RESEARCH

Spatiotemporal variability of soil moisture over Ethiopia and its teleconnections with remote and local drivers

Tamirat B. Jimma[1] · Teferi Demissie[2,5] · Gulilat T. Diro[3] · Kassahun Ture[4] · Tadesse Terefe[1] · Dawit Solomon[2]

Abstract

Soil moisture is one of the essential climate variables with a potential impact on local climate variability. Despite the importance of soil moisture, studies on soil moisture characteristics in Ethiopia are less documented. In this study, the spatiotemporal variability of Ethiopian soil moisture (SM) has been characterized, and its local and remote influential driving factors are investigated. An empirical orthogonal function (EOF) and K_{Means} clustering algorithm have been employed to classify the large domain into homogeneous zones. Complex maximum covariance analysis (CMCA) is applied to evaluate the covariability between SM and selected local and remote variables such as rainfall (RF), evapotranspiration (ET), and sea surface temperature (SST). Inter-comparison among SM datasets highlight that the FLDAS dataset better depicts the country's SM spatial and temporal distribution (i.e., a correlation coefficient $r = 0.95$, $rmsd = 0.04\text{m}^3\text{m}^{-3}$ with observations). Results also indicate that regions located in northeastern Ethiopia are drier irrespective of the season (JJAS, MAM, and OND) considered. In contrast, the western part of the country consistently depicted a wetter condition in all seasons. During summer (JJAS), the soil moisture variability is characterized by a strong east–west spatial contrast. The highest and lowest soil moisture values were observed across the country's central western and eastern parts, respectively. Furthermore, analyses indicate that interannual variability of SM is dictated substantially by RF, though the impact on some regions is weaker. It is also found that ET likely drives the SM in the eastern part of Ethiopia due to a higher atmospheric moisture demand that ultimately invokes changes in surface humidity and rainfall. A composite analysis based on the extreme five wettest and driest SM years revealed a similar spatial distribution of wet SM with positive anomalies of RF across the country and ET over the southern regions. Remote SSTs are also found to have a significant influence on SM distribution. In particular, equatorial central Pacific and western Indian oceans SST anomalies are predominant factors for spatiotemporal SM variations over the country. Major global oceanic indices: Oceanic Nino Index (ONI), Indian Ocean Dipole (IOD), Pacific warm pool (PACWARMPOOL), and Pacific Decadal Oscillations (PDO) are found to be closely associated with the SM anomalies in various parts of the country. The associationship between these remote SST anomalies and local soil moisture is via large-scale atmospheric circulations that are linked to regional factors such as precipitation and temperature anomalies.

✉ Tamirat B. Jimma
atomictamirat@gmail.com

Teferi Demissie
t.demissie@cgiar.org

Gulilat T. Diro
gulilattef@gmail.com

Kassahun Ture
kassahun.ture@aau.edu.et

Tadesse Terefe
tadesse.terefe@aau.edu.et

Dawit Solomon
d.solomon@cgiar.org

[1] IGSSA, Addis Ababa University, King George VI St, P.O. Box 1176, Addis Ababa, Ethiopia

[2] International Livestock Research Institute, Asmera Road, 5689 Addis Ababa, Ethiopia

[3] ESCER Center, University of Quebec at Montreal, 201 President-Kennedy Avenue, Montreal H2X 3Y7, Canada

[4] CES, Addis Ababa University, King George VI St, P.O. Box 1176, Addis Ababa, Ethiopia

[5] Bjerknes Center for Climate Research, NORCE Norwegian Research Centre, P.O. Box 22, 5838 Bergen, Norway

 Springer

1 Introduction

Soil moisture is among the Essential Climate Variables (ECV) (Dorigo et al. 2017; Seneviratne et al. 2010; Wagner et al. 2012; Wang et al. 2016) and it determines the water and energy exchange between the land surface and the atmosphere (An et al. 2016; Li et al. 2011; Liu et al. 2011; Qiu et al. 2016). This is due to its impact on the variability of evapotranspiration, runoff, latent, and sensible heat fluxes (Dorigo et al. 2017; Seneviratne et al. 2010). Besides, soil moisture is an important variable in modulating the spatiotemporal weather variability and precipitation distributions as significant mesoscale atmospheric circulation patterns can be stimulated by anomalous regional soil moisture (Dorigo et al. 2017; Taylor et al. 2012). This is because SM has a slowly varying memory like the ocean that can stay for weeks to months and influence the weather system through surface energy fluxes and evaporation (Koster et al. 2004; Seneviratne et al. 2006). Soil moisture attribute of longer memory, together with a strong influence on surface climate variables, makes soil moisture an attractive source of subseasonal to seasonal predictability. As a result, numerous studies (Drewitt et al. 2012; Hirsch et al. 2014; Koster et al. 2010; Koster et al. 2011; Thomas et al. 2016) highlighted that accurate initialization of SM enhances forecast skills (e.g., precipitation and temperature) significantly for certain regions of the world.

Several studies (Koster et al. 2004; Schär et al. 1999; Zheng and Eltahir 1998) indicated that there is a relationship between soil moisture and rainfall. For instance, Koster et al. (2004) pointed out that the impact of SM on climate is more dominant than the ocean in mid-latitude regions during summer season. Moreover, Schär et al. (1999) also demonstrated that European summertime precipitation mostly depends on the soil moisture state. Soil moisture is also shown to affect precipitation in a range of scales (Taylor et al. 2012). Soil moisture effects on the rainfall could be distinguished as a two-stage process, i.e., SM's influence on ET and ET's influence on RF (Guo et al. 2006; Seneviratne et al. 2010; Wei and Dirmeyer 2012; Diro et al. 2014). However, RF's effect on SM and SM's effects on ET are understandable, but the impact of SM on RF is not straightforward. Therefore, understanding the temporal trends, variability, and spatial distributions of soil moisture is a critical requirement for demystifying the role of land on the overlaying atmosphere (An et al. 2016).

Despite the aforementioned importance of soil moisture, in situ measurements of soil moisture are generally very scarce. This is particularly apparent over most parts of Africa. Consequently, researchers and practitioners try to infer the amount of soil moisture indirectly from other sources such as satellite estimates or atmospheric reanalysis driven offline land surface model (LSM) simulation outputs (Dorigo et al. 2017). It must be noted that these indirect estimates have their own strengths and limitations. For instance, the accuracy of soil moisture estimated from land surface models is affected by both the quality of meteorological forcing fields (Dorigo et al. 2017; An et al. 2016) as well as by the quality of the model itself (Koster et al. 2009). On the other hand, the accuracy of satellite estimates is unreliable over vegetated areas (Peng et al. 2017; Srivastava et al. 2015; Feng and Liu 2015), and satellite estimates can also retrieve soil moisture only for the top few centimeters, whereas both reanalysis and satellite methods could provide a wider spatial coverage and continuous-time span with less missing data. However, one important aspect that is not well addressed is the degree of agreement or contradictions among station, reanalysis, and satellite estimates in representing various aspects of SM characteristics over East Africa. While there are few literatures (e.g., Mekonnen 2009) conducted on soil moisture estimation and validation at catchment level, studies that cover wider spatial and longer temporal scales soil moisture validation study across Ethiopia are still lacking. Therefore, the first objective of this study is to inter-compare satellite, reanalysis, and in situ soil moisture datasets to determine the level of agreement.

Studies (Feng and Liu 2015; Gaur and Mohanty 2013; Liu et al. 2022) argue that rainfall and evapotranspiration are the major driving factors that restrain SM dynamics besides wind, temperature, solar radiation, soil physical properties, topography, and vegetation. On the other hand, Nicolai-Shaw et al. (2016) and Zhu et al. (2020) highlighted that oceanic surface temperature is one of the reasons for global soil moisture persistence and predictability. However, these studies are rather global and some focus on specific regions of the globe, and less emphasis has been given to the East Africa region. Apart from a few attempts to understand SM residue monitoring (e.g., Ayehu et al. 2019) and SM-RF variability (e.g., Ayehu et al. 2020) across a river basin in Ethiopia, obtaining a study that covers detailed spatiotemporal characteristics of a soil moisture and its driving factors spanning across the country is limited. Therefore, evaluating such local and remote factors influencing the East African SM spatiotemporal variability is important. Thus, the second objective of this work is to assess the characteristics of spatiotemporal variability of SM and infer its local and remote drivers. Overall, this work provides a validated SM dataset, a delineation of homogeneous climate zones based on soil wetness in Ethiopia, and a better understanding of the local and remote influential SM drivers. Therefore, this study would benefit weather forecasting, agricultural planning, hydrological, and climatological modeling.

This paper is organized as follows: Sect. 2 describes the data employed, the statistical analysis methodologies applied in this study. Section 3 discusses results and findings. Finally, Sect. 4 presents a summary and draws some conclusions.

2 Materials and methods

2.1 Location and study area

Ethiopia is located in East Africa, $33°E − 48°E$ and $3°N − 15°N$. Figure 1 shows the station's distribution, geographic location, and country's topography. The country has a complex topography ranging from lowland areas of 116-m below sea level to 4600-m-high mountains (Diro et al. 2008; Zeleke et al. 2013; Diro et al. 2009; Fekadu 2015). Most of the country's north, central, partial eastern, and western parts are characterized by a highland plateau. The majority of eastern, northeastern, south, and southwestern parts of the country are lowland areas.

Climatologically, Ethiopia has three main seasons that have classified as February to May (Belg), June to September (Kiremt), and October to January (Bega) (Degefu 1987; Diro et al. 2008; Fekadu 2015; Gissila et al. 2004). Among these seasons, Kiremt is the main rainy season for northern and western parts while Belg is the major rainy season for southern and southeastern of Ethiopia. Most

crop growth and productivity are carried out during these seasons.

2.2 Data

Satellite, reanalysis and station soil moisture data have been used. Below is the description of each of the dataset.

2.2.1 FLDAS

A monthly land surface model simulated soil moisture product from Famine Early Warning Systems Network (FEWS NET) Land Data Assimilation System (FLDAS) (FEWS NET FLDAS, McNally et al. (2017) and McNally (2018)) is employed. FLDAS is forced by a combination of Modern-Era Retrospective analysis for Research and Application v2 (MERRA2, Gelaro et al. 2017) reanalysis and Climate Hazard Group InfraRed Precipitation with Station (CHIRPS, Funk et al. 2015) precipitation. This implies that FLDAS differs from its predecessor by using observation-based precipitation data as a forcing for the land surface model. The model has four layers for soil water content at 10-cm, 40-cm, 100-cm, and 200-cm depths at a spatial resolution of $0.1° × 0.1°$. The first 10-cm depth soil water content is used for this study. This dataset is available from 1982 to present.

2.2.2 ERA5Land

ERA5Land soil moisture data (Muñoz Sabater 2019) is obtained from Climate Data Store (CDS) of Copernicus

Fig. 1 Stations distribution, geographic location, and topography (in meters) of Ethiopia
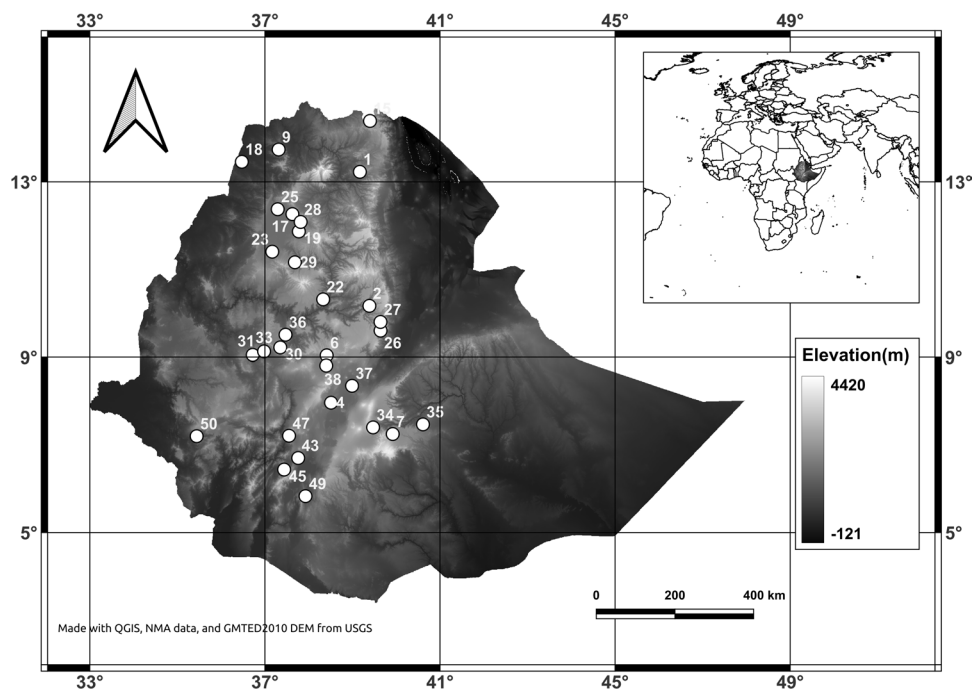
Climate Change Services (C3S) at 0.1° × 0.1° spatial and hourly temporal resolutions starting from 1981 to present. The ERA5Land model outputs are generated from the European Center for Medium-Range Weather Forecast (ECMWF) ERA5 climate reanalysis (Hersbach et al. 2019) by re-integrating its land surface model at a higher resolution than ERA5. Though ERA5Land has four soil levels: i.e., layer 1: 0–7 cm, layer 2: 7–28 cm, layer 3: 28–100 cm, layer 4: 100–289 cm, only the model top surface layer is used for this study.

2.2.3 GLDAS2

Two sets of Global Land Data Assimilation System 2 (GLDAS v2.0 and GLDAS v2.1, (Rodell et al. 2004) have been employed. Similar to FLDAS and ERA5, GLDAS is a global offline land surface modeling system (in this case Noah model) driven with observed/reanalysis meteorological fields to produce a series of surface and subsurface variables. In GLDAS v2.0, the Noah land surface model is simulated by forcing using Princeton meteorological input data and integrated from 1948 till 2014. GLDAS V2.1, on the other hand, uses a combination of both model and observation data and has been integrated from 2000 up to present. Both GLDAS v2.0 and GLDAS v2.1 have 0.25° × 0.25° spatial and monthly temporal resolution (Rodell et al. 2004).

2.2.4 Multi-satellite product

A multi-satellite soil moisture product based on the European Space Agency (ESA) Climate Change Initiative (CCI) satellite observations soil moisture data (Dorigo et al. 2017) are downloaded from Climate Data Store (CDS) of Copernicus Climate Change Services (C3S). It is version 03.3 gridded and combined (i.e., active and passive sensors) multi-satellite observation product (Gruber et al. 2017; Wagner et al. 2012). The merged active and passive products are produced by mixing scatterometer and radiometer soil moisture data, respectively.

The scatterometer is derived from Active Microwave Instrument-Windscat (AMI-WS) and Advanced Scatterometer (ASCAT) (Wagner et al. 2013) sensors while radiometer is derived from Scanning Multichannel Microwave Radiometer (SMMR) (Njoku et al. 1980), Special Sensor Microwave Imager (SSM/I) (Basist et al. 1998), the Tropical Rainfall Measuring Mission (TRMM), Microwave Imager (TMI) (Cashion et al. 2003), Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E) (Njoku et al. 2003), WindSat Spaceborne Polarimetric Microwave Radiometer (WindSat) (Li et al. 2010), Advanced Microwave Scanning Radiometer 2 (AMSR2)

(Imaoka et al. 2012), and Soil Moisture and Ocean Salinity (SMOS) (Kerr et al. 2010) sensors. Moreover, the combined soil moisture data is produced by blending these two merged products. These data are available in daily, dekadal, and monthly temporal resolutions and at 0.25° × 0.25° spatial resolution with soil water content measurement accuracy of 0.04 m^3m^{-3} unbiased root-mean-square-error (Dorigo et al. 2017). Thus, a daily soil moisture product is produced from multiple satellite instruments and microwave sensors. Dekadal and monthly means are calculated from these daily data. These products span from 1978 onwards with global coverage (Wagner et al. 2012).

It must be noted that satellite products have limitations over regions with high vegetation cover, rainfall, water bodies, and organic soils as these surfaces affect the microwave emission (Wang et al. 2016; Dorigo et al. 2017) and result in large uncertainties. Hence, for regions under a dense vegetation canopy, the soil moisture dataset has been masked (C3S SM Product User Guide 2018). This is more apparent over central western part of the Ethiopian rainforest.

2.2.5 Station dataset

Observed station data spanning from 2018 to 2021, obtained from the Ethiopian Meteorological Institute (EMI) Ethiopia, is used to validate offline land surface model data and satellite products discussed in the previous subsections. Due to its high temporal coverage, only 2019 data has been used. Figure 1 shows a map of soil moisture recording stations distribution over the country. These stations are recording SM degrees of saturation (%). Since the satellite and reanalysis soil moisture products are using volumetric units and observation are using SM degree of saturation (%) unit, conversion to the same volumetric (m^3m^{-3}) unit is important. Therefore, a conversion from degree of saturation to volumetric soil moisture has been carried out following the methods adopted from (An et al. 2016; Dorigo et al. 2011) and is given as follows:

$$SM_{vol}\left(m^3m^{-3}\right) = \frac{SM(\%)}{100} \times Porosity_{vol}\left(m^3m^{-3}\right)$$

where SM_{vol} is soil moisture in volumetric, $SM(\%)$ is soil moisture degree of saturation, and $Porosity_{vol}$ is soil porosity. The soil porosity data from the ESA CCI version v04.7 applied for soil moisture unit conversions. A thorough quality control has been performed on the station dataset as the raw data contains a large number of missing data and unrealistic SM values. The final number of stations selected for further analysis are limited to 30 stations (shown in Fig. 1) and for the year 2019.

2.2.6 Other surface datasets

To assess the local and remote factors that affect soil moisture variability, blended satellite and station precipitation dataset from CHIRPS (Funk et al. 2015), surface latent heat flux from ERA5Land (Hersbach et al. 2019; Muñoz Sabater 2019), and ERA5 Sea-Surface Temperature (SST) from Copernicus climate data center (Merchant et al. 2019) are employed. A global sea surface temperature indices: Oceaic Nino Index (ONI) (5°N − 5°S, 120°W − 170°W), Dipole Mode Index (DMI), i.e., an area average of (10°S − 10°N, 50°E − 70°E minus 10°S − 0°N, 90°E − 110°E), Pacific Warmpool (60°E − 170°E, 15°S − 15°N) area averaged, and Pacific Decadal Oscillation (PDO) are downloaded from NOAA Physical Sciences Laboratory (PSL) https://psl.noaa.gov/data/climateindices/list/.

2.3 Methods

To make consistent comparisons, all datasets were mapped to the lowest spatial resolution of (i.e., $0.25° \times 0.25°$) using a first-order conservative remapping (Hanke et al. 2016; Jones 1999) and aggregated to a monthly temporal scale. All analyses were carried out based on the seasonal and annual time scales. The focus seasons are March–April-May (MAM), June-July–August-September (JJAS), and October–November-December (OND).

2.3.1 Validation procedures

After applying data quality control on the station's SM datasets, a 15-min observation was resampled to a monthly temporal scale. Monthly data from each LSM and satellite gridded dataset was extracted by averaging each corresponding grid box nearby the station location. The Spearman correlation skill, root-mean-square difference (RMSD), and standard deviation of model and satellite datasets were evaluated against the observations.

2.3.2 EOF and CMCA analysis

EOF analysis is employed to outline SM homogeneous zones in the study area that could help us analyze other meteorological variables in the respective zones to identify SM characteristics. Furthermore, to gain insights into the local (i.e., RF and ET) and remote (i.e., global SST) drivers of SM variability, a complex rotated maximum covariance analysis (CMCA) has been applied. Following methods discussed by Rieger et al. (2021); Dawson (2016); Hannachi et al. (2009); and Baldwin et al. (2009) original input data is prepared as follows: Firstly, to account for the effect of meridians convergence at a higher latitude that affects the

grid size and results in high sampling variation, the square-root of the cosine of latitude area weight applied on the input data (North et al. 1982).

$$w_i = \sqrt{\cos(\theta)}$$

where w_i is weights of data at each grid point and θ is the latitude of a grid point in radian. Then to remove seasonal trends and minimize the effect of extreme values, a time mean of every grid point, x_j, is calculated, and its anomaly is computed.

$$x_j = x_j - \bar{x}$$

$$z_j = \frac{x_j - \bar{x}}{\sigma}$$

where x_j is the jth observation point and $\bar{x}$ is the time mean for that observation. z_j is the standardized anomaly of observation x_j and σ is its standard deviation.

The EOF and CMCA methods are applied to the spatial data (x_j) anomalies and standardized anomalies, respectively, over a sampling time (t_i), where ($j = 1, ..., r$ and $i = 1, ..., s$). Where **r** and **s** are the extent of spatial and temporal dimensions, respectively. This spatiotemporal data arranged in matrix form as $\mathbf{s} \times \mathbf{r}$ is called S-mode, and applying principal component analysis on it could help us to identify statistically significant temporal soil moisture anomalies (Dommenget and Latif 2002; Hannachi et al. 2009, 2007; Jollife and Cadima 2016).

The CMCA method, based on the Hilbert transformation of the input dataset to its real and imaginary parts and decomposing the resulting covariance matrix in complex space, has been applied to identify the covariability of SM and other meteorological variables. The CMCA method also enables us to evaluate the phase shift and time lags between signals. However, due to the limitations of the CMCA method, the time lags can be identified if certain conditions are fulfilled between the two geophysical variables (Boashash 1992). Mathematically, the Hilbert transformation could be written as follows:

$$\hat{X} = X + iH(X)$$

where H is the Hilbert transform applied on the real signal X and $\hat{X}$ is the resulting analytical signal. A Promax oblique rotation method ($p = 2$) is applied on the first 100 modes to alleviate the orthogonality constraint imposed by MCA as mentioned in (Richman 1986; Rieger et al. 2021). A theta extended model (Fiorucci et al. 2016) with the dominant period of signal 12 (i.e., seasonal cycle) has also been applied to minimize the extent of spectral leakage at the boundary. Similar to (Rieger et al. 2021), all monthly datasets are smoothed by 6 months moving average for

remote variable (i.e., SST); however, local variables (i.e., RF, ET) are smoothed by 12 months to coordinate their temporal variability and circumvent high frequency noise. A comprehensive description of the CMCA method has been discussed in (Rieger et al. 2021).

2.3.3 Data clustering

Data clustering into homogeneous non-overlapping subgroups of similar data points within-cluster is carried out using KMeans iterative clustering algorithm. K_{Means} clustering targets to group data points that have smaller variances within-cluster, i.e., minimize sum of squares in the group. Mathematically:

$$min\left(\sum_{i=1}^{n}\sum_{x_j g_i}(x_j - \bar{x})^2\right)$$

where $\bar{x}$ and x_j's are the mean and data points within the group g_i. The K_{Means} algorithm is employed on nondegenerate leading EOFs with application of the Elbow method, Silhouette criterion, and expert opinion to perform regionalization analysis and divide Ethiopia into five homogeneous soil moisture zones. The KMeans initial centroids seeded by the robust, fast and simple K_{Means}++ algorithm that outperform KMeans in terms of achieving accurate and optimum solution (Arthur and Vassilvitskii 2006; Bahmani et al. 2012) has been applied. We have also employed a wet minus dry composite analysis on SM, RF, and ET based on the extreme five wetter and dryer SM years to evaluate the nonlinear relationship among these variables.

3 Results and discussion

3.1 Spatiotemporal variability of soil moisture datasets (validation)

In this subsection, satellite and reanalysis estimates are validated for 2019, where there is station data over Ethiopia. Figure 2 shows the annual cycles of soil moisture from reanalysis, satellite, and station data for the year 2019 averaged over stations in Ethiopia. Observation data indicates that soil moisture is drier during the winter period and starts to increase slowly till the end of summer and then declines during the autumn months.

This pattern is reproduced well by both reanalysis and satellite estimates. While the rainfall or soil moisture distribution over Ethiopia could be mono-modal or bimodal (as it will be discussed later) depending on the region considered, the mono-modal pattern of soil moisture presented here is due to the fact that majority of the stations are taken from predominantly summer rain receiving regions. All
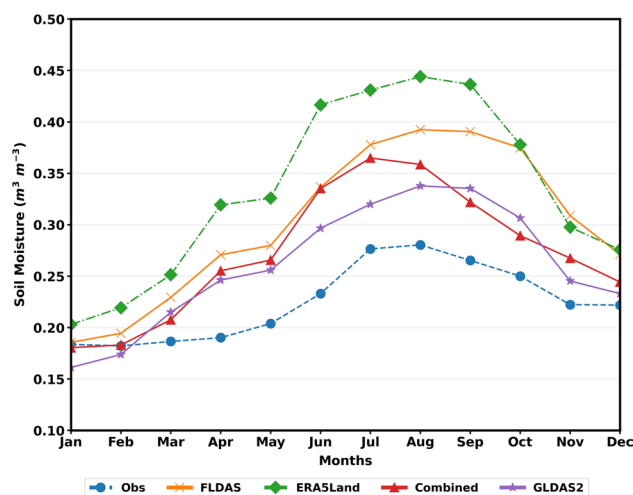

Fig. 2 Annual cycle of area-averaged soil moisture from observations, FLDAS, ERA5Land, combined ESA CCI, and GLDAS2 for 2019

the datasets, except the Combined ESA CCI, showed the peak soil moisture content in August. The Combined satellite data, however, indicated July is the maximum. Compared to the station data both reanalysis and Combined ESA satellite datasets generally overestimate the annual cycle with ERA5Land showing the highest value for most of the months. The difference between station dataset and indirect satellite and reanalysis estimates can be partly related to the difference in the depth of the top soil layer and hence the difference in porosity at a different depth, even if the soil moisture amount presented is a normalized value. The top layer soil depth is much deeper in station data (20 cm) compared to reanalysis (7–10 cm) and satellite (5–10 cm) data.

Figure 3 presents the spatial distribution of the 2019 annual mean soil moisture with an overlay of stations value, to compare observations, reanalysis, and satellite estimates spatial consistencies. Both reanalysis and satellite estimates distinguish the drier lowlands of eastern part of the country from the wetter western region. Reanalysis in particular identified the southwest as the wettest region. It is noticeable that the observations grid box average is proportionately different from respective satellite and reanalysis datasets. It is likely that the top soil layer depth representation difference for absolute soil moisture measurement at each dataset is the reason. This discrepancy could be also attributed to the difference in reanalysis models soil parameters representations and satellite data analysis algorithms. Although ERA5Land values are comparable with that of the station measurements for the mid-central highlands, the FLDAS dataset is the one that consistently captures the soil moisture spatial distributions across the country. This is also proven by Taylor diagram analysis (not shown here, (Taylor 2001)) that FLDAS shown a higher correlation ($r = 0.95$,

Fig. 3 Stations annually averaged grid box data overlayed on each soil moisture dataset climatologies over Ethiopia. **a** FLDAS, **b** ERA5Land, **c** combined, and **d** GLDAS2 for 2019
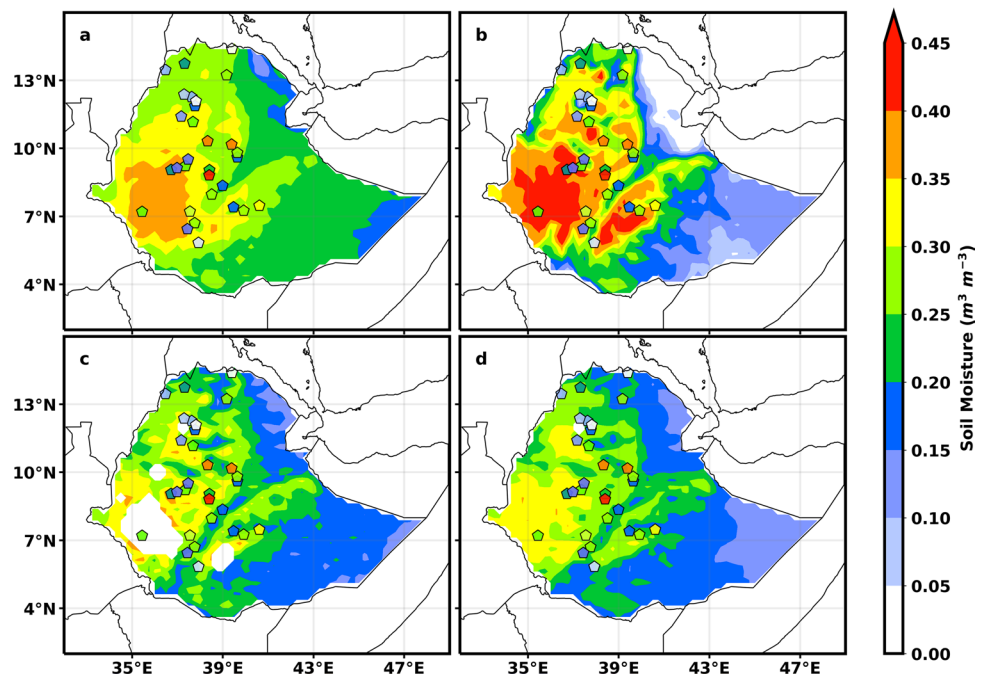
Table 1 Correlation coefficients among the soil moisture datasets and their RMSD from observation

	Obs	FLDAS	ERA5Land	Combined	GLDAS2	RMSD
Obs	1.00	0.95	0.90	0.91	0.93	-
FLDAS	0.95	1.00	0.96	0.93	0.99	0.04
ERA5Land	0.90	0.96	1.00	0.97	0.98	0.05
Combined	0.91	0.93	0.97	1.00	0.94	0.03
GLDAS2	0.93	0.99	0.98	0.94	1.00	0.03

Table 1) and low root-mean-square difference (0.04 $m^3 m^{-3}$) skills. Besides, the spatial resemblance of FLDAS and GLDAS2 datasets are also significant, whereas the GLDAS2 is slightly underestimating the soil wetness distributions over the country. Consequently, FLDAS has been chosen for further analysis, including to identify homogeneous zones.

3.2 Identification of homogeneous soil moisture zones

EOF analysis and K_{Means} clustering methods have been employed to identify homogeneous soil moisture zones at a regional scale. Application of EOF analysis on the FLDAS dataset has resulted in the eigenvalue spectrum shown in the supplementary materials Figure S1. It indicates that the first three EOFs are nondegenerate as they are significant according to North's rule of thumb (North et al. 1982) and do not overlap in 95% confidence interval (Korres et al. 2010). Therefore, the first three EOFs are considered for further analysis to get insight into the total variances in the spatial patterns. EOF-1 possesses a significant spatial pattern that explains variance

of 62.7%, while EOF-2 explained 22% of the total variance, and the least nondegenerate dominant mode EOF-3 accounted for 6.9% variances. These three EOFs' explained 91.6% of the total variances in the data; therefore, the rest of the EOFs can be discarded as noise. Since EOF-1 only comprised more than 50% of the total variances, we mostly used the first rank explained variances for homogeneous zones classification. Thus, five distinct homogeneous regions are classified based on these EOF results and the KMeans clustering algorithm. The supplementary materials Figure S2 (or the middle panel of Fig. 4) presents Ethiopia's five non-overlap homogeneous soil moisture zones.

The homogeneity within the regions and dissimilarity among classified regions evaluated using Pearson's product moment correlation. Table 2 presents the inter- and intra- correlations among the five homogeneous climate regions and it clearly shows that the inter-correlation values within the region is always higher than the intra-correlation values among the regions. This table shows Reg-I has the highest correlation value within the region ($r = 0.80$), while Reg-III has the lowest intra-correlation ($r = 0.58$).

Fig. 4 Homogeneous regions annual cycles for FLDAS, ERA5Land, Combined, and GLDAS2 soil moisture datasets (1982–2020)
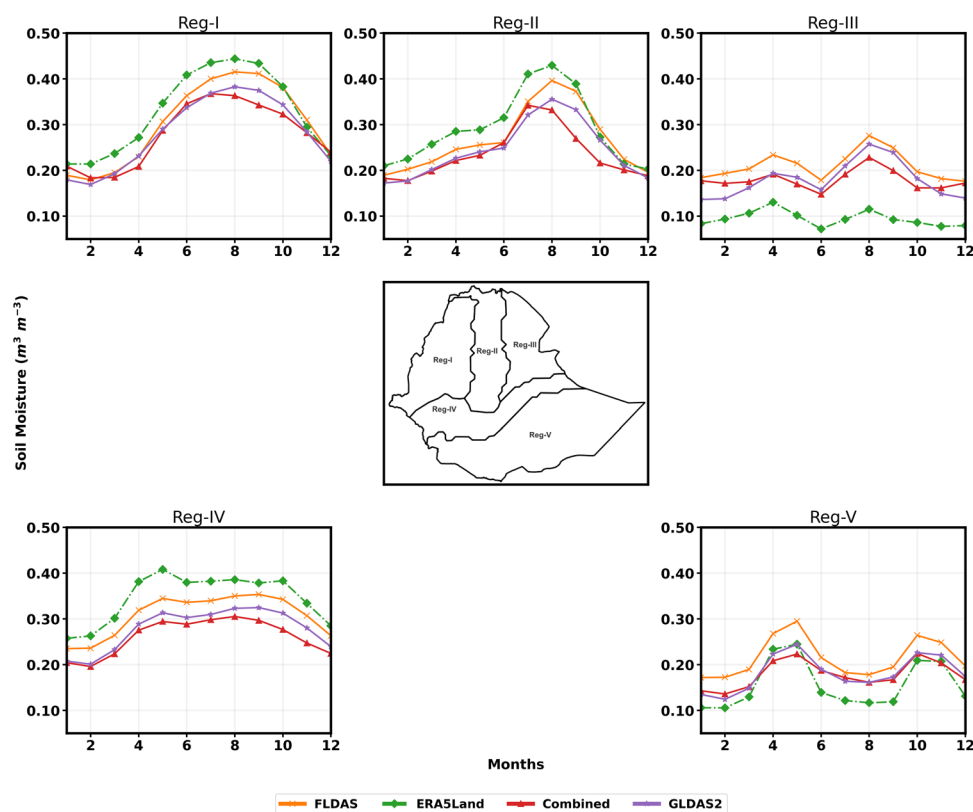
Table 2 Correlation coefficients among homogeneous zones

	Reg-I	Reg-II	Reg-III	Reg-IV	Reg-V
Reg-I	0.80	0.30	0.31	0.30	0.30
Reg-II	0.30	0.71	0.27	0.27	0.27
Reg-III	0.31	0.27	0.58	0.21	0.21
Reg-IV	0.30	0.27	0.21	0.59	0.31
Reg-V	0.30	0.27	0.21	0.31	0.59

Based on this regionalization, Fig. 4 shows annual cycles for individual homogeneous regions to FLDAS, ERA5Land, combined, and GLDAS2 datasets for the 1982–2020 period. Contrary to the country wide annual cycles shown in Fig. 2, homogeneous regions exhibited mono-modal and bimodal patterns of SM depending on the region considered. The bimodal nature of the SM pattern is clearly shown in Reg-III and Reg-V, whereas Reg-I, Reg-II, and Reg-IV are uni-modal. On top of this, SM over Reg-IV is wet almost for the entire year. Besides, Fig. 4 shows the ERA5Land dataset is the wettest of all datasets for Reg-I, Reg-II, and Reg-IV, while FLDAS is the wettest for Reg-III and Reg-V. A closer inspection of Reg-I and Reg-II shows the Combined dataset picked maximum SM in July while other datasets lag one month behind to peak in August. As shown in Fig. 4, the annual soil wetness in Reg-I, Reg-II, and Reg-IV is significantly higher, whereas the lowest is in Reg-III. It is

interesting to note that these spatial patterns align with the country's topographic features.

Figure 5 demonstrates the spatial pattern of annual and seasonal climatologies of SM from FLDAS data. Overlayed are the homogeneous soil moisture regions. It is apparent that the Dallol depression, located in northeastern Ethiopia (Reg-III), is drier irrespective of the season considered. Conversely, the western part of Reg-IV is consistently depicted as wetter in all seasons (JJAS(b), MAM(c), and OND(d)) or in the Annual mean (a).

The SM during Summer (JJAS) is characterized by a strong east–west spatial contrast. The highest SM value is noted over Reg-I(b) and Reg-IV(b). Whereas the lowest wetness is shown at Reg-V and the northeastern edge of Reg-III. During the Belg season (Fig. 5c), the contrast between wet and dry regions is lower and most parts of the country receive values greater than 0.2 m^3m^{-3}. Despite the Bega season (Fig. 5d) being a dry season, most parts of the country with the exception of lowland areas of the north eastern part of the country possess SM values greater than 0.2 m^3m^{-3}, highlighting the longer memory of SM.

### 3.3 Interannual variability of soil moisture datasets in the homogeneous climate zones

Figure 6 presents the interannual variability of annual mean soil moisture from FLDAS, ERA5Land, and GLDAS2

Fig. 5 Annual and seasonal climatologies of soil moisture from FLDAS overlayed on homogeneous zones: **a** annual, **b** JJAS, **c** MAM, and **d** OND (1982–2020)
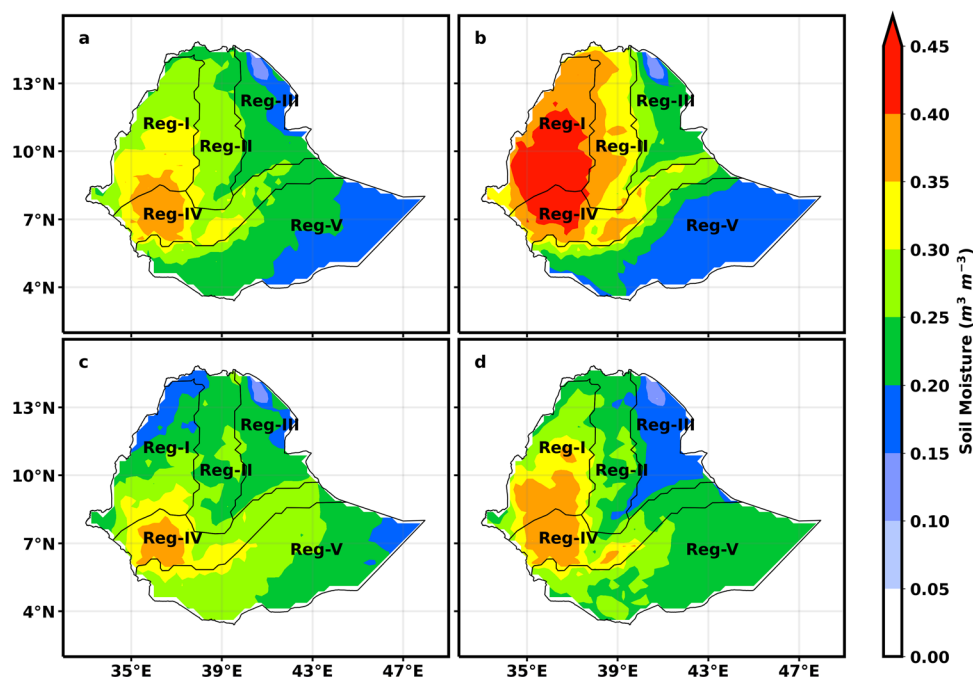

Fig. 6 Time series of standardized anomalies of annual mean soil moisture datasets from FLDAS, ERA5Land, and GLDAS2 for each homogeneous zone (i.e., Reg-I, Reg-II, Reg-III, Reg-IV, and Reg-V)
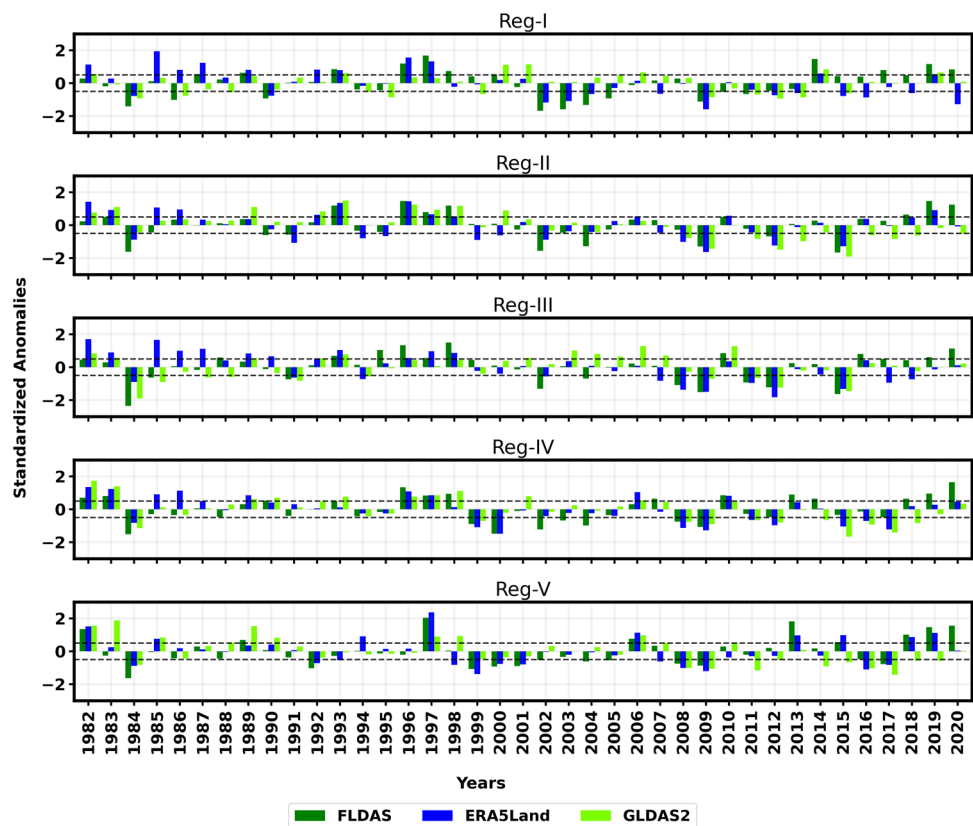
datasets for each homogeneous region. While different homogeneous regions depict slightly different wet and dry years at various levels, some years show the same sign of soil moisture anomalies across all homogeneous zones. For

instance, in 1984 and 2009, negative soil moisture anomalies were noted over all homogeneous zones, whereas, in 1997, wetter soil moisture was observed for all zones, even if the impact is substantially higher for Reg-V. The time series
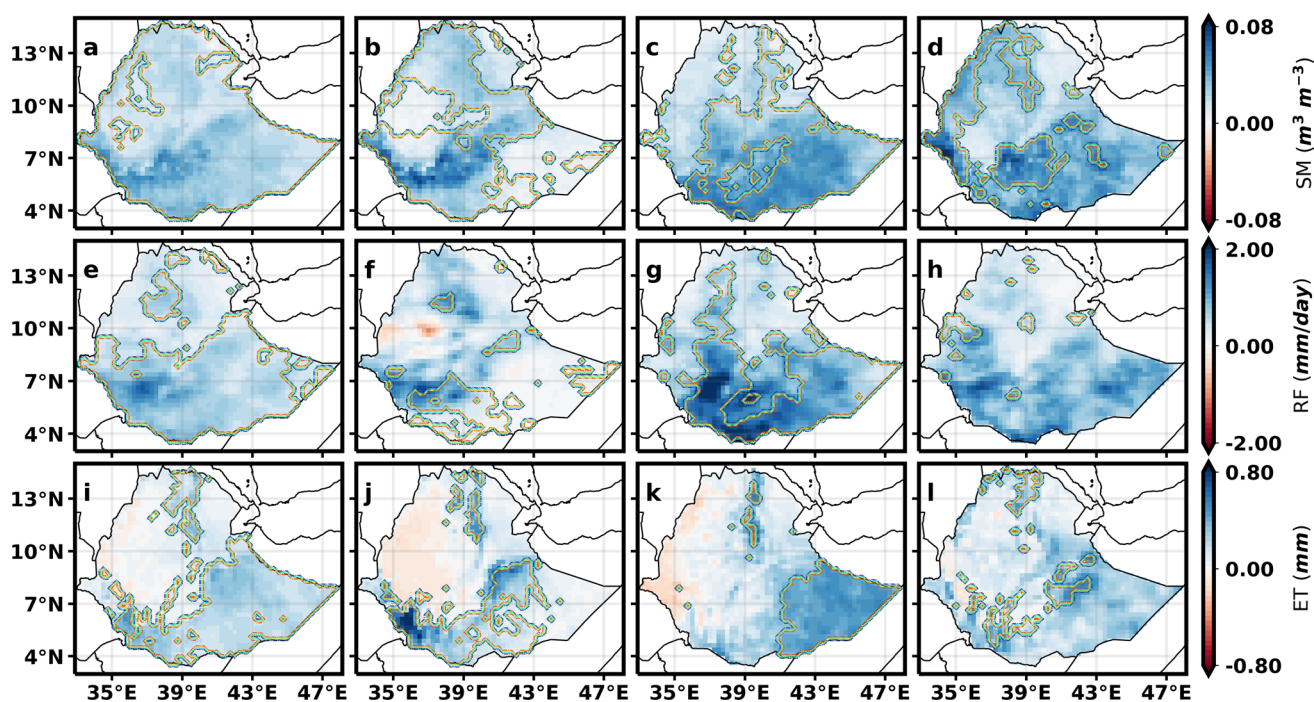
Fig. 7 Composites of SM (top row), RF (second row), and ET (third row) for annual (first column), JJAS (second column), MAM (third column), and OND (fourth column) based on extreme five wettest and drier soil moisture years. Contour lines represent regions that are significant at 0.05 level

is also characterized by low-frequency variability with a period of consecutive dry years and a consecutive period of wet years. This is apparent in the mid-1990s and late 2000s, where there are consecutive wet and dry episodes, respectively, for most of the homogeneous zones. It is interesting to note that not all El Nino years are linked to SM anomalies in a similar fashion. For instance, the 1997 and to some extent the 1982 El Ninos are linked to positive SM anomalies for all zones. In contrast, the 2015 El Nino year left Reg-II and Reg-III severely dry while Reg-I and Reg-V were barely wet.

3.4 Local and global factors affecting soil moisture distributions and variations

Since distinct climatic zones could be affected by various driving factors, the extent of these differences are further investigated against potential local and global SM driving factors in the following subsections.

3.4.1 Local factors associated with soil moisture anomalies

In this sub-section, we analyzed the effects of RF and ET on the variability of SM. Figure 7 shows the spatial distribution of composites of SM (Fig. 7a, b, c, d), RF (Fig. 7e, f, g, h), and ET (Fig. 7i, j, k, l) based on extreme five wettest minus five drier SM years. Hence, composites of country

wide wetter minus dryer SM anomalies (Fig. 7a–d) indicate wetter anomalies over southwestern Ethiopia and over the eastern highland areas south of the rift valley. The spatial distribution of SM anomalies suggests that country average SM anomalies are dictated by regions south of the rift valley in all seasons. The composite analysis asserted the close resemblance of the spatial distribution of wet SM anomalies with the positive anomalies of RF (Fig. 7e) over most parts of the country. The results also demonstrate that drier/ wetter soil moisture is closely related to negative/positive ET (Fig. 7i) over the southern and southeastern parts of the country.

The central and western part of the country is associated with very low and slightly negative insignificant ET anomalies (Fig. 7i, j, k). During the summer season (JJAS), both RF and ET are closely associated with SM anomalies over the southwestern and eastern highlands of Ethiopia. However, a stronger RF-SM association is noted over the central southern region. These results indicated that while ET and RF have slightly negative relationships with SM over the western part of the country, there is a stronger positive link over southwestern, central, and eastern parts of Ethiopia during the summer season. The central, north-, and southeastern part of the country's soil moisture (Fig. 7c) indicated wetter SM anomalies in MAM season that directly followed the corresponding RF pattern in that region (Fig. 7g) except in the northwest. On the other

Fig. 8 Correlation between near-surface air temperature and SM (top row), RF (second row), and ET (third row) for annual (first column), JJAS (second column), MAM (third column), and OND (fourth column). Contour lines represent regions that are significant at 0.05 level
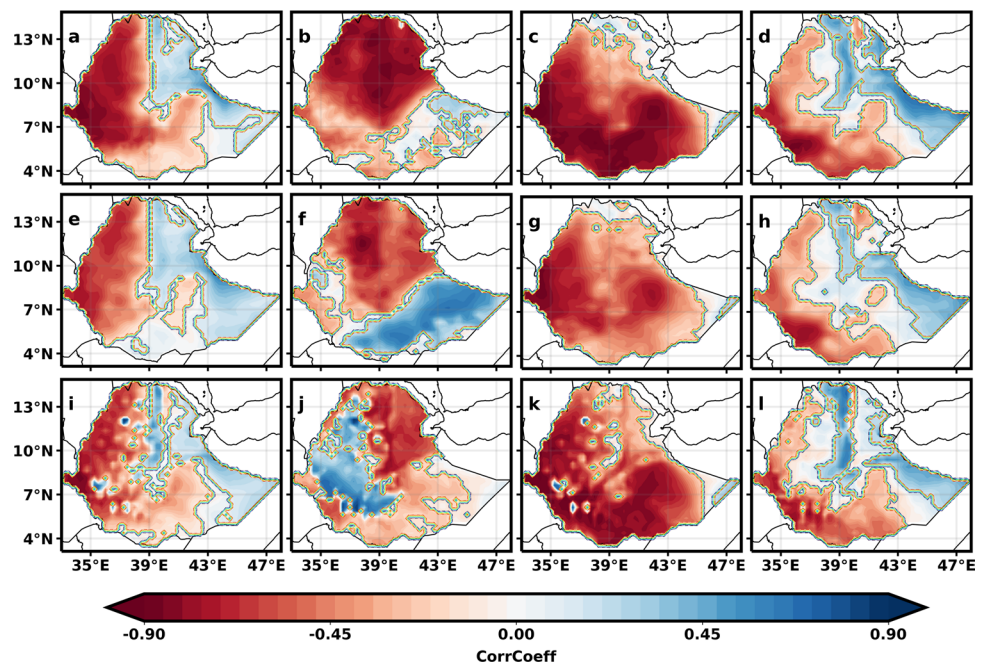
hand, the ET anomaly in (Fig. 7k) highlights insignificant negative anomalies at the western region of the country though the anomalies vanish along the western tips of the country. These findings show that the SM, RF, and ET likely interrelated in the southeastern part of the country.

Similar relationship among the SM, RF, and ET is noted during OND season (Fig. 7d, h, i) suggesting that the spatial pattern of RF and SM anomalies co-located for most of part of the countries, whereas the co-location of SM and ET is concentrated over eastern part of Ethiopia.

Figure 8 shows the spatial pattern of correlation of near-surface air temperature (Tair) with SM, RF, and ET. It is noted that higher Tair is associated with higher ET values over the eastern part of the country and lower values in the western region (Fig. 8i). In the west, the negative correlation implies that the SM plays a role in partitioning more of the net radiation to latent heat flux and lowers the surface temperature via evaporative cooling. This process entails that the land is a driver in the region. In the east, a positive correlation between Tair and ET signifies that a higher atmospheric demand drives an increase in evapotranspiration. The increase in ET leads to more surface humidity, resulting in an increase in wet-bulb temperature; therefore, it lowers the wet-bulb depression in the planetary boundary layer. Thus, this process leads to an increase in RF (Eltahir 1998; Findell and Eltahir 1999) and enhances the surface moisture. It is, therefore, likely that the ET drives the local SM in the eastern and southeastern region. These annual relationships are more dominated by the winter season associationships, whereas the summer (Fig. 8j) and spring (Fig. 8k) seasons correlations suggest that the impact of ET on SM is not conclusive.

The supplementary materials Figure S3 shows the annual and seasonal year-to-year variability of SM, RF, and ET. From the data in Figure S3, it is apparent that the SM interannual variability goes inline with the RF. This finding is consistent with that of (Seneviratne et al. 2010) and Teuling and Troch (2005). Besides, correlations between SM and RF/ET show there are significant relationships in every homogeneous zone except insignificant SM versus ET correlations in Reg-I, Reg-II, and Reg-IV in the summer. RF has significant association with SM in all homogeneous zones and seasons while its magnitude is lower in Reg-I (JJAS), whereas ET demonstrated lower associationship with SM during summer in all zones except Reg-III and Reg-V. Overall, correlation magnitudes are high in most seasons and zones whereas ET indicates weak relationships in some zones and seasons.

We have also analyzed the covariability of SM with RF and ET using CMCA methodology adopted from (Rieger et al. 2021). Hence, Fig. 9 shows the ET-SM phase functions (Fig. 9a–j) and amplitude functions (Fig. 9k–t) of CMCA for the five dominant modes of variabilities. Similarly, Fig. 10 also presents the RF-SM phase functions (Fig. 10a–j) and amplitude functions (Fig. 10k–t) of the five dominant modes.

Mode-1 contributes 32.8% and 28.5% of the covariability between ET and SM (Fig. 9a, f, k, p) and RF and SM (Fig. 10a, f, k, p), respectively. The amplitude functions (Figs. 9 and 10k, p) show the higher amount of variations over the southern part of the country in both RF-SM and ET-SM covariability. It is also noted that these modes of variabilities indicated in-phase functions (Figs. 9 and 10a, f) of RF and ET with SM in this region; hence, it seems

Fig. 9 Phase functions for ET (**a–e**) and SM (**f–j**) relative phase shifts with respect to their corresponding principal components (PCs) and amplitude functions for ET (**k–o**) and SM (**p–t**). The maximum normalized scale has been applied and phase values having less than 0.25 amplitude (i.e., insignificant noisy phase values) have been masked out (Rieger et al. 2021)
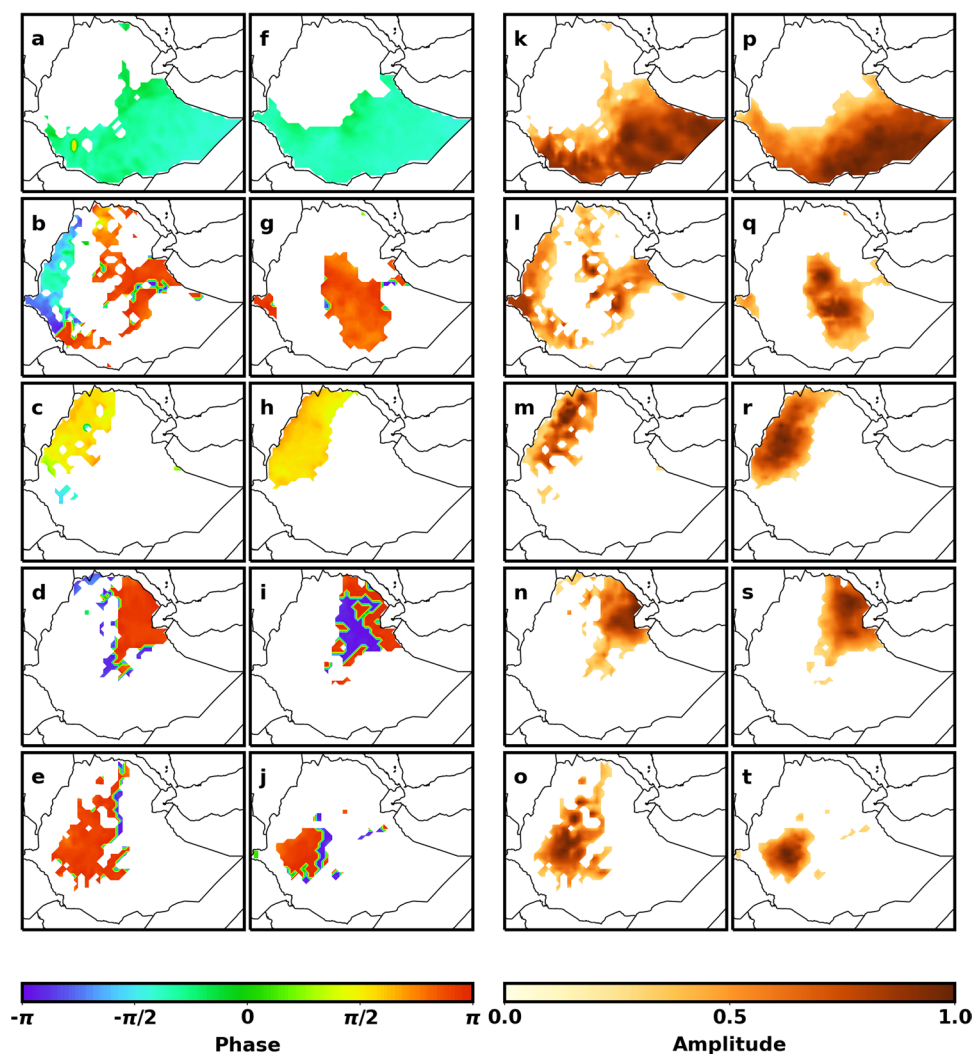
possible that increase in RF leads to increase in SM. The in-phase relationship between ET and SM involves a chain of atmospheric processes by which an increase in atmospheric demand for moisture driven by higher temperature leads to increases in ET and RF, and ultimately SM as discussed in the composite analysis.

Mode-2 contributes 10.4% and 11.7% of co-variations of the ET and SM (Fig. 9b, g, l, q); RF and SM (Fig. 10b, g, l, q), respectively. This ET-SM mode of variabilities represent the central, central-east, and along western tips of the country. The spatial pattern of the second mode of ET-SM covariability is not collocated since the pattern of SM is over the central highland, whereas the signal of ET is at the eastern and western edges of the Ethiopian highland. This result is interesting as the SM over the central highland is associated closely with the remote ET residing over the edges of the highland rather than the ET on the highland itself. This suggests the critical role of regional atmospheric circulation in transporting moist air from lowland to the highland. On the other hand, the second mode of RF and SM (Fig. 10b, g, l,

q) covariability explains the central southern region. This mode of variability shows a complicated phase relationship as it is shifted by $+\pi/2$. Nevertheless, mode-2 of variability also partially overlaps with mode-1, though this outcome is not expected from a rotated EOFs.

Mode-3, on the other hand, represents the northwestern ET-SM (Fig. 9c, h, m, r) and northeastern RF-SM (Fig. 10c, h, m, r) covariability. This mode explains 8.8% (ET) and 10.4% (RF) versus SM covariability. It should be noted that the spatial patterns of mode-3 of ET-SM and mode-4 of RF-SM are interchanged but represent the same northwest region. It has to be noted that mode-4 (RF-SM) showed a positive correlation, whereas mode-3 (ET-SM) demonstrated a complicated relationship that phase shifted by $+\pi/2$. However, mode-4 of ET-SM (Fig. 9d, i, n, s) and mode-3 of RF-SM (Fig. 10c, h, m, r) represent northeastern regions of the country with 7.6% and 6.9% covariability fractions, respectively. Moreover, mode-3 (RF-SM) and mode-4 (ET-SM) identify a direct covariability among these variables that likely supports our hypothesis ET drives the

Fig. 10 The same as Fig. 9 but for phase functions of RF (**a–e**) and SM (**f–j**); amplitude functions for RF (**k–o**) and SM (**p–t**)
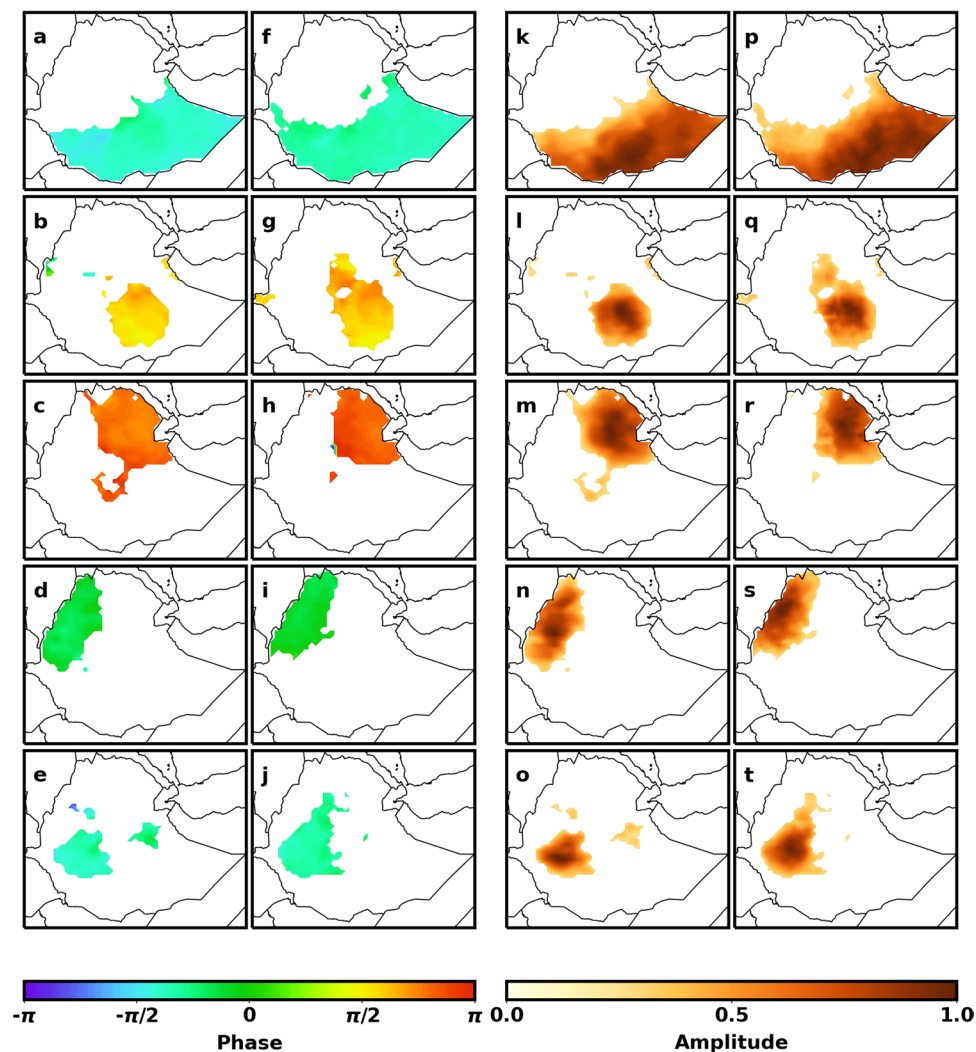
local SM in this region via a cascade of atmospheric processes, as discussed at the beginning of this section.

Mode-5 (Figs. 9 and 10e, j, o, t) pointed out a central western spot that outlines the direct relationship between RF and SM and accounts for 5.6% of covariability. The corresponding mode of ET-SM explains 7.5% covariance. This mode highlights a direct relationship between SM with both RF and ET in this region.

3.4.2 Global climatic variables effects on soil moisture

This subsection analyzes the link between global SST and SM variability over Ethiopia using CMCA. Figures 11, supplementary materials Figure S4, and 12 show the amplitude functions, phase functions, and principal components (PCs), respectively, for the first four leading complex rotated MCA modes of SST and SM. The percentage of covariabilities explained by these modes are 59.7,%, 6.3% 2.1%, and 0.9% for mode-1, mode-2, mode-3, and mode-4 respectively.

The first mode (Figs. 11, S4, and 12a, e) shows the annual cycles of the global SST and SM. The spatial amplitude function (Fig. 11a) showed low equatorial SST variability, while the lowest variability is along the equatorial pacific ocean. Moreover, it is noted that the phase function (Fig. S4: a) identified the anti-correlation between the southern and northern hemisphere SST anomalies that is consistent with findings in (Rieger et al. 2021). On the other hand, the annual SM variation (Fig. 11e) is predominantly higher in the northwestern part of the country. Furthermore, the phase function (Fig. S4: e) of this mode also showed distinct separate SM regions in the northeast, northwestern, and southern parts of the country. These distributions might be related to the summer season (JJAS) rainfall distribution in the country.

The second mode (Figs. 11, S4, and 12b, f) demonstrates that equatorial pacific and western Indian oceans SST variations predominantly affect SM anomalies over the southern and northeastern part of the country. The anti-correlations between the northeastern SM and equatorial

**Fig. 11** Amplitude functions for global SST (**a–d**) and soil moisture (**e–h**)
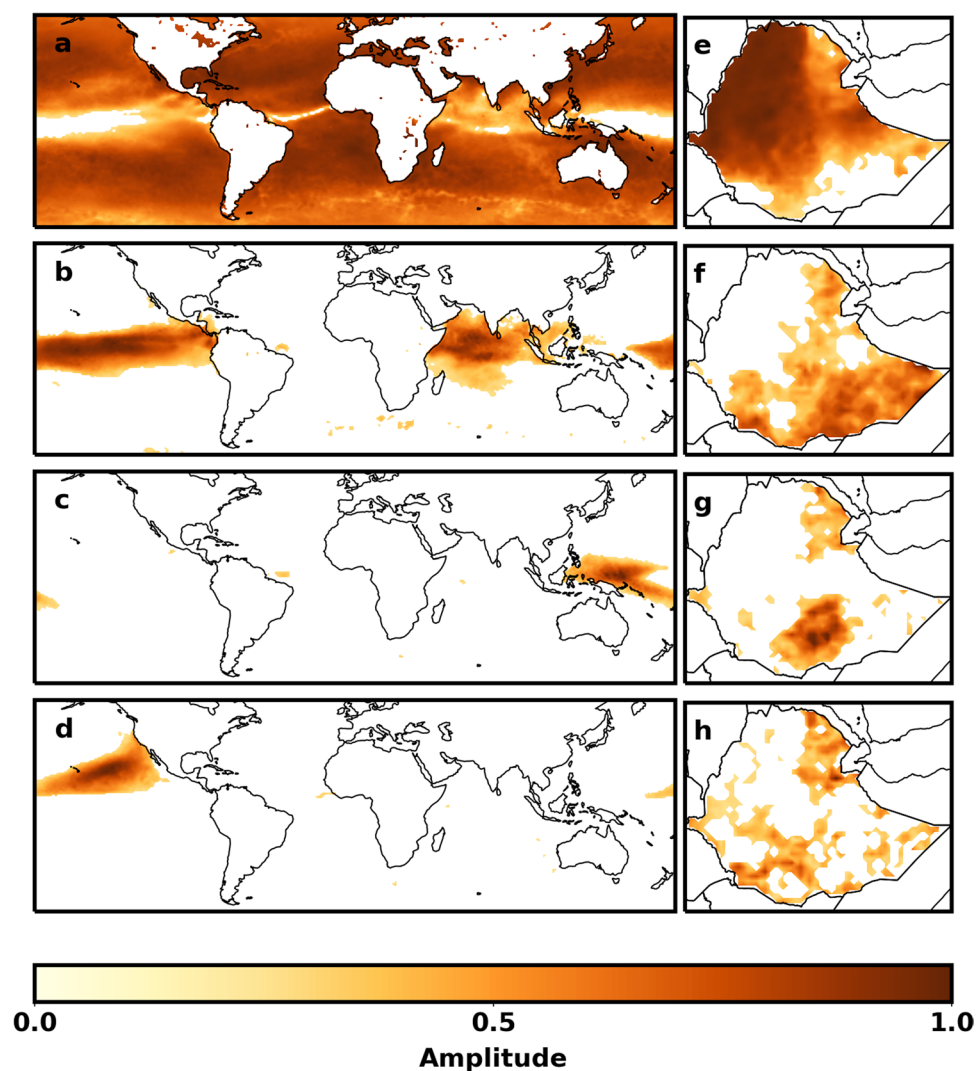
pacific and western Indian oceans SST anomalies are well identified by the phase functions shown in Figure S4 (b, f), whereas the southern part of the country's SM is directly correlated with these oceanic regions. Moreover, the second mode temporal variation is highly associated ($r = 0.72$) with the Oceanic Nino Index (ONI) as shown in Fig. 12b. A positive ENSO (El-Nino Southern Oscillation) phenomena that a higher than normal SST anomalies in the eastern and lower than normal SST anomalies in the western equatorial pacific ocean (El-Nino) or the vice versa (La-Nina) are associated with lower or higher SM anomalies in the northeastern part of the country, whereas the El-Nino condition is directly linked with the southern part of the country. The physical mechanism that links the negative association between the central Pacific Ocean SST anomalies and the SM in the northeastern part of the country involves change in atmospheric features as discussed (Camberlin 1997; Diro et al. 2011; Gleixner et al. 2017; Korecha and Barnston 2007; Segele et al. 2009).

According to these studies, positive ENSO conditions are linked to weakening of Tropical Easterly Jets (TEJ) and weaker upper level divergence. This weakening of TEJ and the upper level divergence leads to a reduction in rainfall and this in turn reduces soil moisture in the region or vice versa. It is also interesting to mention the second mode and the Dipole Mode Index (DMI), an area average of (10°S − 10°N, 50°E − 70°E minus 10°S − 0°N, 90°E − 110°E) Indian Ocean SST anomalies, significant level of relationship is $r = 0.43$ with p-value < 0.0001. It is a well established phenomenon that the Indian Ocean Dipole (IOD) is directly associated with OND rainfall in the south and southeastern part of the country (Bahaga et al. 2015; Behera et al. 2005; Black 2005; Liebmann et al. 2014; Lüdecke et al. 2021; Saji et al. 1999) that possibly reflected on the local SM.

The third mode (Figs. 11, S4, and 12c, g) shows SST anomalies over Oceania that directly affect the northeastern and central southern part of the country. This mode is clearly

Fig. 12 Principal components (PC1, PC2, PC3, PC4) of SST (left) and soil moisture (right) with global oceanic indices overlaid on SST PCs. The correlations between SST PCs and the global oceanic indices are Oceanic Nino Index (ONI) $r = 0.72$, Pacific Warm Pool (PACWARMPOOL) $r = 0.48$, and Pacific Decadal Oscillation (PDO) $r = -0.52$ for p-value < 0.0001. For readability and smoothly overlay the PDO index over SST PC4, we invented PC4 of SST and the corresponding SM PC4
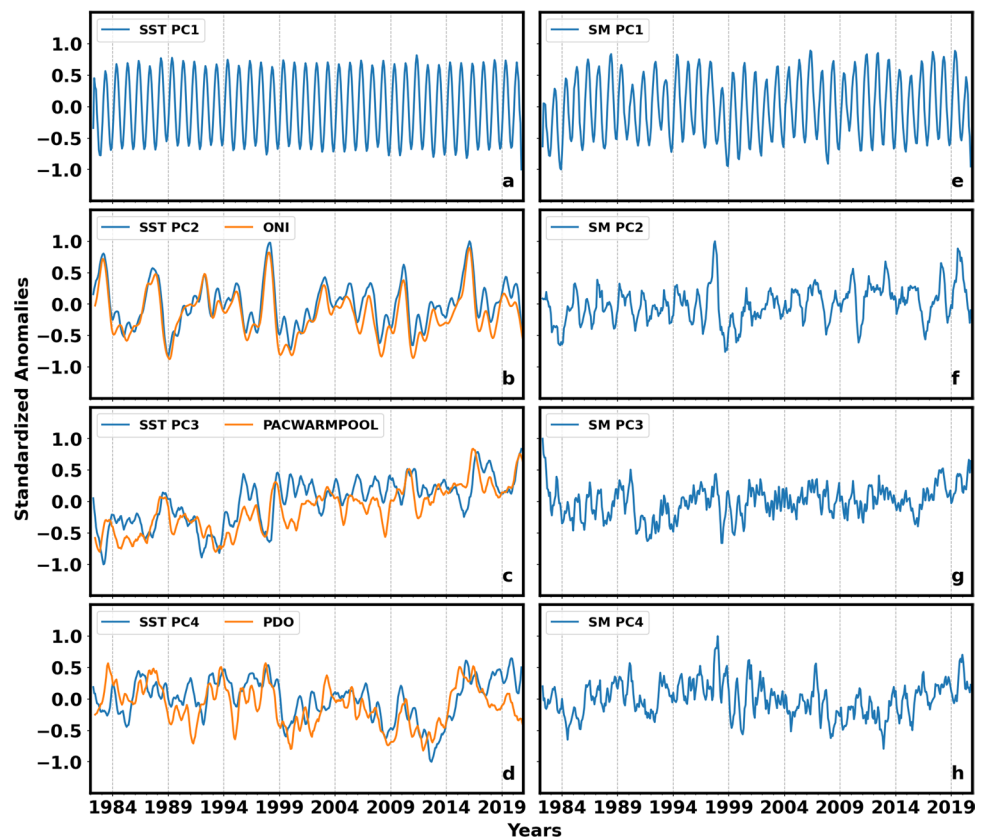
associated with Pacific Warm Pool (PACWARMPOOL) and the correlation between PACWARMPOOL index and SST anomalies time series ($r = 0.48$) over Oceania are presented in Fig. 12c. The pacific warm pool SST anomalies directly vary with the northeastern and central southern regions with slight negative and positive phase shift in these regions, respectively. It is also apparent that the warm pool has a positive link with the western tip of the country (Gambella region). However, unlike studies (Funk et al. 2008; Lyon and DeWitt 2012; Peterson et al. 2012; Williams et al. 2012) indicated the negative associationship between the south-central Indian and west Pacific Ocean SST anomalies and East African rainfall, the southern central and northeastern local SM has positive links with Pacific warm pool.

The fourth mode (Figs. 11, S4, and 12d, h) illustrated a widespread effect of the Pacific Decadal Oscillation (PDO) manifestation across the countrywide SM anomalies that were dominated by negative correlations (Fig. S4: h) all over the country. Though the physical mechanisms that drive this relationship are complex and not well understood, Lüdecke et al. (2021) and Taye and Willems (2012) also found that PDO has negatively correlated with Ethiopian rainfall. However, this mode also presented a direct correlation in the northeastern, some patches of central and southeastern regions, and along the narrow band in the southern tips of

the country. Figure 12d shows correlation between the PDO index and northeastern Pacific SST anomalies (SST PC4). The dominant influence of PDO in the southern region rainfall is discussed by Jury 2010, 2016, and it might also be the case that it reflects on the local SM via their direct relationship.

4 Summary and conclusions

Soil moisture plays a significant role in land–atmosphere interactions as it affects both energy exchange and water balance. Despite its importance, there is a lack of a reliant observational network of in situ SM stations. In the current study, the degree of agreement among station, reanalysis, and satellite estimates of soil moisture in representing the spatiotemporal characteristics is evaluated across Ethiopia. Furthermore, the spatiotemporal variability of soil moisture and its link with local and remote drivers is investigated.

Correlation and root-mean-square-error (RMSE) analyses have been used to select the best representative soil moisture dataset over the region. In addition, an empirical orthogonal functions (EOFs) analysis and K_{Means} clustering algorithm are employed to classify the country

(i.e., Ethiopia) into homogeneous soil moisture zones. Composite analysis and complex maximum covariance analysis (CMCA) are also applied to examine the link between soil moisture variability and the local (i.e., ET and RF) and remote (i.e., SST) driving factors.

Inter-comparison among the different datasets has shown that all soil moisture datasets have captured the spatiotemporal variability of soil moisture in a consistent manner throughout the country. However, the level of agreement differs in their absolute magnitudes and time of the peak. It is noted that the combined satellite datasets, unlike the others, show an earlier peak of the soil moisture. Our comparative analysis demonstrates that the FLDAS is a highly skilled dataset to reasonably represent the spatiotemporal characteristics of soil moisture across the country.

The EOF and K_{Means} analysis has resulted in five non-overlapping climate zones that have considerably high intra- and low inter-correlations among homogeneous climate zones. Annual cycle of soil moisture revealed a mono-modal soil moisture pattern in the western part (Reg-I) and bimodal pattern for the eastern and southern part (Reg-III and Reg-V) of the country. It is also noted that the northeastern tip of Ethiopia (i.e., the Dallol depression) is drier, whereas the central southwestern is wetter throughout the year. These soil moisture patterns are consistent with seasonal evolution of rainfall in the country (REF, (Degefu 1987; Diro et al. 2008; Fekadu 2015; Gissila et al. 2004)).

Our study showed that rainfall and evapotranspiration have strong relationships with soil moisture variation over the southern region as demonstrated in the leading modes of CMCA and composite analysis. Over these regions, higher soil moisture values are associated with higher values of rainfall and evapotranspiration. The significant positive correlation between rainfall and soil moisture imply that rainfall and soil moisture are strongly coupled for most of the southern part of Ethiopia. Correlation between evapotranspiration and temperature also revealed that atmosphere is the leading factor (rather than the soil moisture) in land–atmosphere interaction for the southeastern part of the country. This suggests that the positive correlation between ET and SM is triggered by intense moisture demand as a result of higher temperature and invokes a chain of atmospheric processes, whereby a higher temperature leads to higher evaporation and enhanced convection to increase soil moisture. Furthermore, analysis of the effects of global SST anomalies on local SM demonstrated that oceanic indices: ONI ($r = 0.72$), IOD ($r = 0.43$), PACWARMPOOL ($r = 0.48$), and PDO ($r = -0.52$) found to have significant relationships with local SM by modulating large-scale circulations and other local climatic variables (i.e., rainfall and surface temperature). However, individual oceanic regions vary at various time scales; therefore, we recommend further study to evaluate their respective role in the local SM.

This work is one of the few studies which, among other things, has characterized soil moisture variability in the horn of Africa and identified the most representative reanalysis dataset for the region. One caveat of the validation study is the short length of records in the station soil moisture dataset employed in the analysis. Therefore, further research using longer observational soil moisture data records would be beneficial.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Competing interests The authors declare no competing interests.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no conflict of interest.

References

An R et al (2016) Validation of the ESA CCI soil moisture product in China. Int J Appl Earth Obs Geoinf 48:28–36. https://doi.org/10.1016/j.jag.2015.09.009

Arthur DS, Vassilvitskii (2006) K-means++: The advantages of careful seeding. Stanford InfoLab; Stanford. http://ilpubs.stanford.edu:8090/778/

Ayehu G, Tadesse T, Gessesse B, Yigrem Y (2019) Soil moisture monitoring using remote sensing data and a stepwise-cluster prediction model: the case of Upper Blue Nile Basin, Ethiopia. Remote Sens 11:125. https://doi.org/10.3390/rs11020125

Ayehu G, Tadesse T, Gessesse B (2020) Monitoring residual soil moisture and its association to the long-term variability of rainfall over the Upper Blue Nile Basin in Ethiopia. Remote Sens 12:2138. https://doi.org/10.3390/rs12132138

Bahaga TK, Tsidu GM, Kucharski F, Diro GT (2015) Potential predictability of the sea-surface temperature forced equatorial east african short rains interannual variability in the 20th century. Q J R Meteorol Soc 141:16–26. https://doi.org/10.1002/qj.2338

Bahmani B, Moseley B, Vattani A, Kumar R and Vassilvitskii S (2012) Scalable k-means++. arXiv preprint arXiv:1203.6402. https://doi.org/10.48550/ARXIV.1203.6402

Baldwin MP, Stephenson DB, Jolliffe IT (2009) Spatial weighting and iterative projection methods for EOFs. J Clim 22:234–243

Basist A, Grody NC, Peterson TC, Williams CN (1998) Using the special sensor microwave/imager to monitor land surface temperatures, wetness, and snow cover. J Appl Meteorol. https://doi.org/10.1175/1520-0450(1998)037%3c0888:UTSSMI%3e2.0.CO;2

Behera SK, Luo J-J, Masson S, Delecluse P, Gualdi S, Navarra A, Yamagata T (2005) Paramount impact of the Indian Ocean Dipole on the East African short rains: a CGCM study. J Clim 18:4514–4530. https://doi.org/10.1175/JCLI3541.1

Black E (2005) The relationship between Indian Ocean sea–surface temperature and East African rainfall. Philos Trans R Soc A Math Phys Eng Sci 363:43–47. https://doi.org/10.1098/rsta.2004.1474

Boashash B (1992) Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals. Proc IEEE 80:520–538. https://doi.org/10.1109/5.135376

C3S SM Product User Guide (2018) Product user guide and specification. Land service: Soil moisture ECV, v201812. https://doi.org/10.24381/cds.d7782f18

Camberlin P (1997) Rainfall anomalies in the source region of the Nile and their connection with the Indian summer monsoon. J Clim 10:1380–1392. https://doi.org/10.1175/1520-0442(1997)010-1380:RAITSR-2.0.CO;2

Cashion JE, Lakshmi V, Bosch D (2003) Use of TRMM microwave imager (TMI) to characterize soil moisture for the little river watershed. AGU Fall Meeting Abstracts 2003:H21C–H6

Dawson A (2016) eofs: a library for EOF analysis of meteorological, oceanographic, and climate data. J Open Res Softw 4

Degefu W (1987) Some aspects of meteorological drought in Ethiopia. Cambridge University Press Cambridge, UK

Diro GT, Black E, Grimes DIF (2008) Seasonal forecasting of Ethiopian spring rains. Meteorol Appl A J Forecast Pract Appl Train Tech Model 15:73–83. https://doi.org/10.1002/met.63

Diro GT, Grimes DIF, Black E, O'Neill A, Pardo-Iguzquiza E (2009) Evaluation of reanalysis rainfall estimates over Ethiopia. Int J Climatol A J R Meteorol Soc 29:67–78. https://doi.org/10.1002/joc.1699

Diro GT, Grimes DIF, Black E (2011) Teleconnections between Ethiopian summer rainfall and sea surface temperature: part I observation and modelling. Clim Dyn 37:103–119. https://doi.org/10.1007/s00382-010-0837-8

Diro GT, Martynov SA, Jeong DI, Verseghy D, Winger K (2014) Land-atmosphere coupling over North America in CRCM5. J Geophys Res Atmos 119(21). https://doi.org/10.1002/2014JD021677

Dommenget D, Latif M (2002) A cautionary note on the interpretation of EOFs. J Clim 15:216–225. https://doi.org/10.1175/1520-0442(2002)015%3c0216:ACNOTI%3e2.0.CO;2

Dorigo WA et al (2011) International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. Hydrol Earth Syst Sci 15:1675–1698. https://doi.org/10.5194/hess-15-1675-2011

Dorigo W et al (2017) ESA CCI Soil Moisture for improved Earth system understanding: state-of-the art and future directions. Remote Sens Environ 203:185–215. https://doi.org/10.1016/j.rse.2017.07.001

Drewitt G, Berg AA, Merryfield WJ, Lee WS (2012) Effect of realistic soil moisture initialization on the Canadian CanCM3 seasonal forecast model. Atmos - Ocean 50:466–474. https://doi.org/10.1080/07055900.2012.722910

Eltahir EAB (1998) A soil moisture-rainfall feedback mechanism 1. Theory and Observations. Water Resour Res 34:765–776. https://doi.org/10.1029/97WR03499

Fekadu K (2015) Ethiopian seasonal rainfall variability and prediction using canonical correlation analysis (CCA). Earth Sci 4:112. https://doi.org/10.11648/j.earth.20150403.14

Feng H, Liu Y (2015) Combined effects of precipitation and air temperature on soil moisture in different land covers in a humid basin. J Hydrol 531:1129–1140. https://doi.org/10.1016/j.jhydrol.2015.11.016

Findell KL, Eltahir EAB (1999) Analysis of the pathways relating soil moisture and subsequent rainfall in Illinois. J Geophys Res Atmos 104:31565–31574. https://doi.org/10.1029/1999JD900757

Fiorucci JA, Pellegrini TR, Louzada F, Petropoulos F, Koehler AB (2016) Models for optimising the theta method and their relationship to state space models. Int J Forecast 32:1151–1161. https://doi.org/10.1016/j.ijforecast.2016.02.005

Funk C, Dettinger MD, Michaelsen JC, Verdin JP, Brown ME, Barlow M, Hoell A (2008) Warming of the Indian Ocean threatens eastern and southern African food security but could be mitigated by agricultural development. Proc Natl Acad Sci 105:11081–11086. https://doi.org/10.1073/pnas.0708196105

Funk C et al (2015) The climate hazards infrared precipitation with stations - a new environmental record for monitoring extremes. Scientific Data 2:1–21. https://doi.org/10.1038/sdata.2015.66

Gaur N, Mohanty BP (2013) Evolution of physical controls for soil moisture in humid and subhumid watersheds. Water Resour Res 49:1244–1258. https://doi.org/10.1002/wrcr.20069

Gelaro R, Max MC, Ricardo S, Andrea T, Lawrence M, Cynthia TA, Anton R, Michael D, Rolf B, Krzysztof R, Lawrence W, Richard C, Clara C, Santha D, Virginie A, Austin B, Arlindo CM, Wei d, Gi-Kong G, Randal K, Robert K, Dagmar L, Jon Eric M, Gary N, Steven P, William P, Michele P, Siegfried R, Meta S, Bin S, Zhao (2017) The modern-era retrospective analysis for research and applications version 2 (MERRA-2). J Clim 30(14):5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Gissila T, Black E, Grimes DIF, Slingo JM (2004) Seasonal forecasting of the Ethiopian summer rains. Int J Climatol A J R Meteorol Soc 24:1345–1358. https://doi.org/10.1002/joc.1078

Gleixner S, Keenlyside N, Viste E, Korecha D (2017) The El Niño effect on Ethiopian summer rainfall. Clim Dyn 49:1865–1883. https://doi.org/10.1007/s00382-016-3421-z

Gruber A, Dorigo WA, Crow W, Wagner W (2017) Triple collocation-based merging of satellite soil moisture retrievals. IEEE Trans Geosci Remote Sens 55:6780–6792. https://doi.org/10.1109/TGRS.2017.2734070

Guo Z et al (2006) GLACE: The global land-atmosphere coupling experiment. Part II: Analysis. J Hydrometeorol 7:611–625. https://doi.org/10.1175/JHM511.1

Hanke M, Redler R, Holfeld T, Yastremsky M (2016) YAC 1.2.0: New aspects for coupling software in Earth system modelling. Geosci Model Dev 9:2755–2769. https://doi.org/10.5194/gmd-9-2755-2016

Hannachi A, Jolliffe IT, Stephenson DB (2007) Empirical orthogonal functions and related techniques in atmospheric science: a review. Int J Climatol A J R Meteorol Soc 27:1119–1152

Hannachi A, Unkel S, Trendafilov NT, Jolliffe IT (2009) Independent component analysis of climate data: a new look at EOF rotation. J Clim 22:2797–2812. https://doi.org/10.1175/2008JCLI2571.1

Hersbach H et al (2019) Global reanalysis: Goodbye ERA-Interim, hello ERA5. ECMWF Newsl 159:17–24

Hirsch AL, Pitman AJ, Kala J (2014) The role of land cover change inmodulating the soil moisture-temperature land-atmosphere coupling strength over Australia. Geophys Res Lett 41:5883–5890. https://doi.org/10.1002/2014GL061179

Imaoka K, Maeda T, Kachi M, Kasahara M, Ito N, Nakagawa K (2012) Status of AMSR2 instrument on GCOM-W1. Development, implementation, and characterization II, Earth observing missions and sensors

Jollife IT, Cadima J (2016) Principal component analysis: a review and recent developments. Philos Trans R Soc A Math Phys Eng Sci 374 https://doi.org/10.1098/rsta.2015.0202

Jones PW (1999) First- and second-order conservative remapping schemes for grids in spherical coordinates. Mon Weather Rev 127:2204–2210. https://doi.org/10.1175/1520-0493(1999)127%3c2204:FASOCR%3e2.0.CO;2

Jury MR (2010) Ethiopian decadal climate variability. Theor Appl Climatol 101:29–40. https://doi.org/10.1007/s00704-009-0200-3

Jury MR (2016) Determinants of southeast Ethiopia seasonal rainfall. Dyn Atmos Oceans 76:63–71. https://doi.org/10.1016/j.dynatmoce.2016.08.004

Kerr YH et al (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

Korecha D, Barnston AG (2007) Predictability of June–September rainfall in Ethiopia. Mon Weather Rev 135:628–650. https://doi.org/10.1175/MWR3304.1

Korres W, Koyama CN, Fiener P, Schneider K (2010) Analysis of surface soil moisture patterns in agricultural landscapes using Empirical Orthogonal Functions. Hydrol Earth Syst Sci 14:751–764. https://doi.org/10.5194/hess-14-751-2010

Koster RD et al (2004) Regions of strong coupling between soil moisture and precipitation. Science 305:1138–1140. https://doi.org/10.1126/science.1100217

Koster RD, Guo Z, Yang R, Dirmeyer PA, Mitchell K, Puma MJ (2009) On the nature of soil moisture in land surface models. J Clim 22:4322–4335. https://doi.org/10.1175/2009JCLI2832.1

Koster RD et al (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

Koster RD et al (2010) Contribution of land surface initialization to subseasonal forecast skill: first results from a multi-model experiment. Geophys Res Lett 37, n/a–n/a, https://doi.org/10.1029/2009GL041677

Li L et al (2010) WindSat global soil moisture retrieval and validation. IEEE Trans Geosci Remote Sens 48:2224–2241. https://doi.org/10.1109/TGRS.2009.2037749

Li MX, Ma ZG, Niu G-YY (2011) Modeling spatial and temporal variations in soil moisture in China. Chin Sci Bull 56:1809–1820. https://doi.org/10.1007/s11434-011-4493-0

Liebmann B et al (2014) Understanding recent Eastern Horn of Africa rainfall variability and change. J Clim 27:8630–8645. https://doi.org/10.1175/JCLI-D-13-00714.1

Liu Q et al (2011) The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system. J Hydrometeorol 12:750–765. https://doi.org/10.1175/JHM-D-10-05000.1

Liu Y, Liu Y, Wang W, Fan X, Cui W (2022) Soil moisture droughts in East Africa: spatiotemporal patterns and climate drivers. J Hydrol Reg Stud 40:101013. https://doi.org/10.1016/j.ejrh.2022.101013

Lüdecke H-J, Müller-Plath G, Wallace MG, Lüning S (2021) Decadal and multidecadal natural variability of African rainfall. J Hydrol Reg Stud 34:100795. https://doi.org/10.1016/j.ejrh.2021.100795

Lyon B, DeWitt DG (2012) A recent and abrupt decline in the East African long rains. Geophys Res Lett39 https://doi.org/10.1029/2011GL050337

McNally A et al (2017) A land data assimilation system for sub-Saharan Africa food and water security applications. Scientific Data 4:1–19. https://doi.org/10.1038/sdata.2017.12

McNally A (2018) NASA/GSFC/HSL, FLDAS Noah L and Surface Model L4 Global Monthly 0.1° × 0.1° degree (MERRA-2 and CHIRPS), Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC). 10.5067/5NHC22T9375G

Mekonnen DF (2009) Satellite remote sensing for soil moisture estimation: Gumara catchment. Ethiopia, ITC

Merchant CJ et al (2019) Satellite-based time-series of sea-surface temperature since 1981 for climate applications. Scientific Data 6:223. https://doi.org/10.1038/s41597-019-0236-x

Muñoz Sabater J (2019) ERA5-L and hourly data from 1981 to present. Copernicus Climate Change Cervice (C3S Climate Data Store (CDS). https://doi.org/10.24381/cds.e2161bac.

Nicolai-Shaw N, Gudmundsson L, Hirschi M, Seneviratne SI (2016) Long-term predictability of soil moisture dynamics at the global scale: persistence versus large-scale drivers. Geophys Res Lett 43:8554–8562. https://doi.org/10.1002/2016GL069847

Njoku EG, Stacey JM, Barath FT (1980) The seasat scanning multi-channel microwave radiometer (SMMR): instrument description and performance. IEEE J Oceanic Eng 5:100–115. https://doi.org/10.1109/JOE.1980.1145458

Njoku EG, Jackson TJ, Lakshmi V, Chan TK, Nghiem SV (2003) Soil moisture retrieval from AMSR-E. IEEE Trans Geosci Remote Sens 41:215–229. https://doi.org/10.1109/TGRS.2002.808243

North GR, Bell TL, Cahalan RF, Moeng FJ (1982) Sampling errors in the estimation of empirical orthogonal functions. Mon Weather Rev 110:699–706. https://doi.org/10.1175/1520-0493(1982)110%3c0699:SEITEO%3e2.0.CO;2

Peng J, Loew A, Merlin O, Verhoest NEC (2017) A review of spatial downscaling of satellite remotely sensed soil moisture. Rev Geophys 55:341–366. https://doi.org/10.1002/2016RG000543

Peterson TC, Stott PA, Herring S (2012) Explaining extreme events of 2011 from a climate perspective. Bull Am Meteor Soc 93:1041–1067. https://doi.org/10.1175/BAMS-D-12-00021.1

Qiu J, Gao Q, Wang S, Su Z (2016) Comparison of temporal trends from multiple soil moisture data sets and precipitation: the implication of irrigation on regional soil moisture trend. Int J Appl Earth Obs Geoinf 48:17–27. https://doi.org/10.1016/j.jag.2015.11.012

Richman MB (1986) Rotation of principal components. J Climatol 6:293–335. https://doi.org/10.1002/joc.3370060305

Rieger N, Corral A, Olmedo E, Turiel A (2021) Lagged teleconnections of climate variables identified via complex rotated Maximum Covariance Analysis. J Clim 1:1–59. https://doi.org/10.1175/jcli-d-21-0244.1

Rodell M, Houser U, Jambor J, Gottschalck K, Mitchell CJ, Meng K, Arsenault B, Cosgrove J, Radakovich M, Bosilovich JK, Entin JP,

Walker D, Lohmann D, Toll (2004) The global land data assimilation system. Bull Am Meteorol Soc 85(3):381–394. https://doi.org/10.1175/BAMS-85-3-381

Saji NH, Goswami BN, Vinayachandran PN, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. Nature 401:360–363. https://doi.org/10.1038/43854

Schär C, Lüthi D, Beyerle U, Heise E (1999) The soil-precipitation feedback: a process study with a regional climate model. J Clim 12:722–741. https://doi.org/10.1175/1520-0442(1999)012%3c0722:tspfap%3e2.0.co;2

Segele ZT, Lamb PJ, Leslie LM (2009) Large-scale atmospheric circulation and global sea surface temperature associations with Horn of Africa June–September rainfall. Int J Climatol A J R Meteorol Soc 29:1075–1100. https://doi.org/10.1002/joc.1751

Seneviratne SI et al (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 SI, Corti T, Davin EL, Hirschi M, Jaeger EB, Lehner I, Orlowsky B, Teuling AJ (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

Srivastava PK, Han D, Rico-Ramirez MA, O'Neill P, Islam T, Gupta M, Dai Q (2015) Performance evaluation of WRF-Noah Land surface model estimated soil moisture for hydrological application: synergistic evaluation using SMOS retrieved soil moisture. J Hydrol 529:200–212. https://doi.org/10.1016/j.jhydrol.2015.07.041

Taye MT, Willems P (2012) Temporal variability of hydroclimatic extremes in the Blue Nile basin. Water Resour Res 48 https://doi.org/10.1029/2011WR011466

Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. J Geophys Res Atmos 106:7183–7192. https://doi.org/10.1029/2000JD900719

Taylor CM, de Jeu RAM, Guichard F, Harris PP, Dorigo WA (2012) Afternoon rain more likely over drier soils. Nature 489:423–426. https://doi.org/10.1038/nature11377

Teuling AJ, Troch PA (2005) Improved understanding of soil moisture variability dynamics. Geophys Res Lett 32:1–4. https://doi.org/10.1029/2004GL021935

Thomas JA, Berg AA, Merryfield WJ (2016) Influence of snow and soil moisture initialization on sub-seasonal predictability and forecast skill in boreal spring. Clim Dyn 47:49–65. https://doi.org/10.1007/s00382-015-2821-9

Wagner W, Dorigo W, de Jeu R, Fernandez D, Benveniste J, Haas E, Ertl M et al (2012) Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture. ISPRS Ann Photogramm Remote Sens Spat Inf Sci (ISPRS Ann) 7:315–321

Wagner W et al (2013) The ASCAT soil moisture product: a review of its specifications, validation results, and emerging applications. Meteorol Z 22:1–29. https://doi.org/10.1127/0941-2948/2013/0399

Wang S, Mo X, Liu S, Lin Z, Hu S (2016) Validation and trend analysis of ECV soil moisture data on cropland in North China Plain during 1981. Int J Appl Earth Obs Geoinf 48:110–121. https://doi.org/10.1016/j.jag.2015.10.010

Wei J, Dirmeyer PA (2012) Dissecting soil moisture-precipitation coupling. Geophys Res Lett 39.

Williams AP et al (2012) Recent summer precipitation trends in the Greater Horn of Africa and the emerging role of Indian Ocean sea surface temperature. Clim Dyn 39:2307–2328. https://doi.org/10.1007/s00382-011-1222-y

Zeleke T, Giorgi F, Tsidu GM, Diro GT (2013) Spatial and temporal variability of summer rainfall over Ethiopia from observations and a regional climate model experiment. Theoret Appl Climatol 111:665–681. https://doi.org/10.1007/s00704-012-0700-4

Zheng X, Eltahir EAB (1998) A soil moisture-rainfall feedback mechanism 2. Numerical Experiments. Water Resour Res 34:777–785. https://doi.org/10.1029/97WR03497

Zhu S, Chen H, Dong X, Wei J (2020) Influence of persistence and oceanic forcing on global soil moisture predictability. Clim Dyn 54:3375–3385. https://doi.org/10.1007/s00382-020-05184-8