The Impact of Noah-MP Physical Parameterizations on Modeling Water Availability during Droughts in the Texas–Gulf Region

WEN-YING WU,a ZONG-LIANG YANG,a AND MICHAEL BARLAGEb

a Department of Geological Sciences, Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas
b Research Applications Laboratory, National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 29 July 2020, in final form 3 March 2021)

ABSTRACT: Texas is subject to severe droughts, including the record-breaking one in 2011. To investigate the critical hydrometeorological processes during drought, we use a land surface model, Noah-MP, to simulate water availability and investigate the causes of the record drought. We conduct a series of experiments with runoff schemes, vegetation phenology, and plant rooting depth. Observation-based terrestrial water storage, evapotranspiration, runoff, and leaf area index are used to compare with results from the model. Overall, the results suggest that using different parameterizations can influence the modeled water availability, especially during drought. The drought-induced vegetation responses not only interact with water availability but also affect the ground temperature. Our evaluation shows that Noah-MP with a groundwater scheme produces a better temporal relationship in terrestrial water storage compared with observations. Leaf area index from dynamic vegetation is better simulated in wet years than dry years. Reduction of positive biases in runoff and reduction of negative biases in evapotranspiration are found in simulations with groundwater, dynamic vegetation, and deeper rooting zone depth. Multiparameterization experiments show the uncertainties of drought monitoring and provide a mechanistic understanding of disparities in dry anomalies.

KEYWORDS: Drought; Water budget/balance; Remote sensing; Land surface model; Model evaluation/performance; Model output statistics

1. Introduction

A drought is a prolonged period of abnormally low precipitation over land. Droughts are the second most costly U.S. weather and climate disasters and have frequently occurred in the southern plains, where there are many agriculture and livestock assets (Smith 2020).

The 2011 southern plains drought affected Texas, Oklahoma, and nearby states, incurring damage estimated to be about $14 billion and was responsible for 95 deaths (Smith 2020). The drought was shown to be caused by La Niña, which is associated with cooler-than-average sea surface temperature in the central and eastern tropical Pacific Ocean (Hoerling et al. 2013). The intensification of dryness was led by increases in convective inhibition that suppressed convection (Fernando et al. 2016). Evidence shows the 2011 Texas drought was accompanied by decreased terrestrial water storage from satellite datasets and modeling results (Long et al. 2013). Suppressed vegetation activity was found by satellite-observed solar-induced chlorophyll fluorescence, and its spatial pattern is similar to the drought intensity map from the U.S. Drought Monitor (USDM) (Sun et al. 2015).

Many efforts have been made for drought monitoring, such as the USDM (Svoboda et al. 2002) and the North America Land Data Assimilation System (NLDAS) Drought Monitor (Xia et al. 2012). NLDAS utilizes multiple land surface models (LSMs) to simulate the hydrometeorological fluxes and states over the conterminous United States (CONUS). Studies have suggested drought monitoring using LSMs could be improved by introducing diagnostic and statistical indices (Anderson et al. 2013; Hao et al. 2016; Mo et al. 2011; Xia et al. 2014), calibrating parameters (Arsenault et al. 2018; Santanello et al. 2013), and integrating data assimilation (Kumar et al. 2019, 2014; Li et al. 2019).

Despite the above efforts, modeling water availability during drought remains challenging. Numerous physical processes control soil moisture variability. Several studies show that disagreement of modeled soil moisture gets larger during its dry-down period (Dirmeyer et al. 2006; Niu et al. 2011) because of modeling disparities in the rates of infiltration, evapotranspiration, and runoff. The drying rates also differ at different spatiotemporal scales (McColl et al. 2019). Therefore, it is important to examine the physical process that controls water availability during drought to understand the mechanisms causing the disparities.

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JHM-D-20-0189.s1.

Corresponding author: Zong-Liang Yang, liang@jsg.utexas.edu

DOI: 10.1175/JHM-D-20-0189.1

© 2021 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
Studies have demonstrated the role of groundwater in modulating land surface fluxes and vegetation during droughts (Barlage et al. 2015; Maxwell and Kollet 2008; Fan et al. 2017) and in the partitioning of evapotranspiration (Maxwell and Condon 2016). Vegetation phenology can affect the land surface energy budget, and accounting for vegetation growth and groundwater in coupled land surface and climate models could improve seasonal precipitation forecasts over the central United States in summer (Jiang et al. 2009).

The goal of this study is to understand the impact of the major processes represented in the state-of-the-art Noah-MP LSM on capturing the southern plains drought. Section 2 contains a description of the Noah-MP model, the experimental design, observational datasets, and the study region. In section 3, the impact of groundwater and vegetation on water variability is assessed, and the simulation results are evaluated. Finally, results are summarized, and the study limitations and implications are discussed in section 4.

2. Model, data, and study region

a. Noah-MP

Noah-MP is an LSM with multiple options for representing physical processes (Niu et al. 2011; Yang et al. 2011). It supports a vegetation canopy layer, multiple snow and soil layers, and an optional unconfined aquifer layer for groundwater. Noah-MP is one of the LSMs in the Weather Research and Forecasting (WRF) Model (Skamarock et al. 2008) used for weather and regional climate forecasts. Noah-MP is also used in the National Water Model (NWM; Gochis et al. 2020), which forecasts real-time streamflow conditions.

In this study, we used an offline 2D Noah-MP configuration based on the High-Resolution Land Data Assimilation System (HRLDAS) as implemented in WRF, version 3.8. Hourly atmospheric forcing data (precipitation, air temperature, humidity, surface pressure, wind speed, and surface radiation) were taken from the NLDAS from the National Aeronautics and Space Administration (NASA; Xia et al. 2009). The simulations are conducted for CONUS with a horizontal grid spacing of 0.125°, which is the resolution of the NLDAS-2 forcing data. Land cover is defined using the Modified International Geosphere-Biosphere Programme (IGBP) Moderate Resolution Imaging Spectroradiometer (MODIS) 20-category vegetation (see Fig. 1 in the online supplemental material). All the parameters used are the same as default and not calibrated. The soil layer configuration used in this study is the same as the default Noah-MP, which contains four layers, and the soil thickness is 0.1, 0.3, 0.6, and 1.0 m from the surface to the bottom, respectively.

The simulations were conducted from 1997 to 2017. For the simulations with groundwater, the spinup was achieved by looping through 1997 for 50 times to ensure groundwater reaches equilibrium. The period of 2002–17 was used for analysis.

b. Noah-MP experimental design

To distinguish the physical processes that are important for drought modeling, eight Noah-MP experiments were conducted with a different combination of tests to investigate the contributions of groundwater, leaf area index (LAI), and plant rooting depth (experiments list in Table 1 and other model configurations list in supplementary Table 1).

Groundwater is an essential water resource in Texas. A previous study (Long et al. 2013) estimated groundwater storage changes account for 8%–16% of terrestrial water storage (TWS) decline during the 2011 drought in Texas. For investigating the contribution of groundwater, two runoff and groundwater parameterization options were used: 1) TOPMODEL with a simple unconfined groundwater aquifer [simple groundwater model (SIMGM); Niu et al. 2007] and 2) free drainage from the bottom soil layer (Schaake et al. 1996) (see appendix A for details). SIMGM is a bucket-type groundwater model that accounts for groundwater recharge and discharge. An unconfined aquifer is designed beneath the soil column, which means the water table depth does not go into soil layers, but the groundwater implicitly interacts with soil moisture through groundwater recharge. For elucidating the role of vegetation phenology during drought, three vegetation modeling approaches were used: 1) dynamic vegetation with default parameters (DyVeg), 2) prescribed, but transient, monthly leaf area index (TranLAI), and 3) prescribed climatological LAI, which is without interannual variability (ClimLAI) (Fig. 1). Similar to what is used in NLDAS and NWM, the prescribed climatological LAI is a common modeling approach in many operational systems. In this study, an 18-yr-average (2000–17) monthly LAI was composited using satellite-based LAI products (section 2c(4)). In TranLAI, both seasonal cycles and interannual variability are considered for better representing vegetation response during drought. For both TranLAI and

<table>
<thead>
<tr>
<th>Experiment name</th>
<th>Vegetation</th>
<th>Runoff option</th>
<th>Root zone</th>
</tr>
</thead>
<tbody>
<tr>
<td>TranLAI_GW</td>
<td>Prescribed transient LAI</td>
<td>SIMGM</td>
<td>0–1 m</td>
</tr>
<tr>
<td>ClimLAI_GW</td>
<td>Prescribed climatological LAI</td>
<td>SIMGM</td>
<td>0–1 m</td>
</tr>
<tr>
<td>DyVeg_GW</td>
<td>Dynamic vegetation</td>
<td>SIMGM</td>
<td>0–1 m</td>
</tr>
<tr>
<td>DyVeg_DRt_GW</td>
<td>Dynamic vegetation</td>
<td>SIMGM</td>
<td>0–2 m</td>
</tr>
<tr>
<td>TranLAI</td>
<td>Prescribed transient LAI</td>
<td>Free drainage</td>
<td>0–1 m</td>
</tr>
<tr>
<td>ClimLAI</td>
<td>Prescribed climatological LAI</td>
<td>Free drainage</td>
<td>0–1 m</td>
</tr>
<tr>
<td>DyVeg</td>
<td>Dynamic vegetation</td>
<td>Free drainage</td>
<td>0–1 m</td>
</tr>
<tr>
<td>DyVeg_DRt</td>
<td>Dynamic vegetation</td>
<td>Free drainage</td>
<td>0–2 m</td>
</tr>
</tbody>
</table>
CliLAI, the prescribed LAI varies spatially and is not tied to vegetation classification. This is unlike the default Noah-MP, in which LAI only depends on months and land-cover types through a vegetation table. In DyVeg, LAI is predicted by a dynamic leaf model (Niu et al. 2011; Dickinson et al. 1998) with Ball–Berry-type stomatal resistance. The vegetation cover fraction is static as prescribed in the model (see appendix B for details).

Different plant rooting depths with dynamic vegetation were also tested. Over 78% of the Texas–Gulf region is covered by grassland and cropland, according to the land-cover datasets used in this study. The Noah-MP default plant rooting depth of grassland and cropland is 1 m (bottom of the third layer). However, studies (Fan 2015; Fan et al. 2017) have suggested that the estimated maximum plant rooting depth is 5.6 m in this region (average in supplementary Fig. 2) with large spatial heterogeneity. Therefore, the rooting depth of all land-cover types was increased to the bottom of the soil layer, which is 2 m in the experiments of DyVeg_DRt_GW and DyVeg_DRt.

c. Datasets for evaluation

1) GRACE

The GRACE mission (Tapley et al. 2004) was a collaboration between NASA and the German Aerospace Center. Using the changes in the distance between two satellites, GRACE had been making accurate measurements of Earth’s gravity field anomalies since its launch in March 2002. The accelerometers located on each satellite’s center of mass measured the nongravitational accelerations, providing highly accurate maps of monthly changes in TWS over the land (Swenson et al. 2003) and the ocean (Chambers et al. 2004). Previous studies have compared GRACE observations of TWS to in situ observations (Rodell et al. 2004; Swenson et al. 2006) with reasonable agreement. In this study, the Center for Space Research (CSR) GRACE RL06 Mascon Solutions (Save 2019; Save et al. 2016) were used to evaluate the model.

To compare modeled TWS with GRACE, the modeled TWS was calculated using

\[
\text{TWS} = \text{SWE} + \text{CANS} + \text{SMS} + \text{GWS},
\]

where SWE is the snow water equivalent, CANS is the canopy water storage, SMS is the soil water storage from soil moisture in all four soil layers, and GWS is the groundwater storage. GWS is included only when the groundwater scheme is used. Although modeled SWE and CANS in the study region are small and may be neglected, we still include them in our estimation. GRACE data are anomalies relative to the average of 2004–09. To be comparable to GRACE observations, the average of 2004–09 was removed to calculate the anomalies of TWS. GRACE data were not always available due to instrument issues; therefore, only available months during 2002–17 were shown and used for comparison.

2) FLUXCOM ET

To evaluate modeled temporal variations of latent heat flux or evapotranspiration (ET), the FLUXCOM database (http://www.fluxcom.org/) was used because it integrated remote sensing data from MODIS, FLUXNET eddy covariance tower with machine learning methods (Jung et al. 2019; Tramontana et al. 2016) to a global gridded product. The latent heat flux data were rescaled from 0.5° to 0.125° resolution (to be the same as the model resolution) with bilinear interpolation to calculate the regional average. Here we used the monthly data from 2002 to 2013. The latent heat flux (MJ m\(^{-2}\) day\(^{-1}\)) is converted to evapotranspiration (mm day\(^{-1}\)) by 2.5104 \times 10^6 \text{ J kg}^{-1}.

3) USGS RUNOFF

Monthly runoff data for the Texas–Gulf region (area-averaged runoff) from the U.S. Geological Survey (USGS) were used for evaluating modeled river flow, assuming all surface and subsurface model water eventually makes to the river system (Seaber et al. 1987). The product was derived from flow data from the USGS river gauges with the area-weighted method (Brakebill et al. 2011). Data were obtained from the USGS WaterWatch website (https://waterwatch.usgs.gov/).

4) GLASS LAI

To evaluate the dynamic vegetation scheme, the Global Land Surface Satellites (GLASS) LAI, version 4, product (Xiao et al. 2016), which is based on MODIS surface-reflectance, was used. The data were upscaled from 8-day and 1-km to monthly and 0.125°. These data were also used for the prescribed LAI simulations (Fig. 1).
d. Texas–Gulf region

Though the simulation is conducted for the whole CONUS, our study region is the Texas–Gulf region, including several rivers flowing from northwest to southeast with final drainage into the Gulf of Mexico. The studied region is the watershed boundary of the USGS two-digit hydrologic unit code (HUC-2) of 12 with approximately 471 000 km$^2$ and includes in the monitoring region of the West Gulf River Forecast Center. The mean annual temperature is 20.12°C, and the mean rainfall is 807.8 mm yr$^{-1}$, calculated from the NLDAS-2 meteorological forcing data from 1979 to 2019. There are two major rainfall seasons in the Texas–Gulf region, one from April to May and the other from September to October (Fig. 2). According to the NLDAS forcing, the mean temperature of June–August 2011 (JJA) is 31.8°C, which was 2.8°C higher than the climatology. Mean rainfall is 23.6, 48.7 mm month$^{-1}$ lower than the climatology.

Using the USDM record from 2000, the Texas–Gulf region experienced a 5-yr-long drought from 2010 spring to 2015 summer (Fig. 3a). The year 2011 is the most severe year with the largest drought area. The peak of the drought occurred on 4 October 2011, with 93% of the area under “exceptional drought.” The USDM is a weekly updated map showing drought conditions calculated from several variables (e.g., several indexes, modeled soil moisture percentiles, streamflow conditions). It should be noted that USDM is not equivalent to any modeled variable of water storages or fluxes.

3. Results

a. 2011 Texas drought

GRACE TWS shows large interannual variability in the Texas–Gulf region (Fig. 3b). Starting from 2010 summer, GRACE TWS anomalies decline and reach a minimum of −17.9 cm in September 2011. The negative anomalies sustained for five years until 2015 spring, which ended with the 2015 Memorial Day flood and several flood events afterward. Time series from GRACE have a good agreement with the weekly USDM (Fig. 3a), showing that TWS from GRACE could be a good indicator of drought conditions in this region.

 Compared to GRACE, the Noah-MP simulations capture the temporal pattern reasonably. However, different simulations show distinct dry anomaly magnitudes during the 2010–15 drought. In particular, modeled water storages from different experiments diverge in the dry-down period (Fig. 3b), which indicates the challenges of modeling water availability during drought. This disagreement (or divergence across experiments) during the dry-down period is due to different parameterizations, resulting in different rates of vertical redistribution, drainage, and evapotranspiration (Dirmeyer et al. 2006). Then,
the infiltration from ensuing rainfall events leads to an agreement (or convergence) among the simulations. Overall, the simulations with groundwater (blue lines in Fig. 3b) result in larger dry anomalies than simulations with free drainage (orange and red lines in Fig. 3b) in the 2011 drought.

The spatial patterns of TWS anomalies from GRACE and modeled simulations are similar (supplementary Figs. 3 and 4), which show a widespread negative anomaly in the entire region, with a larger anomaly in southeastern Texas in 2011 summer.

b. Contribution of TWS components during the 2011 drought

Adding groundwater was found to be the dominant factor affecting the TWS simulations by comparing the experiments (comparing Fig. 4 top and bottom panels) in the 2011 drought. Overall, the simulations, including a groundwater aquifer, show larger dry anomalies and are closer to GRACE observations because the additional groundwater storage provides an extra source of water storage anomalies. This implies that groundwater or 1–2-m soil moisture drawdown plays an important role in TWS changes during a severe drought in the Texas–Gulf region. Further analysis was conducted to understand the contribution of different water storages to seasonal TWS in 2011. Figure 4 shows the contribution of upper soil moisture (0–1 m), bottom soil moisture (1–2 m), and groundwater in bar charts, indicating that these three components are all important for TWS anomalies during the 2011 drought. Canopy water storage and snow water storage are too small to be visible in Fig. 4 and can be neglected when computing modeled TWS in Texas. Only experiments with dynamic vegetation schemes (DyVeg, DyVeg_DRT, DyVeg_GW, and DyVeg_DRT_GW) can simulate the observed subseasonal variation of TWS: negative TWS anomalies in summer larger than autumn anomalies in 2011 Texas drought (Figs. 4a,d,e,h,i) but it might be due to phenology being modeled incorrectly (see Fig. 1). These are the results only for the 2011 Texas drought and might not apply to other events and other regions.

The results from the experiments with deeper rooting depth (DyVeg_DRT and DyVeg_DRT_GW) show more contribution from 1–2-m soil moisture than from 0 to 1 m, indicating root water uptake plays a role in the vertical redistribution of soil water. In other words, vegetation dynamics might play an essential role in water availability during drought.

c. Role of vegetation anomalies

To examine the role of vegetation anomalies in water availability, TranLAI was compared to ClimLAI. Figure 5 shows the differences (TranLAI minus ClimLAI) in the experiments in
the Texas–Gulf region during the 2011 summer. Changes are of consistent sign but varying magnitude across Texas. LAI in 2011 JJA is much lower than the climatology as shown in Fig. 5a. Less vegetation leads to less transpiration (Fig. 5e). As a result, less root water uptake leaves more soil moisture behind. This has been well recognized by the LSM community (e.g., Kumar et al. 2019). It can explain why experiments with prescribed climatological LAI have a larger TWS drawdown during drought compared to experiments with transient LAI (Figs. 4f,g).

To support Fig. 5, area-averaged monthly results of TranLAI and CliLAI are shown in Fig. 6. The difference in soil moisture gets larger in summer and autumn (through time) due to the lack of rainfall supply (Fig. 6f). LAI anomalies influence not only the water budget but also the surface energy budget. Compared to CliLAI, TranLAI shows a smaller latent heat flux (Fig. 6d) and a slightly larger sensible heat flux (Fig. 6c), leading to a higher Bowen ratio (sensible heat/latent heat; Fig. 6b). With less cooling effect from latent heat, ground temperature from TranLAI is higher. Area-averaged ground temperature differences are about 0.2 K in summer (Fig. 6h). The results imply that modeling vegetation in different ways can lead to distinct water availability and energy budget during drought.

Fig. 4. Contributions of TWS anomalies for March–May (MAM), July–August (JA), September–November (SON) 2011 from observations and different experiments. The anomalies are relative to their average of 2004-09.
d. Role of root depth with dynamic vegetation modeling

Deeper root results in higher ET and lower runoff (Table 2 and Fig. 8) because more water can be tapped from the bottom soil layer. However, the modification of deeper rooting depth slightly improves temporal variation of leaf area index in both wet and dry years (Table 3). The variation of soil moisture in the deeper layer is small because of damping and delay effects. In other words, the deeper rooting depth results in a smaller variation of beta factor, which is the water stress factor controlling the transpiration rate that ranges from 0 to 1 (supplementary Fig. 5).

e. Evaluation of experiments

The evaluation of the eight experiments is shown in Table 2. Area-averaged monthly time series are evaluated with observations: TWS from GRACE, runoff from USGS, and ET from FLUXCOM.

Experiments with an unconfined aquifer result in better temporal TWS variability with higher $R^2$ and smaller RMSE compared to simulations without groundwater. ClimLAI_GW exhibits the largest $R^2$ because using climatological LAI can lead to more transpiration and larger TWS declines during drought.

All experiments show overestimated runoff (Table 2) with a larger base flow, as shown in the hydrograph (Fig. 7a). In terms of the monthly average, experiments with free drainage overestimated surface runoff during rainy seasons, while peaks from the experiments with groundwater (e.g., TranLAI_GW) fit the observation better (Fig. 7a).

ET from all experiments is smaller than observation. The time series of modeled ET fits observed ET well in most of the seasons except for summer (Fig. 7b). ET from Noah-MP is smaller than observed in years with lower ET; the underestimation is especially large during the 2011 drought. This temporal aridity-dependent performance is similar to the previous study of spatial aridity-dependent performance (Lin et al. 2018).

The long-term underestimated ET bias is proportional to overestimated runoff bias, and experiments with a larger negative bias of ET are associated with a larger positive bias of runoff (Fig. 8). The smaller biases of ET and runoff are associated with experiments with groundwater, dynamic vegetation, and deeper roots. ET bias is smaller when adding groundwater while using dynamic vegetation. However, temporal variabilities of TWS, ET, and runoff are not improved by dynamic vegetation compared to the prescribed LAI (Table 2).

The evaluation of LAI (Table 3 and supplementary Fig. 6) indicates the challenge of modeling dynamic vegetation. The $R^2$ of LAI is low mainly because the modeled seasonal cycle differs from observations (e.g., Fig. 1). The performance of dynamic vegetation depends on both parameterizations and parameters. For example, this study shows the impact of runoff options, stomatal conductance options, and plant rooting depth. LAI from Noah-MP is usually larger than MODIS-based LAI (positive bias in Table 3). The $R^2$ of LAI is lower in dry years compared to wet years, indicating the challenge of vegetation modeling during drought. The dynamic vegetation scheme might be too sensitive to water stress, and the modification of deeper rooting depth improves the $R^2$ of LAI and ET due to the soil moisture damping and memory effects.
FIG. 6. Comparison between TranLAI and CliLAI. Area average of the Texas–Gulf region in 2011. (a) LAI, (b) Bowen ratio, (c) sensible heat, (d) latent heat, (e) transpiration, (f) root-zone soil moisture, (g) transpiration ratio ($T/ET$), and (h) ground temperature. Red line represents the differences (TranLAI minus CliLAI) and corresponds to the y axis on right.
4. Summary and discussion
Understanding and better monitoring drought are important for water management and agricultural productivity. In this study, experiments were conducted with multiple parameterizations in a single land surface model to show the uncertainties in the simulated water availability during drought. Overall, the results suggest that using different parameterizations can influence the modeled water availability, especially during drought.

Even though it is difficult to determine all sources of uncertainties because there are many other potential contributors, multiparameterization options in Noah-MP are demonstrated to be useful for attributing the uncertainties to physical mechanisms. Groundwater and vegetation are found to play an important role not only in the water cycle but also in the energy cycle. The experiments are evaluated with multiple observation-based datasets. Results show that Noah-MP tends to overestimate runoff and underestimate ET in the Texas–Gulf region (same as the results from Zheng et al. 2020), and especially during drought. More effort could be made to improve vegetation dynamics and representing the rooting depth variability in a realistic manner is one of the potential approaches (Arsenault et al. 2018; Niu et al. 2020).

The goal of this study is not to find any “best” experiments. The choices of the parameterizations and the evaluation matrix differ from different model applications or scientific questions. In addition, many other factors are not simulated in this study. For example, water deficit in surface water storage (rivers and lakes) and adaption strategies of water usage (e.g., groundwater pumping for irrigation) also contributes to TWS anomalies but are not considered in this study and these might be the reason why the model underestimates the negative TWS anomalies during severe droughts. Also, the observational errors are found to be relatively large in northwestern Texas due to sparse observational sites (Zheng et al. 2020).

It should be noted that no parameter calibration was performed in this study. Model results are very sensitive to several parameters when using dynamic vegetation schemes (Arsenault et al. 2018). Vegetation parameters could be updated by utilizing datasets, such as the TRY Plant Trait Database (Kattge et al. 2011). Several hard-coded parameters need to be updated (Cuntz et al. 2016) by accounting for spatial heterogeneity. Parameterizations related to vegetation physiological (e.g., stomatal conductance) and hydrometeorological processes, due to their cascading effects, should be considered carefully when using physical models for drought monitoring. Development of dynamic rooting depth (Wang et al. 2018; Niu et al. 2020) and plant hydraulic (Kennedy et al. 2019; Li et al. 2021) are relevant efforts. The sensitivity test with the rooting depth shows its important role in redistributing soil moisture vertically. In situ measurements can provide a comprehensive soil moisture profile and give a hint on root distribution. The rooting depth is also found to be strongly correlated with the water table depth (Fan et al. 2017), and it might be critical during recovery from drought.

The multimodel ensemble approach is commonly used for climate research. In this study, the benefits of a multiparameterization ensemble approach, such as a mechanistic explanation related to targeted physical processes, were shown to be effective.

Besides using different runoff/groundwater schemes, the design of soil layers and thickness in LSMs affects water capacity and thus is a fundamental constraint on the modeled TWS (Swenson and Lawrence 2015). The modeled configuration of the soil layer could be updated with the dataset of soil

<table>
<thead>
<tr>
<th>Experiment name</th>
<th>$R^2$</th>
<th>RMSE (cm)</th>
<th>Bias</th>
<th>$R^2$</th>
<th>RMSE (mm day$^{-1}$)</th>
<th>Bias</th>
<th>$R^2$</th>
<th>RMSE (mm day$^{-1}$)</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>TranLAI_GW</td>
<td>0.86</td>
<td>2.98</td>
<td></td>
<td>0.75</td>
<td>0.22</td>
<td></td>
<td>0.19</td>
<td>0.80</td>
<td>0.53</td>
</tr>
<tr>
<td>ClimLAI_GW</td>
<td>0.89</td>
<td>2.45</td>
<td></td>
<td>0.74</td>
<td>0.23</td>
<td></td>
<td>0.19</td>
<td>0.82</td>
<td>0.51</td>
</tr>
<tr>
<td>DyVeg_GW</td>
<td>0.84</td>
<td>2.93</td>
<td></td>
<td>0.75</td>
<td>0.16</td>
<td></td>
<td>0.11</td>
<td>0.78</td>
<td>0.51</td>
</tr>
<tr>
<td>DyVeg_DRt_GW</td>
<td>0.87</td>
<td>2.63</td>
<td></td>
<td>0.75</td>
<td>0.13</td>
<td></td>
<td>0.05</td>
<td>0.82</td>
<td>0.45</td>
</tr>
<tr>
<td>TranLAI</td>
<td>0.79</td>
<td>3.77</td>
<td></td>
<td>0.86</td>
<td>0.26</td>
<td></td>
<td>0.23</td>
<td>0.81</td>
<td>0.54</td>
</tr>
<tr>
<td>ClimLAI</td>
<td>0.86</td>
<td>3.20</td>
<td></td>
<td>0.87</td>
<td>0.27</td>
<td></td>
<td>0.23</td>
<td>0.83</td>
<td>0.53</td>
</tr>
<tr>
<td>DyVeg</td>
<td>0.76</td>
<td>3.76</td>
<td></td>
<td>0.86</td>
<td>0.19</td>
<td></td>
<td>0.16</td>
<td>0.78</td>
<td>0.53</td>
</tr>
<tr>
<td>DyVeg_DRt</td>
<td>0.81</td>
<td>3.23</td>
<td></td>
<td>0.87</td>
<td>0.13</td>
<td></td>
<td>0.10</td>
<td>0.81</td>
<td>0.47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experiment name</th>
<th>$R^2$</th>
<th>RMSE (mm day$^{-1}$)</th>
<th>Bias</th>
<th>$R^2$</th>
<th>RMSE (mm day$^{-1}$)</th>
<th>Bias</th>
<th>$R^2$</th>
<th>RMSE (mm day$^{-1}$)</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>TranLAI_GW</td>
<td>0.86</td>
<td>2.98</td>
<td></td>
<td>0.75</td>
<td>0.22</td>
<td></td>
<td>0.19</td>
<td>0.80</td>
<td>0.53</td>
</tr>
<tr>
<td>ClimLAI_GW</td>
<td>0.89</td>
<td>2.45</td>
<td></td>
<td>0.74</td>
<td>0.23</td>
<td></td>
<td>0.19</td>
<td>0.82</td>
<td>0.51</td>
</tr>
<tr>
<td>DyVeg_GW</td>
<td>0.84</td>
<td>2.93</td>
<td></td>
<td>0.75</td>
<td>0.16</td>
<td></td>
<td>0.11</td>
<td>0.78</td>
<td>0.51</td>
</tr>
<tr>
<td>DyVeg_DRt_GW</td>
<td>0.87</td>
<td>2.63</td>
<td></td>
<td>0.75</td>
<td>0.13</td>
<td></td>
<td>0.05</td>
<td>0.82</td>
<td>0.45</td>
</tr>
<tr>
<td>TranLAI</td>
<td>0.79</td>
<td>3.77</td>
<td></td>
<td>0.86</td>
<td>0.26</td>
<td></td>
<td>0.23</td>
<td>0.81</td>
<td>0.54</td>
</tr>
<tr>
<td>ClimLAI</td>
<td>0.86</td>
<td>3.20</td>
<td></td>
<td>0.87</td>
<td>0.27</td>
<td></td>
<td>0.23</td>
<td>0.83</td>
<td>0.53</td>
</tr>
<tr>
<td>DyVeg</td>
<td>0.76</td>
<td>3.76</td>
<td></td>
<td>0.86</td>
<td>0.19</td>
<td></td>
<td>0.16</td>
<td>0.78</td>
<td>0.53</td>
</tr>
<tr>
<td>DyVeg_DRt</td>
<td>0.81</td>
<td>3.23</td>
<td></td>
<td>0.87</td>
<td>0.13</td>
<td></td>
<td>0.10</td>
<td>0.81</td>
<td>0.47</td>
</tr>
</tbody>
</table>

### Table 3. Evaluation of LAI based on the area-averaged monthly time series in the Texas–Gulf region.

<table>
<thead>
<tr>
<th>Experiment name</th>
<th>$R^2$</th>
<th>RMSE (cm)</th>
<th>Bias</th>
<th>$R^2$</th>
<th>RMSE (mm day$^{-1}$)</th>
<th>Bias</th>
<th>$R^2$</th>
<th>RMSE (mm day$^{-1}$)</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>DyVeg_GW</td>
<td>0.24</td>
<td>0.95</td>
<td>0.83</td>
<td>0.20</td>
<td>1.08</td>
<td>0.99</td>
<td>0.07</td>
<td>0.81</td>
<td>0.67</td>
</tr>
<tr>
<td>DyVeg_DRt_GW</td>
<td>0.33</td>
<td>0.91</td>
<td>0.82</td>
<td>0.36</td>
<td>1.06</td>
<td>1.00</td>
<td>0.15</td>
<td>0.74</td>
<td>0.63</td>
</tr>
<tr>
<td>DyVeg</td>
<td>0.23</td>
<td>0.90</td>
<td>0.77</td>
<td>0.18</td>
<td>1.05</td>
<td>0.95</td>
<td>0.07</td>
<td>0.74</td>
<td>0.59</td>
</tr>
<tr>
<td>DyVeg_DRt</td>
<td>0.32</td>
<td>0.86</td>
<td>0.76</td>
<td>0.34</td>
<td>1.02</td>
<td>0.95</td>
<td>0.14</td>
<td>0.67</td>
<td>0.56</td>
</tr>
</tbody>
</table>
thickness (Brunke et al. 2016), and the setting of groundwater storage could be updated with the dataset of bedrock depth (Shangguan et al. 2017) or calibrated with streamflow observed from gauges.

This study utilized an offline land surface model in which atmospheric conditions are constrained by atmospheric forcing. However, vegetation dynamics were found to still play a significant role not only on water budget but also on surface energy budget, similar to the effects of deforestation. This implies the vegetation dynamics are important in land–atmosphere interaction and have an impact on regional precipitation prediction, which requires future work with coupled model experiments (Barlage et al. 2015; Meng et al. 2014). The insights gained from this study could help identify the directions of model development. Finally, besides drought monitoring, this study may have implications for weather, subseasonal-to-seasonal, and streamflow predictions.

Acknowledgments. This study was supported by the NOAA MAPP Program (Award NA17OAR4310134). The model simulations were conducted using the Texas Advanced Computing Center (TACC) computing resources. We thank all scientists, software engineers, and administrators who contributed to the development of Noah-MP. This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement 1852977. W.-Y. Wu would like to thank to Hui Zheng, Jingjing Liang, Weijing Chen, and Yu-Chiao Liang for making suggestions to improve the manuscript. We thank three anonymous reviewers and the editor for their comments to improve the manuscript.

Data availability statement. Noah-MP codes based on the High-Resolution Land Data Assimilation System (HRLDAS) are available at https://github.com/NCAR/hrldas-release. Model outputs used in the study are available from the authors upon request. CSR RL06 GRACE mascon data were downloaded from The University of Texas at Austin Center for Space Research at http://www2.csr.utexas.edu/grace. USGS HUC02 runoff was obtained from https://waterwatch.usgs.gov/. GLASS V04 LAI data were downloaded from...
from http://www.glass.umd.edu/. The NLDAS-2 data used in this study were acquired as part of the mission of NASA’s Earth Science Division and archived and distributed by the Goddard Earth Sciences Data and Information Services Center (GES DISC) from https://doi.org/10.5067/6J5LHHOHZHN4.

APPENDIX A

Groundwater and Runoff Parameterizations

a. SIMGM

In the scheme of TOPMODEL, an unconfined aquifer is below a 2-m soil column. The groundwater storage budget is expressed as the water balance,

$$\frac{dW}{dt} = Q - R_{sh},$$  \hspace{1cm} (A1)

where $W$ is the groundwater storage, $Q$ is the groundwater recharge, and $R_{sh}$ is the subsurface runoff or the discharge from the aquifer. The groundwater recharge $Q$ is based on Darcy’s law:

$$Q = K \frac{z_{bot} - z_{w}}{z_{bot} - z_{bot, sat}},$$  \hspace{1cm} (A2)

where $K$ is the hydraulic conductivity of the bottom soil layer, $z_{bot}$ is the midpoint of the bottom layer of the soil column, which is 1.5 m in this study. This equation includes both gravitational drainage and upward capillary force.

The subsurface runoff (discharge) is based on an exponential storage–runoff relationship, which is

$$R_{sh} = (1 - f_{imp}) R_{sh, max} e^{-f(z_{bot} - z_{w})},$$  \hspace{1cm} (A3)

where $f_{imp}$ is the maximum impermeable fraction due to frozen soil, $R_{sh, max}$ is the maximum subsurface runoff, which was estimated to be $5.0 \times 10^{-3}$ m s$^{-1}$. The variable $\lambda$ is the gridcell mean topographic index, which is 10.5 and can be calibrated. The term $f$ is the saturated area decay factor, which was assumed 6.0 m$^{-1}$, and $z_{bot}$ is the depth of the soil column bottom, which is 2 m in this study.

b. Free drainage

Conceptually different from TOPMODEL, the configuration of the free-drainage runoff option does not include any groundwater storage. The surface runoff is the excess precipitation that does not infiltrate into the soil. The subsurface runoff calculation is the gravitational drainage from the bottom soil layer and is parameterized as

$$R_{sh} = S \times K,$$  \hspace{1cm} (A4)

where $S$ is the slope index that can be calibrated, and $K$ is the hydraulic conductivity at the bottom layer.

APPENDIX B

Leaf Dynamics

One of the dynamic vegetation schemes used in this study is option 5 in Noah-MP, version 3.8. The leaf dynamics are based on Dickinson et al. (1998) and Niu et al. (2011). The vegetation fraction of the grid cell is prescribed as inputs (in this option only), but LAI is predicted dynamically. The carbon budgets include vegetation parts (leaf, stem, wood, and root) and soil carbon pool. The leaf carbon mass changes are the function of carbon assimilation rate, death rate, leaf turnover rate, and leaf respiration rate. And LAI is converted from leaf carbon mass with specific leaf area, which is vegetation-type dependent.

REFERENCES


Brunke, M. A., and Coauthors, 2016: Implementing and evaluating variable soil thickness in the Community Land Model, version...


