ABSTRACT: Soil moisture (SM) links the water and energy cycles over the land–atmosphere interface and largely determines ecosystem functionality, positioning it as an essential player in the Earth system. Despite its importance, accurate estimation of large-scale SM remains a challenge. Here we leverage the strength of neural network (NN) and fidelity of long-term measurements to develop a daily multilayer cropland SM dataset for China from 1981 to 2013, implemented for a range of different cropping patterns. The training and testing of the NN for the five soil layers (0–50 cm, 10-cm depth each) yield $R^2$ values of 0.65–0.70 and 0.64–0.69, respectively. Our analysis reveals that precipitation and soil properties are the two dominant factors determining SM, but cropping pattern is also crucial. In addition, our simulations of alternative cropping patterns indicate that winter wheat followed by fallow will largely alleviate the SM depletion in most parts of China. On the other hand, cropping patterns of fallow in the winter followed by maize/soybean seem to further aggravate SM decline in the Huang-Huai-Hai region and southwestern China, relative to prevalent practices of double cropping. This may be due to their low soil porosity, which results in more soil water drainage, as opposed to the case that winter crop roots help maintain SM. This multilayer cropland SM dataset with granularity of cropping patterns provides an important alternative and is complementary to modeled and satellite-retrieved products.

KEYWORDS: Soil moisture; Hydrologic models; Agriculture; Crop growth; Machine learning

1. Introduction

Soil moisture (SM) is an essential component of the Earth system. It affects the variability of the coupled energy (latent and sensible heat fluxes) and water fluxes (runoff and evapotranspiration) by modifying the partitioning of water and energy across the land–atmosphere interface (Seneviratne et al. 2010). The effects of SM on evapotranspiration also impact temperature variability and may intrigue persistent heatwaves (Fischer et al. 2007; Hirschi et al. 2011; Miralles et al. 2014), and they may also affect the intensity, frequency, and distribution of precipitation (Findell et al. 2011; Guillod et al. 2015; Taylor et al. 2012). Further, a shortage or excess of SM could trigger the occurrence of droughts (Wang et al. 2011) or floods (Koster et al. 2010). As such, SM is crucial for weather prediction, climate forecasting and ecosystem dynamics assessment. Moreover, SM is vital for agricultural production since it is the only direct source for crop water demand, and thus SM could also be a crucial factor affecting socioeconomic conditions.

Despite the criticality of SM in the Earth system, accurate estimation of large-scale soil moisture is still a challenge, mainly due to its rapid fluctuations and the lack of sufficient ground truth observations. Currently, most large-scale SM products are either retrieved from satellite data or produced from land surface models (LSMs). As an example of product derived from satellites, the European Space Agency (ESA) Climate Change Initiative (CCI) soil moisture product provides harmonized global daily SM for the period of 1978–2017, by combining various single-sensor active and passive microwave soil moisture products (Dorigo et al. 2017). Meanwhile, LSMs also simulate soil moisture at large scale. For instance, the Variable Infiltration Capacity (VIC) land surface model has been used to estimate global SM (Nijssen et al. 2001). Nonetheless, satellite SM retrievals are usually jeopardized by interference of dense plant canopy and limitations in the depth of detection within the soil column (Entekhabi et al. 2014), as well as sensor limitations (especially before the use of L-band radiometers). Likewise, the fidelity of LSM simulations may be undermined by model errors (e.g., model structure error,
model parameter error), forcing data uncertainties and limited usage of ground-based observations for model calibration and validation (Xia et al. 2014). Thus, efforts that aim to improve large-scale SM product are essential.

With the largest population in the world, SM shortage is an imminent concern in China because of its inherent effects on food production (Khan et al. 2009; Ye et al. 2013) and the fact that food security is one of the upmost important national issues (Brown and Halweil 1998; Wong and Huang 2012). In the past three decades, SM extremes—droughts—have affected an average area of 23.4 million ha yr$^{-1}$ and 15.5% of agricultural planting area in China (National Bureau of Statistics of China 2020). Further, recent studies have revealed that dry extremes are prone to become more frequent and intense in China (Elliott et al. 2014; Trenberth et al. 2014). At the same time, since China spans a vast area, environmental factors such as soil properties (e.g., soil texture, structure, organic matter, depth, density, and salinity) (Shangguan et al. 2013), climate (e.g., precipitation, solar radiation, temperature, etc.) (Chen et al. 2011; Dee et al. 2011), and topography vary greatly across the cropland regions in the country (Danielson and Gesch 2011), as well as the farming practices (Chen et al. 2009; Li and Wenhua 2001). Consequently, cropland SM tends to be highly variable in both space and time, at scale smaller than that captured by remote sensing or LSMS. Thus, improved understanding of the intricate characteristics of cropland SM across China and possible strategies of ameliorating soil desiccation is of great importance and merits further attention.

Although long-term SM observations from agrometeorological stations depict a general picture of SM variations across China from ground truth (Jia et al. 2018; Liu et al. 2015), the observations are sparse in terms of spatial and temporal coverage. Meanwhile, a thorough understanding of SM variations across China and their characteristics requires a reasonable SM product with national coverage and long-term continuity. Since both LSM-simulated and satellite-derived SM products have their own aforementioned shortcomings and usually do not differentiate crop types or cropping patterns (i.e., the yearly sequence and spatial arrangement of crops and fallow on a particular land area), we seek a new way to construct an alternative long-term SM product with information of cropping patterns over all of China. To that end, we leverage the fidelity of ground-truth measurements and the power of machine learning.

Recent years have seen an escalating deployment of machine learning techniques in Earth science (Reichstein et al. 2019). By predicting temporally varying target variables in land, ocean and atmosphere domains from temporally varying features, machine learning has been actively used to study Earth system dynamics. Particularly, compared to previous mechanistic or semiempirical modeling approaches, machine learning methods have been proven to be more powerful and flexible when inferring continental or global estimates from point observations, such as predicting carbon and water fluxes (Jung et al. 2011; Papale and Valentini 2003). Furthermore, those machine-learning-predicted fluxes can be used as important benchmarks to assess land surface and climate models (Bonan et al. 2011; Jung et al. 2010, 2017). Besides, machine learning can help to detect, visualize and understand the patterns of model error, which allows correction of model outputs accordingly (McGovern et al. 2017). In addition, there are recent studies using neural network (NN) to estimate large-scale SM from satellite observations (Fang et al. 2017; Kolassa et al. 2018; Rodriguez-Fernandez et al. 2015). However, those studies do not focus on China, and typically have a short time span because of the limited availability of satellite data. Furthermore, all these studies have not explored the effects of cropping patterns on SM. Thus, an alternative SM product for China which combines the long-term ground-truth measurements and machine learning, with information of cropping patterns, fuels a potential of improving large-scale and long-term SM estimates.

Motivated by the needs of an improved understanding of SM characteristics in the vast cropland area of China and potential strategies in mitigating SM decline, this study has two main objectives. First, we aim to develop a new daily multilayer (0–50 cm, 10-cm depth each layer) SM product for China over 1981–2013, with granularity of cropping patterns. The new SM product will provide an opportunity to explore a variety of research topics such as water and carbon fluxes in agroecosystems, and land–atmosphere interactions. Second, we shed light on the effects of different cropping patterns on SM and seek alternative cropping patterns that may alleviate SM decline. This exploration will prompt decision making in ameliorating soil water stress and inform adaptation plans in an increasingly water-scarce future.

2. Methods

a. Neural network

The neural network used in this work is a deep feed forward NN, which consists of multiple hidden layers and multiple units in each layer (Reichstein et al. 2019). We develop this NN using Tensorflow in Python (Abadi et al. 2016). The schematic diagram in Fig. 1 illustrates the structure of our NN.

Our NN model takes a variety of environmental features (input variables) that closely affect cropland SM, including climatic variables, soil properties, topography, and farming practices (Brady and Weil 2008; Huang et al. 2011; Liu et al. 2015), as detailed in section 2b (see also Table 1), to predict daily SM. Cropping pattern is treated as a (nominal) categorical input in the NN. To account for the ecosystem memory effect on SM (Anderegg et al. 2015; Hughes et al. 2019; Ogle et al. 2015), we also include accumulated precipitation within six previous periods (i.e., previous 1, 2, 3, 4, 5, and 12 months) in the features. Note that for deeper soil layers, we also use the SM prediction in the adjacent upper layer as an additional input feature (e.g., SM prediction at 20–30-cm soil layer is a feature for SM prediction at 30–40-cm soil layer) to account for the strong autocorrelation between those variables (Salvucci and Entekhabi 2011). In addition, date of year is used as an input to track the different measurement dates at different stations, and also to account for the seasonality of SM.
All the valid SM observations (missing data are removed) at each of the five soil layers have been used to train a NN for each corresponding soil layer. In total, five NNs are developed and separately trained on the data pertaining to each of the five soil layers. For a given soil layer, a NN model is trained on 70% of the SM observations, which are randomly sampled as “truth,” while the remaining 30% of the SM observations are used to test the trained model’s performance. Out of the 70% training set, 10% of the data are used as validation set.

The hyper parameters in the NN are tuned to improve its performance in predicting SM. Specifically, we conduct a grid search for key hyperparameters—batch size, learning rate, hidden layers, and number of units in each layer—for each soil layer (see Table 2) (Gupta and Raza 2020). For the grid search, the optional values for each parameter are 1) 128, 512, 1024, and 2048 for batch size; 2) 0.0005, 0.001, 0.005, 0.01, and 0.015 for learning rate; 3) 1, 2, and 3 for hidden layers; and 4) 16, 32, 64, and 96 for number of units in each hidden layer. The set of hyper parameters is selected by choosing the one that achieve the lowest loss function value (mean squared error) on the validation set.

The hyper parameters in the NN are tuned to improve its performance in predicting SM. Specifically, we conduct a grid search for key hyperparameters—batch size, learning rate, hidden layers, and number of units in each layer—for each soil layer (see Table 2) (Gupta and Raza 2020). For the grid search, the optional values for each parameter are 1) 128, 512, 1024, and 2048 for batch size; 2) 0.0005, 0.001, 0.005, 0.01, and 0.015 for learning rate; 3) 1, 2, and 3 for hidden layers; and 4) 16, 32, 64, and 96 for number of units in each hidden layer. The set of hyper parameters is selected by choosing the one that achieve the lowest loss function value (mean squared error) on the validation set.

The hyper parameters in the NN are tuned to improve its performance in predicting SM. Specifically, we conduct a grid search for key hyperparameters—batch size, learning rate, hidden layers, and number of units in each layer—for each soil layer (see Table 2) (Gupta and Raza 2020). For the grid search, the optional values for each parameter are 1) 128, 512, 1024, and 2048 for batch size; 2) 0.0005, 0.001, 0.005, 0.01, and 0.015 for learning rate; 3) 1, 2, and 3 for hidden layers; and 4) 16, 32, 64, and 96 for number of units in each hidden layer. The set of hyper parameters is selected by choosing the one that achieve the lowest loss function value (mean squared error) on the validation set.

The hyper parameters in the NN are tuned to improve its performance in predicting SM. Specifically, we conduct a grid search for key hyperparameters—batch size, learning rate, hidden layers, and number of units in each layer—for each soil layer (see Table 2) (Gupta and Raza 2020). For the grid search, the optional values for each parameter are 1) 128, 512, 1024, and 2048 for batch size; 2) 0.0005, 0.001, 0.005, 0.01, and 0.015 for learning rate; 3) 1, 2, and 3 for hidden layers; and 4) 16, 32, 64, and 96 for number of units in each hidden layer. The set of hyper parameters is selected by choosing the one that achieve the lowest loss function value (mean squared error) on the validation set.

The hyper parameters in the NN are tuned to improve its performance in predicting SM. Specifically, we conduct a grid search for key hyperparameters—batch size, learning rate, hidden layers, and number of units in each layer—for each soil layer (see Table 2) (Gupta and Raza 2020). For the grid search, the optional values for each parameter are 1) 128, 512, 1024, and 2048 for batch size; 2) 0.0005, 0.001, 0.005, 0.01, and 0.015 for learning rate; 3) 1, 2, and 3 for hidden layers; and 4) 16, 32, 64, and 96 for number of units in each hidden layer. The set of hyper parameters is selected by choosing the one that achieve the lowest loss function value (mean squared error) on the validation set.

All the valid SM observations (missing data are removed) at each of the five soil layers have been used to train a NN for each corresponding soil layer. In total, five NNs are developed and separately trained on the data pertaining to each of the five soil layers. For a given soil layer, a NN model is trained on 70% of the SM observations, which are randomly sampled as “truth,” while the remaining 30% of the SM observations are used to test the trained model’s performance. Out of the 70% training set, 10% of the data are used as validation set.

The hyper parameters in the NN are tuned to improve its performance in predicting SM. Specifically, we conduct a grid search for key hyperparameters—batch size, learning rate, hidden layers, and number of units in each layer—for each soil layer (see Table 2) (Gupta and Raza 2020). For the grid search, the optional values for each parameter are 1) 128, 512, 1024, and 2048 for batch size; 2) 0.0005, 0.001, 0.005, 0.01, and 0.015 for learning rate; 3) 1, 2, and 3 for hidden layers; and 4) 16, 32, 64, and 96 for number of units in each hidden layer. The set of hyper parameters is selected by choosing the one that achieve the lowest loss function value (mean squared error) on the validation set.

The hyper parameters in the NN are tuned to improve its performance in predicting SM. Specifically, we conduct a grid search for key hyperparameters—batch size, learning rate, hidden layers, and number of units in each layer—for each soil layer (see Table 2) (Gupta and Raza 2020). For the grid search, the optional values for each parameter are 1) 128, 512, 1024, and 2048 for batch size; 2) 0.0005, 0.001, 0.005, 0.01, and 0.015 for learning rate; 3) 1, 2, and 3 for hidden layers; and 4) 16, 32, 64, and 96 for number of units in each hidden layer. The set of hyper parameters is selected by choosing the one that achieve the lowest loss function value (mean squared error) on the validation set.

The hyper parameters in the NN are tuned to improve its performance in predicting SM. Specifically, we conduct a grid search for key hyperparameters—batch size, learning rate, hidden layers, and number of units in each layer—for each soil layer (see Table 2) (Gupta and Raza 2020). For the grid search, the optional values for each parameter are 1) 128, 512, 1024, and 2048 for batch size; 2) 0.0005, 0.001, 0.005, 0.01, and 0.015 for learning rate; 3) 1, 2, and 3 for hidden layers; and 4) 16, 32, 64, and 96 for number of units in each hidden layer. The set of hyper parameters is selected by choosing the one that achieve the lowest loss function value (mean squared error) on the validation set.

The hyper parameters in the NN are tuned to improve its performance in predicting SM. Specifically, we conduct a grid search for key hyperparameters—batch size, learning rate, hidden layers, and number of units in each layer—for each soil layer (see Table 2) (Gupta and Raza 2020). For the grid search, the optional values for each parameter are 1) 128, 512, 1024, and 2048 for batch size; 2) 0.0005, 0.001, 0.005, 0.01, and 0.015 for learning rate; 3) 1, 2, and 3 for hidden layers; and 4) 16, 32, 64, and 96 for number of units in each hidden layer. The set of hyper parameters is selected by choosing the one that achieve the lowest loss function value (mean squared error) on the validation set.

The hyper parameters in the NN are tuned to improve its performance in predicting SM. Specifically, we conduct a grid search for key hyperparameters—batch size, learning rate, hidden layers, and number of units in each layer—for each soil layer (see Table 2) (Gupta and Raza 2020). For the grid search, the optional values for each parameter are 1) 128, 512, 1024, and 2048 for batch size; 2) 0.0005, 0.001, 0.005, 0.01, and 0.015 for learning rate; 3) 1, 2, and 3 for hidden layers; and 4) 16, 32, 64, and 96 for number of units in each hidden layer. The set of hyper parameters is selected by choosing the one that achieve the lowest loss function value (mean squared error) on the validation set.
2) CLIMATE

Daily climate data including precipitation ($P$), air temperature ($T$), dew temperature, surface net solar radiation ($SSR$), wind speed, and total cloud cover (Table 1) from 1981 to 2013 at a spatial resolution of $0.25^\circ$ are obtained from ERA-Interim (Dee et al. 2011). Dew temperature and air temperature are used together to calculate the vapor pressure deficit (VPD) through the Clausius–Clapeyron equation using the Magnus approximation (Andrews 2010; Lawrence 2005), as VPD is an important factor that affects photosynthesis and evapotranspiration rates and thus influences SM (Jarvis 1976; Medlyn et al. 2011; Zhou et al. 2019). The accumulated precipitation in previous months for a specific day is calculated by aggregating precipitation in each corresponding time period.

3) TOPOGRAPHY

Elevation data from Global Multiresolution Terrain Elevation Data (GMTED) 2010 (Danielson and Gesch 2011) are used to assess the impact of topography and are assumed static in our NN. The original GMTED data are at a resolution of 30 arc s, and they have been converted to $0.25^\circ$ by averaging original elevation in a $30 \times 30$ pixel window that falls within a $0.25^\circ$ grid. This conversion is to make the spatial resolution of the elevation data consistent with other input data.

4) SOIL

Soil properties including pH value, sand, silt, and clay percentage, and bulk density at different soil layers are obtained from the Soil Database of China for Land Surface Modeling (Shangguan et al. 2013). These soil data are used as input features of the NN, whereas soil porosity is excluded in the NN to avoid collinearity, because soil porosity can be determined by soil texture (i.e., sand, silt, and clay percentage) (Huang et al. 2011; Shaxson and Barber 2003). Similar to elevation data, the original soil data are at a spatial resolution of 30 arc s, and they are also converted to a $0.25^\circ$ resolution using the same method. In addition, since the soil layers in the original data do not align exactly with the five soil layers here, we use linear interpolation to get layer-specific data for each of the five 10-cm soil layers.

5) CROPLAND AREA

We extract the cropland area for China from the global agricultural lands in the year 2000 (Ramankutty et al. 2008). As we focus on cropland area only, we set an area fraction threshold of 0.1 for cropland as earlier study did (Lobell and Gourdji 2012). That is, in our NN simulations we only consider the grid cells where the cropland area fraction is equal or larger than 0.1. In addition, we only consider six major crops in our analysis: maize, wheat, soybean, rapeseed, potato, and cotton, which in total account for ~80% total planting area in China (National Bureau of Statistics of China 2019). The global dataset of monthly irrigated and rain-fed crop areas around the year 2000 (MIRCA2000; Portmann et al. 2010) is used here to calculate the area fraction for each of the six main crops in every $0.25^\circ$ grid cell (Fig. S2), and the harvested area of each crop in each grid cell is assumed to be static during our study period (1983–2013).

6) CROPPING PATTERN

Typically the entire China domain is divided into seven geographical regions: the north, northeast, northwest, east, central, south, and southwest (Fu et al. 1999; Ren and Bao 1992; Zheng and Yang 1997). In this study, we separate the Huang-

---

Table 2. Hyperparameters for each soil layer in the neural network and the corresponding $R^2$ between soil moisture predictions and observations ($R^2_{\text{train}}$ and $R^2_{\text{test}}$ are $R^2$ for NN training and testing, respectively).

<table>
<thead>
<tr>
<th>Soil layer</th>
<th>Hidden layers</th>
<th>No. of units</th>
<th>Batch size</th>
<th>Learning rate</th>
<th>$R^2_{\text{train}}$</th>
<th>$R^2_{\text{test}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>32</td>
<td>128</td>
<td>0.0005</td>
<td>0.65</td>
<td>0.64</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>16</td>
<td>128</td>
<td>0.0005</td>
<td>0.67</td>
<td>0.66</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>16</td>
<td>128</td>
<td>0.001</td>
<td>0.68</td>
<td>0.67</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>16</td>
<td>128</td>
<td>0.0005</td>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>16</td>
<td>128</td>
<td>0.0005</td>
<td>0.70</td>
<td>0.69</td>
</tr>
</tbody>
</table>
Huai-Hai (HHH) region (Lei et al. 2018), which is an important agricultural area that yields ~28% of national crop production (Lei et al. 2006), from northern, eastern, and central China (see Fig. 2). This is because the environmental conditions and the prevalent cropping pattern in this region are relatively consistent (Lei et al. 2006, 2018; Yan et al. 2010).

Region-specific major cropping patterns are obtained through extensive literature survey (see Table 3), which exemplify the general prevalent practices for each region. Due to the complexity of topography, climate, soil properties and farming practices, actual cropping patterns could vary from field to field and from year to year. However, spatiotemporally explicit cropping pattern data are currently not available, we thus assume them to be static along time.

c. Experiment design

1) CONTROL EXPERIMENTS AT STATION LEVEL

To tease apart the contribution of the individual factors in determining SM variations, we conduct a series of control experiments with the station level observations. Specifically, we first construct a full NN model with all features (i.e., input variables) as listed in Table 1, and obtain a $R^2$ value between the SM predictions and observations during NN training. Next, we run nine reduced models omitting each of the individual feature and obtain corresponding $R^2$ values, which are lower than the full model. Note that the seven precipitation related features (see Table 1) are removed together for the control experiment of precipitation, likewise, the five soil related features are omitted for the experiment of soil. Afterward, we calculate the difference between the $R^2$ value of the full model and the reduced models.
Table 3. Major cropping pattern year round in the eight regions of China (e.g., fallow + maize stands for fallow in the winter followed by maize in the summer).

<table>
<thead>
<tr>
<th>Region</th>
<th>Primary pattern</th>
<th>Secondary pattern</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northeast (NE)</td>
<td>Fallow + maize</td>
<td>Fallow + soybean</td>
<td>Chen et al. (1997, 2009, 2013); K. Li et al. (2010); Y. Li et al. (2010); Qu et al. (2010); Sui et al. (2012); Wang et al. (2015, 2012); Yan et al. (2010); Yang et al. (2015, 2011); Yang and Chen (2014); Yang et al. (2007); Zhao et al. (2010, 2014)</td>
</tr>
<tr>
<td>North (N)</td>
<td>Fallow + maize</td>
<td>Fallow + soybean</td>
<td></td>
</tr>
<tr>
<td>Central (C)</td>
<td>Winter wheat + maize</td>
<td>Rapeseed + potato/cotton</td>
<td></td>
</tr>
<tr>
<td>South (S)</td>
<td>Rapeseed + maize + potato</td>
<td>Sugarcane + fallow</td>
<td></td>
</tr>
<tr>
<td>Huang-Huai-Hai (HHH)</td>
<td>Winter wheat + maize</td>
<td>Fallow + maize</td>
<td></td>
</tr>
<tr>
<td>East (E)</td>
<td>Winter wheat + maize</td>
<td>Winter wheat + cotton</td>
<td></td>
</tr>
<tr>
<td>Northwest (NW)</td>
<td>Spring/winter wheat + fallow</td>
<td>Potato + fallow</td>
<td></td>
</tr>
<tr>
<td>Southwest (SW)</td>
<td>Winter wheat + maize</td>
<td>Rapeseed + maize</td>
<td></td>
</tr>
</tbody>
</table>

and the nine reduced models, and then rank the $R^2$ difference. The control variables that result in the highest $R^2$ differences mark the most important factors affecting SM.

2) Reference NN SM Product and Ones with Alternative Cropping Patterns

We derive a reference NN SM product based on region-specific major cropping patterns (Table 3) and crop-specific harvested area (Fig. S2), which is constructed through the following steps. First, we conduct six simulations using the full NN model, assuming six respective homogeneous cropping patterns across China [see the six major cropping patterns in section 2b(1)]. For example, for the simulation with cropping pattern across China [see the six major cropping patterns in section 2b(1)]. The identified area where cropping pattern is crucial could be used to inform optimal agricultural practices (e.g., mitigating SM decline without compromising crop production) and relevant policies.

3) Control Experiments at National Level

After evaluating the importance of individual features, similar to developing reference NN SM [see section 2c(2)], we simulate SM with four reduced NN models at national scale for the period of 1981–2013: two models omitting each of the top two climatic factors, another one omitting soil properties, and the last one omitting cropping pattern. This design is to uncover the area where these factors play a prominent role in determining SM, by comparing against the full model (reference NN SM). The identified area where cropping pattern is crucial could be used to inform optimal agricultural practices (e.g., mitigating SM decline without compromising crop production) and relevant policies.

3. Results

a. SM Variations Detected from Observations

SM across the 273 agrometeorological stations exhibits a wide range of trends across China, with both positive and negative trends. Yet, most stations show a negative trend and the overall magnitude of the negative trends is higher than that of positive trends (Fig. 2), emphasizing an overall depletion of soil moisture across China. This wide variety of trends could be due to differences in climate, topography, soil properties, and farming practices across different regions.

The mean of annual SM across all the 273 stations during 1981–2013 at each of the five soil layers presents two notable characteristics (Fig. 3). First, SM in all five layers exhibit a consistent decline, with an average trend of $-1.5 \times 10^{-3} \text{ m}^3 \text{ m}^{-2} \text{ yr}^{-1}$. Second, the SM time series across the five layers are highly correlated (the correlation coefficient $r$ between the top and bottom layer is 0.85), posing opposite evidence to the CMIP5 model simulations which indicate diverging trends between surface and deep soil moisture (Berg et al. 2017). This difference may be due to hydraulic lift, which is not represented in most models (Caldwell et al. 1998; Kennedy et al. 2019).

Nonetheless, SM trends across different cropping patterns exhibit quite different magnitudes and spreads (Fig. S3). Specifically, the median SM trends for stations with cropping patterns of fallow + maize, fallow + soybean, and winter
wheat + maize are negative, while other cropping patterns present slight positive or minimal trend. This does not necessarily mean that the three aforementioned cropping patterns tend to deplete more soil water than the other patterns. This is because stations with the same cropping pattern are distributed across different regions, and thus their SM responses to different climate trends, soil properties, topography, and farming practices may have geographical differences. In addition, the SM trends in cropping patterns of fallow + soybean and rapeseed + cotton/potato/fallow display much less spread. This might be because the pattern of fallow + soybean is mostly located in northeast China (Fig. S1), which exhibits limited variations in environmental conditions. For the pattern rapeseed + cotton/potato/fallow, the low spread in SM trends may be due to the small number of stations collected in our analysis (6 stations), relative to other patterns such as winter wheat + fallow (42 stations).

b. NN training and testing

After a grid search of the NN hyper parameters for each of the five soil layers, the corresponding set of parameters that yields the lowest loss function value are selected for each layer (Table 2). The training and testing of the NN yield satisfactory $R^2$ values of 0.65–0.70 and 0.64–0.69 for the five soil layers, respectively (Table 2). Furthermore, the scatterplots of SM predictions versus observations during the NN training and testing are consistent, emphasizing the avoidance of overfitting (Fig. 4). These results lend us confidence in the performance of our trained NN models. In addition, the SM at
lower layers such as soil layer 5 (40–50 cm) presents better agreement with the observations (e.g., $R^2$ value of 0.70 in the 40–50-cm layer versus 0.65 in the 0–10-cm layer during NN training). This is mainly because: 1) SM at lower layer has less noisy processes and thus is less variable than the upper ones (Salvucci and Entekhabi 1994); and 2) the NN SM in adjacent upper layer is used as an input feature to predict SM at lower layer, the high correlation between SM at upper and lower layers (Fig. 3) has thus improved the SM prediction at lower layers.

c. Attribution of SM variations to controlling factors

Our control experiments using the NN identify precipitation and soil properties as the two dominant factors in determining SM variability (Fig. 5). The $R^2$ difference between the full and the two reduced models omitting $P$ and soil properties are both 0.15, which is substantially higher than other variables, underscoring the importance of $P$ and soil properties. Our results indicate that elevation also exerts some effects on SM, with a $R^2$ difference of 0.09. Excluding air temperature $T$ or VPD or cropping patterns in the inputs results in an $R^2$ loss $\sim$0.07 for all, implying that these three factors have considerable effect on SM. Other meteorological factors such as surface net solar radiation (SSR), cloud cover, and wind speed tend to have lower impacts on SM.

We compare the four NN SM sets from nation-level control experiments [see section 2c(3)], which omit each feature of the four controlling variables ($P$, soil properties, $T$, and cropping patterns), against the reference NN SM. Their differences reflect the cascading effects of a specific controlling factor on SM, and their spatial pattern identifies the area where that factor does matter for SM. Apparently, $P$, $T$, and soil properties have overall significant effects on cropland SM (Figs. 6a–c). In addition, cropping patterns have pronounced effects on SM. That is, omitting cropping patterns in the NN causes an overall overestimation of SM in northern China (except for the eastern part of northwest China) and underestimation in southern China (Fig. S4d). This is particularly apparent in the southern part of northeast China, the northern part of the Huang-Huai-Hai region, the eastern part of the northwest China, and central and east China, highlighting the role of cropping pattern on SM in these regions. The implication here is that in those areas where cropping pattern is important in determining SM, cropping patterns could be modified to reduce the SM decline in an increasingly water-scarce future.

d. SM variations detected from NN

After gaining confidence in our NN performance at site level, we employ it to predict daily SM for croplands at the national scale across the five 10-cm soil layers during 1981–2013. Here we examine the spatial and temporal variations of the reference NN SM product [see section 2c(2)].

Our reference NN SM product presents relatively low SM in the north and high SM in the south (0–10-cm soil layer, Fig. 7a), which could be attributed to the spatial pattern of precipitation—low precipitation in the north and high in the south (Fig. S4a). Moreover, it exhibits an overall negative trend of SM across the cropland area (Fig. 7b). This could be mainly a result of the overall drying and warming trends across croplands of China (Figs. S4b,d), also as suggested by earlier studies (Jia et al. 2018; Liu et al. 2015).

Apparently, the spatial patterns of SM magnitude and trend are generally in accordance with those from the remote sensing based CCI SM product (Dorigo et al. 2017), especially for those areas where cropland is dominant (i.e., gray dots on Fig. 7). The inconsistency in magnitudes and trends between these two products may be due to several causes. First, there is a mismatch due to subpixel land use and land cover. The CCI SM product captures the temporal SM variations over an entire 0.25° pixel which is a mosaic of different land cover (e.g., forests, grasslands, shrublands) and land use (e.g., urban area) types. However, the NN SM provides estimates only for the cropland part of the pixel—when mapped we use a full coverage of the reference cropping pattern, yet the cropland area fraction is particularly low in southern and western China (Fig. S2a). Second, the CCI SM reflects the effects of farming practices such as irrigation on SM. However, the NN SM only represent SM variations on rain-fed cropland and does not consider effects of those farming practices. Third, the depths of soil moisture inference are different between them. While the NN SM at the top 10-cm soil layer represents an average SM of the top 10 cm, the CCI SM product may only detect a depth of up to 5 cm. Fourth, there could be large uncertainties in the retrieval of the CCI SM product from historical high-frequency sensors (e.g., C band; Kolassa et al. 2017).

Despite the abovementioned differences and caveats, the similar patterns of SM magnitude and trend with CCI SM still provides us some confidence in our NN SM product. First of all, there is no “true” product allowing an “apples-to-apples” comparison to corroborate the credibility of our NN product. To the authors’ best knowledge, the CCI SM product is the
only satellite-retrieved dataset with global coverage and time span of more than three decades which allows the comparison, and it is widely used as satellite observation (Boke-Olén et al. 2018; Gómez et al. 2018; Gruber et al. 2019; Jia et al. 2018; Liu et al. 2015; McNally et al. 2016; Naz et al. 2020; Zhang et al. 2019; Zhu et al. 2019). Furthermore, since the SM variations are highly correlated between the top and lower soil layers according to ground truth observations (see Fig. 3), despite the fact that there is a discrepancy in soil depth, the similar patterns of SM magnitude and trend between our NN SM (top 0–10 cm) and the CCI SM reveal the reasonableness of our NN SM product.

The SM variations in dense cropland regions (Fig. S2a), such as the Huang-Huai-Hai region and northeast China, present strong seasonal cycle—relatively high SM in the summer and low SM in the winter (Fig. 8). This is mainly a result of the seasonal distribution of annual precipitation, as most precipitation occurs in the summer in these regions (Zhang and Qian 2003; Zhang et al. 2011). Besides, we find the seasonal patterns of our reference NN SM generally follow those of the CCI SM across different regions (Fig. 8). In addition, the discrepancies in SM magnitudes tend to be much smaller in the aforementioned dense cropland regions. Larger differences in cropland sparse regions (e.g., south China) could be due to the aforementioned mismatch in subpixel land use and land cover between these two products.

e. Effects of cropping patterns on SM

To explore the effects of different cropping patterns on SM, we compare the six SM sets with alternative cropping patterns against the reference NN SM. Overall, the planting of winter wheat + fallow across China tend to mitigate soil moisture decline relative to the reference cropping pattern, especially in the northern part of China (Fig. 9a). In addition, planting winter wheat + maize also tends to relieve SM depletion in the North relative to the reference (Fig. 9b). At the same time, an interesting finding is that the cropping patterns of fallow + maize and fallow + soybean seem to further deplete soil water relative to double cropping in the HHH region and southwest China (Figs. 9e,f). Usually, higher crop production is achieved at the cost of more soil water consumption, and thus double cropping typically consumes more soil water than single cropping (Chen et al. 2015; Guo et al. 2013; Zhao et al. 2018). These unintuitive findings here could be explained by the large role of soil properties (Fig. S5), whose effects on SM are reasonably captured by our NN (Fig. 5). Specifically, soil porosity is particularly low in these two regions (Fig. S5f), and the corresponding water holding capacity will also low...
(Shaxson and Barber 2003). Thus, relative to the case of fallow in winter (which may engender more soil water drainage), the roots of winter wheat will help hold more soil water. Similarly, earlier study also indicates that southwest China presents lower soil moisture relative to central/east China due to its lower soil porosity (Yang et al. 2020). Our finding suggests that rather than leaving cropland fallow in the winter season for the HHH region and southwest China, planting winter crop may tend to be more desirable to maintain soil water.

In terms of seasonal SM differences between the six alternative SM sets and the reference NN SM, the most pronounced positive discrepancies occur in the pattern of winter wheat fallow, with overall higher SM in most regions year-round and particularly higher SM for northwest and northeast China (Fig. 10a). In addition, the pattern of winter wheat + maize also generally presents a slightly higher SM than the reference SM over the course of the year (Fig. 10b). On the other hand, the cropping patterns of fallow + maize and fallow + soybean show apparent lower SM in the Huang-Huai-Hai region and southwestern China than the reference, especially during the winter, as explained above. Due to ecosystem memory (Anderegg et al. 2015; Hughes et al. 2019; Ogle et al. 2015), lower soil water in the winter (with fallow) will cascade to the summer, resulting in overall lower SM in summer season as well in these two regions (Figs. 10e,f). Nonetheless, it seems that only alternative cropping patterns of winter wheat + fallow and winter wheat + maize result in slightly higher national average SM as compared to the reference case, other alternative patterns only lead to minimal changes in national average (Fig. 10).

4. Discussions

a. Implications of alternative cropping patterns on alleviating SM decline

In light of the attribution of SM variations (see section 3c), while most of the controlling factors such as climate, topography and soil properties could not be instantly manipulated by
humans, farming practices such as cropping patterns could be modulated to alleviate SM decline. This could be potentially implemented in areas where cropping patterns significantly affect SM, particularly in the southern part of northeast China, the northern part of the Huang-Huai-Hai region, the eastern part of the northwest China, and central and east China (Fig. 6d).

We find that the cropping pattern of winter wheat + fallow generally ameliorates SM decline in the HHH region, northeastern and northern China where cropping pattern is suggested to be important in determining SM (Figs. 6d and 9a). This is consistent with field experiments conducted in China which indicate the lower water consumption of wheat relative to maize and soybean (He et al. 2012; Liu et al. 2015; Ren and Luo 2004). In addition, we find that alternative cropping patterns may give rise to contrasting effects on SM in different regions—rising SM in some regions and months but declining SM in other regions and months (Fig. 10). For example, the cropping patterns of fallow + maize and fallow + soybean tends to maintain more soil water than current cropping practices in eastern and central China (Figs. 10e,f), as opposed to the Huang-Huai-Hai region and the southwest China. Moreover, the shift of current cropping pattern practices to fallow + maize or fallow + soybean in the Huang-Huai-Hai region and southwestern China seem to cause more SM decline in the winter than in the summer (Figs. 10c,f). These findings could be used to inform optimal farming practices at regional and local scales that aim to mitigate soil water depletion.

**b. Potential use of the SM product**

Our NN SM product is the first long-term daily multi-layer cropland SM dataset for China with granularity of cropping patterns. The product includes a set of reference NN SM, plus six sets of NN SM with six alternative prevalent cropping patterns. The dataset is open access. Potential applications include but are not limited to the following areas:

1) Detecting the impacts of irrigation by comparing satellite SM products with our NN SM product. Satellite products...
capture the impacts of irrigation while our NN SM product does not, and thus any differences may reveal the separate effects of irrigation on SM.

2) Estimating water budgets such as runoff and evapotranspiration (ET) based on water balance equation.

3) Examining crop water use efficiency (WUE), which is defined as the amount of carbon assimilated as biomass or grain produced per unit of water used by the crop. The NN SM product can be used to estimate ET and thus could be used for estimating WUE.

4) Evaluating crop growth and predicting food production in rain-fed region, in light of the direct and intrinsic connection between SM and crops.

5) Simulating cropland carbon fluxes such as soil organic carbon (SOC) decomposition and soil respiration, which are closely affected by SM.

6) Evaluating the effects of alternative cropping pattern on water and carbon fluxes, which may inform regional and local decision making pertaining to environmental and agricultural policies.

c. Limitations and uncertainties

The extrapolation of the NN model outside the range of training conditions may be the major uncertainty source of the NN SM product. The NN SM product during the nongrowing season may be uncertain, especially when the soils are frozen. Since the NN is trained with SM measurements conducted during the growing season, although the NN SM products provide daily SM estimates year around, the NN’s capability in simulating nongrowing season SM is not adequately validated, particularly for regions with extensive cold season like northeastern China. Furthermore, due to the sample bias in cropping patterns at different agrometeorological stations, the extrapolation of SM for one specific cropping pattern to regions where SM observations for that cropping pattern are missing may be biased. Therefore, users need to be cautious...
when interpreting the NN SM. Thus, two quality flags are provided in the dataset: 1) a time flag indicating whether the date of SM estimate is within the temporal range of observations and 2) a cropping pattern flag indicating whether the associated cropping patterns of SM estimate are among the observed cropping patterns. The time flag and cropping pattern flag are determined based on the date range of SM observation and whether the agrometeorological stations in a specific region have the cropping pattern(s) of interest. That is, for a specific region, any dates within the observation time range will be assigned a time flag of 1 (otherwise 0), and the cropping pattern flag will be assigned to 1 if agrometeorological stations in this region do have the cropping pattern(s) of interest (otherwise 0).

Farming practices, such as irrigation, fertilizer use, equipment use, crop cultivars, tillage and mulching, which could significantly affect SM, are not explicitly considered in our NN, mainly due to the lack of relevant data. Nevertheless, all these agricultural practices may have evolved substantially, as agriculture has been greatly intensified since the 1980s (Liu et al. 2015). The cropland area used here corresponds to conditions in the year 2000 and is assumed static from 1981 to 2013, and any contraction or expansion of cropland area is not considered.

Despite these inevitable inherent uncertainties and caveats, the SM product here provide the community an alternative opportunity to scrutinize the characteristics of SM and explore its role in the Earth system, such as in the aspects of water cycling, ecosystem functioning and agricultural management. This product complements LSM and satellite based products, and provides additional information related to soil moisture depth and cropping patterns, and avoids the noise and depth limitation of satellite products. Further evaluation and scrutiny of the SM product with more ground truth observations will be pursued to improve its validity in our future work.

5. Conclusions

Accurately quantifying large-scale soil moisture is a challenge and previously has been primarily done through satellite retrievals or land surface model simulation. In this work, we
leverage high-quality, consistent, long-term ground truth measurements and machine learning (deep neural network), to develop a daily multilayer soil moisture product for the cropland regions of China during 1981–2013, accounting for cropping patterns. The neural network presents good performance ($R^2$ value of 0.64–0.7 between predictions and observations during training and testing) across regions and cropping patterns. This product is open access and provides the research community an alternative opportunity to explore a variety of research topics pertaining to soil moisture.

While it is well known that soil moisture is affected by a myriad of intricate factors, we find that precipitation and soil properties are the two most important environmental factors in determining cropland soil moisture. Cropping pattern is also a critical factor affecting soil moisture. Moreover, through the lens of machine learning, we find that alternative cropping pattern of winter wheat followed by fallow could mitigate soil moisture decline in most part of China. In addition, the current practice of double cropping in the Huang-Huai-Hai region and southwest China seems to favor higher soil water relative to single cropping (fallow in winter), potentially due to the low soil porosity in these two regions and consequently more soil water drainage if the cropland is left fallow.

With the extensive agricultural intensification since 1980s, soil moisture decline in cropland is observed over all depths across China. This tendency might be exacerbated in an increasingly drying future, which poses great challenges to sustainability. Thus, research on mitigating further cropland soil water depletion and at the same time ensuring ever-increasing food security is the natural course along which the scientific community should proceed for the sustainable development of China. Our effort here is a stepping stone to research of this kind that will inform the nation’s march toward sustainability.

Acknowledgments. This research is supported by the Modeling, Analysis, Predictions and Projections (MAPP) program of NOAA NA17OAR4310127 and NASA NNH17ZDA001N-THP to Pierre Gentine. We thank Chinese Meteorological Administration for providing the soil moisture measurements, and also thank our colleague Adam Massmann for proofreading this article.

Data availability statement. The soil moisture dataset developed in this work is freely available at the NASA Aura Validation Data Center (https://avdc.gsfc.nasa.gov/pub/data/project/SMML_CHN/).

REFERENCES


Sui, Y., W. Huang, X. Yang, and M. Li, 2012: Characteristics and adaption of seasonal drought in southern China under the background of global climate change II. Spatiotemporal characteristics of drought for wintering grain and oil crops based on crop water deficit index. *Ying Yong Sheng Tai Xue Bao*, **23**, 2467–2476.


