

Using a Simple Water Balance Framework to Quantify the Impact of Soil Moisture Initialization on Subseasonal Evapotranspiration and Air Temperature Forecasts^①

RANDAL D. KOSTER

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

SIEGFRIED D. SCHUBERT

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, and Science Systems and Applications, Inc., Lanham, Maryland

ANTHONY M. DEANGELIS

Science Systems and Applications, Inc., Lanham, Maryland

ANDREA M. MOLOD

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

SARITH P. MAHANAMA

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, and Science Systems and Applications, Inc., Lanham, Maryland

(Manuscript received 10 January 2020, in final form 27 May 2020)

ABSTRACT

Past studies have shown that accurate soil moisture initialization can contribute significant skill to near-surface air temperature (T2M) forecasts at subseasonal leads. The mechanisms by which soil moisture contributes such skill are examined here with a simple water balance–based model that captures the essence of soil moisture behavior in a state-of-the-art subseasonal-to-seasonal (S2S) forecasting system. The simple model successfully transforms initial soil moisture contents into average “forecast” evapotranspiration (ET) values at 16–30-day lead that agree well, during summer, with the values forecast by the full NASA GEOS S2S system, indicating that soil moisture initialization dominates over forecast meteorological conditions in determining ET fluxes at subseasonal leads. When the simple model’s ET anomalies are interpreted in terms of T2M anomalies, a similar conclusion is reached for T2M: soil moisture initialization explains much (about 50% in the eastern half of the continental United States) of the T2M anomaly values produced by the full GEOS S2S system at 16–30-day lead, and the T2M forecasts produced by the simple model capture about one-half of the skill attained by the full system. The simple model’s framework is particularly conducive to an analysis of uncertainty in forecasts. Drier soils are generally found to induce larger uncertainty in ET (and thus T2M) forecasts, a result linked to the functional form relating ET to soil moisture in the simple model and verified by an analysis of the ensemble spreads within the forecasts produced by the full GEOS S2S system.

1. Introduction

The ability to forecast anomalies in meteorological variables such as precipitation and air temperature can provide substantial benefits to society. Accurate forecasts can provide information with regard to, for example, upcoming water supply deficits or temperature extremes, potentially giving various socioeconomic sectors the time that they need to implement damage-mitigation

^① Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JHM-D-20-0007.s1>.

Corresponding author: Randal Koster, randal.d.koster@nasa.gov

strategies. Complex, coupled numerical models of Earth system processes have accordingly been constructed to produce meteorological forecasts at various lead times. Weather forecast systems, for example, focus on providing meteorological anomalies out to about 10 days, largely deriving their skill from the initialization of the atmosphere (e.g., [Leutbecher and Palmer 2008](#); [Zhang et al. 2019](#)). In contrast, seasonal forecast systems ([Doblas-Reyes et al. 2013](#)) extract most of their skill from the initialization and subsequent modeling of the slower components of the Earth system, such as the ocean circulation ([Shukla 1998](#)).

Subseasonal time scales (forecast leads from ~2 weeks to ~2 months), which lie between the weather and seasonal time scales, are gaining increased attention from forecast system developers (e.g., [Robertson et al. 2015](#); [Vitart et al. 2017](#); [Pegion et al. 2019](#)). Commonly cited sources of subseasonal predictability include climatic phenomena such as the Madden–Julian oscillation (e.g., [Kim et al. 2018](#)) and stratospheric impacts on the North Atlantic Oscillation (e.g., [Scaife et al. 2016](#)). Another important source of skill is soil moisture in the root zone, which has a time scale on the order of a month ([Entin et al. 2000](#); [Seneviratne et al. 2006](#))—longer than standard weather time scales and yet shorter than the time scales of ocean circulation. Through soil moisture memory, an initial soil moisture anomaly in a forecast system can be “remembered” into the forecast and can thereby affect atmospheric processes at weather, subseasonal, and perhaps even seasonal time scales (e.g., [Dirmeyer and Halder 2017](#); [Koster and Suarez 2003](#)). The ability of soil moisture to influence forecasts in this way is cited as a major motivation for recent space-based soil moisture measurement missions ([Entekhabi et al. 2010](#); [Kerr et al. 2010](#)), and it also motivates some key activities in the international World Weather Research Programme/World Climate Research Programme Subseasonal to Seasonal Prediction Project (http://s2sprediction.net/file/documents_reports/P2_Pro.pdf).

The contribution of soil moisture initialization to skill in subseasonal forecasts was addressed extensively in the Global Land Atmosphere Coupling Experiment, Part 2 (GLACE-2; [Koster et al. 2010, 2011](#); [van den Hurk et al. 2012](#)). In GLACE-2, a number of forecast systems performed two suites of hindcast experiments: one that used realistic soil moisture initialization, and the other (which was otherwise identical) using unrealistic initial soil moistures (e.g., from a randomly chosen year). By comparing the forecast skill generated in the two suites of hindcasts, the subseasonal skill derived from soil moisture initialization was isolated and quantified. GLACE-2 revealed a significant contribution of soil moisture initialization to the skill of forecasting 2-m

air temperature (T2M) during boreal summer at leads out to 45 days. The contribution to precipitation forecast skill was much more modest. GLACE-2 further demonstrated that T2M forecast skill increased when soil moisture anomalies at the start of a forecast were particularly large.

The experimental strategy underlying GLACE-2 is effective for quantifying soil moisture impacts on forecast skill in a given forecast system. What the strategy cannot provide, however, is physical insight into why the soil moisture-related skill is large in some places (or days of year) and not in others or why some soil moisture initial states may produce forecasts with smaller uncertainty. To answer such questions, we would need to understand what controls the evolution of soil moisture over the course of a forecast. How do evapotranspiration (ET), drainage, and the infiltration of precipitation modify soil moisture with time, given that the efficiencies of these fluxes are themselves a function of soil moisture content?

One simple approach to examining such soil moisture dynamics is through the characterization of a soil moisture “memory” using a calculated autocorrelation in time (e.g., [Delworth and Manabe 1988](#); [Vinnikov and Yeserkepova 1991](#); [Entin et al. 2000](#)). Such a quantification carries with it the assumption of a linear reservoir, meaning that the rate at which a soil moisture anomaly is reduced (through ET or drainage) is assumed to vary linearly with the soil moisture anomaly itself. There is certainly a benefit to characterizing memory with a single time scale, as it allows a first-order indication of how long a soil moisture anomaly will last into a forecast period and how this duration might vary between locations. [Seneviratne et al. \(2006\)](#) used such an approach to evaluate the soil moisture memory simulated in land models against that derived from in situ soil moisture measurements.

The true dynamics of soil moisture, however, are much more complicated. The ET sink for soil moisture does not vary linearly with soil moisture, instead behaving differently according to hydroclimatic regime. In the drier regime, ET rates are limited by soil moisture availability so that ET increases with increasing soil moisture; in contrast, in the wetter regime, soil moisture is not limiting, and ET is effectively insensitive to soil moisture variations (e.g., [Manabe 1969](#); [Eagleson 1978](#); [Dirmeyer et al. 2006](#)). Runoff production can also vary nonlinearly with soil moisture, with particularly high drainage rates seen when the soil is very wet (e.g., [Salvucci 2001](#)). [See [Shellito et al. \(2018\)](#) and [McColl et al. \(2017\)](#) for additional recent takes on the nuanced character of soil moisture drydown.] This nonlinearity complicates our understanding of soil moisture dynamics,

as does the joint response of ET and runoff to soil moisture variations. Indeed, it is this joint response that largely determines a land surface model's overall behavior (Koster and Milly 1997; Koster and Mahanama 2012). The hypothesis of the present paper is that by considering together the nonlinear relationships that link both ET and runoff to soil moisture, we can shed new light on our understanding of soil moisture impacts on subseasonal forecast skill.

In principle, we could analyze such effects through a careful study of the land model's governing equations. This approach, however, is made intractable by the complexity of the codes used in current forecast systems. The "average" relationships between soil moisture and both ET and runoff within these systems are in fact the net product of multiple interconnected and complex process parameterizations, and their interactions cannot be ascertained a priori with any confidence. Fortunately, the average relationships can be quantified and examined relatively easily by processing model output diagnostics (e.g., Koster et al. 2019; Koster 2015). Once the average relationships are established, their impact on soil moisture evolution can be evaluated by building a simple model around them and then using it to evolve a soil moisture state forward in time, under a range of atmospheric forcing scenarios.

We employ this approach in this paper: we analyze soil moisture and ET behavior during a subseasonal forecast by first quantifying the average relationships underlying the land surface model utilized in the full forecast system and then applying them within a simple model to time series of precipitation and net radiation representative of the forecast period. Clearly, such a simple model cannot capture all facets of the forecast system—it can never capture, for example, the ability of remote sea surface temperatures to provide forecast skill through atmospheric teleconnections. What it can do, though, is capture that part of the full forecast system specifically related to soil moisture initialization and its impacts. Capturing this specific aspect in isolation allows the simple model to illustrate how soil moisture initialization contributes to forecast skill.

Section 2 describes the approach used to estimate the average relationships and the construction of the simple model built around them. Section 3 then follows with results—with a demonstration that the simple model does capture the behavior of the more complex seasonal forecast system and that it can be used to illustrate the impact of soil moisture initialization on subseasonal forecast skill. Section 4 provides a summary and additional discussion.

2. Data and methods

In essence, the analysis below utilizes a full atmospheric reanalysis (MERRA-2; section 2a) to characterize the

behavior of both the land surface and its meteorological drivers, and it then uses this characterization as the basis for a simple water balance framework that can produce estimates of evapotranspiration and air temperature at subseasonal leads (sections 2b and 2c). The results of the simple framework's forecast simulations are in turn used to understand the behavior of a fully coupled subseasonal-to-seasonal forecast system (section 2d).

a. The MERRA-2 reanalysis

The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2), produced by the Global Modeling and Assimilation Office (GMAO) at the National Aeronautics and Space Administration's Goddard Space Flight Center (NASA GSFC), is a state-of-the-art atmospheric reanalysis that provides a comprehensive and self-consistent picture of atmospheric and land surface fields over the period 1980–present (Gelaro et al. 2017). MERRA-2 combines a wealth of conventional and satellite-based measurements (McCarty et al. 2016) together within a global atmosphere–land surface modeling environment [the Global Earth Observing System (GEOS) model; Molod et al. 2015] that serves to translate in space and time the information contained in the observations and also to impose additional physical constraints on the evolution of the meteorological fields. MERRA-2, unique among existing atmospheric reanalyses in its utilization of aerosol information (Randles et al. 2016), has been evaluated extensively against independent observational datasets (Bosilovich et al. 2015).

The present analysis uses daily-averaged MERRA-2 data on a 0.5° latitude \times 0.625° longitude grid. We use in particular the following land surface data: ET, runoff Q , and average profile soil moisture W from GMAO (2015a), precipitation P from GMAO (2015b), net radiation R_{net} from GMAO (2015a,c), and 2-m air temperature (T2M) from GMAO (2015d). The precipitation falling on the land surface during the MERRA-2 integration included corrections from rain gauge networks (Reichle et al. 2017), which are accordingly built into MERRA-2's ET and W fields.

b. The projected evapotranspiration

1) A SIMPLE MODEL FOR ESTIMATING SOIL MOISTURE EVOLUTION

The simple model used in this study is based on the equation for the daily water balance of a soil column:

$$CW_{n+1} = CW_n + P_n - ET_n - Q_n, \quad (1)$$

where C is the water-holding capacity of the soil profile, the subscript n is the day index, W_n is the average soil moisture content (degree of saturation) over the profile

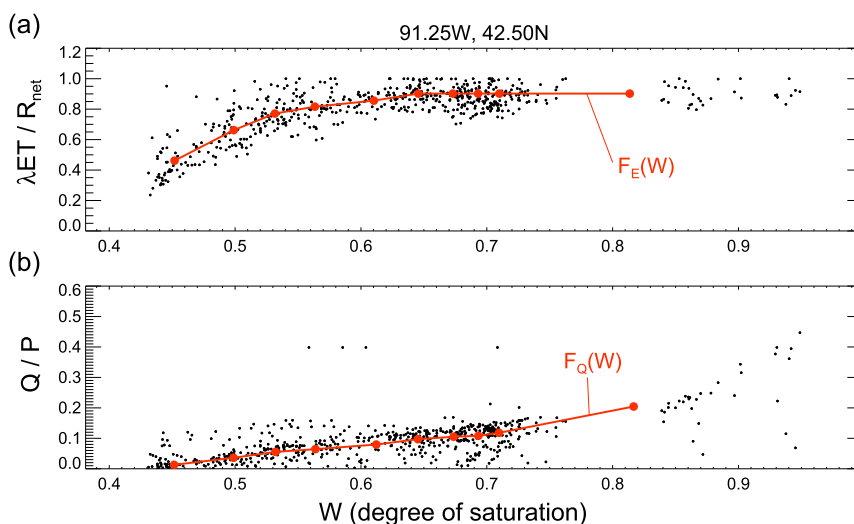


FIG. 1. The (a) evaporation efficiency or (b) runoff efficiency (y axis) plotted against the average degree of saturation in the soil profile (x axis). Each black dot represents data from MERRA-2 at a sample location for a single day between 16 Jun and 16 Jul during the period 1980–99. The red curve shows the binned fit to the data; linear extrapolation is used for soil moisture values lying outside the range captured by the red curve.

at the start of day n , and P_n , ET_n , and Q_n are the accumulated values of P , ET , and Q on day n . (The value for C in the simple model is calculated here as a bedrock depth times a nominal porosity of 0.45, with the bedrock depth distribution taken from the GMAO S2S system. Across the continental United States, the water-holding capacity varies from a few hundred mm to almost 2000 mm, as illustrated in Fig. S1 in the online supplemental material.) We now assume that both ET and Q can be estimated from the current soil moisture content and from the net radiation and precipitation forcing:

$$ET = F_E(W)R_{\text{net}}/\lambda \quad \text{and} \quad (2)$$

$$Q = F_Q(W)P, \quad (3)$$

where λ is the latent heat of vaporization. Substituting Eqs. (2) and (3) into Eq. (1) produces

$$CW_{n+1} = CW_n + P_n - F_E(W_n)R_{\text{net},n}/\lambda - F_Q(W_n)P_n, \quad (4)$$

an equation that allows soil moisture to evolve through the application of only two atmospheric forcing variables (P and R_{net}) and a predetermined description [through $F_E(W)$ and $F_Q(W)$] of the hydrological behavior of the soil column.

The function $F_E(W)$, which is the relationship between evaporation efficiency ($\lambda ET/R_{\text{net}}$) and soil moisture for a given location and day of year, is not assigned an analytical functional form. Rather, it is simply estimated

by plotting daily MERRA-2 values of W and $\lambda ET/R_{\text{net}}$ against each other on a scatterplot and then determining the piecewise linear function that connects 10 binned averages of $\lambda ET/R_{\text{net}}$. An example for a U.S. grid cell in Iowa (91.25°W, 42.5°N) is shown in Fig. 1a. Each point in the plot represents a pair of W and $\lambda ET/R_{\text{net}}$ values for a day lying within the period 15 June–16 July of some year during 1980–98. The piecewise function fitted through the points is the diagnosed form of $F_E(W)$ for 1 July. (We extract data from the 31-day window surrounding this date to increase the sample size from which the function is computed. The 10 soil moisture bins used for the averaging are defined so that each bin contains the same number of points. Such curves are recomputed for each day of the year, each time using a date-specific 31-day window.) Note that the function's shape is consistent with the idea that ET is sensitive to (increases with) soil moisture at the dry end and is relatively insensitive to soil moisture at the wet end, where it is instead controlled by energy availability (section 1). Similar curves were successfully derived from MERRA-2 data for the flash drought analysis of Koster et al. (2019), although without the normalization of ET by net radiation in that application.

Note that in tying ET to average profile soil moisture rather than to root zone moisture, we are assuming, in effect, that the average profile soil moisture represents well the average moisture in the 1-m-deep root zone. This is not a poor assumption for this particular land model, which is known to have strong coupling among its soil water states (e.g., Kumar et al. 2009). We mainly

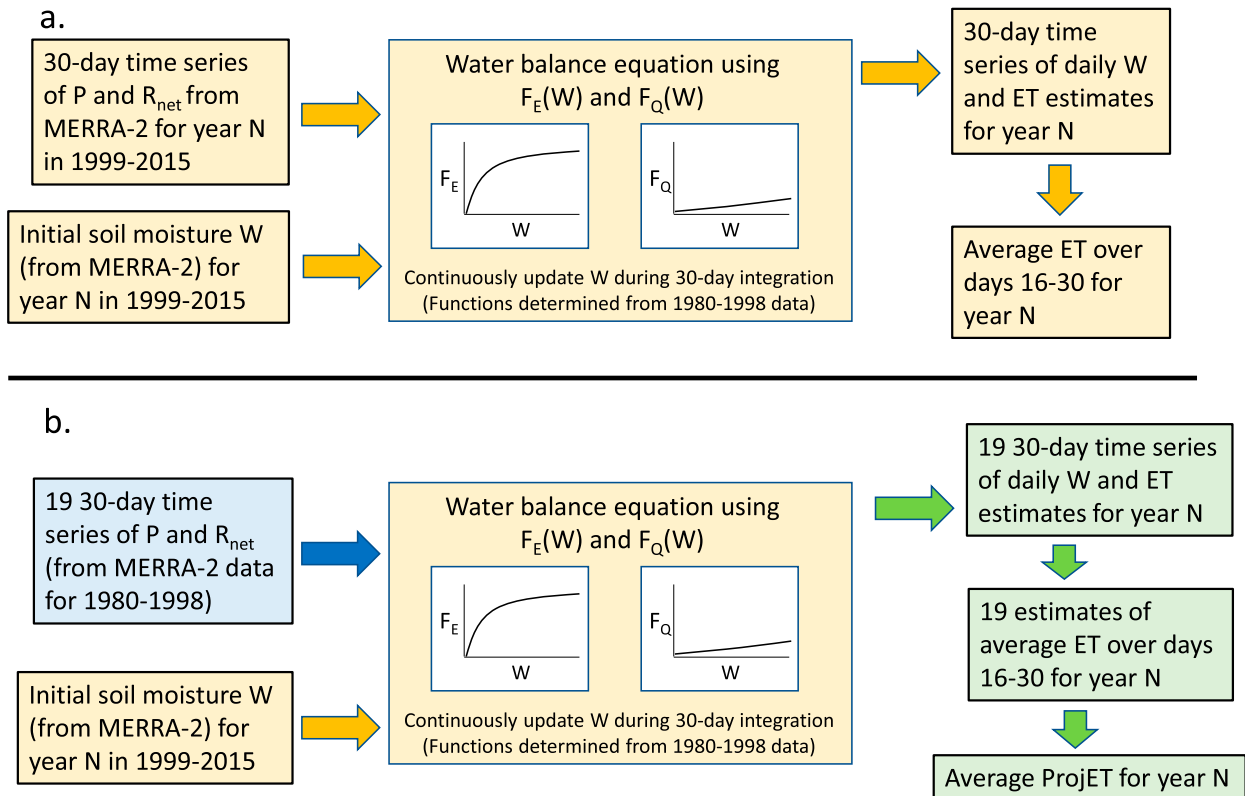


FIG. 2. Schematic of the calculation approach: (a) Approach underlying Eq. (5), allowing the calculation of ET with the simple model when the precipitation and net radiation forcing are perfectly known (from MERRA-2). (b) Approach for computing ProjET via Eq. (6). Notice that no observed meteorological data from within the forecast periods of 1999–2015 are used in the ProjET estimation procedure.

make this assumption here because MERRA-2 data holdings do not include information on moisture flux between the root zone and deeper layers, which precludes the use of the root zone as the control volume in the water balance analysis. We show below that, in any case, our use of profile soil moisture does successfully allow the effective estimation of ET when precipitation and net radiation forcing are known.

Figure 1b illustrates, for the same location and day of year, the corresponding estimation of the soil moisture–runoff ratio relationship $F_Q(W)$ from MERRA-2 data. The binned fit here captures reasonably well an average model relationship underlying the individual data points. The expected nonlinear character of the relationship (concave upward; Koster 2015) can be seen at the wet end of the soil moisture range.

Given the scatter seen in both panels of Fig. 1, it is appropriate to evaluate the effectiveness of these curves. As a simple test, we apply (4) to MERRA-2 data during each May of 1999–2015, using the framework outlined in Fig. 2a. Soil moistures were initialized on 1 May with values for that day from the MERRA-2 product, and Eq. (4) was then used with daily MERRA-2 precipitation

and net radiation data for that May to update the soil moistures over the course of the month. As this integration proceeded, we averaged the ET fluxes over 16–30 May:

$$\text{Estimated ET} = \frac{1}{15} \sum_{n=16}^{30} F_E(W_n) R_{\text{net},n} / \lambda. \quad (5)$$

We then compared the ET estimates so obtained directly with the ET product in MERRA-2 for that period.

If the approximations in Eqs. (2) and (3) are reasonable, the estimated and actual 15-day ET averages should agree. The top panel of Figure 3 shows a map, over the continental United States, of the square of the correlation coefficient r^2 between the 17 (one for each year between 1999 and 2015) 15-day ET averages estimated in this fashion and the corresponding averages computed directly from the MERRA-2 ET product. The bottom panel of Figure 3 shows the corresponding map of the RMSE of the evapotranspiration estimates relative to the MERRA-2 values. The r^2 values lie close to 1 throughout most of the conterminous United States (CONUS), and the RMSE values generally lie below

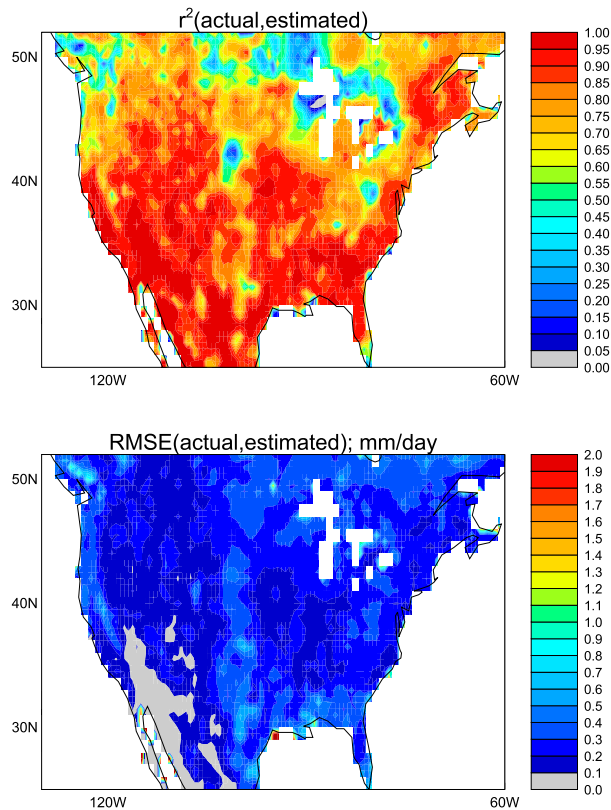


FIG. 3. (top) The square of the correlation coefficient (r^2) between 15-day-average ET values from MERRA-2 and the corresponding ET estimates produced by the simple model when the meteorological forcing (precipitation and net radiation) is itself taken from MERRA-2. The daily ET values were averaged over 16–30 May (for each year during 1999–2015); the simple model was integrated forward starting from soil moisture conditions on 1 May. Any r^2 value greater than 0.46 is significantly different from zero at the 99.9% level. (bottom) As in the top panel, but for the RMSE metric.

0.25 mm day^{-1} . Together, the two maps indicate that, despite the scatter seen in Fig. 1, Eq. (4) is indeed a viable means of estimating ET if the R_{net} and P time series are known and the underlying $F_E(W)$ and $F_Q(W)$ relationships are reasonably estimated.

2) CALCULATING THE PROJET

The “projected evapotranspiration” (hereinafter referred to as the ProjET) is an estimate of the local ET that will occur in a forecast simulation at a predefined lead. Specifically, as examined here, the ProjET is a predetermined (i.e., preforecast) estimate of the average ET (mm day^{-1}) to be experienced on days 16–30 of the forecast. We will use Eq. (4) to produce our ProjET estimates for the 1999–2015 forecast period (see section 2d) after deriving the functional forms $F_E(W)$ and $F_Q(W)$ from independent (1980–98) MERRA-2 data.

Of course, at the start of a forecast, we cannot know the daily values of R_{net} and P that will be experienced by a region. Indeed, forecasts of rainfall at subseasonal time scales are notoriously poor. We do have available, however, a suite of R_{net} and P time series that are consistent with each other and are appropriate for the region and day of year: the time series experienced by that region following the start date in question during each of the 19 years in 1980–98, the subset of MERRA-2 independent of the 1999–2015 forecast period. As indicated in Fig. 2b, we can initialize the simple water balance model with a soil moisture value for, say, 2005 (taken from MERRA-2 2005 holdings) and then integrate Eq. (4) for 30 days 19 different times, each time using a different set of R_{net} and P time series extracted from the 1980–98 period. This procedure results, using Eq. (5), in a set of 19 values of estimated ET, from which a single, average value of ProjET can be computed:

$$\text{ProjET} = \frac{1}{19} \sum_{k=1}^{19} \left[\frac{1}{15} \sum_{n=16}^{30} F_E(W_{n,k}) R_{\text{net},n,k} / \lambda \right], \quad (6)$$

where n is again the day index (covering days 16–30 of a forecast) and k is the year index for the R_{net} and P time series forcing.

This approach not only gives us an estimate (the ProjET) of the average ET occurring at a grid cell during the second 15-day period of the subseasonal forecast, it also provides an indication of the uncertainty of this estimate, as represented by the spread of the 19 contributing values in Eq. (6). (This will be explored further in section 3d.) Notice also that the ProjET estimate relies solely on the initial soil moisture, the estimated $F_E(W)$ and $F_Q(W)$ relationships, and the suite of forcing time series taken from the independent time period—no information from the forecast period itself goes into the calculation. Of course, the actual ET during that 15-day period will differ from the ProjET since it will reflect the R_{net} and P time series that actually occurred during the forecast period. Still, the ProjET can arguably be considered the best guess for this ET flux in the absence of knowledge about these forcing time series.

c. Application of ProjET to temperature anomaly estimation

Variations in ET are known to induce variations in surface air temperature (T2M), particularly in the drier regime, wherein ET is controlled by soil moisture variations rather than by variations in atmospheric demand (Seneviratne et al. 2010). The mechanism is straightforward: more evapotranspiration induces more evaporative cooling of the land surface, which translates

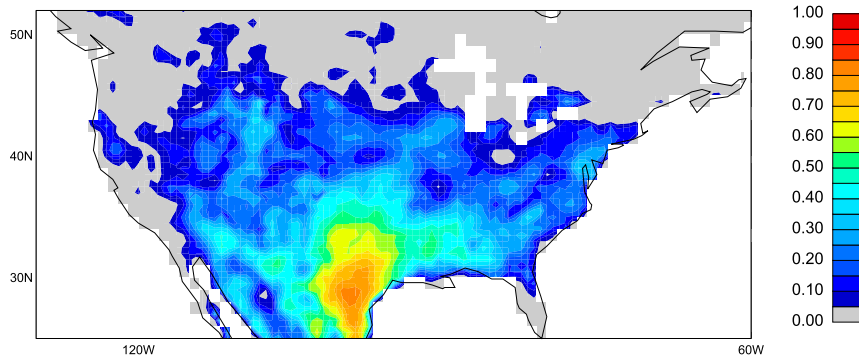


FIG. 4. Square of the correlation coefficient between 15-day-average T2M and 15-day average ET, with the 15-day averages taken from the June–August period during 1999–2015. Values are zeroed if T2M and ET are positively correlated.

to a cooling of the overlying air. Stated another way, for the land surface to experience a positive evapotranspiration anomaly ET' over a period of time Δt , it must utilize additional energy in the amount of $\lambda ET' \Delta t$, and, by the principle of energy conservation, this energy must come from somewhere—from some combination of reduced net radiation, reduced outgoing sensible heat, and the reduced storage of energy at the land surface (i.e., a reduced land surface temperature). Given that a reduced sensible heat leads to a reduced heating of the air and that the air temperature is, in any case, strongly tied to the surface temperature, the latter two reductions in particular suggest that a positive ET anomaly should contribute to a reduced near-surface air temperature relative to climatology.

Our simple framework for producing ProjET is based on a water balance equation and thus ostensibly has little to do with the estimation of temperature. As argued above, however, evapotranspiration and air temperature anomalies should be related. As a crude first approximation, we will assume a proportional relationship between the two:

$$\Delta T2M = \alpha \Delta \text{ProjET}, \quad (7)$$

where α is some unspecified negative constant [whose value, in fact, can remain unspecified given our use (see below) of the r^2 metric] and Δ refers to the anomaly relative to climatology for the given location and day of year. Nonlinearities in the energy balance equation may, of course, strongly limit the accuracy of Eq. (7), as may the fact that ProjET for a given location and day of year varies only with initial soil moisture, so that the approximation in Eq. (7) ignores other potential influences (e.g., warm incoming winds) on air temperature. Such issues, however, would presumably only reduce the effectiveness of Eq. (7).

Arguably, then, any r^2 skill seen in the comparison of α ProjET anomalies with forecast T2M anomalies despite these deficiencies can thus be considered a lower bound on the ability of soil moisture to affect forecast air temperature at subseasonal leads—presumably, an equation that more accurately captured the operating nonlinearities in the energy balance equation would only perform better.

For context, Fig. 4 shows, for June–August, the square of the correlation between 15-day averages of ET and T2M. Data for both variables are taken from MERRA-2, and the seasonal cycles of both are removed prior to computing statistics. Values are zeroed if ET and T2M are positively correlated, which would indicate (for example) that variations in net radiation are driving both atmospheric temperature variations and ET variations (a behavior relevant to the wetter regime) rather than that ET variations are driving temperature variations. The r^2 values in Fig. 4 are nonzero in the southern half of CONUS and are relatively large in the lower third, with values of about 0.6 in Texas. Figure 4 indicates where we might expect our treatment of soil moisture and ET to indeed have some relevance to subseasonal T2M forecasts.

d. GMAO subseasonal forecasts

After analyzing the character of the ProjET estimates in section 3a, we will, in section 3b, apply our findings to the analysis of full subseasonal forecasts. We examine in particular a series of forecasts (actually a series of hindcasts) produced with the GMAO's Subseasonal-to-Seasonal (S2S), version 2, prediction system (Molod et al. 2020), which consists of a fully coupled ocean–atmosphere–land–sea ice model initialized through the application of a weakly coupled Atmosphere–Ocean Data Assimilation System. Two aspects of the land model component of the forecast system [the Catchment

model of Koster et al. (2000)] are worth emphasizing for this study: (i) the Catchment model is also used for MERRA-2, so that the functional forms $F_E(W)$ and $F_Q(W)$ extracted from MERRA-2 data holdings also apply to the S2S system, and (ii) during the analysis that underlies the initialization, the precipitation applied to the Catchment model is scaled to agree with a gauge data product (Reichle et al. 2017), so that the initial soil moisture anomalies used in the forecasts are, to first order, realistic.

The particular series of forecasts (hindcasts) examined in this study are those produced by the GEOS S2S system for the multimodel SubX subseasonal project (Pegion et al. 2019). We focus here mainly on the SubX forecasts initialized on 5 June, 20 June, 5 July, 20 July, 4 August, and 19 August during each of the years between 1999 and 2015. This provides us with 102 independent warm-season (for CONUS) forecast periods, each period covering forecast days 16–30. We focus specifically on the average evapotranspiration rates and T2M values produced in each forecast, averaging their 16–30-day values across the forecast’s four ensemble members after regridding the $0.5^\circ \times 0.5^\circ$ forecast data onto the $0.625^\circ \times 0.5^\circ$ grid used for MERRA-2 diagnostic products.

e. Verification data and skill metrics

For the purposes of evaluating the SubX ET forecasts and the ProjET estimates produced with the simple modeling framework, we will utilize as “observations” the MERRA-2 ET product. In some ways this is inappropriate, since a reanalysis’s ET product is strongly influenced by the land model underlying it, and the same land model is used in both MERRA-2 and the GEOS S2S system. Thus, we do not pretend here to quantify true ET forecast skill. (Note that ET itself is rarely measured directly and that existing gridded ET datasets are themselves subject to significant uncertainties.) Despite this limitation, our ET evaluation has relevance. In effect, the MERRA-2 estimates will differ from the SubX forecasts only through MERRA-2’s application of year-specific time series of net radiation and gauge-corrected precipitation for each 30-day time period in question. Using MERRA-2 ET to evaluate SubX ET and the ProjET estimates will accordingly tell us how much of the MERRA-2 ET product is associated with meteorological forcing during that period as opposed to the soil moisture state at the beginning of that period.

The evaluation of the T2M forecasts will be more stringent. Evaluations against MERRA-2 T2M data, which, as for ET, are limited by the use of a common land model, are supplemented here with evaluations against the independent, fully observations-based CPC

temperature dataset (<https://www.esrl.noaa.gov/psd/data/gridded/data.cpc.globaltemp.html>). The CPC T2M data, provided at $0.5^\circ \times 0.5^\circ$ resolution, are generated from station observations, and these observations are suitably dense over CONUS and neighboring areas, our focus in this paper. Note, however, that only daily minimum and maximum temperatures (T_{\min} and T_{\max} , respectively) are provided in the CPC dataset; in this study, we approximate a given day’s observed average temperature as $0.5(T_{\min} + T_{\max})$.

For each evaluated field, the mean seasonal cycle was computed over the SubX period (1999–2015) and then subtracted from each year’s time series prior to computing statistics. This is done to avoid quantifying “false” skill associated with the reproduction of the seasonal cycle. We use the square of the correlation (r^2) between forecast and observed time series as our main metric of forecast skill.

3. Results

a. Sample ProjET calculation

Figure 5 provides a sample calculation for a grid cell centered in Missouri (37.5°N , 93.12°W) and for a start date of 5 July, so that ProjET is being computed for the period 20 July–3 August. As noted in Fig. 2, 19 different self-consistent sets of daily rainfall and net radiation forcing data covering 5 July–3 August (one set extracted from each year in 1980–98) are available for use in Eq. (4) to estimate ET for the target year. Figure 5a shows the 19 estimates for 20 July–3 August 2009 ET produced through the integration of Eq. (4) with these 19 sets of forcing, each integration being initialized with the known initial soil moisture on 5 July 2009. The ET estimates range from 2.7 mm day^{-1} (based on the “random” set of forcing extracted from 1980) to 5 mm day^{-1} (based on the forcing from 1992). The distribution represented by these estimates is summarized with the box-and-whisker plot in Fig. 5b, with the whiskers providing the maximum and minimum estimates, and with the edges of the box (defining the “spread”) showing the 25th- and 75th-percentile values. The mean of the 19 estimates is shown as a dot; this is the ProjET defined in Eq. (6).

Each year during 1999–2015 has its own 5 July soil moisture value, and thus a histogram like that in Fig. 5a can be produced for each year using the same 19 sets of daily forcing data. A box-and-whisker representation, as in Fig. 5b, can accordingly be derived from each year’s histogram. Figure 5c shows the box-and-whisker representation obtained in this way for each of the 17 years in the period 1999–2015, plotted against that year’s initial (5 July) soil moisture value. The impact of initial soil

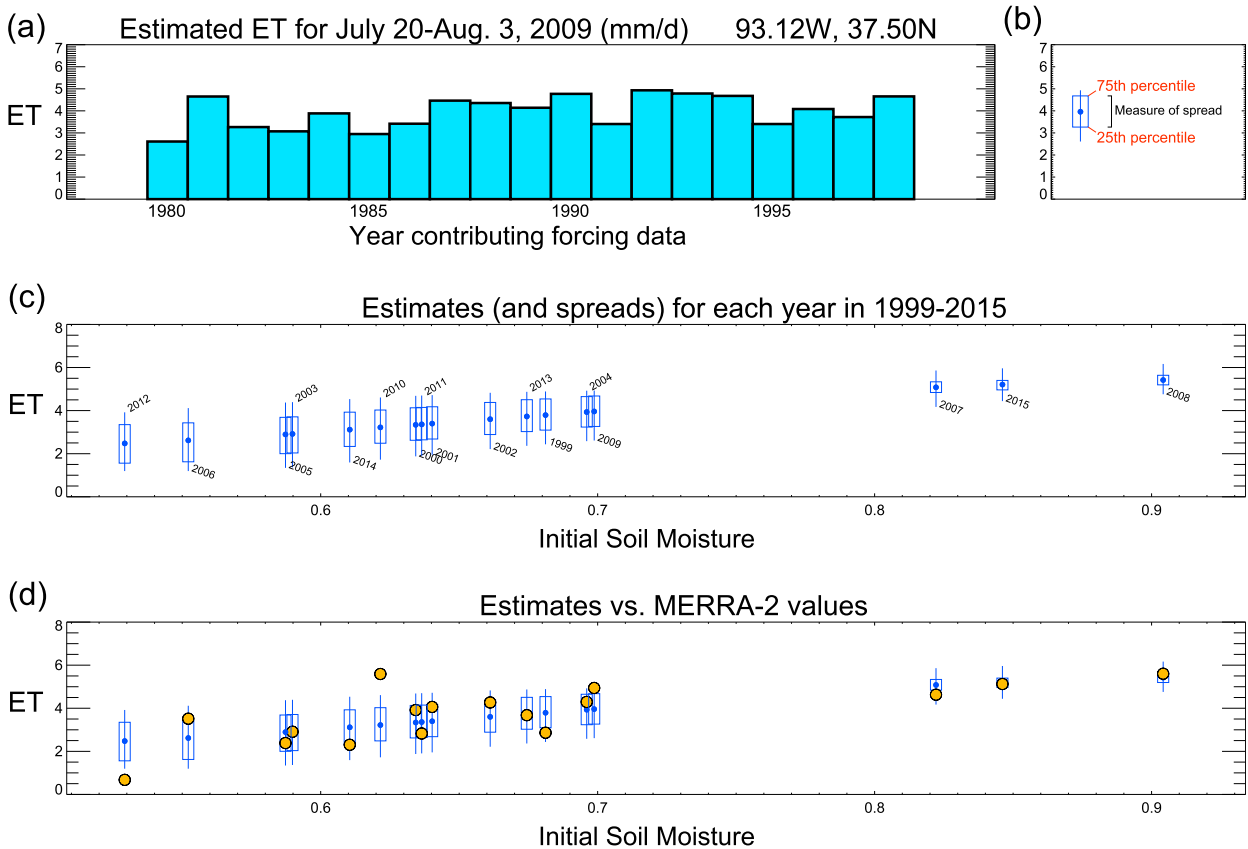


FIG. 5. (a) Histogram showing the different 16–30 Jul 2009 ET estimates for a central U.S. grid cell generated with the simple model when initialized with the cell's 5 Jul 2009 soil moisture. The different estimates arise from the different P and R_{net} forcing time series used; these time series are extracted from an independent MERRA-2 period (1980–98). (b) The histogram in (a) summarized as a box-and-whisker plot, with the line representing the full range, the box representing the 25th and 75th percentiles, and the blue dot representing the mean (i.e., the ProjET). (c) Box-and-whisker plots for the corresponding ET estimates computed for each year in 1999–2015 as a function of the soil moisture content (in degree of saturation; dimensionless) on 5 Jul of that year. (d) As in (c), but with the corresponding MERRA-2 16–30 Jul average ET values overlaid on the plot as orange dots.

moisture on the ET estimates is clear. 5 July soil moistures during 2006 and 2012 were relatively dry, and thus the ProjET (represented by the blue dots) for these years is low, just under 3 mm day^{-1} . In contrast, the 5 July soil moistures for 2007, 2008, and 2015 were relatively wet, and the ProjET for these years is accordingly higher, about 5 mm day^{-1} . As might be expected, ProjET never decreases with increasing initial soil moisture. An additional interesting feature of Fig. 5c is the sensitivity of the range of the ET estimates to the initial soil moisture. The wetter initial conditions produce a distinctly smaller range, which can be interpreted as a reduced uncertainty in the ET estimation using this framework (section 3d).

Figure 5d repeats Fig. 5c but with the addition of the “observed” 20 July–3 August ET (orange dots) for each year during 1999–2015, as extracted from MERRA-2 output. The word “observed” is in quotes because again,

the MERRA-2 estimates are effectively model-based estimates produced by the same model underlying the relationships illustrated in Fig. 1. Still, the comparison does indicate the degree to which the simple framework embodied in Eq. (4), when combined with the range of daily precipitation and net radiation time series extracted from an independent MERRA-2 time period, can capture the ET produced by the MERRA-2 system in a specific year. The results are positive: the ranges represented by the box-and-whisker plots are seen, in general, to encompass the MERRA-2 ET values. The two exceptions are 2010 and 2012, for which the MERRA-2 values lie outside the indicated ranges. A linear fit to the scatterplot (not shown) of the blue dots in Fig. 5d (the ProjET estimates) against the orange dots (the MERRA-2 values) is characterized by an r^2 of close to 0.5, implying that, for this location and day of year, the initial soil

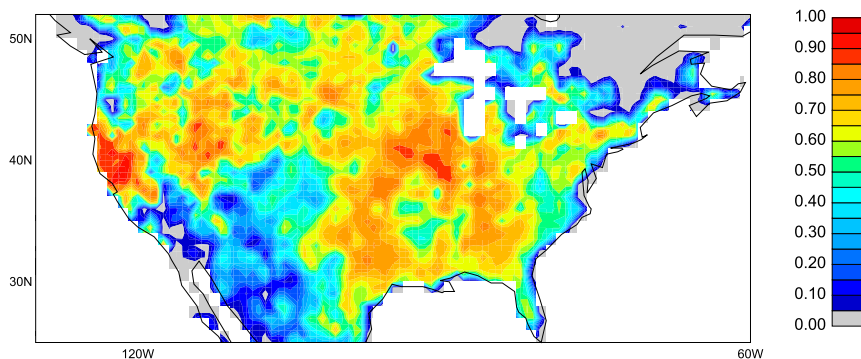


FIG. 6. The square of the correlation between the average ET produced by the full GEOS S2S forecast system over days 16–30 of the forecast and the corresponding ProjET value, which in essence varies only with the initial soil moisture. Statistics at each grid cell are computed from 102 data pairs: six 15-day periods per year (for forecasts initialized in June–August) over 17 years.

moisture explains roughly one-half of the variance of MERRA-2 ET at a lead of 16–30 days. Such agreement is in fact typical across CONUS, as illustrated in section 3b.

b. Relevance of framework to subseasonal ET forecasts

For each SubX forecast with the full GEOS S2S system (section 2d), we can calculate, at each location, the ET forecast over days 16–30. Arguably, we can also interpret ProjET, as defined in Eq. (6), as an ensemble mean forecast of ET, with 19 contributing “ensemble members” that account for 19 different ways that precipitation and net radiation might have manifested themselves during the forecast period. Note that both the SubX and ProjET forecasts make use of information contained in the initial soil moisture content to estimate ET. As a result, the SubX forecasts and ProjET estimates will differ only because the SubX forecasts also make use of information contained in forecast meteorological drivers such as precipitation, wind speed, air temperature, incoming radiation, and specific humidity. Of course, differences could also stem from the simplicity of the energy balance framework relative to the full GEOS land surface model; given Fig. 3, however, we suspect that this would not be a major issue. We expect that if SubX shows low skill in predicting meteorological drivers, the skill in the SubX ET forecasts and ProjET estimates should be similar.

We examine the agreement between the SubX ET forecasts and the ProjET estimates in Fig. 6. Results are shown for the combined 5 June, 20 June, 5 July, 20 July, 4 August, and 19 August start dates. Given the 17 years of the SubX period, this amounts to 102 data pairs going into each r^2 calculation.

The r^2 values across CONUS are high, generally above 0.5 except in the southwest. [The low values in

the southwest may reflect a low soil moisture memory in the region, a feature indicated for an earlier version of this land–atmosphere modeling system (Seneviratne et al. 2006; see the NSIPP panel in their Fig. 5).] The values across CONUS are indeed often above 0.7. Figure 6 indicates that across most of CONUS, the simple framework embodied in Eq. (4) is able to capture the majority of the ET behavior (i.e., it generally explains 50% or more of the variance in ET) at subseasonal leads produced by the much more complex and comprehensive GEOS S2S forecast system. We can thus infer that at least half (and often much more than half) of the forecast ET behavior in the full system comes from soil moisture initialization rather than from the forecast model’s ability to capture variations in the subseasonal meteorological forcing stemming, for example, from remote SST conditions, atmospheric modes like the MJO, and so on.

Whereas Fig. 6 compares the ProjET estimates with the SubX ET forecasts, Fig. 7 compares both with the MERRA-2 reanalysis ET estimates. When MERRA-2 is used as “truth,” ProjET captures essentially all of the subseasonal ET forecast skill produced by the GEOS S2S system. Apparently, at least in terms of predicting evapotranspiration, the full S2S system’s ability to predict explicit meteorological forcing over CONUS during the first month of the forecast does not give it an edge relative to consideration of initial soil moisture alone, through ProjET. The comparison of Figs. 7a and 7b speaks to the overwhelming importance of soil moisture initialization for ET forecasts.

In fact, the skill levels for ProjET often exceed those for the SubX forecasts, as indicated in Fig. 7c. Further testing indicates that the apparent higher skill levels shown for ProjET reflect sampling, i.e., the fact that the ProjET estimates are based on 19 possible manifestations

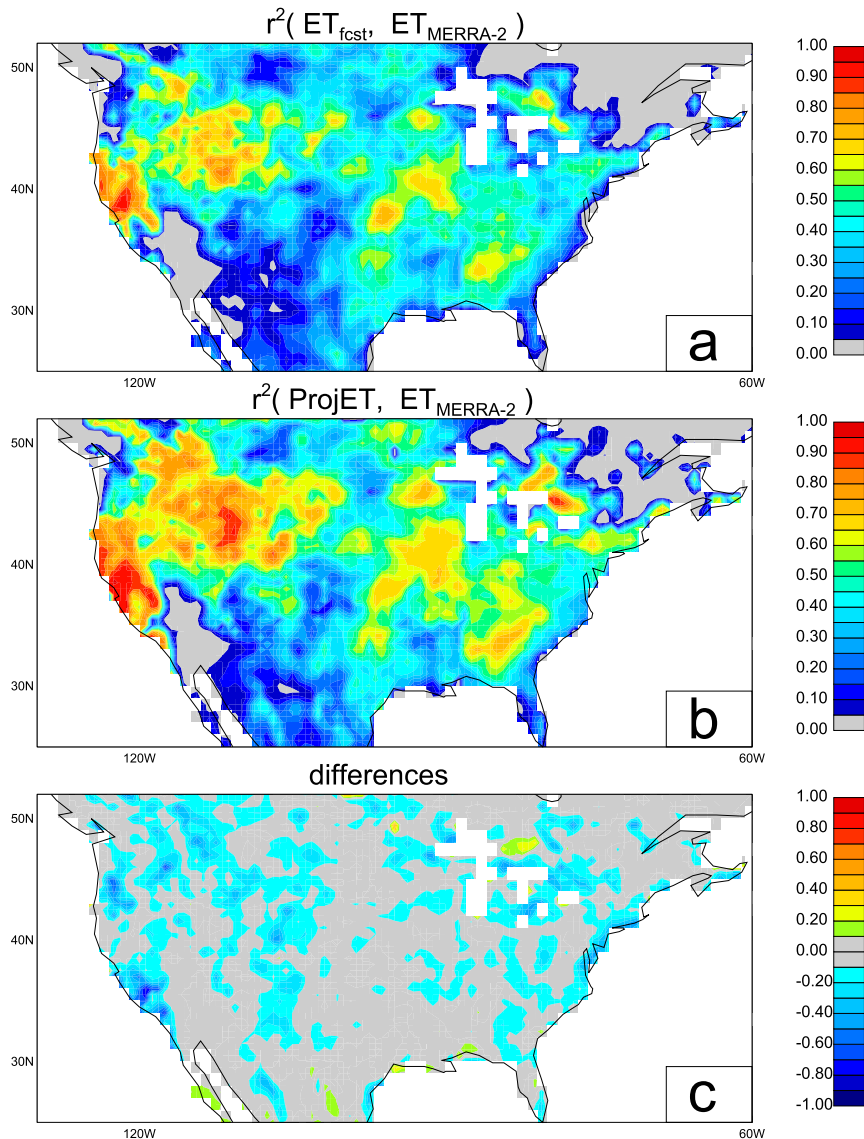


FIG. 7. (a) Square of the correlation coefficient between the average ET produced by the full GEOS S2S forecast system over days 16–30 of the forecast and the corresponding truth, as represented here by MERRA-2 ET estimates. Statistics at each grid cell are computed from 102 data pairs: six 15-day periods per year (for forecasts initialized in June–August) over 17 years. (b) As in (a), but for the comparison between ProjET and the MERRA-2 estimates. (c) Differences: (a) minus (b).

of precipitation and net radiation forcing during the forecast period whereas the SubX estimates are based on only four ensemble members. We state this because when the calculations underlying the ProjET skill in Fig. 7b are instead based on only four precipitation and net radiation time series (i.e., from four years sampled randomly from 1980 to 1998), the agreement between the two maps increases considerably (Fig. S2 in the online supplemental material), and there is even an indication that the full S2S system performs slightly better

than ProjET along the Gulf Coast. This result speaks to the importance of adequate sampling—an adequate number of ensemble members—in S2S forecasts.

c. Relevance of framework to subseasonal air temperature forecasts

Section 2c above describes our rationale for using α ProjET anomalies ($\alpha\Delta$ ProjET) as surrogates for T2M anomalies at 16–30-day lead. Figure 8 shows the square of the correlation between Δ ProjET (and thus $\alpha\Delta$ ProjET)

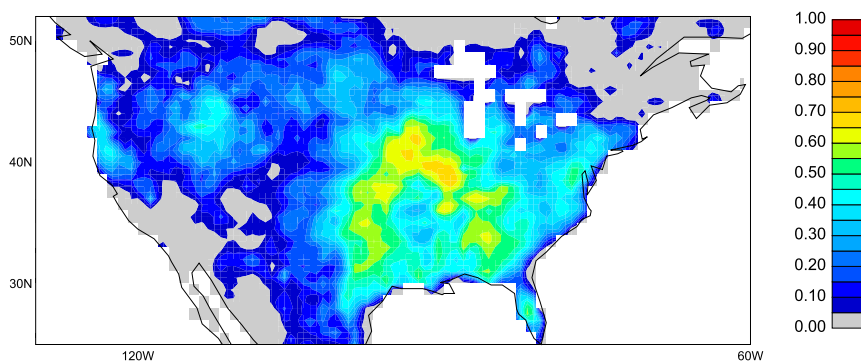


FIG. 8. Square of the correlation coefficient between the average T2M produced by the full GEOS S2S forecast system over days 16–30 of the forecast and the corresponding estimates tied to the computed ProjET values. Statistics at each grid cell are computed from 102 data pairs: six 15-day periods per year (for forecasts initialized in June–August) over 17 years.

and the corresponding forecast T2M anomalies produced by the GEOS system for SubX, based on the same 102 periods spanning June–August of 1999–2015. Higher values (>0.5) are limited to the eastern half of CONUS; in the areas with these higher values, it appears that at least half of the variance of the T2M predictions produced by the GEOS S2S system stems specifically from soil moisture initialization. Obviously, though, other factors also contribute to the T2M forecast by the full system, as very few areas in Fig. 8 show an r^2 exceeding 0.6. Such other factors include the advection of warm or cool air masses from remote regions and variations in cloudiness and associated radiative forcing. Again, a full S2S forecast system can account for such factors, whereas the simple framework embodied in Eq. (4) cannot.

How successful are the SubX forecasts in predicting observed temperature T2M, and to what extent is this skill captured by $\alpha\Delta\text{ProjET}$? The r^2 skill scores presented in Fig. 9, based on the same 102 periods, use two different sources for the T2M truth: 1) T2M from MERRA-2, and 2) T2M from the stations-based CPC dataset (section 2e). Figures 9a and 9b show a significant difference in SubX skill depending on the target, with the SubX estimates matching the MERRA-2 estimates more closely (higher r^2 values of ~ 0.3 – 0.4) than the CPC estimates (higher r^2 values of ~ 0.2). This, in fact, speaks strongly to the model dependence of reanalysis-based surface temperatures. The closer connection of the SubX results to the MERRA-2 estimates very likely reflects the use of the Catchment land model in both systems.

Figures 9c and 9d show the corresponding r^2 skill levels when the $\alpha\Delta\text{ProjET}$ values are used as the temperature forecasts. Information in the initial soil moisture content, as captured by $\alpha\Delta\text{ProjET}$, is seen to account

for a large portion of the SubX temperature forecast skill, regardless of which data source is used for truth. The comparison of Figs. 9a and 9b with Figs. 9c and 9d indeed underlines the importance of soil moisture initialization for subseasonal air temperature forecasts during the warm season. Again, though, soil moisture initialization, as important as it is, is not the only player. By subtracting the $\alpha\Delta\text{ProjET}$ skill levels (Fig. 9c or 9d) from the corresponding SubX skill levels (Fig. 9a or 9b), we obtain the residual skill in SubX T2M prediction (Fig. 9e or 9f), i.e., skill drawn from other facets of the full S2S forecast system. These residual skill distributions in fact look very similar, despite the low levels of significance for the differences. Values of about 0.1 cover much of the southeastern third of CONUS; we can speculate that this reflects the ability of SubX to account explicitly for interannual variations in the advection of warm or cool air onto the continent from the Gulf of Mexico, depending on a given year's SST there. (Such factors may very well have a greater impact on T2M than on ET.) Note that, to emphasize the residual aspect of this difference, negative differences are not highlighted in Figs. 9e and 9f. If they were, we would see that the $\alpha\Delta\text{ProjET}$ skill levels are in fact higher in some parts of the west. In such areas, some aspect of the SubX system—perhaps the small ensemble size—is degrading the ability of the initial soil moisture information to contribute to the temperature forecasts.

d. Relevance of framework to the estimation of forecast uncertainty

Figure 5c shows a clear, soil moisture-dependent distinction in the spread of the ET estimates underlying ProjET at a particular grid cell—a drier soil moisture initialization leads to a larger spread in the individual estimates than does a wetter initialization. This spread

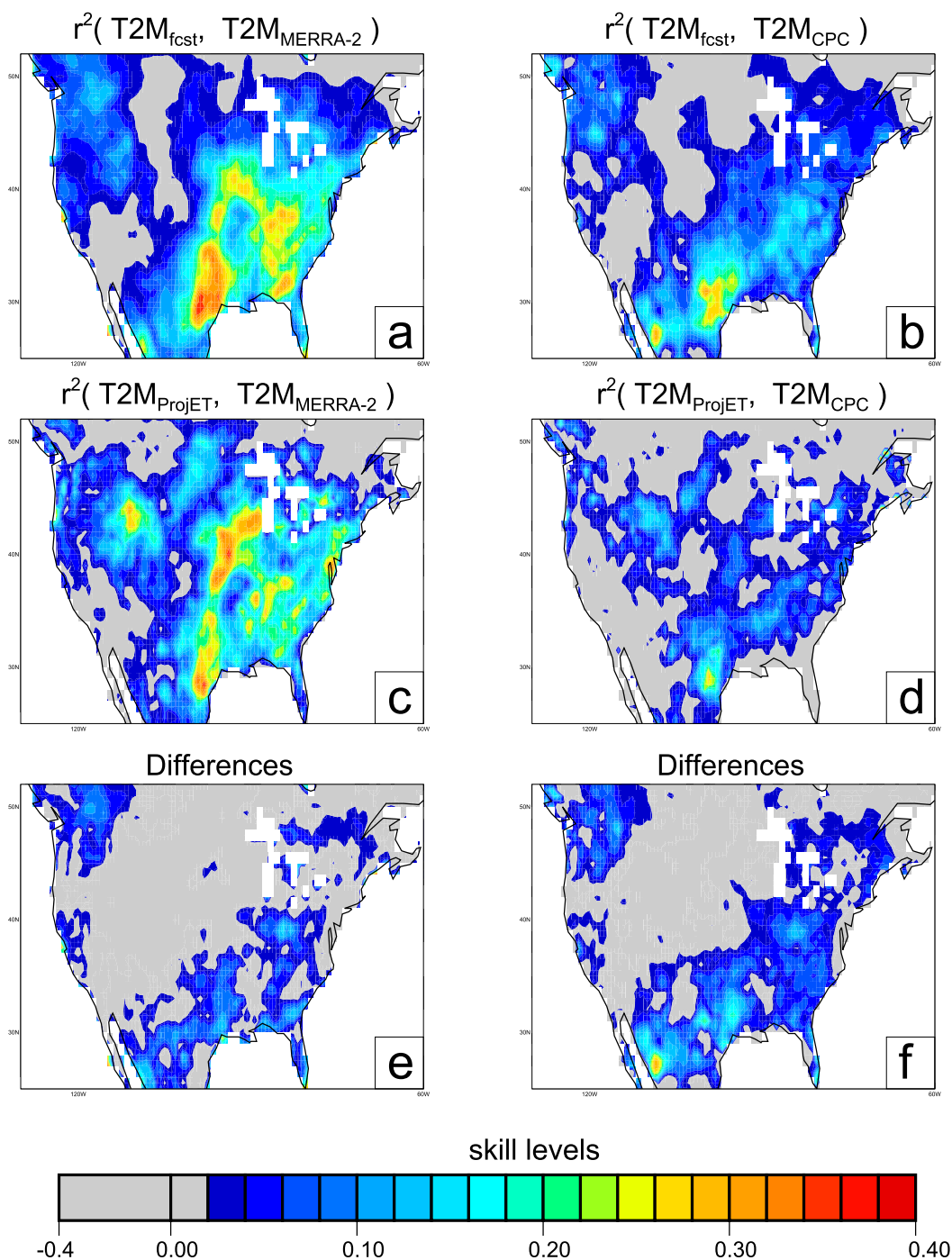


FIG. 9. (a) Square of the correlation coefficient between the average T2M produced by the full GEOS S2S forecast system over days 16–30 of the forecast and the corresponding truth, as represented here by MERRA-2 T2M estimates. (b) As in (a), but with the truth being the gridded CPC station-based T2M observations. (c) As in (a), but comparing the ProjET-based T2M estimates with the MERRA-2 T2M data. (d) As in (b), but comparing the ProjET-based T2M estimates with the gridded CPC T2M data. (e) Differences: (a) minus (c). (f) Differences: (b) minus (d). All r^2 values greater than 0.054 in (a)–(d) are significantly different from zero at the 99% confidence level based on Monte Carlo analysis.

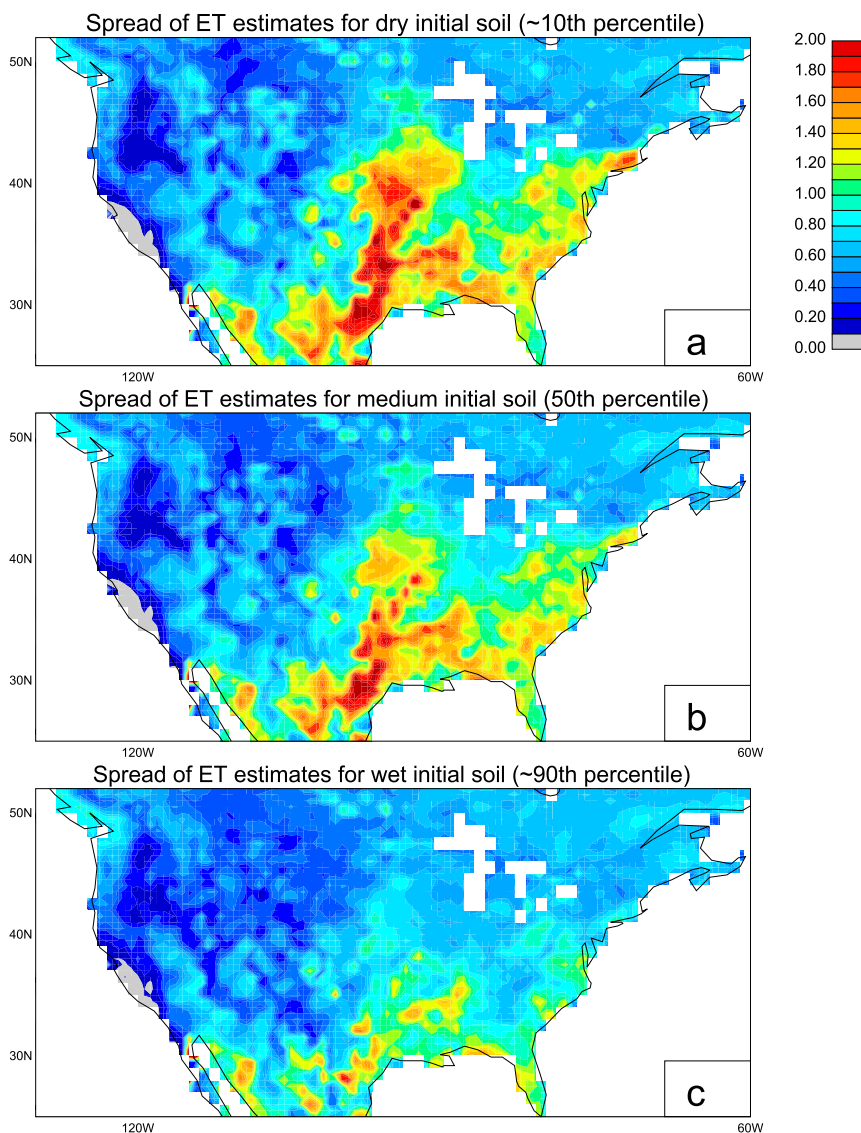


FIG. 10. Spread (see Fig. 5b for definition) of the ET estimates (mm day^{-1}) underlying ProjET for the particular year in 1999–2015 with the (a) second-to-driest, (b) median, and (c) second-to-wettest initial soil moisture content. Each plotted value is an average over six values, one for each forecast start date during June–August.

has distinct relevance for the uncertainty of the ProjET estimate and (based on the results in section 3b) for the uncertainty of the ensemble mean GEOS S2S ET forecast. A larger spread naturally implies a larger uncertainty.

A look at the spread inherent in the ProjET calculations is provided in Fig. 10. Figure 10a shows the distribution of the spread for a relatively dry soil moisture initialization. For each of the six start dates considered for the June–August period, we compute (for a given grid cell) the spread, as defined in Fig. 5b, for the second-driest (corresponding approximately

to the 10th percentile) of the 17 initializations. For example, for the 20 June start date, we find the spread for the second-driest 20 June between 1999 and 2015 at the grid cell in question. We then average the six spreads into a single value for plotting. Figure 10b shows the corresponding average spread for the initialization with the median dryness, and Fig. 10c shows the average spread for the second-wettest (approximately the 90th percentile) of the 17 initializations. The spread clearly decreases in southeastern CONUS and in the central Great Plains with increasing initial soil moisture. Sensitivity of spread to initial soil moisture is

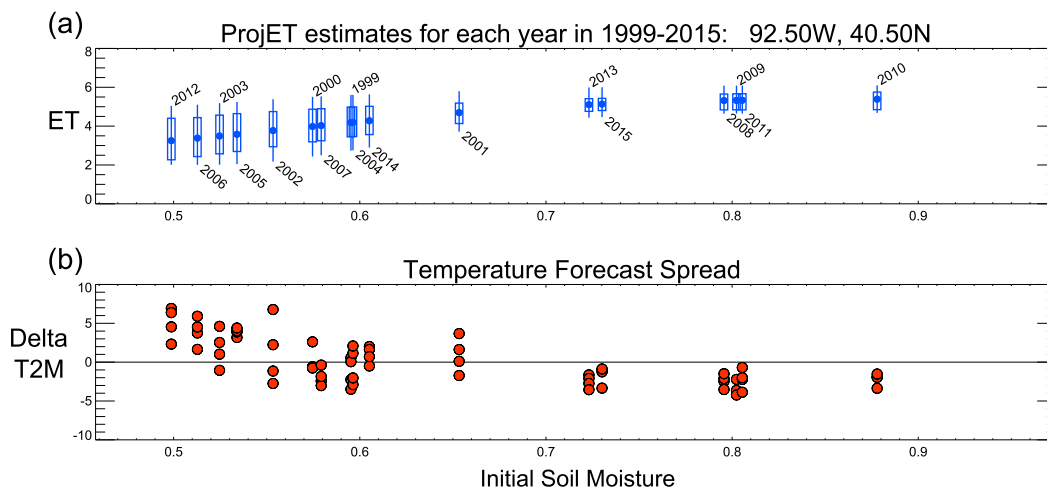


FIG. 11. (a) Box-and-whisker representation of the 20 Jul–3 Aug ET values underlying the ProjET for each year in 1999–2015, plotted against the soil moisture on 5 Jul of that year (i.e., as in Fig. 5c, but for a different grid cell). (b) T2M temperature anomalies for 20 Jul–3 Aug (for forecasts initialized on 5 Jul) as produced by the full GEOS S2S system, plotted against the soil moisture on 5 Jul. Each forecast consists of four ensemble members; the anomaly (relative to climatology) for each ensemble member is plotted separately to provide a rough estimate of the range.

also seen in the east but is relatively absent in the western half of the continent.

Why should spread tend to decrease with increasing initial soil moisture content? The simple answer involves the shape of the $F_E(W)$ relationship in Fig. 1a. Regardless of the time series of P and R_{net} forcing applied, a soil moisture initialized wet might tend, over the course of the 30-day integration of the simple model, to remain mostly in the wet regime, i.e., in the regime for which ET varies little. Different forcing time series would accordingly give similar values for the 16–30-day ET, inducing low spread. In contrast, a soil moisture initialized dry will more likely start in the drier, soil moisture–controlled evapotranspiration regime and thereby produce ET rates that respond much more strongly to the forcing time series, with rainier time series tending to produce higher ET values. The spread would accordingly be high. This idea indeed explains why the variations in Fig. 10 are maximized in the center of the continent, as this is where relatively wet and dry soil moistures for any of the six forecast periods are most likely to lie in the two very different hydrological regimes. The idea is further supported by a supplemental analysis in which the shape of the $F_E(W)$ relationship was forced to be linear—rather than the binned fit shown in Fig. 1a, linear regression was used to fit a single line through the points. Under this simplification, the variations in spread indicated in Fig. 10 were dramatically reduced (Fig. S3 in the online supplemental material).

Because ET and T2M are connected, the variations in spread seen in Fig. 10 have important implications for air temperature forecasts. Presumably, a low spread in the ET estimates underlying ProjET should be consistent

with a low spread underlying the T2M forecast produced by the full S2S system, as characterized by the range of predicted anomalies across the forecast’s ensemble members. Such a connection is illustrated in Fig. 11. Figure 11a shows, in analogy to Fig. 5c, how the spread of the 20 July–3 August ET estimates from the simple framework varies with initial soil moisture at a central U.S. grid cell. Figure 11b shows, for each year, the corresponding SubX T2M forecasts, with the individual prediction from each ensemble member shown separately. The range (across ensemble members) underlying the S2S system’s T2M forecast clearly tends to be larger when the range underlying the ProjET estimates is larger. At this location and day of year, the GEOS S2S system’s “confidence” in its subseasonal temperature forecasts appears to be a strong function of the initial soil moisture content, in a way consistent with the spread underlying the ProjET calculation.

A proper statistical demonstration of this idea must await the advent of the next generation of the GEOS S2S system, which will provide forecasts on certain dates with a much greater number of ensemble members (Molod et al. 2020). Simply put, four ensemble members is inadequate for obtaining robust statistics. Still, even with four ensemble members, we can generate some preliminary results. Figure 12 shows, for the 16–30-day period after the 5 July start date, the correlation between the spread in the ET estimates generated with the simple framework (as defined in Fig. 5b) and the range, across the ensemble members, of the T2M values produced in the GEOS SubX forecasts. The small number of ensemble members naturally induces some noise in

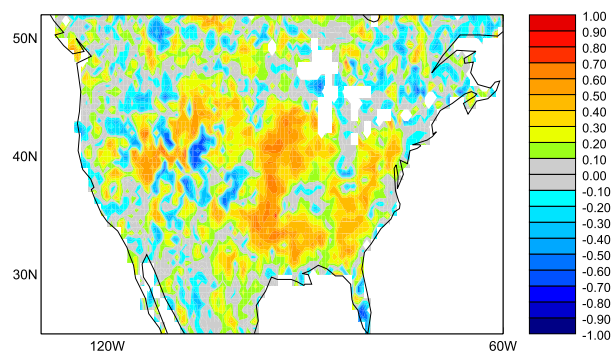


FIG. 12. Correlation, for a representative summertime period (16–30 days after the 5 Jul start date), between the spread underlying the ProjET calculation (as defined in Fig. 5b) and the range (across the four ensemble members) of the SubX T2M forecasts.

the field. Nevertheless, we see generally positive values, particularly along a north–south swath down the center of the continent, largely coincident with the area of high ET spread seen in Fig. 10a.

Together, Figs. 11 and 12 suggest that the spread underlying the ProjET calculation can tell us something about the spread in the ensemble T2M forecasts produced by the GEOS S2S system. Although more work is needed to address this problem, these preliminary calculations suggest that the simple framework could provide information on the uncertainty of the T2M forecasts—how wetter soils, for example, may lead to forecasts with reduced uncertainty.

4. Summary and discussion

We examine in this study how soil moisture influences the skill and uncertainty of the subseasonal ET and T2M forecasts produced by the GEOS S2S, version 2, forecast system. We do this by representing the soil moisture–related aspects of the full system with a simple framework built around the water balance of the soil column. The net behavior of multiple interacting, complex processes within the land model–interactions that are difficult, if not impossible, to understand a priori through an analysis of the land model’s governing equations—are effectively captured with fitted functions that relate evapotranspiration efficiency and runoff efficiency to soil moisture content (Fig. 1). Under the simple framework, these fitted functions are used in conjunction with the water balance equation to evolve a soil moisture state forward in time in response to daily P and R_{net} forcing. To produce the projected evapotranspiration, or ProjET, at a lead of 16–30 days (and thereby the corresponding T2M anomaly at this lead), we use multiple representative, but independent, time

series of P and R_{net} to generate a distribution of ET estimates for days 16–30. We then average these estimates into the single ProjET value.

The ProjET is found to capture very well the subseasonal ET values forecast by the full S2S system (Fig. 6). Because ProjET at a given location and day of year varies only with initial soil moisture, we thus conclude that the subseasonal ET forecasts produced by the S2S system are largely determined by the soil moisture initialization. Other facets of the full system, such as its ability to forecast variations in wind speed, humidity, rainfall, and so on, have an accordingly small overall impact.

The interpretation of ProjET anomalies as proxies for T2M anomalies is based on the simple idea that an anomalously high ET implies a cooling of the local land–atmosphere system; some of the absorbed radiative energy that would normally be transferred to the near-surface atmosphere through the sensible heat flux is instead used to evaporate the land moisture. The assumption in Eq. (7) that ET anomalies are directly proportional to T2M anomalies is meant to be nothing more than a crude approximation, with the potential limits to this approximation clearly indicated in Fig. 4, which shows the square of the correlation coefficient between summertime ET and T2M anomalies in MERRA-2. Still, despite its crudeness, the equation apparently does capture a true link between ET and T2M—the simple framework, through Eq. (7), captures much of the T2M behavior in the full S2S system. We indeed arrive at a conclusion for T2M analogous (though more muted) to that for ET: the subseasonal T2M forecasts produced by the full system during summer are significantly determined by the initial soil moisture content in central and eastern parts of CONUS (Fig. 8).

Proper soil moisture initialization is indeed shown to contribute substantially to summertime T2M forecast skill (Fig. 9). The fact that the skill levels are higher when the T2M validation data are taken from MERRA-2 rather than from the independent CPC station-based dataset indicates that the MERRA-2 T2M data themselves are significantly influenced by the land model parameterization. This raises an interesting question: could the CPC station-based data be used to “tune” the functional relationships $F_E(W)$ and $F_Q(W)$ in the model? That is, could we find forms of these relationships that maximize the skill levels that would be plotted in Fig. 9d, and could those tuned relationships then be used to guide further land model development, leading to a land model with a more realistic connection to air temperature? We leave such questions for future research, noting only that the benefits of such an analysis could be substantial.

Another potentially fruitful line of inquiry involves forecast uncertainty. A range of ET estimates underlies any given ProjET value produced by the simple framework, one estimate for each representative P and R_{net} time series applied. The spread in these estimates has relevance to uncertainty estimation, particularly because the spread in many locations is seen to vary with initial soil moisture content due to the nonlinear character of the soil moisture–evaporation efficiency function, $F_E(W)$. The preliminary calculation underlying Fig. 12 provides hope that, through further analyses with the simple framework, particularly in conjunction with the next generation S2S forecast system, we can begin assigning sensible, explainable uncertainty estimates to individual subseasonal T2M forecasts.

One can imagine additional studies built around the simple framework in Fig. 2b that could further elucidate S2S forecast system behavior without the expense of running a full series of S2S hindcasts. For example, the impact of water-holding capacity on the evolution of soil moisture during a forecast and on associated forecast skill could be examined through imposed modifications in the assigned water-holding capacity C ; an extremely high value of C would even allow the simple framework to mimic an S2S system that artificially persists initial soil moistures through a forecast—an approach relevant to some early systems. Potential correlations between initial soil moisture and subsequent meteorological forcing (e.g., to represent land–atmosphere feedback) could be artificially imposed on the precipitation and net radiation time series applied, allowing us to quantify the impacts of such correlations on the skill of forecast ET and air temperature. Given the importance of accurate forecasts in a number of socioeconomic sectors (with forecast ET, for example, showing increasing relevance for agricultural applications such as crop projections), a basic understanding of what controls forecast skill and its uncertainty would be invaluable. Simply put, it should help us to extract as much skill as possible from these systems. The simple framework outlined here provides a useful means for analyzing and understanding how soil moisture contributes to forecast skill at subseasonal leads.

Acknowledgments. This work was supported by the NASA MAP (NNG17HP01C and WBS802678.02.17.01.33) and NOAA MAPP (NA14OAR4310221) programs. The SubX forecasts and the MERRA-2 reanalysis were produced by the Global Modeling and Assimilation Office of NASA GSFC, which is funded by the NASA MAP program. The authors thank Yehui Chang and Hailan Wang for useful discussions and help with the data.

Data availability statement: MERRA-2 data are disseminated through the Goddard Earth Science Data and Information Services Center (GES DISC). CPC Global Temperature data were obtained online from the NOAA/OAR/ESRL Physical Sciences Division (now the Physical Sciences Laboratory) (<https://www.esrl.noaa.gov/psd/>). SubX output may be obtained from the Data Library of the International Research Institute for Climate and Society at Columbia University (<http://iridl.ldeo.columbia.edu/SOURCES/Models/SubX/>).

REFERENCES

- Bosilovich, M. G., and Coauthors, 2015: MERRA-2: Initial evaluation of the climate. NASA Tech. Memo. NASA/TM-2015-104606/Vol. 43, 145 pp., <https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich803.pdf>.
- Delworth, T. L., and S. Manabe, 1988: The influence of potential evaporation on the variabilities of simulated soil wetness and climate. *J. Climate*, **1**, 523–547, [https://doi.org/10.1175/1520-0442\(1988\)001<0523:TIOPEO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1988)001<0523:TIOPEO>2.0.CO;2).
- Dirmeyer, P. A., and S. Halder, 2017: Application of the land–atmosphere coupling paradigm to the operational Coupled Forecast System, version 2 (CFSv2). *J. Hydrometeorol.*, **18**, 85–108, <https://doi.org/10.1175/JHM-D-16-0064.1>.
- , R. D. Koster, and Z. Guo, 2006: Do global models properly represent the feedback between the land and atmosphere? *J. Hydrometeorol.*, **7**, 1177–1198, <https://doi.org/10.1175/JHM532.1>.
- Doblas-Reyes, F. J., J. Garcia-Serrano, F. Lienert, A. P. Biescas, and L. R. L. Rodrigues, 2013: Seasonal climate predictability and forecasting: Status and prospects. *Wiley Interdiscip. Rev.: Climate Change*, **4**, 245–268, <https://doi.org/10.1002/wcc.217>.
- Eagleson, P. S., 1978: Climate, soil and vegetation: 4. The expected value of annual evapotranspiration. *Water Resour. Res.*, **14**, 731–739, <https://doi.org/10.1029/WR014i005p00731>.
- Entekhabi, D., and Coauthors, 2010: The Soil Moisture Active Passive (SMAP) mission. *Proc. IEEE*, **98**, 704–716, <https://doi.org/10.1109/JPROC.2010.2043918>.
- Entin, J. K., A. Robock, K. Y. Vinnikov, S. E. Hollinger, S. Liu, and A. Namkhai, 2000: Temporal and spatial scales of observed soil moisture variations in the extratropics. *J. Geophys. Res.*, **105**, 11 865–11 877, <https://doi.org/10.1029/2000JD900051>.
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- GMAO, 2015a: tavg1_2d_Ind_Nx hourly collection. Goddard Space Flight Center Distributed Active Archive Center, accessed 27 September 2019, <https://doi.org/10.5067/RKPHT8KC1Y1T>.
- , 2015b: tavg1_2d_flux_Nx hourly collection. Goddard Space Flight Center Distributed Active Archive Center, accessed 27 September 2019, <https://doi.org/10.5067/7MCPBJ41Y0K6>.
- , 2015c: tavg1_2d_rad_Nx hourly collection. Goddard Space Flight Center Distributed Active Archive Center, accessed 27 September 2019, <https://doi.org/10.5067/Q9QMY5PBNV1T>.
- , 2015d: tavg1_2d_slv_Nx hourly collection. Goddard Space Flight Center Distributed Active Archive Center, accessed 27 September 2019, <https://doi.org/10.5067/VJAFPL1ICSIV>.
- Kerr, Y. H., and Coauthors, 2010: The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proc. IEEE*, **98**, 666–687, <https://doi.org/10.1109/JPROC.2010.2043032>.

- Kim, H., F. Vitart, and D. E. Waliser, 2018: Prediction of the Madden-Julian oscillation: A review. *J. Climate*, **31**, 9425–9443, <https://doi.org/10.1175/JCLI-D-18-0210.1>.
- Koster, R. D., 2015: Efficiency space, A framework for evaluating joint evaporation and runoff behavior. *Bull. Amer. Meteor. Soc.*, **96**, 393–396, <https://doi.org/10.1175/BAMS-D-14-00056.1>.
- , and P. C. D. Milly, 1997: The interplay between transpiration and runoff formulations in land surface schemes used with atmospheric models. *J. Climate*, **10**, 1578–1591, [https://doi.org/10.1175/1520-0442\(1997\)010<1578:TIBTAR>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<1578:TIBTAR>2.0.CO;2).
- , and M. J. Suarez, 2003: Impact of land surface initialization on seasonal precipitation and temperature prediction. *J. Hydrometeorol.*, **4**, 408–423, [https://doi.org/10.1175/1525-7541\(2003\)4<408:IOLSIO>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)4<408:IOLSIO>2.0.CO;2).
- , and S. P. P. Mahanama, 2012: Land surface controls on hydroclimatic means and variability. *J. Hydrometeorol.*, **13**, 1604–1620, <https://doi.org/10.1175/JHM-D-12-050.1>.
- , M. J. Suarez, A. Ducharme, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *J. Geophys. Res.*, **105**, 24 809–24 822, <https://doi.org/10.1029/2000JD900327>.
- , and Coauthors, 2010: Contribution of land surface initialization to subseasonal forecast skill, First results from a multi-model experiment. *Geophys. Res. Lett.*, **37**, L02402, <https://doi.org/10.1029/2009GL041677>.
- , and Coauthors, 2011: The second phase of the Global Land-Atmosphere Coupling Experiment: Soil moisture contributions to subseasonal forecast skill. *J. Hydrometeorol.*, **12**, 805–822, <https://doi.org/10.1175/2011JHM1365.1>.
- , S. D. Schubert, H. Wang, S. P. Mahanama, and A. M. DeAngelis, 2019: Flash drought as captured by reanalysis data: Disentangling the contributions of precipitation deficit and excess evapotranspiration. *J. Hydrometeorol.*, **20**, 1241–1258, <https://doi.org/10.1175/JHM-D-18-0242.1>.
- Kumar, S., R. H. Reichle, R. D. Koster, W. T. Crow, and C. Peters-Lidard, 2009: Role of subsurface physics in the assimilation of surface soil moisture observations. *J. Hydrometeorol.*, **10**, 1534–1547, <https://doi.org/10.1175/2009JHM1134.1>.
- Leutbecher, M., and T. N. Palmer, 2008: Ensemble forecasting. *J. Comput. Phys.*, **227**, 3515–3539, <https://doi.org/10.1016/j.jcp.2007.02.014>.
- Manabe, S., 1969: Climate and the ocean circulation: I. The atmospheric circulation and the hydrology of the Earth's surface. *Mon. Wea. Rev.*, **97**, 739–774, [https://doi.org/10.1175/1520-0493\(1969\)097<0739:CATOC>2.3.CO;2](https://doi.org/10.1175/1520-0493(1969)097<0739:CATOC>2.3.CO;2).
- McCarty, W., L. Coy, R. Gelaro, A. Huang, D. Merkova, E. B. Smith, M. Seinkiewicz, and K. Wargan, 2016: MERRA-2 input observations: Summary and assessment. NASA Tech. Rep. NASA/TM-2016-104606/Vol. 46, 64 pp., <https://gmao.gsfc.nasa.gov/pubs/docs/McCarty885.pdf>.
- McColl, K. A., W. Wang, B. Peng, R. Akbar, D. J. Short Gianotti, H. Lu, M. Pan, and D. Entekhabi, 2017: Global characterization of surface soil moisture drydowns. *Geophys. Res. Lett.*, **44**, 3682–3690, <https://doi.org/10.1002/2017GL072819>.
- Molod, A., and Coauthors, 2020: GEOS-S2S version 2: The GMAO high-resolution coupled model and assimilation system for seasonal prediction. *J. Geophys. Res. Atmos.*, **125**, e2019JD031767, <https://doi.org/10.1029/2019JD031767>.
- Molod, A. M., L. Takacs, M. Suárez, and J. Bacmeister, 2015: Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA2. *Geosci. Model Dev.*, **8**, 1339–1356, <https://doi.org/10.5194/gmd-8-1339-2015>.
- Pegion, K., and Coauthors, 2019: The Subseasonal Experiment (SubX): A multimodel subseasonal prediction experiment. *Bull. Amer. Meteor. Soc.*, **100**, 2043–2060, <https://doi.org/10.1175/BAMS-D-18-0270.1>.
- Randles, C. A., and Coauthors, 2016: The MERRA-2 aerosol assimilation. NASA Tech. Rep. NASA/TM-2016-104606/Vol. 45, 153 pp., <https://gmao.gsfc.nasa.gov/pubs/docs/Randles887.pdf>.
- Reichle, R. H., Q. Liu, R. D. Koster, C. S. Draper, S. P. P. Mahanama, and G. S. Partyka, 2017: Land surface precipitation in MERRA-2. *J. Climate*, **30**, 1643–1664, <https://doi.org/10.1175/JCLI-D-16-0570.1>.
- Robertson, A. W., A. Kumar, M. Peña, and F. Vitart, 2015: Improving and promoting subseasonal to seasonal prediction. *Bull. Amer. Meteor. Soc.*, **96**, ES49–ES53, <https://doi.org/10.1175/BAMS-D-14-00139.1>.
- Salvucci, G. D., 2001: Estimating the moisture dependence of root zone water loss using conditionally averaged precipitation. *Water Resour. Res.*, **37**, 1357–1365, <https://doi.org/10.1029/2000WR900336>.
- Scaife, A. A., and Coauthors, 2016: Seasonal winter forecasts and the stratosphere. *Atmos. Sci. Lett.*, **17**, 51–56, <https://doi.org/10.1002/asl.598>.
- Seneviratne, S., and Coauthors, 2006: Soil moisture memory in AGCM simulations: Analysis of Global Land-Atmosphere Coupling Experiment (GLACE) data. *J. Hydrometeorol.*, **7**, 1090–1112, <https://doi.org/10.1175/JHM533.1>.
- Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Sci. Rev.*, **99**, 125–161, <https://doi.org/10.1016/j.earscirev.2010.02.004>.
- Shellito, P. J., E. E. Small, and B. Livneh, 2018: Controls on surface soil drying rates observed by SMAP and simulated by the Noah land surface model. *Hydrol. Earth Syst. Sci.*, **22**, 1649–1663, <https://doi.org/10.5194/hess-22-1649-2018>.
- Shukla, J., 1998: Predictability in the midst of chaos: A scientific basis for climate forecasting. *Science*, **282**, 728–731, <https://doi.org/10.1126/science.282.5389.728>.
- van den Hurk, B., F. Doblas-Reyes, G. Balsamo, R. D. Koster, S. Seneviratne, and H. Camargo Jr., 2012: Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe. *Climate Dyn.*, **38**, 349–362, <https://doi.org/10.1007/s00382-010-0956-2>.
- Vinnikov, K. Ya., and I. B. Yeserkepova, 1991: Soil moisture: Empirical data and model results. *J. Climate*, **4**, 66–79, [https://doi.org/10.1175/1520-0442\(1991\)004<0066:SMEDAM>2.0.CO;2](https://doi.org/10.1175/1520-0442(1991)004<0066:SMEDAM>2.0.CO;2).
- Vitart, F., and Coauthors, 2017: The Subseasonal to Seasonal (S2S) prediction project database. *Bull. Amer. Meteor. Soc.*, **98**, 163–173, <https://doi.org/10.1175/BAMS-D-16-0017.1>.
- Zhang, F., Y. Q. Sun, L. Magnusson, R. Buizza, S.-J. Lin, J.-H. Chen, and K. Emanuel, 2019: What is the predictability limit of midlatitude weather? *J. Atmos. Sci.*, **76**, 1077–1091, <https://doi.org/10.1175/JAS-D-18-0269.1>.