

## PERSPECTIVE OPEN



## Seasonal to multi-year soil moisture drought forecasting

Musa Esit<sup>1,2</sup>, Sanjiv Kumar<sup>1✉</sup>, Ashutosh Pandey<sup>3</sup>, David M. Lawrence<sup>4</sup>, Imtiaz Rangwala<sup>5,6</sup> and Stephen Yeager<sup>4</sup>

Soil moisture predictability on seasonal to decadal (S2D) continuum timescales over North America is examined from the Community Earth System Modeling (CESM) experiments. The effects of ocean and land initializations are disentangled using two large ensemble datasets—initialized and uninitialized experiments from the CESM. We find that soil moisture has significant predictability on S2D timescales despite limited predictability in precipitation. On sub-seasonal to seasonal timescales, precipitation variability is an order of magnitude greater than soil moisture, suggesting land surface processes, including soil moisture memory, reemergence, land–atmosphere interactions, transform a less predictable precipitation signal into a more predictable soil moisture signal.

*npj Climate and Atmospheric Science* (2021)4:16; <https://doi.org/10.1038/s41612-021-00172-z>

## INTRODUCTION

Drought is a complex, multifaceted phenomenon involving climate, water resources, and socioeconomic drivers and impacts<sup>1–4</sup>. It causes billions of dollars in economic losses and severe stress on ecosystem productivity<sup>5–7</sup>. While significant efforts are underway to develop drought monitoring and forecasting systems, the ability to forecast drought is limited due to inherent uncertainties in precipitation forecasts at long lead times (months to years)<sup>8–14</sup>. For example, the NOAA's Seasonal Climate Outlook did not predict the 2012 agricultural drought (also commonly referred to as soil moisture drought) in the central US which resulted in \$30 billion economic losses<sup>15,16</sup>. In this study, we assess the initialized<sup>16</sup> and uninitialized<sup>5</sup> large ensemble experiments of the Community Earth System Model (CESM) for North America and show that recent advances in earth system modeling<sup>17</sup> combined with an improved understanding of long-memory land surface processes<sup>18,19</sup> can enable skillful predictions of soil moisture drought several months in advance (Fig. 1) despite limited skills in the precipitation forecast.

The soil moisture variability represents an integrated effect of climate, vegetation, and soil processes. The Community Land Model (CLM)<sup>20,21</sup>, which is the land component of the CESM, simulates a range of biophysical and biogeochemical processes, including land surface heterogeneity, radiation scheme, momentum, energy balance, hydrology, photosynthesis, stomatal conductance, and prognostic vegetation phenology<sup>22,23</sup>. The CLM simulates soil moisture variability with high fidelity at par with the observationally constrained remote-sensing estimates (Supplementary Fig. 1). Previous studies have compared soil moisture variability with various drought metrics and found the soil moisture metric's superiority for agricultural drought applications<sup>24,25</sup>. The soil moisture variability is correlated significantly with the Palmer Drought Severity Index in the CESM experiment (Supplementary Fig. 2).

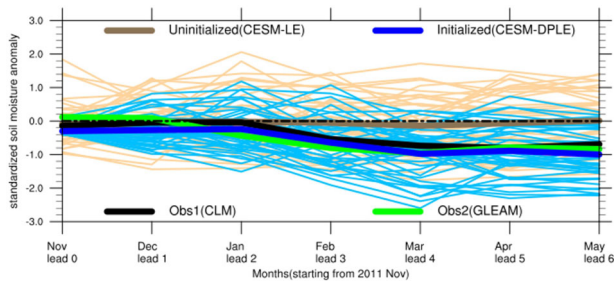
The slowly varying ocean anomalies, e.g., El Niño–Southern Oscillation (ENSO)<sup>12,13</sup>, Atlantic Multi-decadal Variability<sup>26,27</sup>, and Pacific Decadal Variability (PDV)<sup>28–31</sup> affect land hydroclimate at seasonal to decadal (S2D) timescales through atmospheric teleconnection processes<sup>32</sup>. Multi-season to multi-year memory from deep soil moisture and groundwater may also contribute to

S2D predictability of hydroclimate features such as drought<sup>33–37</sup>, pluvials<sup>31</sup>, and wildfire<sup>38</sup>. External climate forcing from greenhouse gas emissions and anthropogenic aerosol emissions can provide predictability on multi-year timescales<sup>32,39</sup>.

We quantify and disentangle the effects of ocean and land initializations on S2D hydroclimate predictability, and demonstrate that land surface processes (i.e., soil moisture memory, reemergence, and land–atmosphere interactions) can contribute to the improvement of forecast skills on S2D timescales by transforming a weak precipitation signal into a predictable soil moisture signal. This finding is an advancement compared with the previous studies that have documented soil moisture as a source of predictability on sub-seasonal to seasonal timescales, only<sup>32,40</sup>. We also develop an observational constraint on the predictability estimates that can motivate future development of S2D prediction systems. Finally, we aim to bring fundamental advances by (a) developing the predictability estimates on the S2D timescale continuum instead of multi-year only<sup>41,42</sup> or seasonal only timescale<sup>43</sup> and (b) assessing hydroclimate predictability in the mid-latitude regions where the signal to noise ratio is known to be smaller than in equatorial regions<sup>12,44</sup> and therefore where predicting drought is more challenging<sup>13</sup>.

The CESM Decadal Prediction Large Ensemble (CESM-DPLE) uses observation equivalent ocean and sea ice states derived from the forced ocean–sea ice (FOSI) simulations<sup>45,46</sup>, and land condition from a selected realization of the Large Ensemble simulation<sup>9</sup> for the given time to initialize S2D earth system prediction<sup>47</sup> (Table 1). Each initialization consists of a 40-member ensemble forecast generated by randomly perturbing the initial atmospheric condition, and each forecast is integrated for 10 years. In contrast, the CESM Large Ensemble (CESM-LE)<sup>9</sup> is a fully coupled 40-member ensemble climate simulations where neither the ocean initial states nor the land initial states are constrained at the start of each forecast period. Additionally, the CESM-LE allows disentangling the effect of land initialization by contrasting the effect of the initialized land anomalies with the remaining 39 ensembles not used for initializations (Table 1). Hence, CESM-DPLE and its uninitialized parallel CESM-LE provide a rich dataset to estimate the potential of drought predictability on the S2D timescales.

<sup>1</sup>Earth System Science Program, School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL, USA. <sup>2</sup>Department of Hydrology, Gaziantep University, Gaziantep, Turkey. <sup>3</sup>Center for Atmospheric Sciences, Indian Institute of Technology, New Delhi, India. <sup>4</sup>National Center for Atmospheric Research, Boulder, CO, USA. <sup>5</sup>Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA. <sup>6</sup>North Central Climate Adaptation Science Center, University of Colorado, Boulder, CO, USA. ✉email: [szk0139@auburn.edu](mailto:szk0139@auburn.edu)



**Fig. 1 Skillful prediction of the 2012 soil moisture drought in the California–Nevada region (30°–37° north latitudes and 110°–120° west longitudes).** The CESM–DPLE drought forecast, initialized in November 2011, and the uninitialized CESM–LE forecast are compared with two observation equivalent data (CLM and GLEAM). The thin lines show anomalies from individual ensemble members, and the thick lines show ensemble mean.

**Table 1.** List of climate experiments employed to investigate S2D hydroclimate predictability.

Ex. type	Experiment name and design details	Reference
Initialized	CESM–DPLE	Yeager et al. <sup>47</sup>
	Ocean IC—from FOSI	
	Land IC—from CESM–LE # 34	
	Atmosphere IC—perturbed	
	Ensemble size: 40	
	Initialization date: Nov. 1 of 1980, 1981, ..., 2015	
Uninitialized	CESM–LE	Kay et al. <sup>9</sup>
	Internally generated weather/climate variability under historical climate forcing	
	ICs—N/A	
	Ensemble size: 40	
Observation equivalent	CLM4—with GSWP3 meteorological forcing	Lawrence et al. <sup>20,21</sup>
	GLEAM—remote sensing-based soil moisture data	Marten et al. <sup>69</sup>

*CESM–DPLE* Community Earth System Model–Decadal Prediction–Large Ensemble, *IC* initial condition, *FOSI* forced ocean–sea ice simulation, *CESM–LE* Community Earth System Model–Large Ensemble, *CESM–LE #34* Ensemble member number 34, *N/A* not applicable, *CLM4* Community Land Model version 4, *GLEAM* Global Land Evaporation Amsterdam Model, *GSWP3* Global Soil Wetness Project Phase 3.

## PREDICTABILITY OF SOIL MOISTURE ON THE S2D TIMESCALES

Our analysis demonstrates that soil moisture predictability is much higher and spatially extensive than for precipitation on S2D timescales. Figure 2 compares the root zone (0–0.5 m) soil moisture predictability with the precipitation over North America for 36 initialization dates (1980, 1981, ..., 2015). We use a recent period (1980–2015) in our analysis because of the significant trend in the longer period (1955–2015) (not shown). We employ a signal to noise ratio metric that compares the variance of the ensemble average forecast (signal) to that of the ensemble spread (noise) to determine the effects of the initializations<sup>48</sup> (see the “Methods” section). This is a measure of the potential predictability (referred to as predictability hereafter) in the perfect model world. Actual realized forecast skill is likely to be lower than the predictability due to model biases and uncertain initial conditions<sup>43,49</sup>.

The spatial extent of the statistically significant signal is three times greater for soil moisture than for precipitation. Averaged over year 1–10 forecasts, 51% of the North American land area shows a statistically significant soil moisture predictability than 18.5% area for the precipitation. The DPLE soil moisture

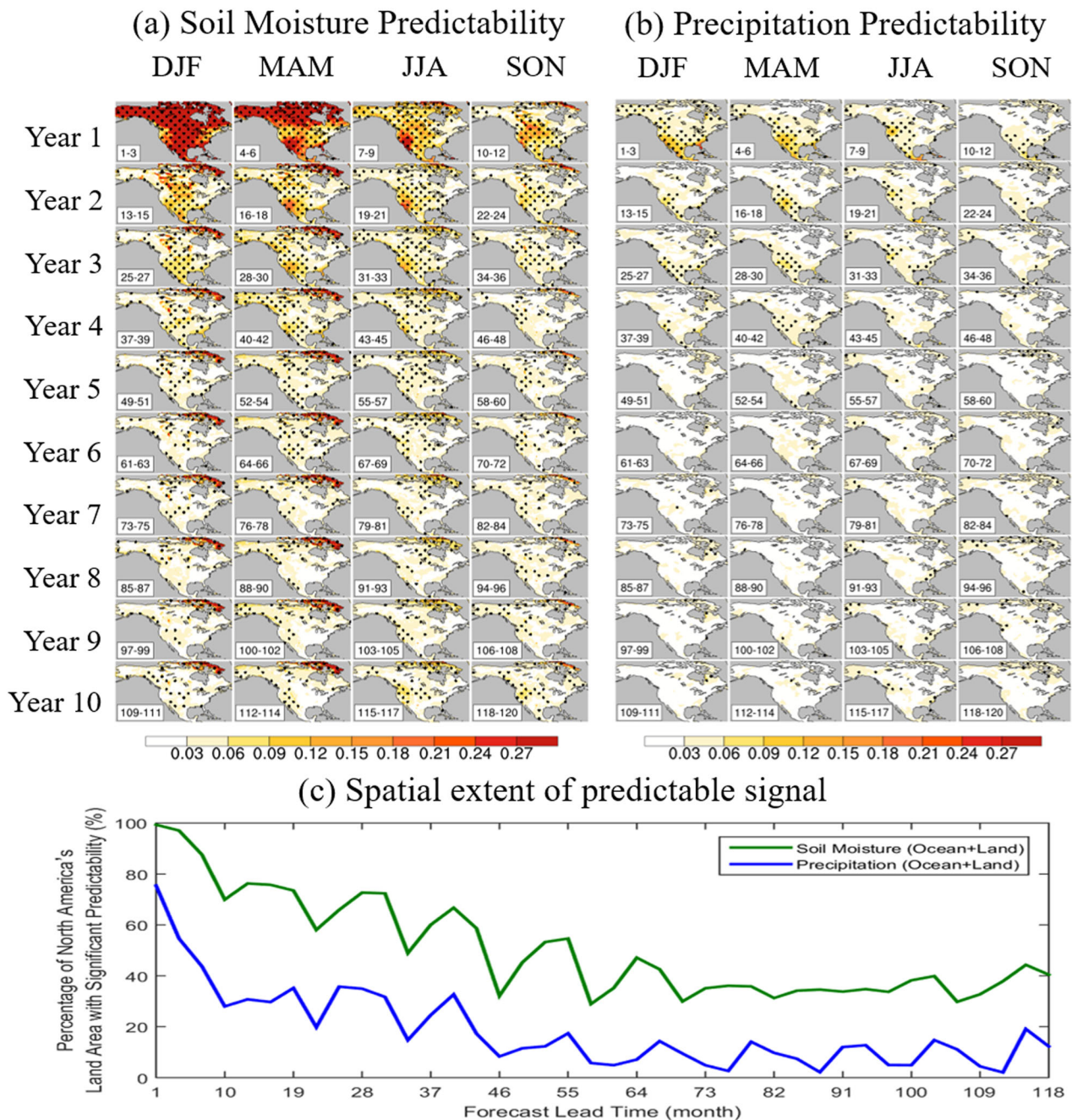
predictability is about four times higher than the soil moisture persistence model forecast (Supplementary Fig. 3). Furthermore, there is predictability for soil moisture beyond the timescale of the precipitation predictability. For example, there is almost minimal area of predictability beyond year 5 for precipitation, whereas there is predictability seen beyond year 5 in soil moisture in about 36% of land area that includes many areas in Western North America, which has a history of the mega-drought<sup>50</sup>. A cyclic behavior in the forecast skill (Fig. 2c) can be due to differences in skill across the annual cycle<sup>43</sup> because of the Nov1 initialization date (Supplementary Fig. 4).

Spring (MAM) has the highest soil moisture predictability (55%) averaged over the 10 years of the forecast, while fall (SON) has the lowest predictability (39%). Soil moisture predictability for winter (DJF) and summer (JJA) are 50% and 53%, respectively (Fig. 3). A previous study<sup>18</sup> has found soil moisture anomalies reemergence in the root zone from deeper soil layer anomalies stored in the preceding seasons. A stronger land–atmosphere coupling in spring and summer can enhance the predictability due to memory stored in the land surface<sup>51</sup>. As expected, soil moisture predictability decreases with increasing lead times<sup>52</sup>, but it remains considerably higher than precipitation throughout the forecast period (Fig. 2). The majority of the land area (68% of North America) shows significant forecast skill for soil moisture by the end of year-2, and decreasing to 37% land area by year 10.

External climate forcing, e.g., greenhouse gas emissions, contribute to a signal in the decadal prediction<sup>32</sup>, i.e., the global warming trend is a part of the signal. Hence, we did not remove the trend from our predictability analysis. However, a global warming trend can also counteract the effects of the initialization. For example, a forecast initialized during the positive phase of the PDV would often exhibit a wetting signal in the US southwest<sup>53</sup> that can be muted at long lead times by the gradual soil moisture drying due to the global warming trend<sup>4</sup>. A removal of trend slightly improves the predictability at longer lead-time, particularly for soil moisture with an accentuated annual cycle in years 7–10 (Supplementary Fig. 5).

Land initialization can bring considerable improvements in S2D soil moisture forecasts. To disentangle the effects of land initialization, we computed the anomaly correlations between initial condition total soil moisture anomaly (past 12 months average before November 1) and the ensemble average forecast anomalies for the 36 initialization date (1980–2015) (Fig. 3). We compared it with the remaining 39 LE ensembles’ anomaly correlations that were not used to initialize the DPLE land condition (thin gray line in Fig. 3). If the spatial extent of significant anomaly correlation with the initialized land anomalies is greater than that for the remaining 39 ensembles, it demonstrates a significant impact of land initialization on soil moisture predictability. This methodology gives an improved result than a previous method that assessed the effect of land initialization by correlating the DPLE forecast anomalies with the evolution of soil moisture in uninitialized LE#34<sup>54</sup> (Supplementary Fig. 6), where soil moisture anomalies evolved from a different SST initial condition than the DPLE forecast.

Land initialization brings statistically significant improvement in predictability for lead times exceeding one year in three out of four seasons, i.e., effects of initialized soil moisture anomalies are significantly higher than the remaining 39 ensembles not used in initializations (Fig. 3). For DJF, the impact of land initialization extends out to 7 years. This is a noteworthy result in at least two respects: (1) previous studies have reported improved predictability due to soil moisture and/or land initialization is generally limited to sub-seasonal to seasonal timescales<sup>32,52</sup>. Our results show that land initialization can bring improvement in predictability for multiple years. (2) CESM–DPLE uses a synthetic land initialization taken from a selected ensemble of the uninitialized experiment (Table 1). We are not aware of any observationally



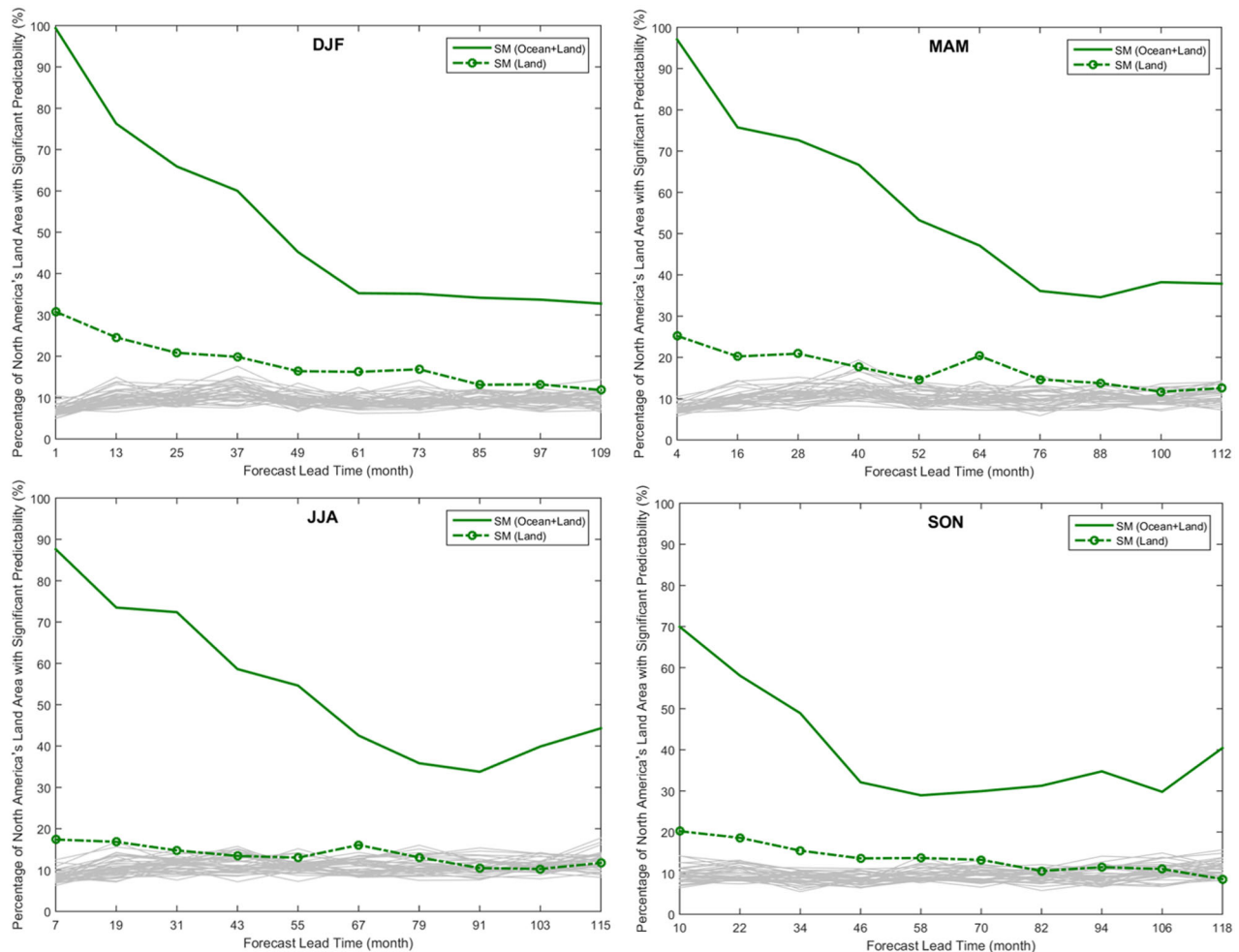
**Fig. 2** The S2D predictability of (a) root zone (0–0.5 m) soil moisture and its comparison with the (b) precipitation predictability in the CESM-DPLE, and using signal to (signal + noise) ratio metric. Stippling shows statistical significance at a 95% confidence level. The bottom panel (c) shows areas of the significant predictability for soil moisture and precipitation in North America.

constrained land initialization technique conceptually similar to FOSI simulations<sup>45,46</sup> for the ocean. We hope that our results provide the necessary motivation for developing the observationally constrained land initialization technique.

Why is soil moisture more predictable? We compared the predictability of soil moisture and precipitation averaged over the annual timescale. The areal extent of predictability is similar between soil moisture and precipitation when averaged annually (Fig. 4). Significant predictability is found for most of the North American land areas (>50%) from year 1 to year 10, both for precipitation and soil moisture, but not always co-located particularly at the longer lead time, e.g., lead 10-year suggesting different climate processes contributing to the predictability. Since

the low-frequency (inter-annual) precipitation variability component is smaller (7%) compared with the sub-seasonal (43%) and seasonal (42%) variability components in North America (Supplementary Fig. 7); annually averaged precipitation predictability does not contribute to improving sub-seasonal to seasonal precipitation predictability (e.g., Fig. 2). Also, the magnitude of the signal strength is weaker for precipitation than soil moisture in years 1 and 2 (Fig. 4, and Supplementary Fig. 8).

The sub-seasonal to seasonal variability in precipitation is an order of magnitude higher than the soil moisture variability (Fig. 5), i.e., a highly variable sub-seasonal to seasonal precipitation is less predictable in the DPLE experiment. However, the soil moisture memory, reemergence, and



**Fig. 3 Comparison of the total soil moisture predictability area in North America due to ocean and land initialization effects (Ocean + Land) with that due to land-only initialization effect (Land).** Thin gray lines show the expected statistical distribution if the land-only were not initialized. We determined the land-only effects by computing the anomaly correlations between initialized total soil moisture anomalies and the ensemble average forecast anomalies for the root zone soil moisture. Thin gray lines are obtained by correlating the uninitialized soil moisture initial condition (remaining 39 LE initial conditions) with the ensemble average DPLE forecast.

land–atmosphere interactions processes can boost the signal strengths<sup>18,55,56</sup> and contribute to a much higher predictability for soil moisture on those timescales. For example, soil moisture variability is smaller (22% in CLM and 7% in GLEAM) at sub-seasonal timescale despite a higher precipitation variability at sub-seasonal scale (43%). In contrast, inter-annual soil moisture variability is higher (16% in CLM, and 29% in GLEAM) than inter-annual precipitation variability (7%) (Supplementary Fig. 7), suggesting contribution from the memory effect<sup>55,56</sup>.

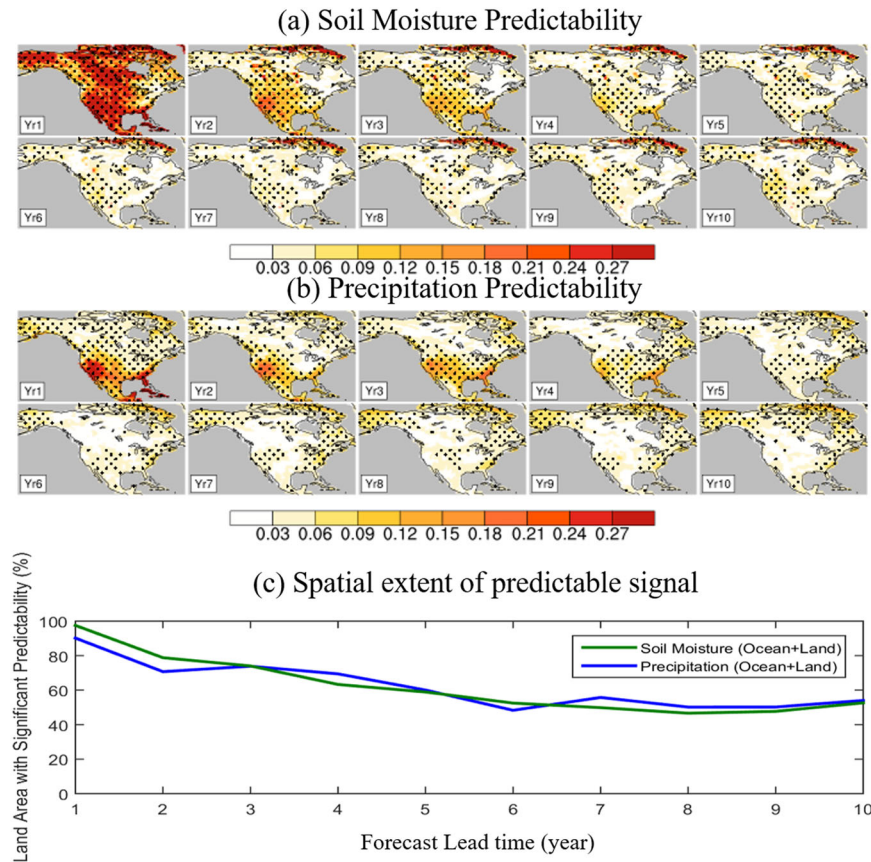
The soil moisture forecast skill seen in Figs 1–4 should be considered conservative estimates of the predictability compared to the observational estimates shown in Fig. 6. The CESM exhibits weak soil moisture to the precipitation feedback loop. The CESM-DPLE precipitation forecasts do not show significant improvements due to soil moisture initialization, i.e., the anomaly correlation between initial soil moisture anomalies and ensemble average precipitation forecast anomalies in DPLE are not statistically distinguishable from the anomaly correlations between the remaining 39 ensemble initial soil moisture anomalies and the precipitation forecasts (Supplementary Fig. 9). Since the CESM-DPLE has retrospective forecasts, it provides an opportunity to test those forecasts against observations.

We compare the anomaly correlations for CLM's soil moisture and precipitation, which are considered here as a proxy for

observations, with that of the CESM-DPLE forecasts (Fig. 6). We find that the anomaly correlations between CLM initial condition soil moisture and the succeeding observed precipitation are considerably higher than those for the DPLE forecast during the first year (Fig. 6a). For example, the observations show statistically significant correlations for 30% land area during the first 6 months, whereas the DPLE does not show a significant anomaly correlation. This analysis suggests that weaker feedbacks between antecedent soil moisture and precipitation<sup>57</sup> contribute to smaller improvements to soil moisture predictability due to land initialization effects in the DPLE (Fig. 6b).

### PERSPECTIVE ON SEASONAL TO MULTI-YEAR DROUGHT FORECASTING

Predictability of drought based primarily on precipitation forecast has low skill, particularly in the mid-latitudes<sup>10,12,13,58</sup> to effectively support drought early warning and response. Land surface integrates the stochastic tendency of weather at longer timescales within the memory of the deep soil layers<sup>18,33</sup>, and land–atmosphere coupling effects by decreasing the humidity gradient between land and atmosphere, and therefore increasing the soil moisture memory<sup>55,59</sup>. The new soil moisture predictability analysis presented here suggests that opportunities may exist to



**Fig. 4** The annual average predictability of (a) the root zone (0–0.5 m) soil moisture and its comparison with (b) the precipitation predictability in the CESM-DPLE, and using signal to (signal + noise) ratio metric. Twelve-month average anomalies (non-overlapping) are calculated for each year of the forecast lead time, and before calculating the signal to total ratio. Stippling shows statistical significance at a 95% confidence level. The bottom panel (c) shows areas of the significant predictability area for soil moisture and precipitation in North America.

develop skillful drought prediction systems focused on soil state, which would be highly relevant for agricultural drought predictions<sup>24,25</sup>.

Land initialization contributes to approximately a third of total soil moisture predictability, with the remainder arising from ocean conditions (Fig. 3). Previous studies have made considerable efforts to develop observationally constrained ocean initializations, e.g., FOSI simulations<sup>45,46</sup>. This study's results imply that a similar effort to develop observationally constrained land-initial conditions could be beneficial. The recent availability of multi-scale multi-source soil moisture data, including in situ<sup>60,61</sup> and remote-sensing measurements<sup>62,63</sup>, may provide an opportunity to develop the land initializations.

Why did the ocean initializations provide skill in the soil moisture forecast but limited skill in the precipitation forecast on the S2D timescale (Fig. 2)? Soil moisture variability has a redder spectrum than precipitation<sup>64</sup>. A recent study hypothesized a new climate process, soil moisture reemergence, in which moisture anomalies stored deep in the soil column can provide year-to-year soil moisture memory<sup>18</sup>. In parallel research, ocean reemergence studies<sup>65–67</sup> have found that much of the PDV results from the reemergence, which acts to redden the ENSO signal in the North Pacific<sup>19</sup>. It will be worth investigating further whether the reddened soil moisture signal is driven by atmospheric teleconnections only<sup>32</sup> or if the land processes contribute to the redness by integrating the ENSO signal in the same way as the North Pacific does.

The soil moisture-precipitation feedback is weak in the climate model used for this study (CESM1) (Fig. 6). When compared to

observations and reanalysis data, Mei and Gulling<sup>57</sup> found a smaller effect of antecedent soil moisture conditions on precipitation in an earlier version of the CESM1 model (CAM4-CLM4). Infanti and Kirtman<sup>68</sup> have also found a weak land initialization effect on precipitation in CAM4-CLM4. A weak soil moisture to precipitation feedback contributes to a muted effect of land initialization on soil moisture predictability compared with the observation (Fig. 6). Clearly, there exists an opportunity to develop and improve the seasonal to multi-year soil moisture drought forecasting system based on recent advances in earth system modeling and our improved understanding of long-memory land surface processes.

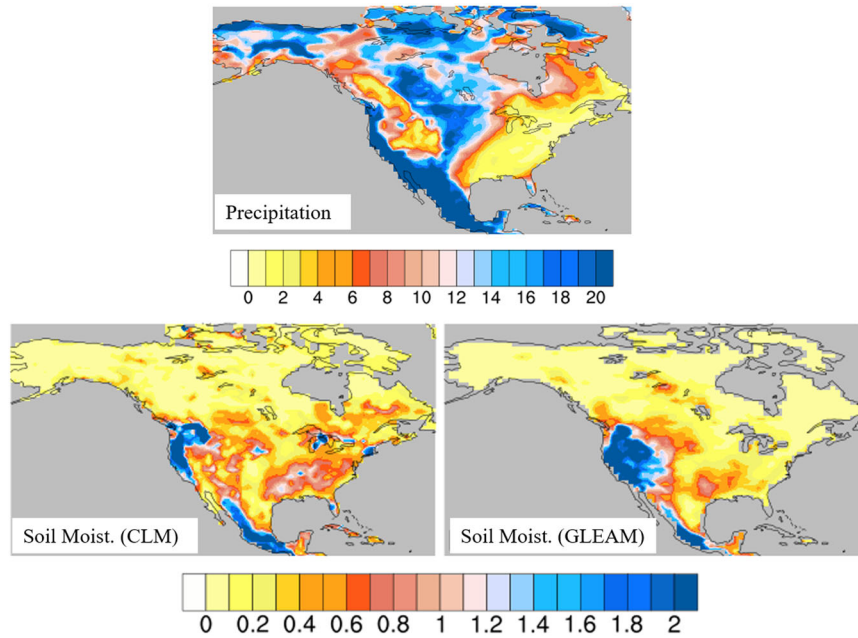
## METHODS

### Forecast anomalies calculation

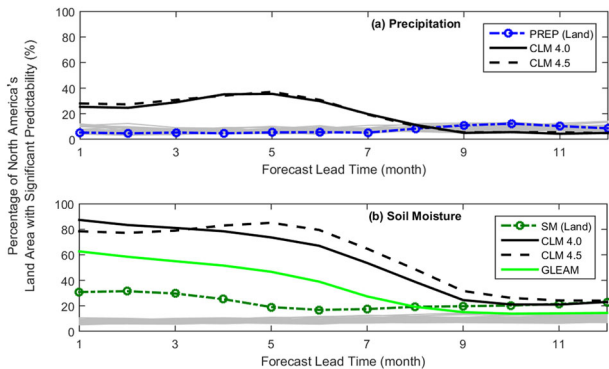
The DPLE forecast anomalies are computed with respect to its lead time-dependent forecast climatology as described below<sup>69</sup>:

$$fA_{i,l,y} = f_{i,l,y} - \frac{1}{p} \sum_{y=1}^p \left( \frac{1}{n} \sum_{i=1}^n f_{i,l,y} \right) \quad (1)$$

where  $f_{i,l,y}$  is the DPLE forecast initialized in year  $y$  (1980, 1981, ..... 2015), at lead months  $l$  (1, 2, .... 122), and  $i$ th ensemble (1, 2, ...40) for a given variables soil moisture or precipitation;  $fA_{i,l,y}$  is the corresponding forecast anomalies.



**Fig. 5 Apportionment entropy (AE) for the observed precipitation and soil moisture data.** Please note that the color scale for the AE of the precipitation is one order magnitude higher than that of the soil moisture. AE<sup>71</sup> is a non-parametric measure of sub-seasonal to seasonal variability in the data—higher AE represents a higher variability in the data (see the “Methods” section).



**Fig. 6 The scope of improving the forecast skill by developing models that show comparable effect of the antecedent soil moisture on precipitation forecast as found in the observational estimates.** Figure compares the effect of land initialization on precipitation and soil moisture forecasts in the CESM-DPLE with that of the observational estimates from CLM4 and CLM4.5. The figure shows anomaly correlations between initial condition total soil moisture anomalies and the ensemble average forecast (colored lines) or observations (black lines) in the respective experiments. The GLEAM root zone soil moisture data results is also shown.

The LE anomalies are computed with respect to its ensemble average climatology for the given month as below:

$$LEA_{i,m,y} = LE_{i,m,y} - \frac{1}{p} \sum_{y=1}^p \left( \frac{1}{n} \sum_{i=1}^n LE_{i,m,y} \right) \quad (2)$$

where  $LE_{i,m,y}$  is the LE simulations for year  $y$  (1980, 1981, ..... 2015), and the given month  $m$  (1, 2, .... 12), and  $i$ th ensemble (1, 2, ...., 40) for a given variables soil moisture or precipitation;  $LEA_{i,m,y}$  is the corresponding LE anomalies.

The observation (CLM) anomalies are computed with respect to observation climatology:

$$OA_{m,y} = O_{m,y} - \frac{1}{p} \sum_{y=1}^p (O_{m,y}) \quad (3)$$

where  $O_{m,y}$  is the CLM land-only simulations for year  $y$  (1980, 1981, ..... 2015), and the given month  $m$  (1, 2, ...., 12) for a given variables soil moisture or precipitation;  $OA_{m,y}$  is the corresponding observation anomalies. Figure 1 demonstrates that the methodology outlined here successfully removed the initial drift in the DPLE forecast<sup>47,69</sup>, and gives comparable results among three experiments: DPLE, LE, and Observations (CLM).

**Signal to total ratio**

To diagnose potential predictability, we determined the squared signal-to-noise ratio (SNR)<sup>48</sup> at a given lead time as

$$SNR = \frac{V_s}{V_n} = \frac{\frac{1}{p} \sum_p (fA_{y,l} - fA_{\cdot,l})^2}{\frac{1}{np} \sum_p \sum_i (fA_{iy,l} - fA_{y,l})^2} \quad (4)$$

where  $fA_{\cdot,l} = \frac{1}{np} \sum_{y=1}^p \sum_{i=1}^n fA_{iy,l}$  and  $f_{y,l} = \frac{1}{n} \sum_{i=1}^n fA_{iy,l}$ .  $V_s$  represents the variability of the ensemble mean, which is a potentially predictable signal due to initializations;  $V_n$  is variability about the ensemble mean, or the noise term. The null hypothesis of no predictability can be rejected at the 95% level if  $SNR \geq F_{p-1,p(n-1)}^{0.05} * \frac{p-1}{p(n-1)}$ , where  $F_{p-1,p(n-1)}^{0.05}$  is the upper 5% threshold for the  $F$ -distribution with  $p-1$  and  $p(n-1)$  degrees of freedom. The “signal-to-total ratio” is then defined as  $STR = \frac{SNR}{SNR+1}$ , which varies from 0 to 1. Note that for an infinite-member perfect-model ensemble, we may measure forecast skill of the predicted signal by the anomaly correlation, which can be shown is equivalent to the square of STR<sup>70</sup>; that is, STR is a measure of predictability. We computed STR for each forecast month (1–122 months lead forecasts), and their 3-month average values, representing different seasons, are shown in Fig. 2.

## Disentangling the effects of land initialization

We determined initial condition soil moisture anomalies from the LE ens# 34, averaged over the previous year. For example, for the forecast start date on November 1, 1980, we computed the initial condition soil moisture anomalies averaged from November 1, 1979 to October 31, 1980, and over the entire soil depth (0–3.8 m). We tested other permutation of the initial condition anomalies, e.g., the last 3 months, and the root zone anomalies. We found that the past 12 months and total depths gives the best result for soil moisture predictability at greater than one-year lead-time. We found that antecedent condition from the past 3 months give a higher anomaly correlation during the first year, but its influence decreases after one year (Supplementary Fig. 6). Then we computed anomaly correlations between initial condition soil moisture anomalies and the ensemble average forecast anomalies at the given lead and for 36 start date (1980–2015). For observational analysis (black lines in Fig. 6), we shifted the analysis from 1970 to 2005, so that for November 1, 2005, soil moisture anomalies, the observations are available at 10 years to lead time, i.e., in October 2015.

## Apportionment entropy (AE)

AE is a non-parametric measure to assess sub-seasonal to seasonal variability in the data. We applied a modified form of the AE formula taken from Konapala et al.<sup>71</sup>.

$$AE = \log 12 + \sum_{i=1}^{12} (x_i/X) \log(x_i/X) \quad (5)$$

$$X = \sum_{i=1}^{12} x_i \quad (6)$$

where  $x_i$  is the monthly precipitation or soil moisture climatology. If the total annual precipitation or soil moisture is equally distributed across the 12 months i.e.  $x_i/X = 1/12$  then  $AE = 0$ . Whereas, if all the precipitation fall in one month only, then  $AE = \log 12$ , i.e., AE ranges between 0 and  $\log 12$ , with the smaller values representing less sub-seasonal to seasonal variability and larger values representing a higher sub-seasonal to seasonal variability. We have used natural logarithm (base  $e$ ) and multiplied the AE by an arbitrary number 100 for better number representation in Fig. 5.

## DATA AVAILABILITY

The CESM DPLE (<https://www.cesm.ucar.edu/projects/community-projects/DPLE/>) and CESM LE (<https://www.cesm.ucar.edu/projects/community-projects/LENS/>) data are available from NCAR.

## CODE AVAILABILITY

The analysis codes can be made available upon reasonable request to the corresponding author.

Received: 2 October 2020; Accepted: 8 February 2021;

Published online: 12 March 2021

## REFERENCES

- Vicente-Serrano, S. M., Quiring, S. M., Peña-Gallardo, M., Yuan, S. & Domínguez-Castro, F. A review of environmental droughts: Increased risk under global warming? *Earth-Sci. Rev.* **201**, 102953 (2019).
- Mishra, A. K. & Singh, V. P. A review of drought concepts. *J. Hydrol.* **391**, 202–216 (2010).
- Ault, T. R. On the essentials of drought in a changing climate. *Science* **368**, 256–260 (2020).
- Cook, B. I., Mankin, J. S. & Anchukaitis, K. J., Climate change and drought: from past to future. *Curr. Clim. Change Rep.* **4**, 164–179 (2018).
- Asner, G. P. et al. Progressive forest canopy water loss during the 2012–2015 California drought. *Proc. Natl Acad. Sci. USA* **113**, E249–E255 (2016).
- Howitt, R., MacEwan, D., Medellín-Azuara, J., Lund, J. & Sumner, D. *Economic Analysis of the 2015 Drought for California Agriculture*. pp. 16 (Center for Watershed Sciences, University of California - Davis, Davis, CA, 2015).
- Smith, A. B. & Matthews, J. L. Quantifying uncertainty and variable sensitivity within the US billion-dollar weather and climate disaster cost estimates. *Nat. Hazards* **77**, 1829–1851 (2015).
- Hobbins, M. T. et al. The evaporative demand drought index. Part I: Linking drought evolution to variations in evaporative demand. *J. Hydrometeorol.* **17**, 1745–1761 (2016).
- Kay, J. E. et al. The Community Earth System Model (CESM) large ensemble project a community resource for studying climate change in the presence of internal climate variability. *Bull. Am. Meteorol. Soc.* **96**, 1333–1349 (2015).
- Livneh, B. & Hoerling, M. P. The physics of drought in the U.S. Central Great Plains. *J. Climatol.* **29**, 6783–6804 (2016).
- Otkin, J. A. et al. Assessing the evolution of soil moisture and vegetation conditions during the 2012 United States flash drought. *Agric. For. Meteorol.* **218**, 230–242 (2016).
- Schubert, S. D. et al. Global meteorological drought: a synthesis of current understanding with a focus on SST drivers of precipitation deficits. *J. Climatol.* **29**, 3989–4019 (2016).
- Seager, R. & Hoerling, M. Atmosphere and ocean origins of North American droughts. *J. Climatol.* **27**, 4581–4606 (2014).
- Sheffield, J. et al. A drought monitoring and forecasting system for sub-Saharan African Water Resources and Food Security. *Bull. Am. Meteorol. Soc.* **95**, 861 (2014).
- Rippey, B. R. The U.S. drought of 2012. *Weather Clim. Extrem.* **10**, 57–64 (2015).
- Hoerling, M. et al. Causes and predictability of the 2012 Great Plains drought. *Bull. Am. Meteorol. Soc.* 130819115119008 (2014) <https://doi.org/10.1175/BAMS-D-13-00055.1>.
- Deser, C. et al. Insights from Earth system model initial-condition large ensembles and future prospects. *Nat. Clim. Change* **10**, 277–286 (2020).
- Kumar, S., Newman, M., Wang, Y. & Livneh, B. Potential reemergence of seasonal soil moisture anomalies in North America. *J. Climatol.* **32**, 2707–2734 (2019).
- Newman, M. et al. The Pacific decadal oscillation, revisited. *J. Climatol.* **29**, 4399–4427 (2016).
- Lawrence, D. M. et al. The Community Land Model version 5: description of new features, benchmarking, and impact of forcing uncertainty. *J. Adv. Model. Earth Syst.* (2019).
- Lawrence, D. M. et al. Parameterization improvements and functional and structural advances in version 4 of the Community Land Model. *J. Adv. Model. Earth Syst.* **3**, 27 (2011).
- Thornton, P. E. et al. Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. *Agric. For. Meteorol.* **113**, 185–222 (2002).
- Thornton, P. E. & Rosenbloom, N. A. Ecosystem model spin-up: estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. *Ecol. Model.* **189**, 25–48 (2005).
- Keyantash, J. & Dracup, J. A. The quantification of drought: an evaluation of drought indices. *Bull. Am. Meteorol. Soc.* **83**, 1167–1180 (2002).
- Tian, L., Yuan, S. & Quiring, S. M. Evaluation of six indices for monitoring agricultural drought in the south-central United States. *Agric. For. Meteorol.* **249**, 107–119 (2018).
- McCabe, G. J., Palecki, M. A. & Betancourt, J. L. Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States. *Proc. Natl Acad. Sci. USA* **101**, 4136–4141 (2004).
- Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D. & Bacmeister, J. T. Causes of long-term drought in the US Great Plains. *J. Clim.* **17**, 485–503 (2004).
- Cole, J. E., Overpeck, J. T. & Cook, E. R. Multiyear La Niña events and persistent drought in the contiguous United States. *Geophys. Res. Lett.* **29**, 25-21-25-24 (2002).
- Hoerling, M. & Kumar, A. The perfect ocean for drought. *Science* **299**, 691–694 (2003).
- Seager, R., Kushnir, Y., Herweijer, C., Naik, N. & Velez, J. Modeling of tropical forcing of persistent droughts and pluvials over Western North America: 1856–2000\*. *J. Clim.* **18**, 4065–4088 (2005).
- Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D. & Bacmeister, J. T. Potential predictability of long-term drought and pluvial conditions in the US Great Plains. *J. Climatol.* **21**, 802–816 (2008).
- Merryfield, W. J. et al. Current and emerging developments in subseasonal to decadal prediction. *Bull. Am. Meteorol. Soc.* **101**, E869–E896 (2020).
- Amenu, G. G., Kumar, P. & Liang, X. Z. Interannual variability of deep-layer hydrologic memory and mechanisms of its influence on surface energy fluxes. *J. Climatol.* **18**, 5024–5045 (2005).

34. Bierkens, M. F. P. & van den Hurk, B. J. J. M. Groundwater convergence as a possible mechanism for multi-year persistence in rainfall. *Geophys. Res. Lett.* **34**, 5 (2007).
35. Entekhabi, D., Rodriguez-Iturbe, I. & Castelli, F. Mutual interaction of soil moisture state and atmospheric processes. *J. Hydrol.* **184**, 3–17 (1996).
36. Xia, Y. L. et al. Evaluation of multi-model simulated soil moisture in NLDAS-2. *J. Hydrol.* **512**, 107–125 (2014).
37. Bellucci, A. et al. Advancements in decadal climate predictability: the role of nonoceanic drivers. *Rev. Geophys.* **53**, 165–202 (2015).
38. Parisien, M. A. & Moritz, M. A. Environmental controls on the distribution of wildfire at multiple spatial scales. *Ecol. Monogr.* **79**, 127–154 (2009).
39. Smith, D. et al. Robust skill of decadal climate predictions. *npj Clim. Atmos. Sci.* **2**, 1–10 (2019).
40. Dirmeyer, P. A., Halder, S. & Bombardi, R. On the harvest of predictability from land states in a global forecast model. *J. Geophys. Res.* **123**, 13,111–113,127 (2018).
41. Zhu, E., Yuan, X. & Wu, P. Skillful decadal prediction of droughts over large-scale river basins across the globe. *Geophys. Res. Lett.* **47**, e2020GL089738 (2020).
42. Solaraju-Murali, B., Caron, L.-P., Gonzalez-Reviriego, N. & Doblas-Reyes, F. J. Multi-year prediction of European summer drought conditions for the agricultural sector. *Environ. Res. Lett.* **14**, 124014 (2019).
43. Becker, E., Kirtman, B. P. & Pegion, K. Evolution of the North American multi-model ensemble. *Geophys. Res. Lett.* **47**, e2020GL087408 (2020).
44. Jensen, L., Eicker, A., Stacke, T. & Dobslaw, H. Predictive skill assessment for land water storage in CMIP5 decadal hindcasts by a global reconstruction of GRACE satellite data. *J. Climatol.* **33**, 9497–9509 (2020).
45. Danabasoglu, G. et al. North Atlantic simulations in Coordinated Ocean–ice Reference Experiments phase II (CORE-II). Part II: Inter-annual to decadal variability. *Ocean Model.* **97**, 65–90 (2016).
46. Yeager, S. G., Karspeck, A. R. & Danabasoglu, G. Predicted slowdown in the rate of Atlantic sea ice loss. *Geophys. Res. Lett.* **42**, 10704–10713 (2015).
47. Yeager, S. et al. Predicting near-term changes in the Earth System: a large ensemble of initialized decadal prediction simulations using the Community Earth System Model. *Bull. Am. Meteorol. Soc.* **99**, 1867–1886 (2018).
48. Guo, Z. C., Dirmeyer, P. A. & DelSole, T. Land surface impacts on subseasonal and seasonal predictability. *Geophys. Res. Lett.* **38**, <https://doi.org/10.1029/2011gl049945> (2011).
49. Hurrell, J. et al. A unified modeling approach to climate system prediction. *B Am. Meteorol. Soc.* **90**, 1819–1832 (2009).
50. Ault, T. R. et al. A robust null hypothesis for the potential causes of megadrought in western North America. *J. Climatol.* **31**, 3–24 (2018).
51. Guo, Z. C., Dirmeyer, P. A., DelSole, T. & Koster, R. D. Rebound in atmospheric predictability and the role of the land surface. *J. Climatol.* **25**, 4744–4749 (2012).
52. Dirmeyer, P. A. et al. Model estimates of land-driven predictability in a changing climate from CCSM4. *J. Clim.* **26**, 8495–8512 (2013).
53. Kam, J., Sheffield, J. & Wood, E. F. Changes in drought risk over the contiguous United States (1901–2012): the influence of the Pacific and Atlantic Oceans. *Geophys. Res. Lett.* **41**, 5897–5903 (2014).
54. Lovenduski, N. S., Bonan, G. B., Yeager, S. G., Lindsay, K. & Lombardozi, D. L. High predictability of terrestrial carbon fluxes from an initialized decadal prediction system. *Environ. Res. Lett.* **14**, 124074 (2019).
55. Kumar, S. et al. The GLACE-hydrology experiment: effects of land–atmosphere coupling on soil moisture variability and predictability. *J. Climatol.* **33**, 6511–6529 (2020).
56. Stacke, T. & Hagemann, S. Life time of soil moisture perturbations in a coupled land–atmosphere simulation. *Earth Syst. Dyn.* **7**, 1–19 (2016).
57. Mei, R. & Wang, G. Summer land–atmosphere coupling strength in the United States: comparison among observations, reanalysis data, and numerical models. *J. Hydrometeorol.* **13**, 1010–1022 (2012).
58. Yuan, X. & Wood, E. F. Multimodel seasonal forecasting of global drought onset. *Geophys. Res. Lett.* **40**, 4900–4905 (2013).
59. Vargas Zeppetello, L. R., Battisti, D. S. & Baker, M. B. The origin of soil moisture evaporation “Regimes”. *J. Climatol.* **32**, 6939–6960 (2019).
60. Dorigo, W. A. et al. The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. *Hydrol. Earth Syst. Sci.* **15**, 1675–1698 (2011).
61. Quiring, S. M. et al. THE NORTH AMERICAN SOIL MOISTURE DATABASE development and applications. *Bull. Am. Meteorol. Soc.* **97**, 1441–+ (2016).
62. Entekhabi, D. et al. *SMAP Handbook—Soil Moisture Active Passive: Mapping Soil Moisture and Freeze/thaw from Space* (JPL Publication, Pasadena, CA, 2014).
63. Reichle, R. H. et al. Assessment of the SMAP level-4 surface and root-zone soil moisture product using in situ measurements. *J. Hydrometeorol.* **18**, 2621–2645 (2017).
64. Ghannam, K. et al. Persistence and memory timescales in root-zone soil moisture dynamics. *Water Resour. Res.* **52**, 1427–1445 (2016).
65. Alexander, M. & Deser, C. A mechanism for the recurrence of wintertime mid-latitude SST anomalies. *J. Phys. Oceanogr.* **25**, 122–137 (1995).
66. Alexander, M., Deser, C. & Timlin, M. S. The reemergence of SST anomalies in the North Pacific Ocean. **12**, 2419–2433 (1999) [https://doi.org/10.1175/1520-0442\(1999\)012<2419:TROSAI>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<2419:TROSAI>2.0.CO;2).
67. Namias, J. & Born, R. M. Temporal coherence in North Pacific sea-surface temperature patterns. *J. Geophys. Res.* **75**, 5952–& (1970).
68. Infanti, J. M. & Kirtman, B. P. Prediction and predictability of land and atmosphere initialized CCSM4 climate forecasts over North America. *J. Geophys. Res.* **121**, 12,690–612,701 (2016).
69. Kumar, S., Dirmeyer, P. A. & Kinter, J. Usefulness of ensemble forecasts from NCEP Climate Forecast System in sub-seasonal to intra-annual forecasting. *Geophys. Res. Lett.* **41**, 3586–3593 (2014).
70. Sardeshmukh, P. D., Compo, G. P. & Penland, C. Changes of probability associated with El Niño. *J. Climatol.* **13**, 4268–4286 (2000).
71. Konapala, G., Mishra, A. K., Wada, Y. & Mann, M. E. Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. *Nat. Commun.* **11**, 1–10 (2020).

## ACKNOWLEDGEMENTS

We thank Matt Newman (NOAA Physical Science Laboratory) for his discussions on this topic. This research is supported by funding from USDA-NIFA Grant Number: 2020-67021-32476 to S.K. and I.R. M.E. was partially supported by the Presidential Awards for Interdisciplinary Research at Auburn University. The CESM experiments data are obtained from NCAR, which is supported by the National Science Foundation. We would like to acknowledge high-performance computing support from Cheyenne (<https://doi.org/10.5065/D6RX99HX>) provided by NCAR’s Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

## AUTHOR CONTRIBUTIONS

S.K. conceived the idea and analysis methods. M.E. and A.P. performed the analysis. All authors contributed to the concept development and manuscript writing.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41612-021-00172-z>.

**Correspondence** and requests for materials should be addressed to S.K.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021