before analyzing predictability. For example, the anomalies for June, July, August 1999 have zero mean when averaged over June, July, August 1999. EOFs were computed from the two-week mean data and then the mean of each season was removed from each PC. Reversing these steps by removing the local seasonal mean from the gridpoint data and then calculating the EOFs gave virtually identical PC time series and EOF spatial patterns.

All observed indices are preprocessed by removing a trend and 3 annual harmonics, and then calculating two week averages. The local seasonal mean is then removed.

d. *Canonical Correlation Analysis (CCA)*

A procedure called Canonical Correlation Analysis (CCA) is used to quantify the relation between variables. Given a vector $\mathbf{x}(t)$ and a vector $\mathbf{y}(t)$, CCA finds a linear combination of $\mathbf{x}(t)$ and a linear combination of $\mathbf{y}(t)$ that maximizes their correlation. More generally, CCA decomposes the data into pairs of variates (time series) such that the first pair has the maximum possible correlation in the data set, the second pair has the maximum correlation uncorrelated to the first pair, and so on, with each pair of variates uncorrelated to all of the variates preceding them. The n'th variate pair has correlation ρ_n called the n'th canonical correlation. Each pair of variates also has a pair of loading vectors (spatial patterns) associated with it. More details of this standard CCA procedure can be found in DelSole and Tippett (2022).

In this work, CCA is applied to two temperature fields, $\mathbf{x}(t)$ and $\mathbf{y}(t)$, where t is a time index. In this work, $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are 2-week means separated by a fixed lag. The precise endpoints for the 2-week means are listed in Table 1.

	start of $\mathbf{x}(t)$	end of $\mathbf{x}(t)$	start of $\mathbf{y}(t)$	end of $\mathbf{y}(t)$
weeks 1-2	day -13	day 0	day 1	day 14
weeks 3-4	day -13	day 0	day 15	day 28

TABLE 1. Start day and end day of the 2-week averaging windows for weeks 1-2 and weeks 3-4 prediction.

e. Selection Criterion - MIC

In climate applications, it is standard practice to reduce the dimension of $\mathbf{x}(t)$ and $\mathbf{y}(t)$ by projecting them onto their leading EOFs. The question naturally arises as to how many EOFs should be chosen. Studies using CCA generally have not used a selection criterion for the number of EOFs used. In this work, we use a selection criterion called Mutual Information Criterion (MIC) (DelSole and Tippett 2021), which is similar to information criteria like Akaike's Information Criterion, except generalized to selection of random predictors and predictands. Following DelSole and Tippett (2022), MIC is calculated as:

$$MIC(T_X, T_Y) = N \log\, \Lambda + P(N, T_X, T_Y), \tag{1}$$

where N is the total number of $\{\mathbf{x}(t), \mathbf{y}(t)\}$ pairs, T_X and T_Y are the number of PCs included in $\mathbf{x}(t)$ and $\mathbf{y}(t)$, respectively, $P(N, T_X, T_Y)$ is a penalty function defined as

$$P(N, T_X, T_Y) = N\left(\frac{(T_X + T_Y)(N+1)}{N - T_X - T_Y - 2} - \frac{T_X(N+1)}{N - T_X - 2} - \frac{T_Y(N+1)}{N - T_Y - 2}\right), \tag{2}$$

and

$$\Lambda = (1-\rho_1^2)(1-\rho_2^2)\ \ldots\ (1-\rho^2_{min(T_X,T_Y)}). \tag{3}$$

As the number of EOFs increases, Λ decreases, reflecting the increase in predictability, but the penalty term increases, reflecting the uncertainty from estimating more parameters. The minimum value of MIC gives us the selection criterion for T_X and T_Y.

f. CCA Critical Values

Statistical significance of the canonical correlations is assessed using Monte Carlo techniques. The significance of the first canonical correlation is determined as follows: Random numbers drawn from a normal distribution are used to populate two matrices $\mathbf{X}$ and $\mathbf{Y}$ of size $T_X \times N$ and $T_Y \times N$, respectively, where N is the sample size and T_X and T_Y are determined by MIC. To ensure consistency, the same preprocessing steps (i.e., removal of the trend, three harmonics, and the seasonal mean) are applied to the random data as to the original data. CCA is then performed on the random matrices to compute the sample canonical correlations. This process is repeated 10,000 times to construct an empirical distribution of the canonical correlations under the null hypothesis of independent $\mathbf{X}$ and $\mathbf{Y}$. The 95th percentile of the leading canonical correlation from the Monte Carlo simulations is taken as the significance threshold at the 5% level.

For the second canonical correlation, the $\mathbf{X}$ and $\mathbf{Y}$ matrices are generated as described above, except this time one (arbitrary) row of the $\mathbf{Y}$-matrix is set equal to a row of the $\mathbf{X}$-matrix, thereby generating a component with a population correlation of 1. The remainder of the procedure is the same as described above. This tests the hypothesis that all canonical correlations except one are 0. Using a population correlation of 1 for the first PC corresponds to a ”worst-case scenario” for the null hypothesis and leads to a conservative estimate of the significance level for the second canonical correlation.

The test for the 3rd correlation is similar, except that two rows of the $\mathbf{Y}$-matrix are set equal to two rows of the $\mathbf{X}$-matrix, and so on.

g. Multivariate Granger Causality

After identifying a predictable relation, we assess whether it is driven by a known climate process (e.g., ENSO or the PNA). Suppose the climate process is represented by an index F. In this case, F can be regressed out of both $\mathbf{X}$ and $\mathbf{Y}$, and CCA applied to the resulting residuals. If F is unrelated to $\mathbf{X}$ and $\mathbf{Y}$, regressing out F should have little effect on the canonical correlations. However, if F drives the relationship between $\mathbf{X}$ and $\mathbf{Y}$, regressing out F should reduce or eliminate at least one of the canonical correlations. The significance of the correlations can be evaluated by incorporating the regress-F-out step into the Monte Carlo procedure described earlier.

The method described above is closely related to Granger Causality (Granger 1969). To understand this connection, recall that a time series F is said to Granger-cause $\mathbf{Y}$ if predictions based on both antecedent $\mathbf{Y}$ and F are more skillful than predictions based on antecedent $\mathbf{Y}$ alone. In practice, Granger Causality is evaluated using the regression model

$$Y = LX + cF + E, \tag{4}$$

where L is a propagator, c is a coefficient, and E is random noise. Whether F Granger-causes $\mathbf{Y}$ depends on c. If the hypothesis $c = 0$ cannot be rejected, then F does not improve the prediction of $\mathbf{Y}$ beyond what can be achieved using $\mathbf{X}$ alone. Conversely, if c is statistically significant, then including F improves the prediction of $\mathbf{Y}$. Therefore, demonstrating that F Granger-causes $\mathbf{Y}$ is equivalent to showing that c is statistically significant.

The two methods are equivalent due to a close connection between CCA and linear regression. Specifically, DelSole and Chang (2003) demonstrate that if each canonical component is predicted separately and then summed across all components, the result is identical to the prediction obtained from multivariate linear regression. This indicates that CCA and linear regression capture the same predictability but express it in different forms. Moreover, by the Frisch-Waugh-Lovell theorem (Frisch and Waugh 1933; Lovell 2008), the regression matrix L is identical to the matrix obtained when F is regressed out of both $\mathbf{X}$ and $\mathbf{Y}$ and fitted to a linear model. Consequently, determining whether c is significant in equation (4) is equivalent to evaluating whether the canonical correlations change after regressing F out of $\mathbf{X}$ and $\mathbf{Y}$.

No procedure can fully guarantee the correct identification of causality, and the above approach is no exception. For instance, suppose both F and $\mathbf{Y}$ are influenced by another climate process, Z. In this case, the coefficient associated with F may still be nonzero, leading the analysis to conclude that F causes $\mathbf{Y}$, when in reality it is Z that causes $\mathbf{Y}$. One way to address this issue is to test multiple climate processes. If more than one process is found to be causal, we may then formulate further hypotheses about the ordering and underlying structure of the causal relationships.

h. How many PCs to regress out when there isn't an index

Soil moisture does not have a standard index associated with it. We compute EOFs of soil moisture over the United States, and then we need to decide how many EOFs of soil moisture we should regress out for Granger Causality. MIC, described in Section 2.d.e, can be used as an objective method to determine how many EOFs to use for testing Granger Causality.

Following Equation 21 of DelSole and Tippett (2021), the appropriate equation is:

$$MIC(X;Y|F) = MIC(XF;Y) - MIC(F;Y) \tag{5}$$

where **X** is the 2m temperature PCs at the initial time, **Y** is the 2m temperature PCs at the response time, and F is the leading PCs of the variable being investigated, at the initial time. To understand Equation 5, recall that MIC is a measure of the degree of predictability. $MIC(XF;Y)$ predicts **Y** using both **X** and F, while $MIC(F;Y)$ predicts **Y** using only F. The difference of these terms tells us how well **X** is able to predict **Y** independent of F (that is, while F is held constant). This gives us $MIC(X;Y|F)$, which is a function of the number of PCs of F. The number of PCs of F to include is determined by the minimum of $MIC(X;Y|F)$.

3. Verifying Predictability in Independent Data

Verifying predictability in independent data is particularly challenging in subseasonal studies, which often involve small sample sizes. Our approach is novel and distinct from the more standard methods discussed in the previous section, so it will be discussed separately in this section.

The sample estimate of the leading canonical correlation is biased upward due to overfitting. Overfitting is a common limitation of statistical optimization methods. This bias becomes pronounced when the true population correlation is small and the sample size is small (Lee 2007).

What is perhaps less widely recognized is that projecting a canonical component onto independent data typically results in an underestimation of the population correlation. This is intuitively reasonable–since CCA tends to overestimate the correlation by incorporating noise into the predictive model, the noise only degrades the predictive value of the model when applied to independent data. As a result, CCA is *expected* to yield upward-biased in-sample correlations and downward-biased out-of-sample correlations, even when both samples come from the same population. Our goal is to quantify these two biases.

It appears to have gone unnoticed that Monte Carlo techniques can be used to estimate both upward *and downward* biases in canonical correlations. The procedure begins as outlined previously: random numbers drawn from a normal distribution are used to fill two matrices, **X** and **Y**, of size $T_X \times N$ and $T_Y \times N$, respectively. By construction, **X** and **Y** are independent. Next, **Y** is modified to include a correlation by setting the first PC of $\mathbf{Y}$, denoted Y_1, to

$$Y_1 = \rho X_1 + \sqrt{(1-\rho^2)} * Z, \tag{6}$$

where X_1 is the first PC of $\mathbf{X}$, and Z is independently drawn from a standard Gaussian distribution. This modification ensures that the population correlation between the first PCs of $\mathbf{X}$ and $\mathbf{Y}$ is ρ, while all other PCs remain independent. CCA is then performed, and we expect at least one sample canonical correlation to be close to ρ. Having performed CCA, we obtain the canonical projection vectors associated with the leading canonical correlation. Applying these vectors to $\mathbf{X}$ and $\mathbf{Y}$ will yield time series with a correlation exactly equal to the leading sample canonical correlation. To validate this relation on independent data, we generate new independent matrices $\mathbf{X}'$ and $\mathbf{Y}'$, in the same manner as described above (particularly using equation (6)), but with a sample size N' matching our verification data. Applying the previously computed projection vectors to $\mathbf{X}'$ and $\mathbf{Y}'$ and computing the correlation gives a realization of the possible correlation that could occur in independent data from the same population. This process is repeated 1,000 times for a given ρ to determine the quantiles of both in-sample and out-of-sample canonical correlations. The procedure is then repeated for different values of ρ, allowing us to estimate the distribution of in-sample and out-of-sample correlations as a function of the population canonical correlation.

An example of the distributions of in-sample and out-of-sample correlations is shown in Figure 1. For each population correlation, the red points represent the mean leading in-sample canonical correlation over the Monte Carlo simulations, with error bars indicating two standard deviations. The black line shows the $x = y$ line for reference. The fact that the red points are above the $x = y$ line illustrates the overfitting discussed earlier, with the largest upward bias occurring when the population correlation is small.

The corresponding blue points represent the mean correlation when the leading in-sample canonical component is projected onto independent data, with error bars also showing two standard deviations. The fact that the blue points lie below the $x = y$ line highlights the tendency to underestimate the population correlation in independent data. While this phenomenon may have been recognized by others, it does not appear to have been previously quantified. Additionally, the in-sample error bars (red) are smaller than the out-of-sample error bars (blue) because the sample

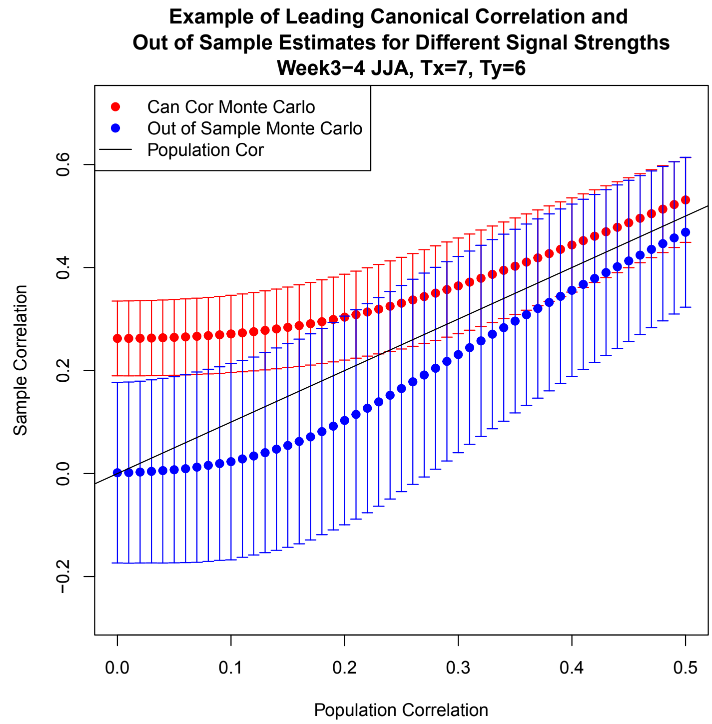


FIG. 1. Estimates of leading canonical correlations (red dots) and their corresponding out of sample correlations (blue dots) for population correlations ranging from 0 to 0.5. These estimates are for the case when $T_X = 7$ and $T_Y = 6$, which corresponds to the number of EOFs used for JJA.

size for the in-sample data is larger than that for the out-of-sample data, as it was chosen to match the actual data length.

The above procedure can be used to derive more comprehensive uncertainty estimates for the canonical correlation that incorporate out-of-sample information. To illustrate this, we use a specific example. In Section 4.b.4, we find that the leading canonical correlation for week 3-4 prediction in JJA is 0.38, while the out-of-sample correlation for this mode is 0.07. Figure 2 presents the same estimated distributions of in-sample and out-of-sample correlations as Figure 1, but with the leading canonical correlation for JJA (0.38, marked as the horizontal red line) and the out-of-sample correlation (0.07, marked as the horizontal blue line) overlaid. For this mode, 7 PCs were included as predictors for $\mathbf{X}$, and 6 were included as predictors for $\mathbf{Y}$; these values were used in the Monte Carlo simulations. The uncertainty of the leading canonical correlation is represented by the horizontal red error bar at the bottom of the figure. This was obtained by calculating the

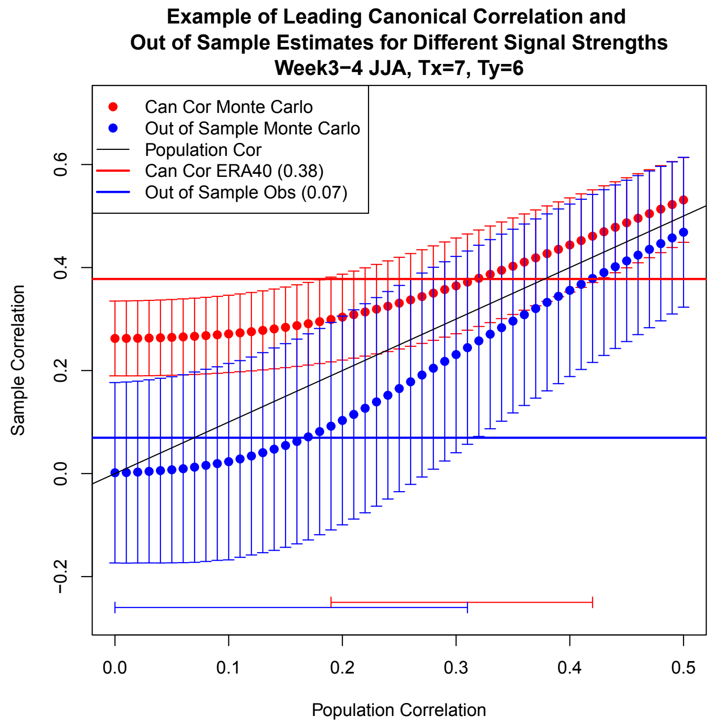


FIG. 2. As in Figure 1, but additionally showing the leading canonical correlation (0.38) as the horizontal red line. The correlation of the leading mode when projected onto an independent sample (0.07) is shown as the horizontal blue line. The bracketed lines at the bottom are the confidence interval for each correlation.

standard errors of the simulated canonical correlations that overlap with 0.38, the observed leading canonical correlation. Similarly, the uncertainty for the out-of-sample correlation is shown by the horizontal blue error bar at the bottom, based on the standard errors of the simulated out-of-sample correlations that overlap with 0.07, the observed out-of-sample correlation. Because these two intervals overlap, we conclude that the in-sample and out-of-sample correlation estimates are consistent with each other. The range of population correlations that overlap (0.19-0.31) represents the interval of population coefficients that is consistent with the 95% confidence intervals of both the in-sample and out-of-sample results.

The above analysis produces an unconventional uncertainty range, as it does not encompass either the in-sample or out-of-sample correlations individually. However, the Monte Carlo simulations demonstrate that any population correlation within the interval (0.19-0.31) could generate results consistent with both the in-sample and out-of-sample correlations derived from observations.

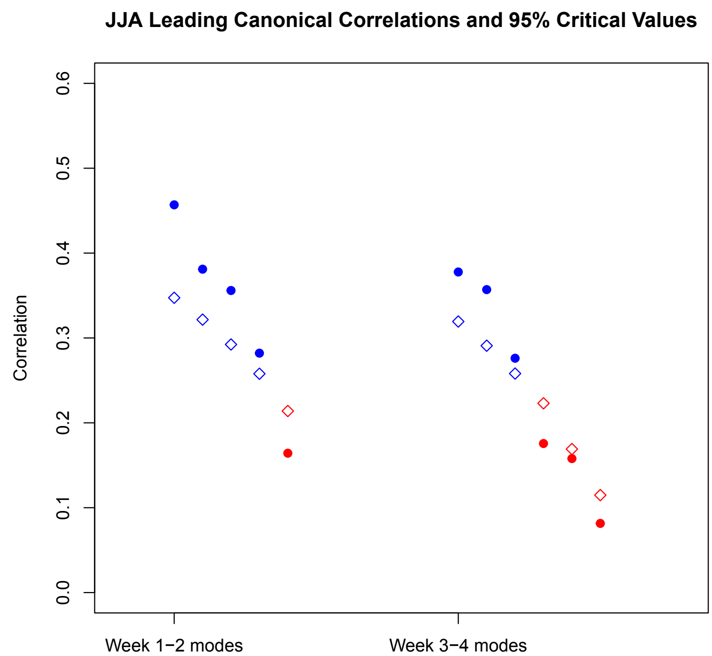


FIG. 3. Canonical correlations and 95% critical values for JJA. Circles are canonical correlations and diamonds are the 95% critical values. Blue points indicate that the correlation is statistically significant while red indicates that it is statistically insignificant, at the 5% level. Correlations when CCA is done at weeks 1-2 is on the left, correlations when CCA is done at weeks 3-4 is on the right.

4. Results

We now present the results of the CCA analysis aimed at identifying the most predictable mode of 2-week mean CONUS temperature during boreal summer. As a reminder, the seasonal mean has been removed to focus exclusively on subseasonal predictability.

Our main finding is that we detect predictable subseasonal modes for both weeks 1-2 and weeks 3-4. Figure 3 shows the leading JJA canonical correlations for weeks 1-2 and weeks 3-4. In this figure, the points are the correlations and the diamonds are the 95% critical values. Correlations above the critical value are statistically significant. As a visual aid, significant correlations are indicated in blue and insignificant correlations are indicated in red.

We next diagnose the structure of the leading modes.

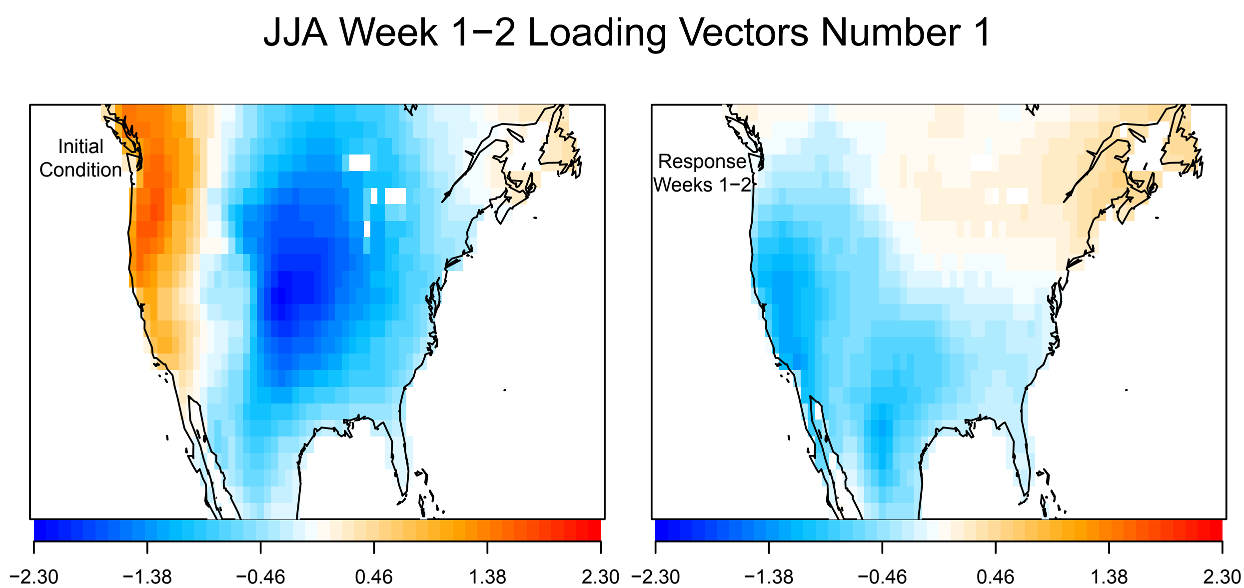


FIG. 4. 1st leading loading vector for JJA at weeks 1-2. The left panel is the initial condition, and the right panel is the week 1-2 response.

a. JJA Weeks 1-2: Leading Mode

1) Loading Vectors

The loading vectors associated with the leading mode for JJA weeks 1-2 are shown Figure 4. The initial condition (left panel of Figure 4) is characterized by a dipole pattern with anomalies of one sign concentrated along the west coast and anomalies of the opposite sign distributed throughout the rest of the US. At the week 1-2 response (right panel of Figure 4), the west coast anomalies have changed sign and propagated to eastern Canada, with most of the interior CONUS remaining the same sign.

2) Relation to Known Sources of Predictability

If a correlation becomes insignificant when a climate index is regressed out, we can conclude that the index that was removed Granger Causes this mode. The canonical correlations after regressing out various climate indices one at a time from the temperature PCs are shown in Figure 5. The red, orange, gold, green, and blue points show the results after removing the Nino 3.4 index, NAO,

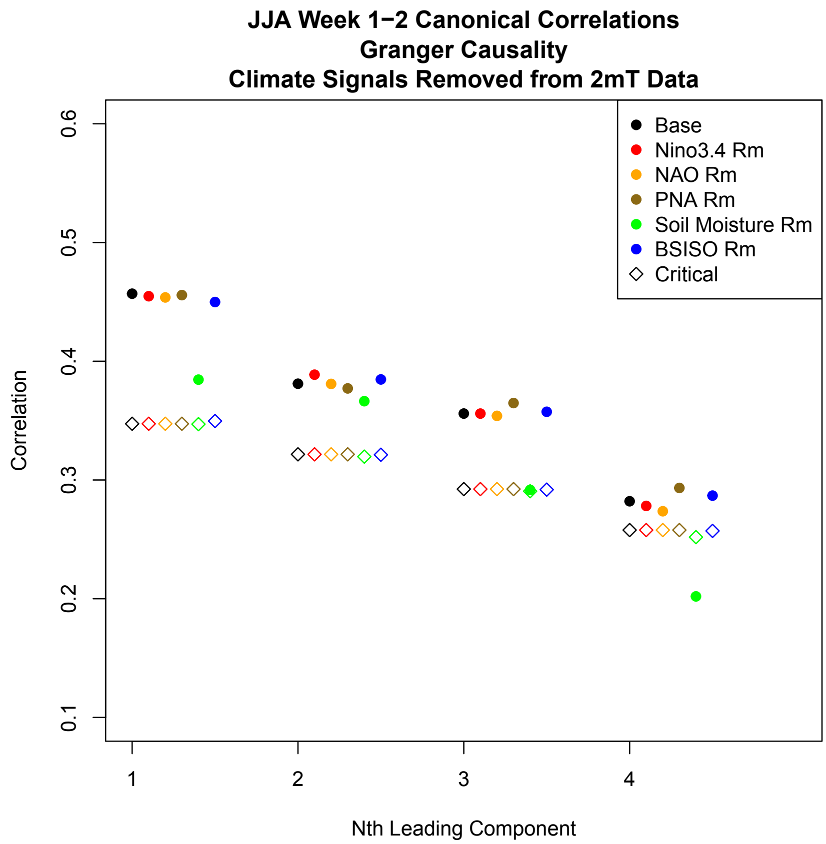


FIG. 5. The statistically significant canonical correlations for JJA at weeks 1-2 after the time series of common climate indices are removed from the 2m temperature PCs. The black points are when no signal is removed and is the same as the week 1-2 correlations in Figure 3. The red, orange, gold, green, and blue points are when the Nino 3.4 index, NAO, PNA, surface soil moisture, and BSISO indices are removed, respectively.

PNA, surface soil moisture, and BSISO indices, respectively. The corresponding critical values for 5% significance are shown as diamonds. For reference, the canonical correlations of the original temperature PCs are shown as black points, reproduced from Figure 3.

Except for the case of soil moisture, the leading canonical correlation remains largely unchanged when the other climate indices are removed. However, when the soil moisture PCs are removed (represented by blue points in Figure 5), the correlation of the first mode decreases, although it remains statistically significant. Notably, the canonical correlation with soil moisture removed is consistent with the second mode's base canonical correlation (compare the leading mode's blue point with the second mode's black point, and similarly, the second mode's blue point with the third mode's black point, and so on). This suggests that regressing out the soil moisture signal effectively eliminates the first mode, causing the second mode to become the new leading mode.

To verify if the modes are indeed the same, we compare their time series. The correlation between the X variates of the leading mode when soil moisture is removed (blue point for the leading mode) and the X variates of the second mode when no signals are removed (black point for the second mode) is 0.62. Similarly, the correlation between the Y variates of these modes is 0.61. Given the uncertainties, these correlations are effectively equal, indicating that the original first mode has been fully removed, and the second mode has shifted into its place. However, this analysis does not determine whether the soil moisture signal itself is a response to other phenomena not represented among our climate indices.

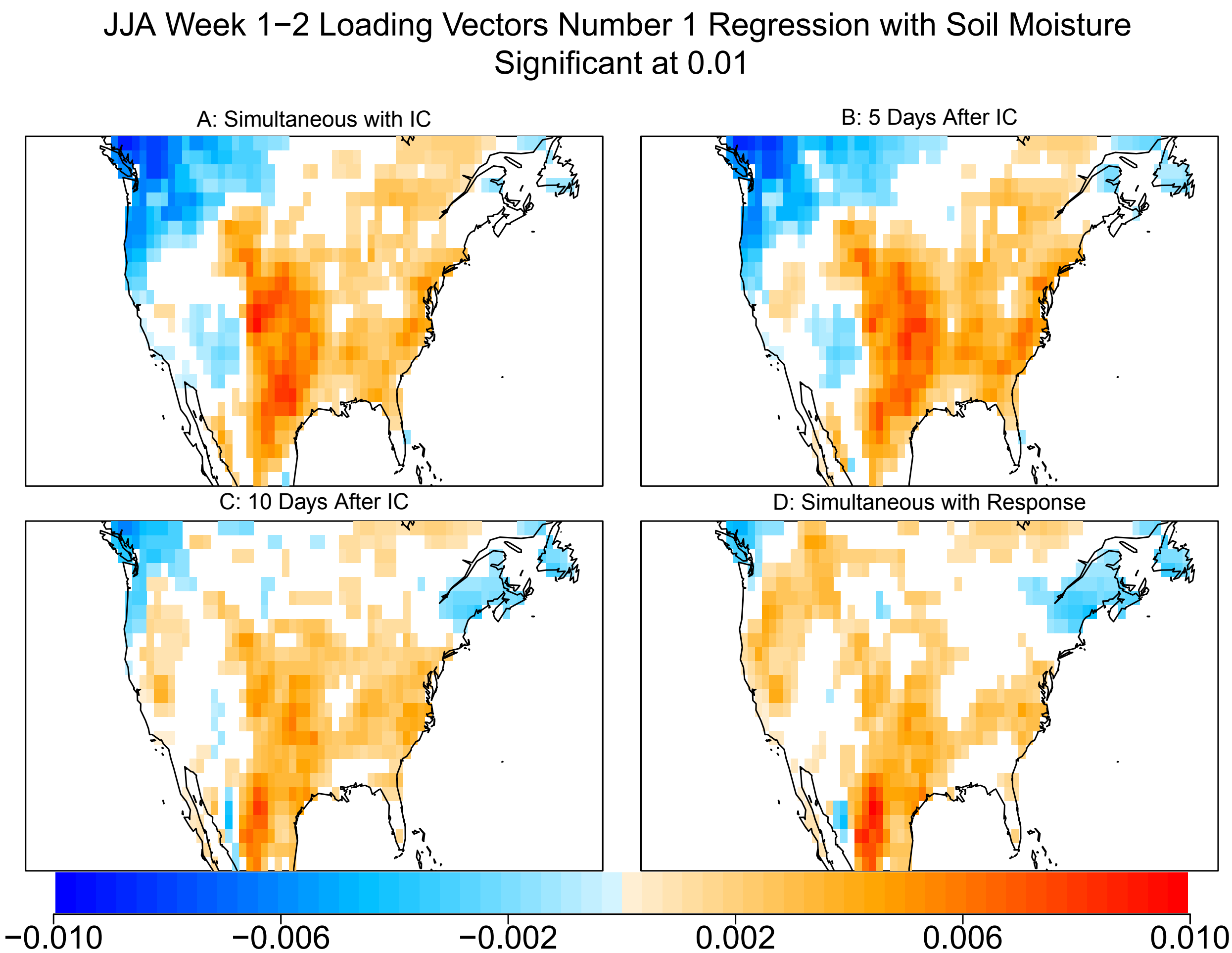
The removal of the other signals does not change the correlation of any of the other modes. This means that we can conclude that the second, third, and fourth modes are not Granger caused by the associated climate mechanisms.

3) Regression Maps

The structure and evolution of each mode, as well as its relationship to other physically relevant variables (denoted Z), will be diagnosed through lagged regression maps. Each predictable mode has an initial condition $\mathbf{X}$ and a response $\mathbf{Y}$. For week 1-2 predictions, $\mathbf{X}$ and $\mathbf{Y}$ represent the same variable, lagged by 14 days. Therefore, a lagged regression map between $X(t)$ and $Z(t+5)$ corresponds to the same day for Z as a lagged regression map between $Y(t)$ and $Z(t-9)$. Since these two regression maps are broadly similar, only one will be presented in the analysis.

As a general rule, before calculating the regression, Z is converted to 2-week means, and the local seasonal mean is removed at each grid point.

Since Granger causality indicates that the leading mode is caused by surface soil moisture, we will start with regression maps of soil moisture. Lagged regression maps between soil moisture and leading mode variates are shown in Figure 6. By comparing Panel A with the initial condition loading vector (left panel in Figure 4), we can see that the soil moisture anomalies are the opposite sign as the loading vector. Comparing Panel D with the week 1-2 response loading vector (right panel of Figure 4), again the soil moisture anomalies and the loading vector are the opposite signs. The conclusion that the temperature anomalies and soil moisture anomalies are anti-correlated



393 FIG. 6. The regression patterns between the leading mode's variates in JJA at weeks 1-2 and surface soil moisture, where each panel shows a regression pattern that is lagged in time. A) Soil moisture is simultaneous with the initial condition, so 14 days prior to the week 1-2 response; B) Soil moisture is 5 days after the initial condition, (9 days prior to the week 1-2 response); C) Soil moisture is 10 days after the initial condition, (4 days prior to the response); D) Soil moisture is 14 days after the initial condition, (simultaneous with the week 1-2 response). The colored grid points are significant at the 0.01 level.

makes physical sense–warmer temperatures will evaporate some of the moisture, and then the lower moisture content means more energy will go into sensible heat than latent heat which will raise the temperature. This suggests that the soil moisture anomalies act to persist the temperature anomalies. This can be seen in the central and southern Great Plains, the southeast, and in the

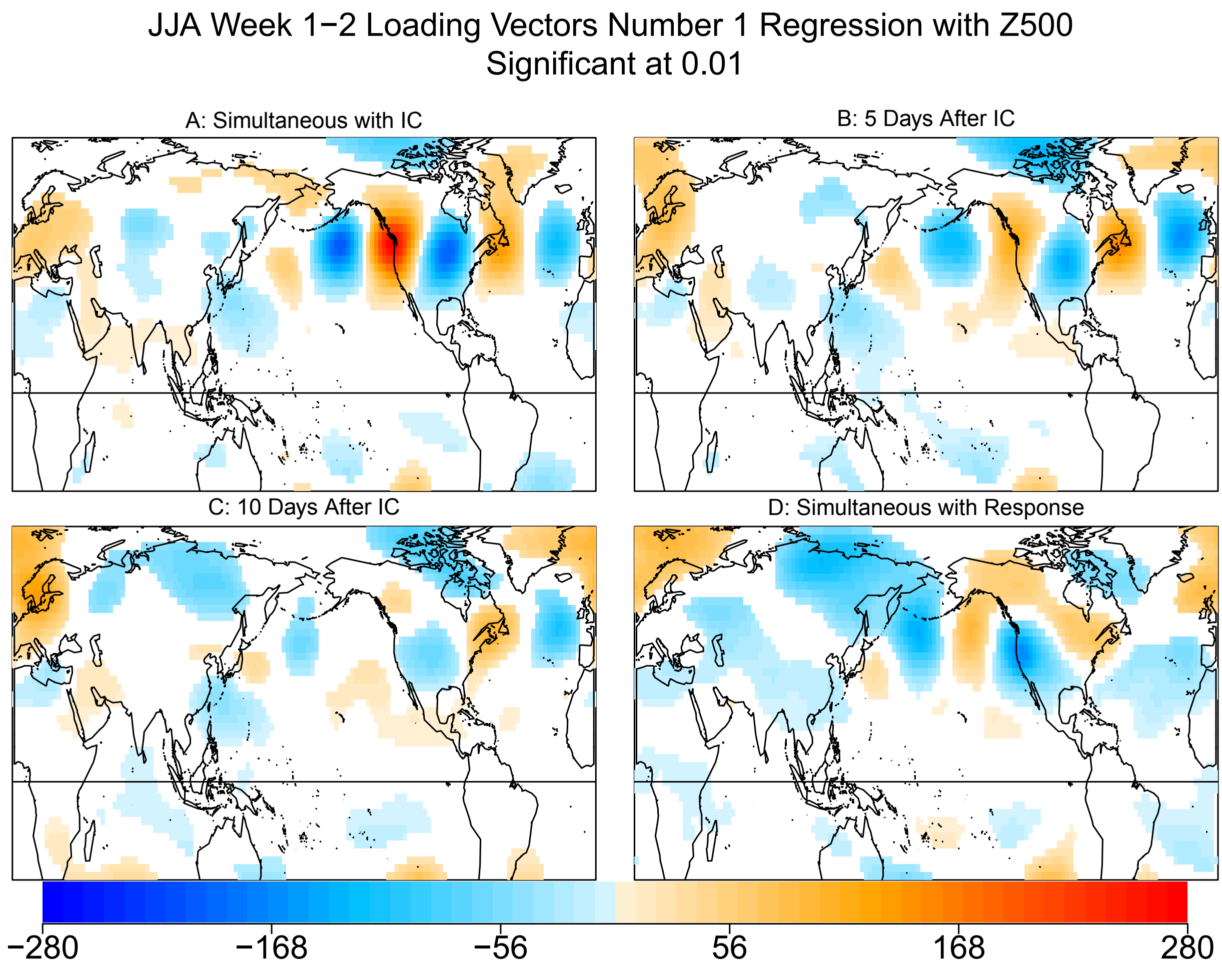


FIG. 7. As in Figure 6, but the regression maps between the leading week 1-2 JJA mode and 500mb geopotential height.

Pacific northwest. It is in these locations that the temperature anomalies remain the same sign from the initial condition to the week 1-2 response (Figure 4), and it is also in these locations that the soil moisture anomalies have the largest amplitude at the initial condition (Panel A of Figure 6).

To investigate the possibility that there may be an atmospheric component to this mode, the variates are regressed onto 500mb height. These lagged regression maps are shown in Figure 7. Simultaneous with the initial condition (Panel A), there is a clear wave originating from the western Pacific. Notably, this wave is oriented zonally. While the path of a Rossby wave typically has a large meridional component in addition to a zonal component, the anomalies in Figure 7 match

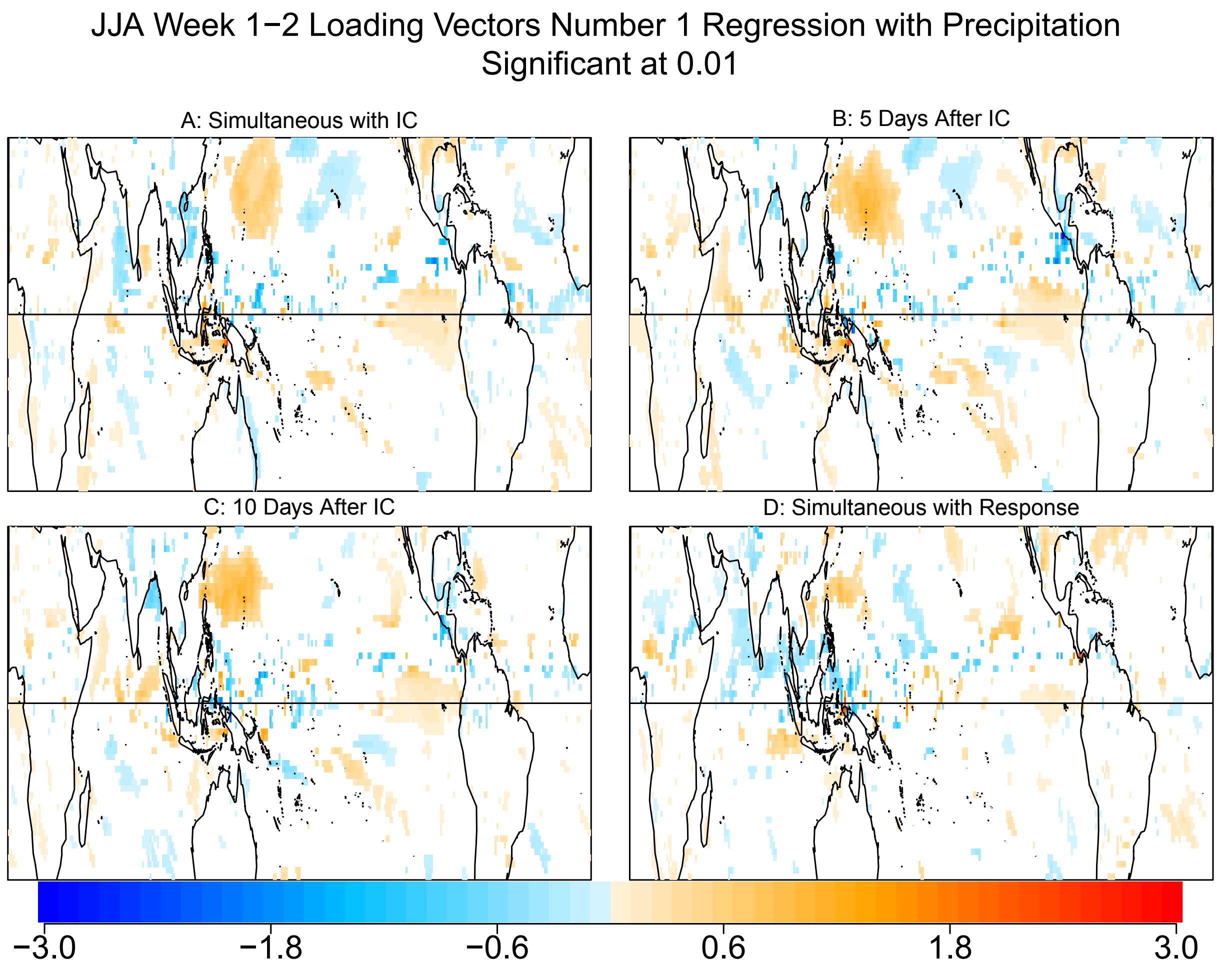


FIG. 8. As in Figure 6, but the regression maps between the leading week 1-2 JJA mode and tropical precipitation.

that of a Rossby wave that is trapped by the climatological jet (Teng and Branstator 2019). These waves are usually called ”circumglobal teleconnections” (Ding and Wang 2005; Branstator 2002), or occasionally ”waveguide teleconnections” (Teng and Branstator 2019).

As the mode progresses in time, the Rossby wave diminishes in amplitude and appears to shift to the west (Panels B-D). This suggests that there is a source of the Rossby wave in the western Pacific. To investigate this, the variates were regressed onto tropical precipitation. These lagged regression maps are shown in Figure 8. At the location where the Rossby wave appears to originate, there is a relatively large-scale precipitation anomaly that persists until the week 1-2 response.

Taken together, these results suggest that the western Pacific precipitation anomaly sets up the Rossby wave, trapped by the jet, which impacts the United States. As the Rossby wave shifts in space, the soil moisture modifies the atmospheric response, causing some temperature anomalies to persist.

The large-scale precipitation anomaly in the western Pacific suggests a possible connection to the Boreal Summer Intraseasonal Oscillation (BSISO). However, several key differences make it challenging to definitively attribute this predictable mode to the BSISO. One difference is that the large-scale precipitation anomaly associated with this mode is farther north than is typical of most descriptions of the BSISO. The precipitation anomaly associated with the predictable mode extends from about 15N to 30N, while the BSISO typically does not extend past 20N (Kikuchi 2021; Chen and Wang 2021). On the other hand, Lee and Wang (2016) show a BSISO extending up to 30N by decomposing the BSISO into Indian Ocean and Western Pacific modes. Additionally, the BSISO has both a positive and a negative precipitation anomaly over the Indian Ocean and western Pacific (although some phases are dominated by anomalies of one sign). In contrast, the predictable mode only has the single positive anomaly. The teleconnections associated with the BSISO are very similar to the Rossby wave generated by this precipitation anomaly (Moon et al. 2013), although that is to be expected given their similar locations. One study by Kerns and Chen (2020) tracked individual large-scale precipitation events in the tropical Pacific. They found that individual MJO events do not always project cleanly onto the MJO indices. However, they also found that large-scale precipitation events poleward of 30N were relatively common, but did not fit their criteria to be defined as an MJO or BSISO event. Due to the differences between the precipitation patterns associated with this mode and the typical BSISO, we cannot definitively conclude that this mode is driven by the BSISO. However, we also cannot rule out the possibility that it may be related to the BSISO.

4) Uncertainty Range of the Canonical Correlation

The description of the out-of-sample correlation test is described in Section 3. The leading canonical correlation in JJA at weeks 1-2 is 0.46, and the correlation of the leading mode in independent data is 0.43. The results of the out of sample correlation test as applied to the leading week 1-2 mode in JJA is shown in Figure 9. Because the two confidence intervals overlap, we

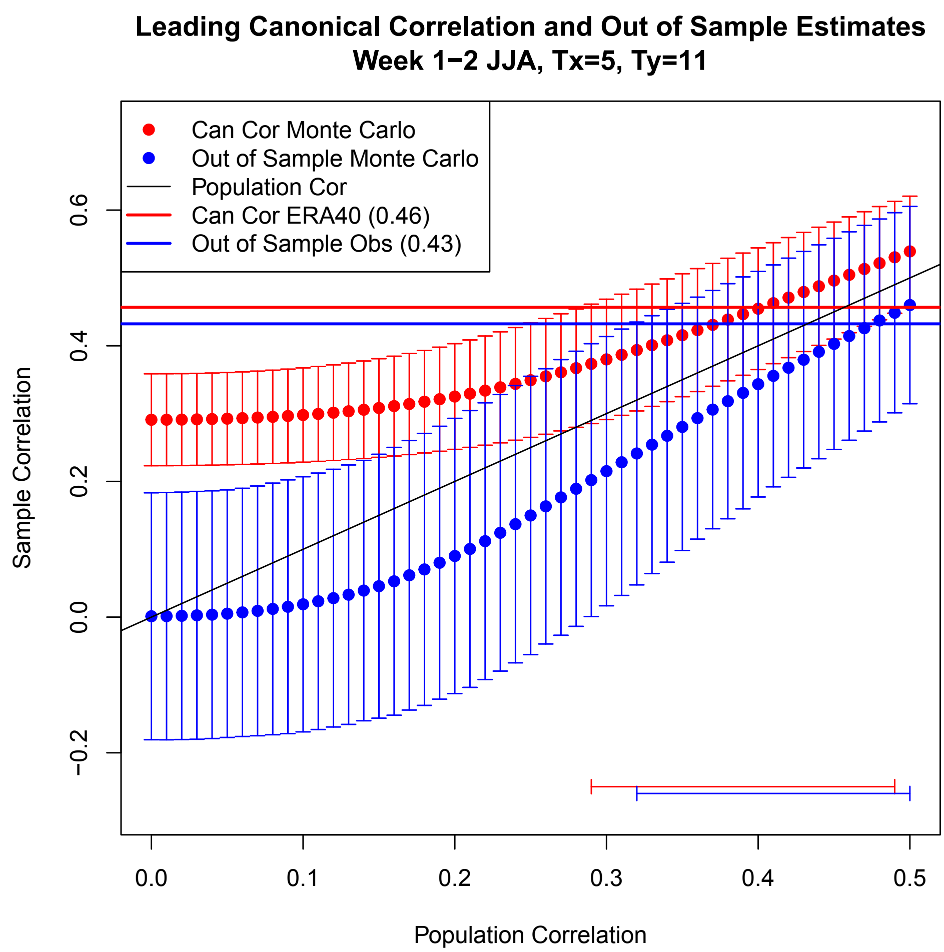


FIG. 9. Estimates of leading canonical correlations (red dots) and their corresponding out of sample correlations (blue dots) for population correlations ranging from 0 to 0.5. These estimates are for the case when $T_X = 5$ and $T_Y = 11$, which corresponds to the number of EOFs used for JJA at week 1-2. The black line shows the 1:1 line for reference. The leading canonical correlation for JJA at weeks 1-2 is shown as the red horizontal line. The correlation of the leading mode when projected onto an independent sample is shown as the horizontal blue line. The bracketed lines at the bottom are the confidence interval for each correlation.

conclude that the leading canonical correlation and out-of-sample correlation are consistent with a population correlation in the range $\rho \in [0.32, 0.49]$.

b. JJA Weeks 3-4

The previous section examined the predictability of temperature at weeks 1-2. However, skillful predictions for weeks 3-4 could potentially be even more valuable to society (White et al. 2017). Therefore, it is important to investigate the most predictable modes at weeks 3-4. The results of this analysis are summarized in the Supplemental Document. Briefly, we detect a predictable mode at weeks 3-4 characterized by (1) an in-sample correlation of 0.38, (2) an input-response loading pair that are largely the same pattern but of opposite sign, suggesting an oscillatory-type predictable pattern, and (3) an associated regression pattern in 500hPa height that strongly resembles the Rossby wave present at the initial condition of the leading week 1-2 JJA mode.

5. CFSv2

The previous analysis presents the predictable subseasonal modes identified by CCA in observational data sets. The next natural question is whether dynamical forecast models capture these predictable subseasonal modes. To address this question, we project the leading mode of each season onto reforecasts of NCEP's dynamical model CFSv2. Then, using the test discussed in Section 3, we assess if the lagged correlations of the leading modes from CFSv2 reforecasts are consistent with observations.

a. Model Data

1) CFSv2 Preprocessing

The reforecasts of the NCEP CFSv2 model (Saha et al. 2014) were evaluated as to whether it was able to capture the subseasonal modes. The reforecasts are available daily from January 1999 to December 2020, excluding 2016. In order to only use data that is independent of the ERA40 data, only reforecasts from 2002 and later are included in this analysis. Each daily reforecast from CFSv2 is 44 days. Anomalies of the ensemble mean were calculated according to Pegion et al. (2019). To get the week 3-4 forecast from each day, forecast days 15-28 were averaged together. Likewise, week 1-2 forecasts were calculated by averaging forecast days 1-14 together. To get the intraseasonal component of the reforecasts, the mean of each season was removed. The data was interpolated onto the ERA40 grid in order to project the loading vectors.

The two-week mean forecasts were selected such that only non-overlapping forecasts were included in the calculation. For example, for boreal summer the two week means beginning on June 1, June 15, July 1, July 15, August 1, and August 15 were selected.

2) Initial Condition Data

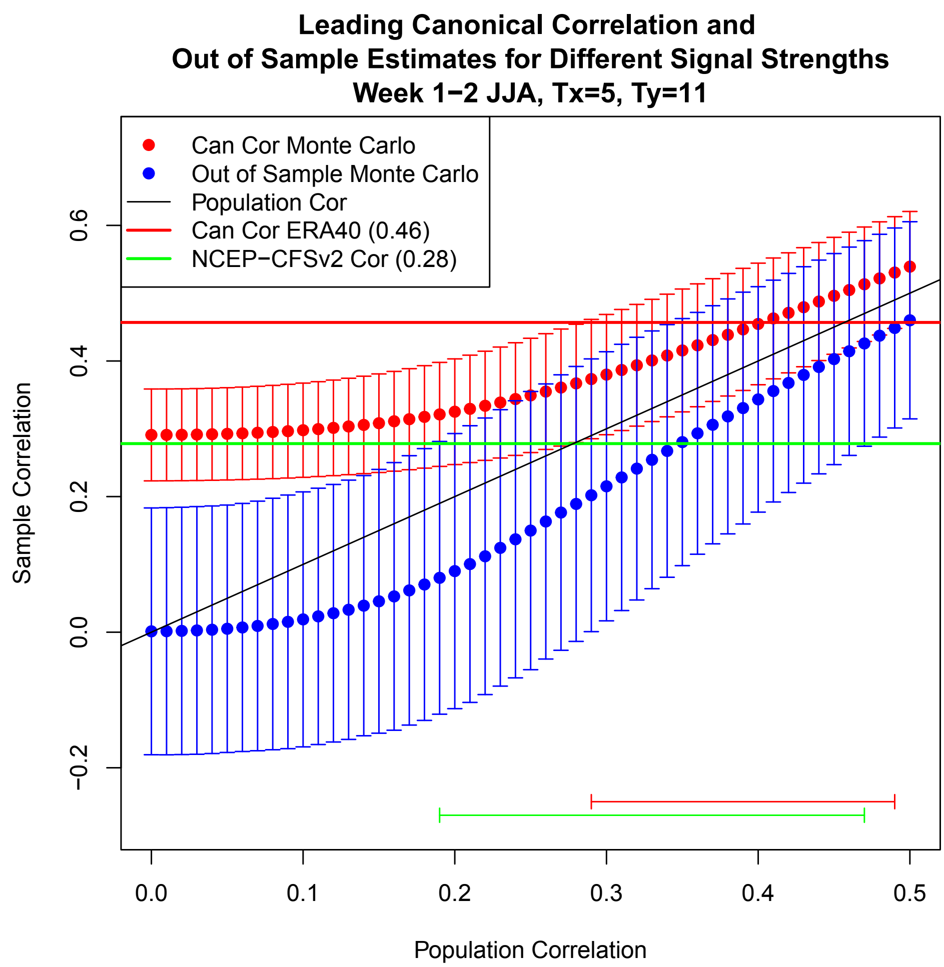
Observed two-week mean 2m temperature from the NOAA ESRL, described in Section 2, was projected onto the X loading vector to obtain the X-variate. CFSv2 re-forecasts at the appropriate leads were projected onto the Y loading vector to obtain the Y-variate.

b. Evaluating CFSv2 for JJA weeks 1-2

The correlation of the prediction of the leading week 1-2 JJA mode in the CFSv2 model is $\rho = 0.28$. The leading canonical correlation is $\rho = 0.46$. Figure 10 shows the canonical correlation, CFSv2 prediction correlation, and the Monte Carlo in-sample and out-of-sample estimates. The confidence intervals overlap, so we conclude that the CFSv2 reforecasts capture this mode.

c. Evaluating CFSv2 for JJA weeks 3-4

The correlation of the prediction of the leading week 3-4 JJA mode by the CFSv2 model is $\rho = -0.12$. The leading canonical correlation is $\rho = 0.38$. Figure 11 shows the canonical correlation, CFSv2 prediction correlation, and the Monte Carlo in-sample and out-of-sample estimates. The confidence intervals do not overlap, so we conclude that the CFSv2 reforecasts do not capture this mode.



504 FIG. 10. As in Figure 9, but for the leading week 1-2 mode in JJA in CFSv2 forecast data. In this case, $T_X = 5$
505 and $T_Y = 11$.

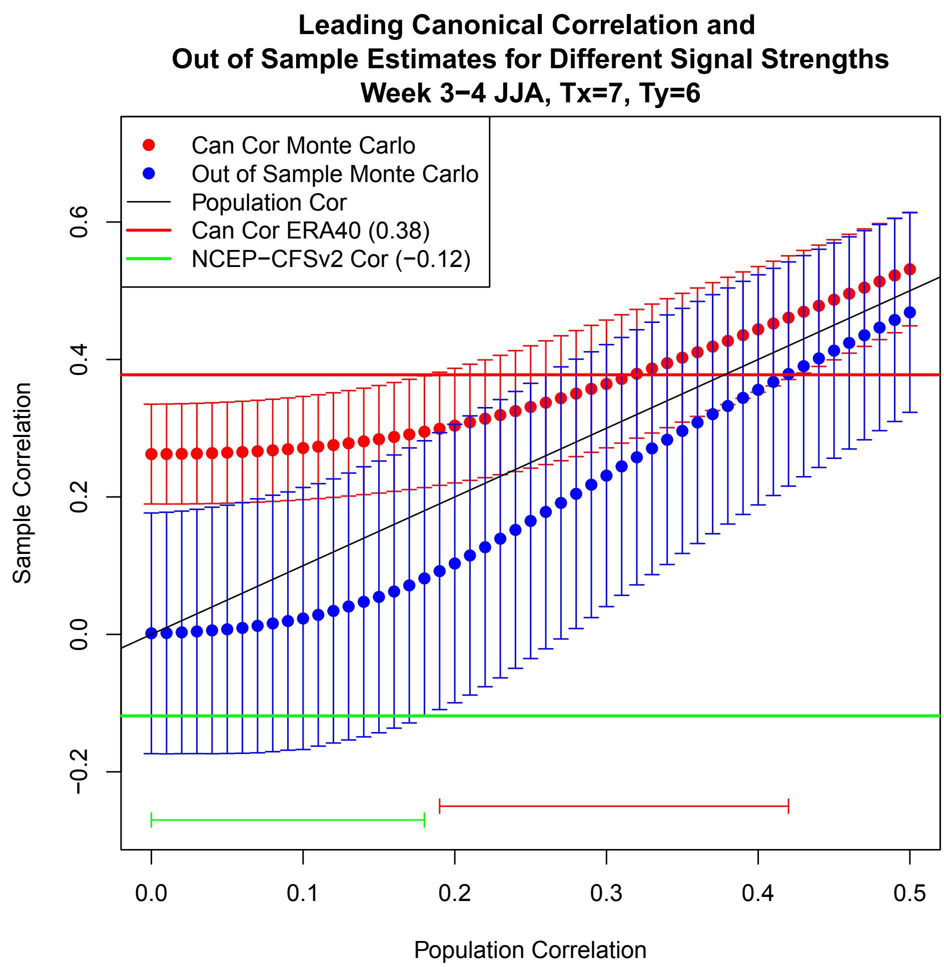


FIG. 11. As in Figure S5, but for the leading week 3-4 mode in JJA in CFSv2 forecast data. In this case, $T_X = 7$ and $T_Y = 6$.

6. Conclusions

The overarching goal of this paper is to identify new sources of subseasonal predictability. To achieve this, we developed a novel methodology that involves removing ENSO influences by subtracting seasonal means from a variable, and then applying Canonical Correlation Analysis (CCA) to time lags of the resulting intraseasonal variable. The rationale underlying this method is that any source of subseasonal predictability that influences a variable should impart a temporal correlation over a few weeks. CCA is an ideal method for identifying such temporal correlations because it finds the indices at the initial and final times that maximize correlation. A major contribution of this work is the development of a rigorous significance test for deciding if the resulting canonical correlations are statistically significant, particularly when validating predictable modes in independent data.

We applied this method to 2-week mean temperature over the United States and identified predictable modes at week 1-2 leads and week 3-4 leads, in JJA. To ascertain if these modes are related to known sources of subseasonal predictability, we applied a Granger Causality test and examined lagged regression maps of variables related to the general circulation. We concluded that the leading JJA modes in weeks 1-2 and 3-4 are new sources of subseasonal predictability.

This mode is associated with a precipitation anomaly in the western Pacific that sets up a Rossby wave, which uses the jet as a waveguide, impacting the United States. As time progresses, the precipitation anomaly switches sign, which sets up a different Rossby wave. A lagged correlation analysis reveals that soil moisture influences the predictable mode in the later stages of its evolution. We suspect that soil moisture in the southern Great Plains modifies the expected atmospheric response by causing the temperature anomalies in the southern Great Plains to persist longer than it otherwise would. In our analysis, the week 3-4 response is the same as the week 1-2 initial condition, which means that the combination of the modes may extend predictability out to week 5-6. Each of these mechanisms has been discussed in the literature, although the chain of mechanisms and their evolution in time has not been presented together before.

One aspect of this mode that we were unable to determine is if the precipitation in the western Pacific is due to the Boreal Summer Intraseasonal Oscillation (BSISO). The BSISO is an oscillation of convection over the Indian Ocean and western tropical Pacific during the boreal summer. It is characterized by northward as well as eastward propagation over the western Pacific (Kikuchi

2021; Chen and Wang 2021). As the precipitation associated with this mode is also in the western Pacific, the BSISO is the natural phenomenon to compare it to. However, there are some differences between our mode and the BSISO. For instance, the precipitation associated with our predictable mode extends about 10 degrees further north than conventional BSISO indices. Furthermore, the BSISO is associated with a large-scale precipitation anomaly of the opposite sign, in contrast to our predictable mode (Section 4.a.3).

We examined if the above predictable modes were captured by the CFSv2 dynamical forecast model. We conclude that it does capture the leading mode of week 1-2 predictability but not the leading mode of week 3-4 predictability. Our results might provide clues about how to improve CFSv2's representation of subseasonal predictability. For instance, CFSv2 was unable to capture the leading week 3-4 JJA mode. The regression maps of this mode show that it is associated by anomalous precipitation in the western Pacific that sets up a Rossby wave impacting the United States (discussed in Section 4.b). This could mean that the CFSv2 model does not have a sufficiently realistic representation of western Pacific precipitation, which could be in the representation of the magnitude or the variability of the precipitation. Another explanation may be that the model does not have a sufficiently good representation of the tropical-extratropical teleconnections, either in setting up the Rossby wave or in its propagation.

One limitation of using CCA to identify predictable temperature patterns is that some climate mechanisms may not impact the temperature during both the initial condition and the response at weeks 1-2 or 3-4, and yet the mechanisms themselves may be predictable that far in advance. For example the MJO has a relatively small direct impact on wintertime temperature over the United States, with only phases four through six producing statistically significant, large scale temperature anomalies (Zhou et al. 2012). However, it has been shown that dynamical models can accurately forecast the state of the MJO four weeks in advance (Pegion et al. 2019; Du et al. 2024). This means that while we may be able to forecast the direct impact of the MJO using dynamical models, CCA applied in the manner described above would not be able to capture that predictability owing to the weak teleconnection.

It might come as some surprise that these new sources of predictability were found using CCA, which has been used to study aspects of the climate for decades. We posit that this is for three reasons. The first is that CCA rarely is applied to lagged temperature fields. Most studies that use CCA have used it to find the temperature pattern that is most correlated with some other field, often SST. As a result, those studies limit themselves to the temperature response from the other variable. In our method, we were not limited to finding only the response from one variable. The second reason is that we were able to employ a relatively new criterion, MIC (DelSole and Tippett 2021), to objectively determine the number of EOFs to use for both the initial condition and the response. Without MIC, prior studies have had to justify the number of EOFs used. This was often based on the total amount of variance explained by the EOFs and the cutoff differed from study to study. The third reason is that we have developed a novel significance test for the leading mode based on a Monte Carlo procedure and by using independent data for validation of the correlation.

While this work has focused on subseasonal predictability, the methodology developed here is broadly applicable to any time series, offering new pathways for uncovering and understanding previously unrecognized sources of predictability.

Acknowledgments. We thank Paul Dirmeyer and Kathy Pegion for valuable guidance throughout the years of this research, and Davis Straus for insightful comments on an early version of this work. This research was supported primarily by the National Science Foundation (AGS-1822221). Additional support was provided from National Science Foundation (AGS-1338427), National Aeronautics and Space Administration (NNX14AM19G), the National Oceanic and Atmospheric Administration (NA14OAR4310160). The views expressed herein do not necessarily reflect the views of these agencies.

Data availability statement. The observed daily 2m temperature is from the NOAA Earth System Research Laboratories (ESRL) and was accessed from `https://www.esrl.noaa.gov/psd/data/gridded/data.cpc.globaltemp.html`. The observed time series of the PNA and NAO are from NOAA's Climate Prediction Center (CPC) and was accessed from `https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/pna.shtml`. The EEOFs of OLR that make up the BSISO indices are from the International Pacific Research Center and was accessed from `https://iprc.soest.hawaii.edu/users/kazuyosh/Bimodal_ISO.html`. The ERA40 reanalysis is a ECMWF product and can be accessed from `https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-40-years`. A set of R codes for performing the analyses described in this paper can be found at `https://github.com/PaulBuchmann/CCA-Summer`.

References

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. *Monthly Weather Review*, **115 (9)**, 1825 – 1850, https://doi.org/10.1175/1520-0493(1987)115⟨1825:OALOMA⟩2.0.CO;2.

Barnston, A. G., and T. M. Smith, 1996: Specification and prediction of global surface temperature and precipitation from global sst using cca. *Journal of Climate*, **9 (11)**, 2660 – 2697, https://doi.org/10.1175/1520-0442(1996)009⟨2660:SAPOGS⟩2.0.CO;2.

Branstator, G., 2002: Circumglobal teleconnections, the jet stream waveguide, and the north atlantic oscillation. *Journal of Climate*, **15 (14)**, 1893 – 1910, https://doi.org/10.1175/1520-0442(2002)

015⟨1893:CTTJSW⟩2.0.CO;2.

Buchmann, P., 2024: Assessing the limits of improving subseasonal predictability indices. Ph.D. thesis, George Mason University.

Chen, G., and B. Wang, 2021: Diversity of the boreal summer intraseasonal oscillation. *Journal of Geophysical Research: Atmospheres*, **126 (8)**, e2020JD034 137, https://doi.org/https://doi.org/10.1029/2020JD034137.

DelSole, T., and P. Chang, 2003: Predictable component analysis, canonical correlation analysis, and autoregressive models. *J. Atmos. Sci.*, **60**, 409–416.

DelSole, T., and M. Tippett, 2022: *Statistical Methods for Climate Scientists*. Cambridge University Press, https://doi.org/10.1017/9781108659055.

DelSole, T., and M. K. Tippett, 2021: A mutual information criterion with applications to canonical correlation analysis and graphical models. *Stat*, **10 (1)**, e385, https://doi.org/10.1002/sta4.385.

Ding, Q., and B. Wang, 2005: Circumglobal teleconnection in the northern hemisphere summer. *Journal of Climate*, **18 (17)**, 3483 – 3505, https://doi.org/10.1175/JCLI3473.1.

Domeisen, D. I., and Coauthors, 2022: Advances in the subseasonal prediction of extreme events: Relevant case studies across the globe. *Bulletin of the American Meteorological Society*, https://doi.org/10.1175/BAMS-D-20-0221.1.

Du, D., A. C. Subramanian, W. Han, W. E. Chapman, J. B. Weiss, and E. Bradley, 2024: Increase in mjo predictability under global warming. *Nature Climate Change*, **14 (1)**, 68–74, https://doi.org/10.1038/s41558-023-01885-0.

Frisch, R., and F. V. Waugh, 1933: Partial time regressions as compared with individual trends. *Econometrica*, **1 (4)**, 387–401.

Ghil, M., and Coauthors, 2002: Advanced spectral methods for climatic time series. *Reviews of Geophysics*, **40 (1)**, 3–1–3–41, https://doi.org/10.1029/2000RG000092.

Granger, C. W. J., 1969: Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, **37 (3)**, 424–438.

He, S., X. Li, L. Trenary, B. A. Cash, T. DelSole, and A. Banerjee, 2021: Learning and dynamical models for sub-seasonal climate forecasting: Comparison and collaboration. *arXiv [Preprint]*, https://doi.org/10.48550/ARXIV.2110.05196.

Huth, R., 2002: Statistical downscaling of daily temperature in central europe. *Journal of Climate*, **15 (13)**, 1731 – 1742, https://doi.org/10.1175/1520-0442(2002)015⟨1731:SDODTI⟩2.0.CO;2.

Hwang, J., P. Orenstein, J. Cohen, K. Pfeiffer, and L. Mackey, 2019: Improving subseasonal forecasting in the western u.s. with machine learning. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Association for Computing Machinery, 2325–2335, KDD '19, https://doi.org/10.1145/3292500.3330674.

Jenney, A. M., K. M. Nardi, E. A. Barnes, and D. A. Randall, 2019: The seasonality and regionality of mjo impacts on north american temperature. *Geophysical Research Letters*, **46 (15)**, 9193–9202, https://doi.org/10.1029/2019GL083950.

Johnson, N. C., D. C. Collins, S. B. Feldstein, M. L. L'Heureux, and E. E. Riddle, 2014: Skillful wintertime north american temperature forecasts out to 4 weeks based on the state of enso and the mjo. *Weather and Forecasting*, **29 (1)**, 23–38, https://doi.org/10.1175/WAF-D-13-00102.1.

Kerns, B. W., and S. S. Chen, 2020: A 20-year climatology of madden-julian oscillation convection: Large-scale precipitation tracking from trmm-gpm rainfall. *Journal of Geophysical Research: Atmospheres*, **125 (7)**, e2019JD032 142, https://doi.org/10.1029/2019JD032142.

Kikuchi, K., 2020: Extension of the bimodal intraseasonal oscillation index using jra-55 reanalysis. *Climate Dynamics*, **54 (1)**, 919–933, https://doi.org/10.1007/s00382-019-05037-z.

Kikuchi, K., 2021: The boreal summer intraseasonal oscillation (bsiso): A review. *Journal of the Meteorological Society of Japan*, **99 (4)**, 933–972, https://doi.org/10.2151/jmsj.2021-045.

Krishnamurthy, V., and Coauthors, 2021: Sources of subseasonal predictability over conus during boreal summer. *Journal of Climate*, **34 (9)**, 3273 – 3294, https://doi.org/10.1175/JCLI-D-20-0586.1.

Lee, H.-S., 2007: Canonical correlation analysis using small number of samples. *Communications in Statistics - Simulation and Computation*, **36 (5)**, 973–985, https://doi.org/10.1080/03610910701539443.

Lee, S.-S., and B. Wang, 2016: Regional boreal summer intraseasonal oscillation over indian ocean and western pacific: comparison and predictability study. *Climate Dynamics*, **46 (7)**, 2213–2229, https://doi.org/10.1007/s00382-015-2698-7.

Lovell, M. C., 2008: A simple proof of the fwl theorem. *The Journal of Economic Education*, **39 (1)**, 88–91, https://doi.org/10.3200/JECE.39.1.88-91.

Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical pacific. *Journal of Atmospheric Sciences*, **28 (5)**, 702 – 708, https://doi.org/https://doi.org/10.1175/1520-0469(1971)028⟨0702:DOADOI⟩2.0.CO;2.

McGovern, A., D. J. Gagne, J. K. Williams, R. A. Brown, and J. B. Basara, 2014: Enhancing understanding and improving prediction of severe weather through spatiotemporal relational learning. *Machine Learning*, **95 (1)**, 27–50, https://doi.org/10.1007/s10994-013-5343-x.

Moon, J.-Y., B. Wang, K.-J. Ha, and J.-Y. Lee, 2013: Teleconnections associated with northern hemisphere summer monsoon intraseasonal oscillation. *Climate Dynamics*, **40 (11)**, 2761–2774, https://doi.org/10.1007/s00382-012-1394-0.

National Academies of Sciences Engineering and Medicine, 2016: *Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts*. The National Academies Press, https://doi.org/10.17226/21873, URL https://nap.nationalacademies.org/catalog/21873/next-generation-earth-system-prediction-strategies-for-subseasonal-to-seasonal.

National Research Council, 2010: *Assessment of Intraseasonal to Interannual Climate Prediction and Predictability*. The National Academies Press, Washington, DC, https://doi.org/10.17226/12878, URL https://www.nap.edu/catalog/12878/assessment-of-intraseasonal-to-interannual-climate-prediction-and-predictability.

Nowak, K., I. M. Ferguson, J. Beardsley, and L. D. Brekke, 2020: Enhancing western united states sub-seasonal forecasts: Forecast rodeo prize competition series. *AGU Fall Meeting Abstracts*, Vol. 2020, A188–0008.

Pegion, K., and Coauthors, 2019: The subseasonal experiment (subx): A multimodel subseasonal prediction experiment. *Bulletin of the American Meteorological Society*, **100 (10)**, 2043 – 2060, https://doi.org/10.1175/BAMS-D-18-0270.1.

Robertson, A. W., and F. Vitart, Eds., 2019: *Sub-Seasonal to Seasonal Prediction The Gap Between Weather and Climate Forecasting*. Elsevier, https://doi.org/10.1016/C2016-0-01594-2.

Saha, S., and Coauthors, 2014: The ncep climate forecast system version 2. *Journal of Climate*, **27 (6)**, 2185 – 2208, https://doi.org/10.1175/JCLI-D-12-00823.1.

Teng, H., and G. Branstator, 2019: Amplification of waveguide teleconnections in the boreal summer. *Current Climate Change Reports*, **5 (4)**, 421–432, https://doi.org/10.1007/s40641-019-00150-x.

Trenary, L., and T. DelSole, 2023: Skillful statistical prediction of subseasonal temperature by training on dynamical model data. *Environmental Data Science*, **2**, e7, https://doi.org/10.1017/eds.2023.2.

Vitart, F., and Coauthors, 2022: Outcomes of the wmo prize challenge to improve subseasonal to seasonal predictions using artificial intelligence. *Bulletin of the American Meteorological Society*, **103 (12)**, E2878 – E2886, https://doi.org/10.1175/BAMS-D-22-0046.1.

White, C. J., and Coauthors, 2017: Potential applications of subseasonal-to-seasonal (s2s) predictions. *Meteorological Applications*, **24 (3)**, 315–325, https://doi.org/https://doi.org/10.1002/met.1654.

White, C. J., and Coauthors, 2021: Advances in the application and utility of subseasonal-to-seasonal predictions. *Bulletin of the American Meteorological Society*, 1 – 57, https://doi.org/10.1175/BAMS-D-20-0224.1.

Zhou, S., M. L'Heureux, S. Weaver, and A. Kumar, 2012: A composite study of the mjo influence on the surface air temperature and precipitation over the continental united states. *Climate Dynamics*, **38 (7)**, 1459–1471, https://doi.org/10.1007/s00382-011-1001-9.