Soil Moisture Cloud Precipitation Feedback in the Lower Atmosphere from Functional Decomposition of Satellite Observations

Yifu Gao[1], Clement Guilloteau[2], Efi Foufoula-Georgiou[1], chonggang xu[3], Xiaoming Sun[4], and Jasper A. Vrugt[1]

[1]University of California, Irvine
[2]UC Irvine
[3]lanl
[4]Los Alamos National Laboratory

May 28, 2024

Abstract

The feedback of topsoil moisture (SM) content on convective clouds and precipitation is not well understood and represented in the current generation of coupled cloud physics and land-surface models. Here, we use functional decomposition of satellite-derived SM (SMAP/L4) and cloud vertical profiles (CVP: GPM/DPR/L2A) in the central US to quantify the relationship between SM and the vertical distribution of cloud water. High-dimensional model representation disentangles the contributions of SM and other land-surface and atmospheric variables to the CVP. Results show the sign and strength of this feedback varies with cloud height and time lag and displays a large spatial variability. Positive anomalies in the antecedent 7-hour SM and land-surface temperature can increase reflectivity up to 4 dBZ in the lower atmosphere (1-3 km above the surface). The presented approach brings new insights into observational understanding of SM-precipitation feedback and possesses the potential for diagnosing cloud models regarding land-atmosphere coupling representation.

Soil Moisture Cloud Precipitation Feedback in the Lower Atmosphere from Functional Decomposition of Satellite Observations

Yifu Gao*, Clément Guilloteau*, Efi Foufoula-Georgiou*, Chonggang Xu†, Xiaoming Sun†, and Jasper A. Vrugt*‡

May 22, 2024

Key points

1. We present a CPU-friendly functional decomposition of satellite-measured soil moisture (SM) and cloud vertical profiles
2. The sign and strength of SM's feedback vary with height, time lag, and geographic locations, which agrees with more qualitative studies
3. The presented approach exhibits potential implications for diagnosing cloud models, particularly in the context of land-atmosphere coupling

* Department of Civil and Environmental Engineering, University of California, Irvine, California, USA.
† Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA.
‡ Corresponding author: jasper@uci.edu

Abstract

The feedback of topsoil moisture (SM) content on convective clouds and precipitation is not well understood and represented in the current generation of coupled cloud physics and land-surface models. Here, we use functional decomposition of satellite-derived SM (SMAP/L4) and cloud vertical profiles (CVP: GPM/DPR/L2A) in the central US to quantify the relationship between SM and the vertical distribution of cloud water. High-dimensional model representation disentangles the contributions of SM and other land-surface and atmospheric variables to the CVP. Results show the sign and strength of this feedback varies with cloud height and time lag and displays a large spatial variability. Positive anomalies in the antecedent 7-hour SM and land-surface temperature can increase reflectivity up to 4 dBZ in the lower atmosphere (1-3 km above the surface). The presented approach brings new insights into observational understanding of SM-precipitation feedback and possesses the potential for diagnosing cloud models regarding land-atmosphere coupling representation.

Plain Language Summary

This paper focuses on the observational analysis of how soil moisture (SM) influences the vertical cloud-water distribution throughout the day. By analyzing data from Soil Moisture Active Passive (SMAP) and Dual-frequency Precipitation Radar (DPR), we gain insights into how antecedent SM levels impact cloud-water reflectivity at different heights in the lower atmosphere. Our data-driven approach produces spatial maps of SM's contribution to cloud reflectivity and rainfall in the central US conditioned on cloud height and SM time lag. The results will help diagnose coupled land-atmosphere models.

1 Introduction

The feedbacks between soil moisture (SM) and precipitation play a critical role in regulating regional hydroclimatic variability. Such feedbacks are governed by a plethora of variables and processes, such as (variations in) land surface temperature (Koster et al., 2006), energy partitioning (Golaz et al., 2001; Fast et al., 2019; Sakaguchi et al., 2022), planetary boundary layer (PBL) development (Ek & Holtslag, 2004; Han et al., 2019) and the initiation of convective clouds and precipitation (Ferguson & Wood, 2011; Taylor et al., 2011; Cioni & Hohenegger, 2017). These feedbacks take place across a continuum of spatiotemporal scales, spanning distances from several to thousands of kilometers and time span of days to seasons (Trenberth, 1999; Duerinck et al., 2016; Liu et al., 2022). Moreover, SM-precipitation feedbacks exhibit substantial regional variability in both their sign and magnitude as a result of the large sensitivity of evapotranspiration and atmospheric conditions to SM and latent heat fluxes, respectively (Guo et al., 2006). In this paper, we focus our attention on diurnal SM-cloud-precipitation feedbacks, abbreviated SMCPF, which control in part the vertical cloud-water distribution, thereby influencing weather conditions (Koster et al., 2004) and regional hydroclimatology (Krakauer et al., 2010; Yin et al., 2014; Ford et al., 2023). Future climate projections suggest further that SMCPFs may play an increasing role in determining changes in mean temperature and extremes as a result of larger SM deficits under higher evaporative demands (Dirmeyer et al., 2013; Seneviratne et al., 2013; Taylor, 2015).

Given the importance of the SMCPF in regulating local and regional weather, much research has been devoted to estimating its sign, causality, and physical linkage. That research may be divided into simulation-based analysis (Schär et al., 1999; Findell & Eltahir, 2003a; G. Wang et al., 2007; Hohenegger et al., 2009; Schlemmer et al., 2012; Tawfik et al., 2015; Gentine et al., 2013), observation-based studies (Taylor & Ellis, 2006; Santanello et al., 2009; Taylor et al., 2010; Ferguson & Wood, 2011; Taylor et al., 2011; Ford, Rapp, Quiring, & Blake, 2015; Guillod et al., 2015) and a combination thereof (Seneviratne et al., 2006; Santanello et al., 2013; Miralles et al., 2014; Spennemann et al., 2018; Baker, Castilho de Souza, et al., 2021; Baker, Garcia-Carreras, et al., 2021). Although numerical models of land-atmosphere interactions have advanced considerably in recent decades, the diurnal impact of SM on cloud formation and composition is still not particularly well understood. The mechanisms governing the sign and strength of the simulated SMCPFs are subject to a large uncertainty depending for example on the choice of boundary conditions (Hohenegger et al., 2009) and sub-grid scale process representation (Deardorff, 1980; Thompson et al., 2004, 2008). In observational studies, on the other hand, it is difficult to filter out the effects of synoptic variability. Moreover, in the absence of high-quality spatiotemporal measurements of SM and cloud vertical profiles, past studies have mainly focused on how (gradients of) SM affect convection initiation, the PBL height, and precipitation probability (Frye & Mote, 2010; Findell et al., 2011; Taylor, 2015; Su & Dickinson, 2017; Graf et al., 2021; Yuan et al., 2020; Ford et al., 2023) without recourse to mesoscale diurnal relationships between antecedent SM and the cloud water distribution. Advances in our understanding of SM-cloud relationships should improve the diagnosis of weather and climate models and enhance the accuracy of their future projections (Williams, 2019).

Fortunately, remote-sensing data products of SM and the cloud vertical profile from polar-orbiting Earth-observing satellites have advanced considerably in the past decades and have the potential to substantially advance our understanding of SM-cloud-precipitation relationships. Specifically, the 3-hr/9 km Soil Moisture Active Passive (SMAP/L4) and 1.5-hr/5 km Global Precipitation Measurement Dual-Frequency Precipitation Radar (GPM/DPR/L2A) provide high-resolution estimates of the topsoil moisture content and the vertical distribution of hydrometeors within and above the PBL, respectively, at a global coverage. Many studies have confirmed the accuracy and reliability of SMAP/L4 (X. Zhang et al., 2017; Reichle et al., 2017; L. Zhang et al., 2017; Koster et al., 2018; Tavakol et al., 2019) and GPM/DPR/L2A (Lasser et al., 2019; Pejcic et al., 2020; Liao & Meneghini, 2022) data products.

In this paper, we demonstrate how functional decomposition of a large database of SMAP/L4 surface SM and GPM/DPR/L2A cloud vertical profiles (CVP) provides valuable insights into the relationship between antecedent SM and cloud water distribution and reflectivity in the lower troposphere. Specifically, we use high-dimensional model representation (HDMR) (Li & Rabitz, 2010; Li & Rabitz, 2012; Gao et al., 2023) to disentangle the intricate and dynamic web of land-surface and atmospheric variables and interactions that give rise to the SMCPF. HDMR is a generalization of the analysis of variance (ANOVA) to dependent input factors and uses a superposition of linear multiples of first-, second-, and higher-order component functions to parse out the structural and correlative contributions of SM and other land-surface variables to the CVP. The expansion coefficients of the component functions are determined from a training data set of collocated SMAP/L4 and GPM/DPR/L2A measurements across the central US using linear least squares and D-MORPH regression (Li & Rabitz, 2010). We are mainly interested in the first-order component functions as they quantify the direct contribution of each land-surface variable to the CVP. The method is CPU-efficient and yields spatial maps of the SM contribution to cloud reflectivity and rainfall for our study region as a function of cloud height and SM time lag.

This paper is organized as follows. Section 2 discusses the SMAP/L4 SM and GPM/DPR/L2A satellite products and study region. Section 3 summarizes the data preprocessing steps and HDMR functional decomposition. Section 4 presents the results of our analysis and documents the relationship between SM and the CVP as a function of cloud height, time lag, and spatial coordinates in our study region. Section 5 summarizes our main findings and presents suggestions for future work.

2 Data and Experimental Region

We use the publicly available 3-hour/9 km SMAP/L4 and 1.5-hour/5 km GPM/DPR/L2A data products and single out samples from our study region in the warm seasons (April to October) of 2016 to 2019 with convective precipitation in the afternoon hours until midnight (14:00-24:00 CDT). The altitude spans 1 to 5 km, with the 1-3 km zone identified by Findell and Eltahir (2003a) as a critical region for convective triggering, and in the 3-5 km zone above this region resides the free atmosphere. We succinctly discuss the SMAP/L4 and GPM/DPR/L2A products and our study region. A more detailed description of the satellite data products is found in Text S1.

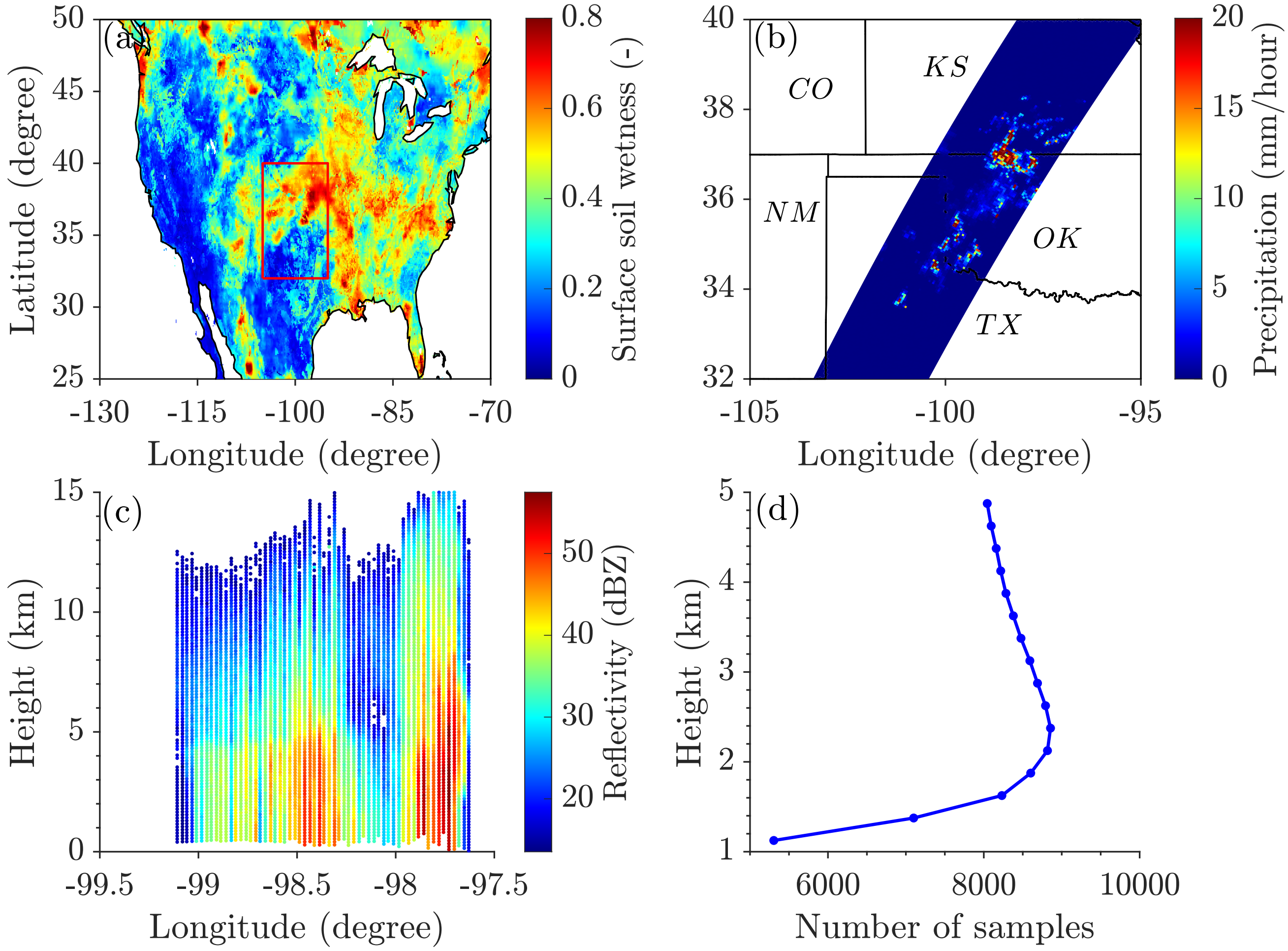

Figure 1: August 7, 2016: (a) SMAP/L4 surface SM (3-hour, 9 km, 19:30 CDT) over CONUS and GPM/DPR/L2A measured (b) surface precipitation (1.5-hour, 5 km, 21:51:10-23:23:44 CDT) and (c) cloud reflectivity profiles (97.5°W - 99.5°W, 36.7°N) for our study region (red rectangle) in the central United States. Graph (d) in the bottom right corner displays the number of samples n we have left at each DPR measurement height after data preprocessing.

The SMAP mission Level 4 SM (L4_SM) product gives 3-hourly estimates of surface and root-zone SM at 9-km spatial resolution and global coverage (Reichle et al., 2015). The 3-hour time-averaged 9-km geophysical data product (SPL4SMGP) provides estimates of the wetness (0-1) of the top soil layer (0-5 cm) (see Figure 1a) and other land-surface variables. Hourly estimates of low-level atmospheric temperature (AT) and total precipitable water (TPW) from $0.25^\circ \times 0.25^\circ$ ERA-5 reanalysis convey the stability and humidity of the antecedent atmosphere and are precursors to mesoscale convective events (Sherwood, 1999; Findell & Eltahir, 2003a; Holloway & Neelin, 2010). In our functional decomposition, we use the mean AT for the critical region, 1-3 km above the soil surface, which roughly corresponds to levels $P_{\mathrm{surf}} - 100$ and $P_{\mathrm{surf}} - 300$ hPa. Section 3.2 discusses in more detail our selection of auxiliary land-surface and atmospheric variables.

The GPM/DPR/L2A product (GPM_2ADPR) provides a swath of precipitation profiles (see Figure 1b) every 1.5 hours at a spatial resolution of 5 km and vertical increment of 125 m. The major data fields zFactorFinal (dBZ) and typePrecip provide vertical profiles of the Ka-band cloud reflectivity factor (see Figure 1c) and an 8-digit precipitation type ID, for individual pixels. We only use samples classified as convective precipitation and work with 250-m averaged Ka-band cloud reflectivities to suppress measurement errors.

Our study region in Figure 1a (95°W-105°W, 32°N-40°N) is a hot spot for SM-precipitation coupling (Findell & Eltahir, 2003b; Koster et al., 2004; Ford et al., 2023) with large spatial variability in climatological sign and strength of the SMCPF (Frye & Mote, 2010; Findell et al., 2011; Su & Dickinson, 2017; Yuan et al., 2020; Ford et al., 2023). This central region of the US offers an excellent demonstration of our method and possibility to benchmark the inferred patterns of the SMCPF sign and magnitude against literature findings.

3 Method

3.1 Data Preprocessing

We extract the GPM/DPR/L2A swaths that overpass our study region and use only those samples classified as convective precipitation in the 'typePrecip' data field. This type classification is an important byproduct of DPR instruments and crucial to an accurate characterization of the antecedent atmosphere using ERA-5 reanalysis AT and TPW data. To avoid water from interception evaporation, we discard all samples which received more than 0.5 mm of precipitation in the 18 hours preceding the DPR's scan according to the Multi-Radars Multi-Sensors (MRMS) Gauge-corrected Quantitative Precipitation Estimates (J. Zhang et al., 2016). This should also reduce the impacts of large-scale synoptic systems

(Findell et al., 2011). Next, we collocate SMAP/L4 and ERA-5 data and GPM/DPR/L2A measured cloud profiles using linear interpolation and time lags $\Delta t = t_{\text{dpr}} - t_{\text{smap}}$ of 7 and 10 hours. In doing so, we allow for a 2-hour grace period so as to maximize the sample size. For example, SM data with a time lag $6.01 \leq \Delta t \leq 7.99$ are pooled together in the 7-hour time lag. Figure 1d displays the number of DPR-measured cloud reflectivities n for the months of April-October (2016-2019) as a function of cloud height. Not all heights have the same sample size due to for instance the absence of clouds, radar detection threshold, and path attenuation (Iguchi et al., 2010). The pooled samples of April-October guarantee a sufficiently large sample size at each cloud height. Next, we decompose this final collection of SMAP/L4 - GPM/DPR/L2A samples using HDMR and expand the DPR-measured cloud reflectivities at each separate cloud height as a sum of first- and higher-order structural and correlative contributions of SM and the auxiliary variables.

3.2 High-Dimensional Model Representation

SMCPFs are notoriously challenging to observe and study outside of model environments (Ford et al., 2023), hence innovative analytical approaches are required to study them (Koster et al., 2004; Seneviratne et al., 2006; Findell et al., 2011; Berg et al., 2013; Guillod et al., 2014; Knist et al., 2017). HDMR is particularly appealing in the present context as it expresses all variable interactions in a system in a hierarchical order. This allows us to quantify the individual contribution of SM to the CVP.

Suppose we group all land-surface and atmospheric variables that govern the cloud reflectivity $y = f(\mathbf{x})$ at a given cloud height in a $d \times 1$ vector $\mathbf{x} = (x_1, \ldots, x_d)^\top$. HDMR builds on the finite multivariable function expansion of Sobol′ (1993) and decomposes the output, $y = f(\mathbf{x})$, of the scalar-valued square-integrable function, $f \in L^2(\mathbb{K}^d)$, on the d-dimensional unit cube, $\mathbb{K}^d = \{\mathbf{x} | 0 \leq x_i \leq 1;\ i = 1, \ldots, d\}$, into summands of component functions, $f_i(x_i)$, $f_{ij}(x_i, x_j)$, $\ldots$, $f_{12\ldots d}(x_1, x_2, \ldots, x_d)$, to yield (Li & Rabitz, 2012)

$$y = f_0 + \sum_{i=1}^{n_1} f_i(x_i) + \sum_{1 \le i < j \le d}^{n_2} f_{ij}(x_i, x_j) + \sum_{1 \le i < j < k \le d}^{n_3} f_{ijk}(x_i, x_j, x_k) + \cdots + f_{12\ldots d}(x_1,\, x_2, \ldots, x_d) + \epsilon, \tag{1}$$

where f_0 is the mean output and the residual $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is assumed to be zero-mean normally distributed with a constant variance, σ_ϵ^2. The $n_1 = d$ first-order functions, $f_i(x_i)$, characterize the individual effects of the input variables on the model output. The $n_2 = d(d-1)/2$ second-, $f_{ij}(x_i, x_j)$, $n_3 = d(d-1)(d-2)/6$ third-, $f_{ijk}(x_i, x_j, x_k)$, up to the d^{th}-order component functions, $f_{12\ldots d}(x_1,\, x_2, \ldots, x_d)$, characterize the cooperative contribution of two, three, up to all land-surface variables combined to the cloud reflectivity y. As third- and higher-order independent and cooperative effects are usually negligible in most physical systems (Rabitz & Aliş, 1999; Kucherenko et al., 2011; H. Wang et al., 2017; Falchi et al., 2018; Shereena & Rao, 2019; Gao et al., 2023), our function expansion of the CVP considers only the $n_{12} = n_1 + n_2$ first- and second-order component functions

$$y = f_0 + \sum_{u=1}^{n_{12}} f_u + \epsilon, \tag{2}$$

where subscript u is the index of the component function rather than its order as in equation (1). Thus, $f_1, \ldots, f_d$, signify the first-order component functions and $f_{d+1}, \ldots, f_{d+d(d-1)/2}$ correspond to the second-order component functions. In our implementation, f_0 signifies the mean reflectivity in units of dBZ and the component functions f_u quantify the individual and bivariate contributions of the land-surface and atmospheric variables to the cloud reflectivity.

The component functions must satisfy hierarchical orthogonality to exactly delineate the independent (structural) and cooperative (correlative) contributions of individual and groups of input variables to y (Li & Rabitz, 2012; Gao et al., 2023). This is enforced through a so-called relaxed vanishing condition (Hooker, 2007)

$$\int_0^1 w_u(\mathbf{x}_u) f_u(\mathbf{x}_u) \mathrm{d}x_i = 0 \qquad \text{for all } u \subseteq \{1, \ldots, d\} \text{ and } i \in u, \tag{3}$$

where u is a subset of superset $U = \{1, \ldots, d\}$, $\mathbf{x}_u$ denote the dimensions u of the input vector and

$w_u(\mathbf{x}_u)$ signifies the probability density function (pdf) of $\mathbf{x}_u$. For a second-order component function, the vanishing condition of equation (3) dictates that $f_{ij}(x_i, x_j)$ should be orthogonal to its lower order component functions, $f_i(x_i)$ and $f_j(x_j)$. The component functions are constructed using the extended bases orthonormalized polynomials and associated linear expansion coefficients. D-MORPH regression (Li & Rabitz, 2010) enforces hierarchical orthogonality of the component functions in pursuit of the optimum expansion coefficients. This method is described in Text S2.

The statistical significance of a given component function is readily determined by comparing the performance of the function expansion with and without this component function. Suppose SSR_1 is the sum of squared residuals of the function $y = y_0 + \sum_{i=1}^{d-1} f_i(x_i)$ with $l_1 = (d-1)p$ expansion coefficients and SSR is the same quantity for the same function $y = y_0 + \sum_{i=1}^{d} f_i(x_i)$ expanded with $f_d(x_d)$ and $l = l_1 + p$ coefficients. To reject the null hypothesis, "$\mathcal{H}_0 : f_d(x_d)$ is insignificant", the F-statistic

$$F = \frac{(\mathrm{SSR}_1 - \mathrm{SSR})/(l - l_1)}{\mathrm{SSR}_1/(n - l_1)}, \tag{4}$$

must exceed $F_{\mathrm{crit}} = F_{\mathcal{F}}^{-1}(1-\alpha | l_1 - l, n - l_1)$ where $F_{\mathcal{F}}^{-1}(p_\alpha | \nu_1, \nu_2)$ is the quantile function of the Fisher-Snedecor distribution with ν_1 and ν_2 degrees of freedom at the critical value $p_\alpha = 1 - \alpha$ and significance level $\alpha \in (0, 1)$. The magnitude of the F-statistic conveys the importance of $f_d(x_d)$ in explaining the CVP and, thus, can be interpreted as a measure of the feedback strength.

Now that we have finished discussing the building blocks of our HDMR data decomposition method, we are left with the selection of land-surface and atmospheric variables $(x_2, \ldots, x_d)$ which complement SM, x_1, in explaining the measured cloud reflectivities, y. We tested many different variables in our analysis and settled on land-surface temperature (LST), leaf area index (LAI), atmospheric temperature (AT), and total precipitable water (TPW) as auxiliary variables. This equates to a 5×1 input vector $\mathbf{x} = (x_1, \ldots, x_5)^\top = (\mathrm{SM}, \mathrm{LST}, \mathrm{LAI}, \mathrm{AT}, \mathrm{TPW})^\top$. LAI and LST modulate evapotranspiration under SM-limited or energy-limited regimes (Seneviratne et al., 2010) and AT and TPW convey information for the SMCPF at synoptic scales about atmospheric preconditioning (Ford, Quiring, et al., 2015; Tuttle

& Salvucci, 2017). This explicit treatment of atmospheric conditions can only raise our confidence in any causal links that are found between SM and CVP. Figure S1 presents a correlogram of the five input variables. Note that we do not consider variables such as the latent heat flux. This derivative product depends on SM, hence would only trouble our inference of the relationships and variables that govern the CVP.

4 Results

4.1 Cloud Height and Temporal Lag of SMCPF

Figure 2 displays the F-statistics of the (a) SM, (b) LST, and (c) LAI component functions as a function of cloud height (1 to 5 km) and time lag ($\Delta t = 7$ and 10 h). The solid line denotes the mean of 1,000 bootstrap trials each with a different selection of $r = 0.75n$ training samples and the light-colored regions portray the associated 95% confidence intervals. The dashed black line in each graph corresponds to the critical F-value at each cloud height using $\alpha = 0.05$. The value of the F-statistic is not constant but altitude dependent. The influence SM, LST and LAI exert on the CVP is dependent on cloud height. In case of SM in panel (a) this equates to a height-dependent SMCPF with a bottom-heavy relationship between SM and CVP. The SMCPF is most pronounced in the lower atmosphere at about 1-3 km above the surface. Above this level, the impact of SM on the CVP decreases rapidly with altitude. As we will shown in Section 4.2, the first-order SM component function $f_1(x_1)$ displays a positive feedback due to a wet soil. A higher SM implies a larger evaporative fraction, promoting moderate PBL growth (see Figure S2) and moisture accumulation (Yin et al., 2015). The CVP at higher altitudes is less dependent on surface SM and controlled more by the upper atmosphere at levels of about 3 km and beyond (Findell & Eltahir, 2003a). Furthermore, a capping inversion layer can inhibit the upward movement of warm, moist air from the surface to the free atmosphere (Findell & Eltahir, 2003b). Indeed, the HDMR-inferred relationship between SM and CVP as articulated by the F-statistic is corroborated by simulation analyses (Findell & Eltahir, 2003a; Koukoula et al., 2019). This physical underpinning inspires confidence in the ability of our methodology to back out SM-cloud feedbacks at different heights. The strong agreement in the results of the two time-lags is a result of SM autocorrelation. The $\Delta t = 7$ hour time lag displays the largest influence on the CVP at all altitudes but the largest two cloud heights near 5 km.

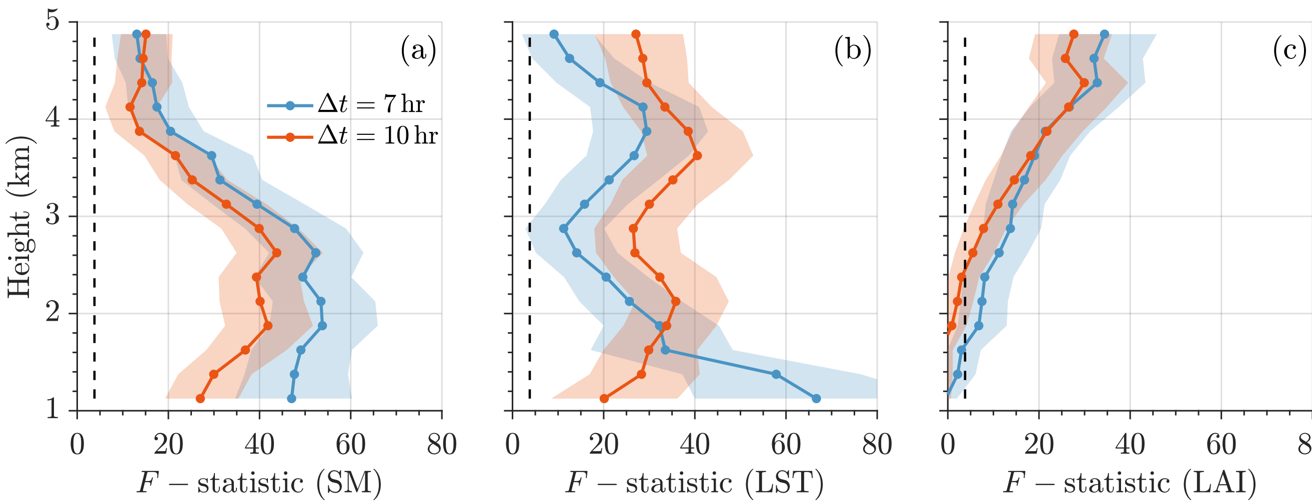

Figure 2: Vertical profiles of the mean F-statistic of the first-order component functions of (a) SM: $f_1(x_1)$, (b) LST: $f_2(x_2)$, and (c) LAI: $f_3(x_3)$ computed from 1,000 bootstrap iterations. Solid blue and red lines differentiate between temporal lags ($\Delta t = 7$ and 10 hours) and black dashed lines represent the critical value at significance level $\alpha = 0.05$, F_{crit}. The light blue and red regions correspond to the 95% bootstrap confidence intervals.

Compared to SM, LST exerts control on CVP across a wider vertical range (in Figure 2b), whose F-statistic shows a bimodal relationship with height, peaking close to the surface with $\Delta t = 7$ hours and at a higher altitude of 3.5-4.0 km with $\Delta t = 10$ hours. As discussed in the next section, $f_2(x_2)$ exhibits a positive correlation with LST, suggesting that positive LST anomalies (or dry soil) play a crucial role in shaping CVP. Therefore, the fact that low-level (1.0-2.5 km) CVP is responsive to LST comes in qualitative agreement with the pathway of negative SMCPF, driven by the effect of positive LST anomalies in catalyzing higher sensible heat flux, convective triggering potential (CTP), and rapid PBL growth. We further support this finding by comparison with the ERA5 reanalysis PBL height in Figure S2. Such observed response of PBL height to wet and dry surface exhibits strong consistency with prior simulation-based and observational studies (Findell & Eltahir, 2003a; Xu et al., 2021; Ford et al., 2023), which indicates two mechanisms for initiating convection: significant moistening of the PBL (over wet soil) and rapid growth of the PBL (over dry soil). In addition, the predictability of LST decreases first at 3.0 km and increases again at 3.5-4.0 km. The reason why LST is significant at a higher altitude may be twofold. On the one hand, the LST anomalies favor strong CTP where air parcels can overcome convective inhibition and reach the level of free convection (Taylor et al., 2012). If we intuitively consider $f_2(x_2)$ the contribution of near-surface air to the cloud reflectivity conditioned on a specific height and time lag, its F-statistic (in Figure 2b) somehow approximates the dynamics of the thermal updraft such that the largest F-statistic value shifts from $\Delta t = 7$ hours to $\Delta t = 10$ hours with height changing from 1.0 km to 5.0 km. On the other hand, local LST may also reflect certain atmospheric conditions such as the melting layer, which typically resides between 3.0-5.0 km above the surface during pre-monsoon and monsoon seasons in the central United States (Song et al., 2021).

The F-statistic of the LAI component, $f_3(x_3)$, informs its poor predictive power in the lower atmosphere, primarily due to the governing effects of SM, LST, and AT (see Figure S3) on initiating convection and the subsequent formation of cloud/precipitation. In contrast, the modest, albeit statistically significant influence of LAI in higher-level CVP can be attributed to its seasonal variations (Savoy & Mackay, 2015) and correlation with the atmospheric conditions (see Figure S1). In Text S3 and Figures S3-S4, we elaborate on our findings in terms of atmospheric controls on CVP which demonstrate a comparable physical underpinning with the land-surface variables.

4.2 The SMCPF across Space

In this section, we focus our attention on the spatial pattern of the SMCPF within the study region. We reiterate that we conduct functional decomposition of the cloud reflectivity using all the samples of April-October (2016-2019) for a specific time lag and cloud height, to guarantee an adequate number of samples and storm events. Our goal here is to present a 4-year averaged spatial distribution of the derived component functions and determine locations of positive and negative SMCPF rather than focusing on interannual and/or cross-season variations.

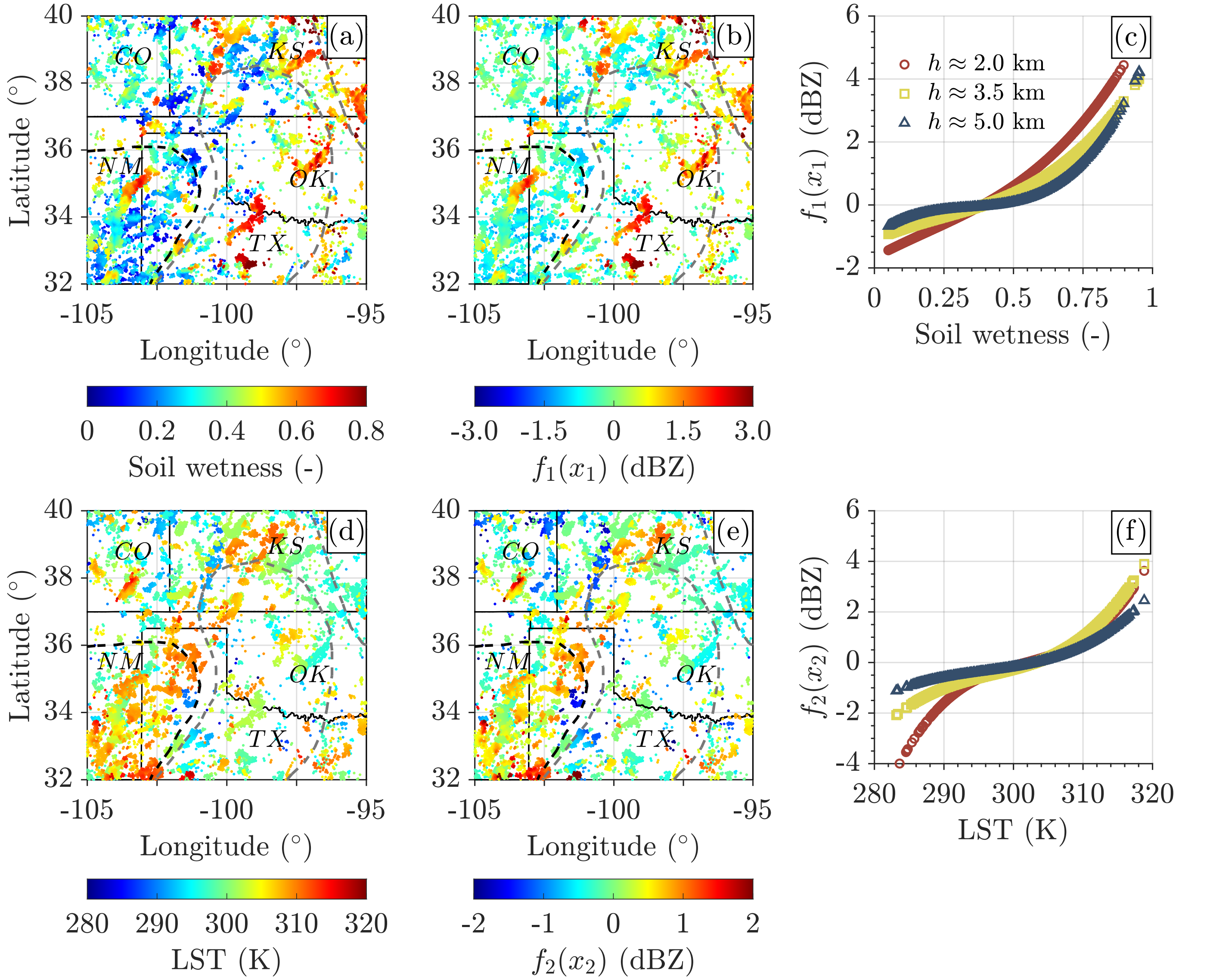

Figure 3: The central United States (95°W-105°W, 32°N-40°N) with (a) antecedent 7-hr SMAP/L4 soil wetness (-) collocated at coordinates of the GPM/DPR/L2A samples and (b) first-order component function of soil wetness, $f_1(x_1)$ (dBZ), evaluated at approximately 2.0 km height. Solid black lines delineate the state borders while dashed black and grey lines depict the negative feedback and transitional regions proposed by Findell and Eltahir (2003b). Panel (c) displays the scatter plots of the samples of antecedent 7-hour SM against the corresponding $f_1(x_1)$ (dBZ), evaluated at three separate heights, 2.0 km (red circles), 3.5 km (yellow squares), and 5.0 km (blue triangles). The bottom row of panels presents the same content as panels (a-c) but for (d) SMAP/L4 LST and (e,f) its associated component function, $f_2(x_2)$.

Figure 3a-b presents the spatial distribution of the antecedent 7-hour SMAP/L4 soil wetness at the top layer (0-5 cm), collocated at the coordinates of the GPM/DPR/L2A samples, alongside the corresponding first-order component function, $f_1(x_1)$ (dBZ), evaluated at 2.0 km. This examination of SM's feedback strength, conditioned on an altitude of 2.0 km and a 7-hour time lag, is of particular interest upon our prior analysis of the F-statistic in Figure 2a. Panels (b-c) reveal the positive feedback from SM represented by $f_1(x_1)$. With a degree of saturation exceeding 0.4, wet soil could increase cloud reflectivity by up to 4 dBZ. The fact that the absolute value of $f_1(x_1)$ decreases with height in Panel (c) again lends support to our inferred height-dependent SMCPF in Section 4.1, underscoring the stronger coupling between SM and CVP in the low-level atmosphere. As a byproduct, we demonstrate in Text S4 and Figure S5 the application of the Marshall-Palmer formula (Marshall & Palmer, 1948) to the transformation of $f_1(x_1)$ (dBZ) into estimates of rainfall rate.

Significant positive feedback of SM is evident in regions such as northern Texas, central Oklahoma, northwestern and southeastern Kansas, and northeastern New Mexico. All these areas, with the exception of northeastern New Mexico, are located inside or close to the 'transitional regions' delineated by dashed grey lines as categorized by Findell and Eltahir (2003b). The middle transitional region, spanning from the semi-arid southwestern to the humid southeastern parts of the central United States, is influenced by both dry and wet soil advantage regimes. Hence, this dual influence explicates the observable positive feedback in the central and eastern sections of the transitional region and negative feedback in the southwestern part (detailed below). These local wet soil anomalies can be attributed to early warm-season mesoscale convective systems (MCSs) and non-MCS rainfall. Typically, the early warm-season MCSs were reported a dominant source of the summer SMCPF (Hu et al., 2021), which are initiated upwind near the Rocky Mountains Foothills and propagate eastward to the central United States (Feng et al., 2019).

Since SM can indirectly exert feedback on cloud and precipitation through heating or cooling the surface (Duerinck et al., 2016), we further delve into examining spatially the samples of antecedent 7-hour LST (K) and their contribution to cloud, $f_2(x_2)$ (dBZ), and rainfall, ΔR (mm/hour), in Figures 3d-f and S6, respectively. $f_2(x_2)$ exhibits a non-linear dependence on LST where LST anomalies exert the most significant influence. From Figure 3d-e, it is suggested that LST above 305 K accounts for an increase of at most 4.0 dBZ in the cloud reflectivity and 2.0 mm/hour in rainfall rate (see Figure S6) at both 2.0 and 3.5 km. On the contrary, the samples with a cooler surface (LST<290 K) seem to foster a more stable atmospheric state, thereby reducing the cloud reflectivity, especially in the near-surface atmosphere ($h \approx 2.0$ km). The underlying LST-driven mechanisms were discussed in the previous section.

Geographically, the most significant effects of these anomalies are evident and clustered in the southwest of the study region, delineated by 101°W-105°W and 32°N-36°N. Within this area, we find a moderate negative correlation ($R = -0.41$, shown in Figure S7a) between surface SM and the LST component function, $f_2(x_2)$. Moreover, we illustrate in Figure S7b that LST contributes to CVP preferentially over dry soil with saturation between 0.1 and 0.4. These findings underscore the presence of the intrinsic SM-LST coupling nested within the SMCPF pathways (Seneviratne et al., 2010), and we can conveniently interpret $f_2(x_2)$ as a proxy for the indirect and negative feedback of SM on CVP. Notably, our identified negative feedback region (101°W-105°W, 32°N-36°N) is consistent with the one proposed by Findell and Eltahir (2003b) (represented by the black dashed line in Figure 3d-e). Several factors can play a role when it comes to the sources of convective clouds and precipitation over the dry soil. For instance, the monsoonal moisture incursion into New Mexico can bring up local humidity and offset the reduced evapotranspiration from the local dry soils (Wallace et al., 1999; Klein & Taylor, 2020). Besides, the Great Plains Low-Level Jet (GPLLJ) can transport abundant moisture southerly from the Gulf of Mexico into the central United States (Ford, Rapp, & Quiring, 2015; Feng et al., 2016).

5 Discussion and Conclusion

This study presents a data-driven approach that uses the functional decomposition of a large database of satellite-measured SM (SMAP/L4) and CVP (GPM/DPR/L2A) for disentangling and quantifying SMCPF in the central United States. Results show that the signs and strengths of the feedback differ among cloud heights and geographical locations. A significant positive feedback is observed in the lower atmosphere, particularly between 1.0 and 3.0 km with a temporal lag of 7 hours. With a degree of saturation over 0.4, wet soil can potentially increase the cloud reflectivity and rainfall rate by up to 4.0 dBZ and 2.0 mm/hour at $h \approx 2.0$ km, evidently in northern Texas, central Oklahoma, northwestern and southeastern Kansas. The negative feedback, indirectly interpreted by the anomalies of LST, is effective with a wider vertical extension from 1.0 km to 4.0 km and a time lag of 7-10 hours. These LST anomalies can explain comparable increments in cloud reflectivity and rainfall rate to SM but in northwestern Texas and southeastern and eastern New Mexico. The identified patterns of SMCPF align qualitatively with previous studies that utilize simulations and observations to investigate the underlying mechanisms and regional categorizations of the feedback (Findell & Eltahir, 2003a, 2003b; Qian et al., 2013; Sathyanadh et al., 2017; Su & Dickinson, 2017; Koukoula et al., 2019; Hu et al., 2021; Ford et al., 2023).

Our approach brings new insights into the observational understanding of the SMCPF characterized by cloud height, time lag, and location and possesses the potential for coupled land-atmosphere model diagnosis. Despite this, certain limitations are highlighted. Even though a decent amount of samples was obtained, they can hardly support extensive analyses over seasonal, interannual, or localized scales due to the substantial downsampling. Another possible limitation is the selection of only five land and atmospheric variables as inputs of the HDMR emulator. We reiterate that this decision is strategically aimed at maximizing the capture of the nonlinear relationship and causal link between cloud and SM. Nonetheless, it concurrently overlooks other pertinent variables that could play a significant role in the SMCPF pathways.

For future work, it is important to conduct a comprehensive analysis employing cloud model simulations and/or reanalysis data sets as inputs of HDMR. This will help diagnose the representativeness of the current-generation coupled land-atmosphere models. We should also build robust HDMR emulators to be integrated with state-of-the-art cloud models for more accurate prediction of convective clouds and precipitation. This necessitates the incorporation of more predictors such as SM gradient (Taylor, 2015; Zhou et al., 2021; Graf et al., 2021; Chug et al., 2023) and evaporative fraction (Taylor et al., 2013; Ford et al., 2023), along with atmospheric variables like wind speed and water vapor mixing ratio (Raymond & Sessions, 2007; Seneviratne et al., 2010). Last but not least, with the advancement of a variety of reanalysis datasets, the methodology can be useful for examining the changes in SMCPF under increasing hydroclimatic extremes at the regional and global scales.

Acknowledgments

The authors would like to acknowledge the support of NASA through the Global Precipitation Measurement Mission program (Grant 80NSSC22K0597) and the Weather and Atmospheric Dynamics program (Grant 80NSSC23K1304), as well as the support of the NSF Division of Information and Intelligent Systems (ExpandAI2ES Grant IIS2324008). The author also wishes to thank the support of the UCI Engineering-Los Alamos National Laboratory (LANL) Fellowship, provided by the UCI Samueli School of Engineering and LANL.

Data and Software Availability

The SMAP/L4 (L4_SM) product is obtained from the National Snow and Ice Data Center at https://nsidc.org/data/spl4smau/versions/7 (Reichle et al., 2022). The GPM/DPR/L2A product (GPM_2ADPR) is obtained from the Goddard Earth Sciences Data and Information Services Center at https://disc.gsfc.nasa.gov/datasets/GPM_2AKaENV_07/summary (Iguchi et al., 2010). MATLAB postprocessing software will be archived in Zenodo along with the final data set of collocated SMAP and DPR samples. A copy of this data set is provided for review in the supporting information.

References

Baker, J. C. A., Castilho de Souza, D., Kubota, P. Y., Buermann, W., Coelho, C. A. S., Andrews, M. B., Gloor, M., Garcia-Carreras, L., Figueroa, S. N., & Spracklen, D. V. (2021). An assessment of land–atmosphere interactions over South America using satellites, reanalysis, and two global climate models. *Journal of Hydrometeorology*, *22*(4), 905–922. https://doi.org/10.1175/JHM-D-20-0132.1

Baker, J. C. A., Garcia-Carreras, L., Buermann, W., De Souza, D. C., Marsham, J. H., Kubota, P. Y., Gloor, M., Coelho, C. A. S., & Spracklen, D. V. (2021). Robust Amazon precipitation projections in climate models that capture realistic land–atmosphere interactions. *Environmental Research Letters*, *16*(7), 074002. https://doi.org/10.1088/1748-9326/abfb2e

Berg, A., Findell, K., Lintner, B. R., Gentine, P., & Kerr, C. (2013). Precipitation sensitivity to surface heat fluxes over north america in reanalysis and model data. *Journal of Hydrometeorology*, *14*(3), 722–743. https://doi.org/10.1175/JHM-D-12-0111.1

Chug, D., Dominguez, F., Taylor, C. M., Klein, C., & Nesbitt, S. W. (2023). Dry-to-Wet Soil Gradients Enhance Convection and Rainfall over Subtropical South America. *arXiv preprint arXiv:2304.04630*. https://doi.org/10.1175/JHM-D-23-0031.1

Cioni, G., & Hohenegger, C. (2017). Effect of soil moisture on diurnal convection and precipitation in large-eddy simulations. *Journal of Hydrometeorology*, *18*(7), 1885–1903. https://doi.org/10.1175/JHM-D-16-0241.1

Deardorff, J. W. (1980). Stratocumulus-capped mixed layers derived from a three-dimensional model. *Boundary-layer meteorology*, *18*, 495–527. https://doi.org/10.1007/BF00119502

Dirmeyer, P. A., Jin, Y., Singh, B., & Yan, X. (2013). Trends in land–atmosphere interactions from CMIP5 simulations. *Journal of Hydrometeorology*, *14*(3), 829–849. https://doi.org/10.1175/JHM-D-12-0107.1

Duerinck, H. M., Van der Ent, R. J., Van de Giesen, N. C., Schoups, G., Babovic, V., & Yeh, P. J. F. (2016). Observed soil moisture–precipitation feedback in Illinois: A systematic analysis over different scales. *Journal of Hydrometeorology*, *17*(6), 1645–1660. https://doi.org/10.1175/JHM-D-15-0032.1

Ek, M. B., & Holtslag, A. A. M. (2004). Influence of soil moisture on boundary layer cloud development. *Journal of hydrometeorology*, *5*(1), 86–99. https://doi.org/10.1175/1525-7541(2004)005⟨0086:IOSMOB⟩2.0.CO;2

Falchi, A., Minisci, E., Kubicek, M., Vasile, M., & Lemmens, S. (2018). HDMR-based sensitivity analysis and uncertainty quantification of GOCE aerodynamics using DSMC. *Stardust Final Conference*, 301–323. https://doi.org/10.1007/978-3-319-69956-1_18

Fast, J. D., Berg, L. K., Feng, Z., Mei, F., Newsom, R., Sakaguchi, K., & Xiao, H. (2019). The impact of variable land-atmosphere coupling on convective cloud populations observed during the 2016 HI-SCALE field campaign. *Journal of Advances in Modeling Earth Systems*, *11*(8), 2629–2654. https://doi.org/10.1029/2019MS001727

Feng, Z., Houze Jr, R. A., Leung, L. R., Song, F., Hardin, J. C., Wang, J., Gustafson Jr, W. I., & Homeyer, C. R. (2019). Spatiotemporal characteristics and large-scale environments of mesoscale convective systems east of the Rocky Mountains. *Journal of Climate*, *32*(21), 7303–7328. https://doi.org/10.1175/JCLI-D-19-0137.1

Feng, Z., Leung, L. R., Hagos, S., Houze, R. A., Burleyson, C. D., & Balaguru, K. (2016). More frequent intense and long-lived storms dominate the springtime trend in central US rainfall. *Nature communications*, *7*(1), 13429. https://doi.org/10.1038/ncomms13429

Ferguson, C. R., & Wood, E. F. (2011). Observed land-atmosphere coupling from satellite remote sensing and reanalysis. *Journal of Hydrometeorology*, *12*(6), 1221–1254. https://doi.org/10.1175/2011JHM1380.1

Findell, K. L., & Eltahir, E. A. B. (2003b). Atmospheric controls on soil moisture–boundary layer interactions. Part II: Feedbacks within the continental United States. *Journal of Hydrometeorology*, *4* (3), 570–583. https://doi.org/10.1175/1525-7541(2003)004⟨0570:ACOSML⟩2.0.CO;2

Findell, K. L., & Eltahir, E. A. B. (2003a). Atmospheric controls on soil moisture–boundary layer interactions. Part I: Framework development. *Journal of Hydrometeorology*, *4*, 552–569. https://doi.org/10.1175/1525-7541(2003)004⟨0552:ACOSML⟩2.0.CO;2

Findell, K. L., Gentine, P., Lintner, B. R., & Kerr, C. (2011). Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation. *Nature Geoscience*, *4* (7), 434–439. https://doi.org/10.1038/ngeo1174

Ford, T. W., Quiring, S. M., Frauenfeld, O. W., & Rapp, A. D. (2015). Synoptic conditions related to soil moisture-atmosphere interactions and unorganized convection in Oklahoma. *Journal of Geophysical Research: Atmospheres*, *120* (22), 11–519.

Ford, T. W., Rapp, A. D., & Quiring, S. M. (2015). Does afternoon precipitation occur preferentially over dry or wet soils in Oklahoma? *Journal of Hydrometeorology*, *16* (2), 874–888. https://doi.org/10.1175/JHM-D-14-0005.1

Ford, T. W., Rapp, A. D., Quiring, S. M., & Blake, J. (2015). Soil moisture–precipitation coupling: Observations from the Oklahoma Mesonet and underlying physical mechanisms. *Hydrology and Earth System. Sciences*, *19*, 3617–3631. https://doi.org/10.5194/hess-19-3617-2015

Ford, T. W., Steiner, J., Mason, B., & Quiring, S. M. (2023). Observation-Driven Characterization of Soil Moisture-Precipitation Interactions in the Central United States. *Journal of Geophysical Research: Atmospheres*, *128* (12), e2022JD037934. https://doi.org/10.1029/2022JD037934

Frye, J. D., & Mote, T. L. (2010). Convection initiation along soil moisture boundaries in the southern Great Plains. *Monthly weather review*, *138* (4), 1140–1151. https://doi.org/10.1175/2009MWR2865.1

Gao, Y., Sahin, A., & Vrugt, J. A. (2023). Probabilistic sensitivity analysis with dependent variables: Covariance-based decomposition of hydrologic models. *Water Resources Research*, *59*(4), e2022WR032834. https://doi.org/10.1029/2022WR032834

Gentine, P., Holtslag, A. A., D'Andrea, F., & Ek, M. (2013). Surface and atmospheric controls on the onset of moist convection over land. *Journal of Hydrometeorology*, *14*, 1443–1462. https://doi.org/10.1175/JHM-D-12-0137.1

Golaz, J. C., Jiang, H., & Cotton, W. R. (2001). A large-eddy simulation study of cumulus clouds over land and sensitivity to soil moisture. *Atmospheric research*, *59*, 373–392. https://doi.org/10.1016/S0169-8095(01)00113-2

Graf, M., Arnault, J., Fersch, B., & Kunstmann, H. (2021). Is the soil moisture precipitation feedback enhanced by heterogeneity and dry soils? A comparative study. *Hydrological Processes*, *35*(9), e14332. https://doi.org/10.1002/hyp.14332

Guillod, B. P., Orlowsky, B., Miralles, D., Teuling, A. J., Blanken, P. D., Buchmann, N., Ciais, P., Ek, M., Findell, K. L., Gentine, P., Lintner, B. R., Scott, R. L., Van den Hurk, B., & I. Seneviratne, S. (2014). Land-surface controls on afternoon precipitation diagnosed from observational data: Uncertainties and confounding factors. *Atmospheric Chemistry and Physics*, *14*(16), 8343–8367. https://doi.org/10.5194/acp-14-8343-2014

Guillod, B. P., Orlowsky, B., Miralles, D. G., Teuling, A. J., & Seneviratne, S. I. (2015). Reconciling spatial and temporal soil moisture effects on afternoon rainfall. *Nature communications*, *6*(1), 6443. https://doi.org/10.1038/ncomms7443

Guo, Z., Dirmeyer, P. A., Koster, R. D., Sud, Y. C., Bonan, G., Oleson, K. W., Chan, E., Verseghy, D., Cox, P., Gordon, C. T., et al. (2006). GLACE: the global land–atmosphere coupling experiment. Part II: analysis. *Journal of Hydrometeorology*, *7*(4), 611–625. https://doi.org/10.1175/JHM511.1

Han, C., Brdar, S., & Kollet, S. (2019). Response of convective boundary layer and shallow cumulus to soil moisture heterogeneity: A large-eddy simulation study. *Journal of advances in modeling earth systems*, *11*(12), 4305–4322. https://doi.org/10.1029/2019MS001772

Hohenegger, C., Brockhaus, P., Bretherton, C. S., & Schär, C. (2009). The soil moisture-precipitation feedback in simulations with explicit and parameterized convection. *Journal of Climate*, *22*(19), 5003–5020. https://doi.org/10.1175/2009JCLI2604.1

Holloway, C. E., & Neelin, J. D. (2010). Temporal relations of column water vapor and tropical precipitation. *Journal of the Atmospheric Sciences*, *67*(4), 1091–1105. https://doi.org/10.1175/2009JAS3284.1

Hooker, G. (2007). Generalized functional ANOVA diagnostics for high-dimensional functions of dependent variables. *Journal of Computational and Graphical Statistics*, *16*(3), 709–732. https://doi.org/10.1198/106186007X237892

Hu, H., Leung, L. R., & Feng, Z. (2021). Early warm-season mesoscale convective systems dominate soil moisture–precipitation feedback for summer rainfall in central United States. *Proceedings of the National Academy of Sciences*, *118*(43), e2105260118. https://doi.org/10.1073/pnas.2105260118

Iguchi, T., Seto, S., Meneghini, R., Yoshida, N., Awaka, J., Le, M., Chandrasekar, V., & Kubota, T. (2010). GPM/DPR level-2 algorithm theoretical basis document. *NASA Goddard Space Flight Center*.

Klein, C., & Taylor, C. M. (2020). Dry soils can intensify mesoscale convective systems. *Proceedings of the National Academy of Sciences*, *117*(35), 21132–21137. https://doi.org/10.1073/pnas.2007998117

Knist, S., Goergen, K., Buonomo, E., Christensen, O. B., Colette, A., Cardoso, R. M., Fealy, R., Fernández, J., García-Díez, M., Jacob, D., Kartsios, S., Katragkou, E., Keuler, K., Mayer, S., van Meijgaard, E., Nikulin, G., Soares, P. M. M., Sobolowski, S., Szepszo, G., ... Simmer, C. (2017). Land-atmosphere coupling in euro-cordex evaluation experiments. *Journal of Geophysical Research: Atmospheres*, *122*(1), 79–103. https://doi.org/10.1002/2016JD025476

Koster, R. D., Liu, Q., Mahanama, S. P. P., & Reichle, R. H. (2018). Improved hydrological simulation using SMAP data: Relative impacts of model calibration and data assimilation. *Journal of Hydrometeorology*, *19*(4), 727–741. https://doi.org/10.1175/JHM-D-17-0228.1

Koster, R. D., Sud, Y. C., Guo, Z., Dirmeyer, P. A., Bonan, G., Oleson, K. W., Chan, E., Verseghy, D., Cox, P., Davies, H., et al. (2006). GLACE: the global land–atmosphere coupling experiment. Part I: overview. *Journal of Hydrometeorology*, *7*(4), 590–610. https://doi.org/10.1175/JHM510.1

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C. T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C.-H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., . . . Yamada, T. (2004). Regions of strong coupling between soil moisture and precipitation. *Science*, *305*(5687), 1138–1140. https://doi.org/10.1126/science.1100217

Koukoula, M., Nikolopoulos, E. I., Kushta, J., Bartsotas, N. S., Kallos, G., & Anagnostou, E. N. (2019). A numerical sensitivity analysis of soil moisture feedback on convective precipitation. *Journal of Hydrometeorology*, *20*(1), 23–44. https://doi.org/10.1175/JHM-D-18-0134.1

Krakauer, N. Y., Cook, B. I., & Puma, M. J. (2010). Contribution of soil moisture feedback to hydroclimatic variability. *Hydrology and Earth System Sciences*, *14*(3), 505–520. https://doi.org/10.5194/hess-14-505-2010

Kucherenko, S., Feil, B., Shah, N., & Mauntz, W. (2011). The identification of model effective dimensions using global sensitivity analysis. *Reliability Engineering & System Safety*, *96*(4), 440–449. https://doi.org/10.1016/j.ress.2010.11.003

Lasser, M., O, S., & Foelsche, U. (2019). Evaluation of GPM-DPR precipitation estimates with WegenerNet gauge data. *Atmospheric Measurement Techniques*, *12*(9), 5055–5070. https://doi.org/10.5194/amt-12-5055-2019

Li, G., & Rabitz, H. (2010). D-MORPH regression: application to modeling with unknown parameters more than observation data. *Journal of Mathematical Chemistry*, *48*, 1010–1035. https://doi.org/10.1007/s10910-010-9722-2

Li, G., & Rabitz, H. (2012). General formulation of HDMR component functions with independent and correlated variables. *Journal of Mathematical Chemistry*, *50*(1), 99–130. https://doi.org/10.1007/s10910-011-9898-0

Liao, L., & Meneghini, R. (2022). GPM DPR retrievals: Algorithm, evaluation, and validation. *Remote Sensing*, *14*(4), 843. https://doi.org/10.3390/rs14040843

Liu, W., Zhang, Q., Li, C., Xu, L., & Xiao, W. (2022). The influence of soil moisture on convective activity: A review. *Theoretical and Applied Climatology*, *149*(1-2), 221–232. https://doi.org/10.1007/s00704-022-04046-z

Marshall, J., & Palmer, W. M. (1948). The distribution of raindrops with size. *Journal of meteorology*, *5*, 166.

Miralles, A. J., Diego G. Teuling, van Heerwaarden, C. C., & Vilà-Guerau de Arellano, J. (2014). Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature Geoscience*, *7*, 345–349. https://doi.org/10.1038/ngeo2141

Pejcic, V., Saavedra Garfias, P., Mühlbauer, K., Trömel, S., & Simmer, C. (2020). Comparison between precipitation estimates of ground-based weather radar composites and GPM's DPR rainfall product over Germany. *Meteorologische Zeitschrift*, *29*(6), 451–466. https://doi.org/10.1127/metz/2020/1039

Qian, Y., Huang, M., Yang, B., & Berg, L. K. (2013). A modeling study of irrigation effects on surface fluxes and land–air–cloud interactions in the Southern Great Plains. *Journal of Hydrometeorology*, *14*(3), 700–721. https://doi.org/10.1175/JHM-D-12-0134.1

Rabitz, H., & Aliş, Ö. F. (1999). General foundations of high-dimensional model representations. *Journal of Mathematical Chemistry*, *25*(2), 197–233. https://doi.org/10.1023/A:1019188517934

Raymond, D. J., & Sessions, S. L. (2007). Evolution of convection during tropical cyclogenesis. *Geophysical research letters*, *34*(6). https://doi.org/10.1029/2006GL028607

Reichle, R. H., De Lannoy, G., Koster, R. D., Crow, W. T., Kimball, J. S., & Liu, Q. (2022). SMAP L4 global 3-hourly 9 km EASE-Grid surface and root zone soil moisture analysis update, version 7 [Dataset]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/LWJ6TF5SZRG3

Reichle, R. H., De Lannoy, G. J. M., Liu, Q., Koster, R. D., Kimball, J. S., Crow, W. T., Ardizzone, J. V., Chakraborty, P., Collins, D. W., Conaty, A. L., et al. (2017). Global assessment of the SMAP Level-4 surface and root-zone soil moisture product using assimilation diagnostics. *Journal of hydrometeorology*, *18*(12), 3217–3237. https://doi.org/10.1175/JHM-D-17-0130.1

Reichle, R. H., Lucchesi, R. A., Ardizzone, J. V., Kim, G.-K., Smith, E. B., & Weiss, B. H. (2015). *Soil moisture active passive (SMAP) mission level 4 surface and root zone soil moisture (L4_SM) product specification document* (tech. rep.). NASA Goddard Space Flight Center. Greenbelt, MD, United States. https://nsidc.org/sites/default/files/reichle789.pdf

Sakaguchi, K., Berg, L. K., Chen, J., Fast, J., Newsom, R., Tai, S. .-., Yang, Z., Gustafson Jr, W. I., Gaudet, B. J., Huang, M., Pekour, M., Pressel, K., & Xiao, H. (2022). Determining spatial scales of soil moisture—Cloud coupling pathways using semi-idealized simulations. *Journal of Geophysical Research: Atmospheres*, *127*(2), e2021JD035282. https://doi.org/10.1029/2021JD035282

Santanello, J. A., Peters-Lidard, C. D., Kennedy, A., & Kumar, S. V. (2013). Diagnosing the nature of land-atmosphere coupling: A case study of dry/wet extremes in the US southern Great Plains. *Journal of Hydrometeorology*, *14*(1), 3–24. https://doi.org/10.1175/JHM-D-12-023.1

Santanello, J. A., Peters-Lidard, C. D., Kumar, S. V., Alonge, C., & Tao, W.-K. (2009). A modeling and observational framework for diagnosing local land–atmosphere coupling on diurnal time scales. *Journal of Hydrometeorology*, *10*(3), 577–599. https://doi.org/10.1175/2009JHM1066.1

Sathyanadh, A., Prabha, T. V., Balaji, B., Resmi, E. A., & Karipot, A. (2017). Evaluation of WRF PBL parameterization schemes against direct observations during a dry event over the Ganges valley. *Atmospheric Research*, *193*, 125–141. https://doi.org/10.1016/j.atmosres.2017.02.016

Savoy, P., & Mackay, D. S. (2015). Modeling the seasonal dynamics of leaf area index based on environmental constraints to canopy development. *Agricultural and Forest Meteorology*, *200*, 46–56. https://doi.org/10.1016/j.agrformet.2014.09.019

Schär, C., Lüthi, D., Beyerle, U., & Heise, E. (1999). The soil–precipitation feedback: A process study with a regional climate model. *Journal of Climate*, *12*(3), 722–741. https://doi.org/10.1175/1520-0442(1999)012⟨0722:TSPFAP⟩2.0.CO;2

Schlemmer, L., Hohenegger, C., Schmidli, J., & Schär, C. (2012). Diurnal equilibrium convection and land surface–atmosphere interactions in an idealized cloud-resolving model. *Quarterly Journal of the Royal Meteorological Society*, *138*, 1526–1539. https://doi.org/10.1002/qj.1892

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., & Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, *99*(3), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Seneviratne, S. I., Lüthi, D., Litschi, M., & Schär, C. (2006). Land–atmosphere coupling and climate change in europe. *Nature*, *443*, 205–209. https://doi.org/10.1038/nature05095

Seneviratne, S. I., Wilhelm, M., Stanelle, T., van den Hurk, B., Hagemann, S., Berg, A., Cheruy, F., Higgins, M. E., Meier, A., Brovkin, V., Claussen, M., Ducharne, A., Dufresne, J.-L., Findell, K. L., Ghattas, J., Lawrence, D. M., Malyshev, S., Rummukainen, M., & Smith, B. (2013). Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment. *Geophysical Research Letters*, *40*(19), 5212–5217. https://doi.org/10.1002/grl.50956

Shereena, O. A., & Rao, B. N. (2019). HDMR-Based Bayesian Structural System Identification. In *Recent advances in structural engineering, volume 1* (pp. 453–464). Springer. https://doi.org/10.1007/978-981-13-0362-3_36

Sherwood, S. C. (1999). Convective precursors and predictability in the tropical western Pacific. *Monthly Weather Review*, *127*(12), 2977–2991. https://doi.org/10.1175/1520-0493(1999)127⟨2977:CPAPIT⟩2.0.CO;2

Sobol′, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiment*, *1*(4), 407–414.

Song, J. I., Yum, S. S., Park, S. H., Kim, K. H., Park, K. J., & Joo, S. W. (2021). Climatology of melting layer heights estimated from cloud radar observations at various locations. *Journal of Geophysical Research: Atmospheres*, *126*(17), e2021JD034816. https://doi.org/10.1029/2021JD034816

Spennemann, P. C., Salvia, M., Ruscica, R. C., Sörensson, A. A., Grings, F., & Karszenbaum, H. (2018). Land-atmosphere interaction patterns in southeastern South America using satellite products and climate models. *International journal of applied earth observation and geoinformation*, *64*, 96–103. https://doi.org/10.1016/j.jag.2017.08.016

Su, H., & Dickinson, R. E. (2017). On the spatial gradient of soil moisture–precipitation feedback strength in the April 2011 drought in the Southern Great Plains. *Journal of Climate*, *30*(3), 829–848. https://doi.org/10.1175/JCLI-D-13-00185.1

Tavakol, A., Rahmani, V., Quiring, S. M., & Kumar, S. V. (2019). Evaluation analysis of NASA SMAP L3 and L4 and SPoRT-LIS soil moisture data in the United States. *Remote Sensing of Environment*, *229*, 234–246. https://doi.org/10.1016/j.rse.2019.05.006

Tawfik, A. B., Dirmeyer, P. A., & Santanello, J. A. (2015). The heated condensation framework.: Part i: Description and southern great plains case study. *Journal of Hydrometeorology*, *16*(5), 1929–1945. Retrieved July 25, 2023, from http://www.jstor.org/stable/24915510

Taylor, C. M. (2015). Detecting soil moisture impacts on convective initiation in Europe. *Geophysical Research Letters*, *42*(11), 4631–4638. https://doi.org/10.1002/2015GL064030

Taylor, C. M., Birch, C. E., Parker, D. J., Dixon, N., Guichard, F., Nikulin, G., & Lister, G. M. S. (2013). Modeling soil moisture-precipitation feedback in the Sahel: Importance of spatial scale versus convective parameterization. *Geophysical Research Letters*, *40*(23), 6213–6218. https://doi.org/10.1002/2013GL058511

Taylor, C. M., de Jeu, R. A. M., Guichard, ., Harris, P. P., & Dorigo, W. A. (2012). Afternoon rain more likely over drier soils. *Nature*, *489*(7416), 423–426. https://doi.org/10.1038/nature11377

Taylor, C. M., & Ellis, R. J. (2006). Satellite detection of soil moisture impacts on convection at the mesoscale. *Geophysical Research Letters*, *33*(3). https://doi.org/10.1029/2005GL025252

Taylor, C. M., Gounou, A., Guichard, F., Harris, P. P., Ellis, R. J., Couvreux, F., & De Kauwe, M. (2011). Frequency of Sahelian storm initiation enhanced over mesoscale soil-moisture patterns. *Nature Geoscience*, *4*(7), 430–433. https://doi.org/10.1038/ngeo1173

Taylor, C. M., Harris, P. P., & Parker, D. J. (2010). Impact of soil moisture on the development of a Sahelian mesoscale convective system: A case-study from the AMMA special observing period. *Quarterly Journal of the Royal Meteorological Society*, *136*(S1), 456–470. https://doi.org/10.1002/qj.465

Thompson, G., Field, P. R., Rasmussen, R. M., & Hall, W. D. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly weather review*, *136*(12), 5095–5115. https://doi.org/10.1175/2008MWR2387.1

Thompson, G., Rasmussen, R. M., & Manning, K. (2004). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Monthly Weather Review*, *132*(2), 519–542. https://doi.org/10.1175/1520-0493(2004)132⟨0519:EFOWPU⟩2.0.CO;2

Trenberth, K. E. (1999). Atmospheric moisture recycling: Role of advection and local evaporation. *Journal of Climate*, *12*(5), 1368–1381. https://doi.org/10.1175/1520-0442(1999)012⟨1368:AMRROA⟩2.0.CO;2

Tuttle, S. E., & Salvucci, G. D. (2017). Confounding factors in determining causal soil moisture-precipitation feedback. *Water Resources Research*, *53*(7), 5531–5544. https://doi.org/10.1002/2016WR019869

Wallace, C. E., Maddox, R. A., & Howard, K. W. (1999). Summertime convective storm environments in central Arizona: Local observations. *Weather and Forecasting*, *14*(6), 994–1006. https://doi.org/10.1175/1520-0434(1999)014⟨0994:SCSEIC⟩2.0.CO;2

Wang, G., Kim, Y., & Wang, D. (2007). Quantifying the strength of soil moisture–precipitation coupling and its sensitivity to changes in surface water budget. *Journal of Hydrometeorology*, *8*(3), 551–570. https://doi.org/10.1175/JHM573.1

Wang, H., Chen, L., Ye, F., & Chen, L. (2017). Global sensitivity analysis for fiber reinforced composite fiber path based on D-MORPH-HDMR algorithm. *Structural and Multidisciplinary Optimization*, *56*(3), 697–712. https://doi.org/10.1007/s00158-017-1681-9

Williams, I. N. (2019). Evaluating soil moisture feedback on convective triggering: Roles of convective and land-model parameterizations. *Journal of Geophysical Research: Atmospheres*, *124*(1), 317–332. https://doi.org/10.1029/2018JD029326

Xu, Z., Chen, H., Guo, J., & Zhang, W. (2021). Contrasting effect of soil moisture on the daytime boundary layer under different thermodynamic conditions in summer over China. *Geophysical Research Letters*, *48*(3), e2020GL090989. https://doi.org/10.1029/2020GL090989

Yin, J., Albertson, J. D., Rigby, J. R., & Porporato, A. (2015). Land and atmospheric controls on initiation and intensity of moist convection: CAPE dynamics and LCL crossings. *Water Resources Research*, *51*(10), 8476–8493. https://doi.org/10.1002/2015WR017286

Yin, J., Porporato, A., & Albertson, J. (2014). Interplay of climate seasonality and soil moisture-rainfall feedback. *Water Resources Research*, *50*(7), 6053–6066. https://doi.org/10.1002/2013WR014772

Yuan, S., Wang, Y., Quiring, S. M., Ford, T. W., & Houston, A. L. (2020). A sensitivity study on the response of convection initiation to in situ soil moisture in the central United States. *Climate Dynamics*, *54*, 2013–2028. https://doi.org/10.1007/s00382-019-05098-0

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., et al. (2016). Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bulletin of the American Meteorological Society*, *97*(4), 621–638. https://doi.org/10.1175/BAMS-D-14-00174.1

Zhang, L., He, C., & Zhang, M. (2017). Multi-scale evaluation of the SMAP product using sparse in-situ network over a high mountainous watershed, Northwest China. *Remote Sensing*, *9*(11), 1111. https://doi.org/10.3390/rs9111111

Zhang, X., Zhang, T., Zhou, P., Shao, Y., & Gao, S. (2017). Validation analysis of SMAP and AMSR2 soil moisture products over the United States using ground-based measurements. *Remote Sensing*, *9*(2), 104. https://doi.org/10.3390/rs9020104

Zhou, S., Williams, A. P., Lintner, B. R., Berg, A. M., Zhang, Y., Keenan, T. F., Cook, B. I., Hagemann, S., Seneviratne, S. I., & Gentine, P. (2021). Soil moisture–atmosphere feedbacks mitigate declining water availability in drylands. *Nature Climate Change*, *11*(1), 38–44. https://doi.org/10.1038/s41558-020-00945-z

Soil Moisture Cloud Precipitation Feedback in the Lower Atmosphere from Functional Decomposition of Satellite Observations

Yifu Gao*, Clément Guilloteau*, Efi Foufoula-Georgiou*, Chonggang Xu†, Xiaoming Sun†, and Jasper A. Vrugt*‡

May 22, 2024

Key points

1. We present a CPU-friendly functional decomposition of satellite-measured soil moisture (SM) and cloud vertical profiles
2. The sign and strength of SM's feedback vary with height, time lag, and geographic locations, which agrees with more qualitative studies
3. The presented approach exhibits potential implications for diagnosing cloud models, particularly in the context of land-atmosphere coupling

* Department of Civil and Environmental Engineering, University of California, Irvine, California, USA.
† Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA.
‡ Corresponding author: jasper@uci.edu

Abstract

The feedback of topsoil moisture (SM) content on convective clouds and precipitation is not well understood and represented in the current generation of coupled cloud physics and land-surface models. Here, we use functional decomposition of satellite-derived SM (SMAP/L4) and cloud vertical profiles (CVP: GPM/DPR/L2A) in the central US to quantify the relationship between SM and the vertical distribution of cloud water. High-dimensional model representation disentangles the contributions of SM and other land-surface and atmospheric variables to the CVP. Results show the sign and strength of this feedback varies with cloud height and time lag and displays a large spatial variability. Positive anomalies in the antecedent 7-hour SM and land-surface temperature can increase reflectivity up to 4 dBZ in the lower atmosphere (1-3 km above the surface). The presented approach brings new insights into observational understanding of SM-precipitation feedback and possesses the potential for diagnosing cloud models regarding land-atmosphere coupling representation.

Plain Language Summary

This paper focuses on the observational analysis of how soil moisture (SM) influences the vertical cloud-water distribution throughout the day. By analyzing data from Soil Moisture Active Passive (SMAP) and Dual-frequency Precipitation Radar (DPR), we gain insights into how antecedent SM levels impact cloud-water reflectivity at different heights in the lower atmosphere. Our data-driven approach produces spatial maps of SM's contribution to cloud reflectivity and rainfall in the central US conditioned on cloud height and SM time lag. The results will help diagnose coupled land-atmosphere models.

1 Introduction

The feedbacks between soil moisture (SM) and precipitation play a critical role in regulating regional hydroclimatic variability. Such feedbacks are governed by a plethora of variables and processes, such as (variations in) land surface temperature (Koster et al., 2006), energy partitioning (Golaz et al., 2001; Fast et al., 2019; Sakaguchi et al., 2022), planetary boundary layer (PBL) development (Ek & Holtslag, 2004; Han et al., 2019) and the initiation of convective clouds and precipitation (Ferguson & Wood, 2011; Taylor et al., 2011; Cioni & Hohenegger, 2017). These feedbacks take place across a continuum of spatiotemporal scales, spanning distances from several to thousands of kilometers and time span of days to seasons (Trenberth, 1999; Duerinck et al., 2016; Liu et al., 2022). Moreover, SM-precipitation feedbacks exhibit substantial regional variability in both their sign and magnitude as a result of the large sensitivity of evapotranspiration and atmospheric conditions to SM and latent heat fluxes, respectively (Guo et al., 2006). In this paper, we focus our attention on diurnal SM-cloud-precipitation feedbacks, abbreviated SMCPF, which control in part the vertical cloud-water distribution, thereby influencing weather conditions (Koster et al., 2004) and regional hydroclimatology (Krakauer et al., 2010; Yin et al., 2014; Ford et al., 2023). Future climate projections suggest further that SMCPFs may play an increasing role in determining changes in mean temperature and extremes as a result of larger SM deficits under higher evaporative demands (Dirmeyer et al., 2013; Seneviratne et al., 2013; Taylor, 2015).

Given the importance of the SMCPF in regulating local and regional weather, much research has been devoted to estimating its sign, causality, and physical linkage. That research may be divided into simulation-based analysis (Schär et al., 1999; Findell & Eltahir, 2003a; G. Wang et al., 2007; Hohenegger et al., 2009; Schlemmer et al., 2012; Tawfik et al., 2015; Gentine et al., 2013), observation-based studies (Taylor & Ellis, 2006; Santanello et al., 2009; Taylor et al., 2010; Ferguson & Wood, 2011; Taylor et al., 2011; Ford, Rapp, Quiring, & Blake, 2015; Guillod et al., 2015) and a combination thereof (Seneviratne et al., 2006; Santanello et al., 2013; Miralles et al., 2014; Spennemann et al., 2018; Baker, Castilho de Souza, et al., 2021; Baker, Garcia-Carreras, et al., 2021). Although numerical models of land-atmosphere interactions have advanced considerably in recent decades, the diurnal impact of SM on cloud formation and composition is still not particularly well understood. The mechanisms governing the sign and strength of the simulated SMCPFs are subject to a large uncertainty depending for example on the choice of boundary conditions (Hohenegger et al., 2009) and sub-grid scale process representation (Deardorff, 1980; Thompson et al., 2004, 2008). In observational studies, on the other hand, it is difficult to filter out the effects of synoptic variability. Moreover, in the absence of high-quality spatiotemporal measurements of SM and cloud vertical profiles, past studies have mainly focused on how (gradients of) SM affect convection initiation, the PBL height, and precipitation probability (Frye & Mote, 2010; Findell et al., 2011; Taylor, 2015; Su & Dickinson, 2017; Graf et al., 2021; Yuan et al., 2020; Ford et al., 2023) without recourse to mesoscale diurnal relationships between antecedent SM and the cloud water distribution. Advances in our understanding of SM-cloud relationships should improve the diagnosis of weather and climate models and enhance the accuracy of their future projections (Williams, 2019).

Fortunately, remote-sensing data products of SM and the cloud vertical profile from polar-orbiting Earth-observing satellites have advanced considerably in the past decades and have the potential to substantially advance our understanding of SM-cloud-precipitation relationships. Specifically, the 3-hr/9 km Soil Moisture Active Passive (SMAP/L4) and 1.5-hr/5 km Global Precipitation Measurement Dual-Frequency Precipitation Radar (GPM/DPR/L2A) provide high-resolution estimates of the topsoil moisture content and the vertical distribution of hydrometeors within and above the PBL, respectively, at a global coverage. Many studies have confirmed the accuracy and reliability of SMAP/L4 (X. Zhang et al., 2017; Reichle et al., 2017; L. Zhang et al., 2017; Koster et al., 2018; Tavakol et al., 2019) and GPM/DPR/L2A (Lasser et al., 2019; Pejcic et al., 2020; Liao & Meneghini, 2022) data products.

In this paper, we demonstrate how functional decomposition of a large database of SMAP/L4 surface SM and GPM/DPR/L2A cloud vertical profiles (CVP) provides valuable insights into the relationship between antecedent SM and cloud water distribution and reflectivity in the lower troposphere. Specifically, we use high-dimensional model representation (HDMR) (Li & Rabitz, 2010; Li & Rabitz, 2012; Gao et al., 2023) to disentangle the intricate and dynamic web of land-surface and atmospheric variables and interactions that give rise to the SMCPF. HDMR is a generalization of the analysis of variance (ANOVA) to dependent input factors and uses a superposition of linear multiples of first-, second-, and higher-order component functions to parse out the structural and correlative contributions of SM and other land-surface variables to the CVP. The expansion coefficients of the component functions are determined from a training data set of collocated SMAP/L4 and GPM/DPR/L2A measurements across the central US using linear least squares and D-MORPH regression (Li & Rabitz, 2010). We are mainly interested in the first-order component functions as they quantify the direct contribution of each land-surface variable to the CVP. The method is CPU-efficient and yields spatial maps of the SM contribution to cloud reflectivity and rainfall for our study region as a function of cloud height and SM time lag.

This paper is organized as follows. Section 2 discusses the SMAP/L4 SM and GPM/DPR/L2A satellite products and study region. Section 3 summarizes the data preprocessing steps and HDMR functional decomposition. Section 4 presents the results of our analysis and documents the relationship between SM and the CVP as a function of cloud height, time lag, and spatial coordinates in our study region. Section 5 summarizes our main findings and presents suggestions for future work.

2 Data and Experimental Region

We use the publicly available 3-hour/9 km SMAP/L4 and 1.5-hour/5 km GPM/DPR/L2A data products and single out samples from our study region in the warm seasons (April to October) of 2016 to 2019 with convective precipitation in the afternoon hours until midnight (14:00-24:00 CDT). The altitude spans 1 to 5 km, with the 1-3 km zone identified by Findell and Eltahir (2003a) as a critical region for convective triggering, and in the 3-5 km zone above this region resides the free atmosphere. We succinctly discuss the SMAP/L4 and GPM/DPR/L2A products and our study region. A more detailed description of the satellite data products is found in Text S1.

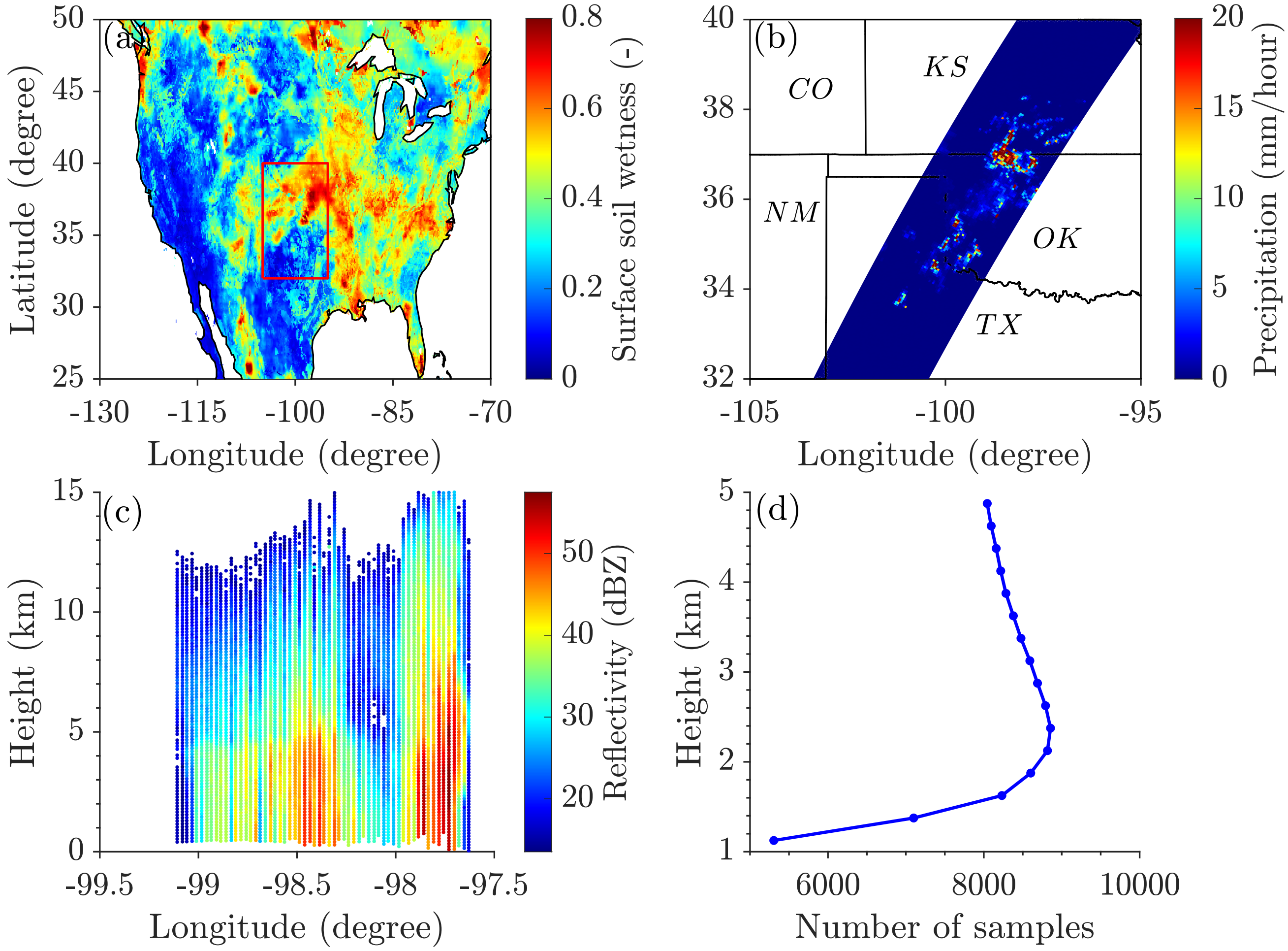

Figure 1: August 7, 2016: (a) SMAP/L4 surface SM (3-hour, 9 km, 19:30 CDT) over CONUS and GPM/DPR/L2A measured (b) surface precipitation (1.5-hour, 5 km, 21:51:10-23:23:44 CDT) and (c) cloud reflectivity profiles (97.5°W - 99.5°W, 36.7°N) for our study region (red rectangle) in the central United States. Graph (d) in the bottom right corner displays the number of samples n we have left at each DPR measurement height after data preprocessing.

The SMAP mission Level 4 SM (L4_SM) product gives 3-hourly estimates of surface and root-zone SM at 9-km spatial resolution and global coverage (Reichle et al., 2015). The 3-hour time-averaged 9-km geophysical data product (SPL4SMGP) provides estimates of the wetness (0-1) of the top soil layer (0-5 cm) (see Figure 1a) and other land-surface variables. Hourly estimates of low-level atmospheric temperature (AT) and total precipitable water (TPW) from $0.25° \times 0.25°$ ERA-5 reanalysis convey the stability and humidity of the antecedent atmosphere and are precursors to mesoscale convective events (Sherwood, 1999; Findell & Eltahir, 2003a; Holloway & Neelin, 2010). In our functional decomposition, we use the mean AT for the critical region, 1-3 km above the soil surface, which roughly corresponds to levels $P_{\mathrm{surf}} - 100$ and $P_{\mathrm{surf}} - 300$ hPa. Section 3.2 discusses in more detail our selection of auxiliary land-surface and atmospheric variables.

The GPM/DPR/L2A product (GPM_2ADPR) provides a swath of precipitation profiles (see Figure 1b) every 1.5 hours at a spatial resolution of 5 km and vertical increment of 125 m. The major data fields zFactorFinal (dBZ) and typePrecip provide vertical profiles of the Ka-band cloud reflectivity factor (see Figure 1c) and an 8-digit precipitation type ID, for individual pixels. We only use samples classified as convective precipitation and work with 250-m averaged Ka-band cloud reflectivities to suppress measurement errors.

Our study region in Figure 1a (95°W-105°W, 32°N-40°N) is a hot spot for SM-precipitation coupling (Findell & Eltahir, 2003b; Koster et al., 2004; Ford et al., 2023) with large spatial variability in climatological sign and strength of the SMCPF (Frye & Mote, 2010; Findell et al., 2011; Su & Dickinson, 2017; Yuan et al., 2020; Ford et al., 2023). This central region of the US offers an excellent demonstration of our method and possibility to benchmark the inferred patterns of the SMCPF sign and magnitude against literature findings.

3 Method

3.1 Data Preprocessing

We extract the GPM/DPR/L2A swaths that overpass our study region and use only those samples classified as convective precipitation in the 'typePrecip' data field. This type classification is an important byproduct of DPR instruments and crucial to an accurate characterization of the antecedent atmosphere using ERA-5 reanalysis AT and TPW data. To avoid water from interception evaporation, we discard all samples which received more than 0.5 mm of precipitation in the 18 hours preceding the DPR's scan according to the Multi-Radars Multi-Sensors (MRMS) Gauge-corrected Quantitative Precipitation Estimates (J. Zhang et al., 2016). This should also reduce the impacts of large-scale synoptic systems

(Findell et al., 2011). Next, we collocate SMAP/L4 and ERA-5 data and GPM/DPR/L2A measured cloud profiles using linear interpolation and time lags $\Delta t = t_{\text{dpr}} - t_{\text{smap}}$ of 7 and 10 hours. In doing so, we allow for a 2-hour grace period so as to maximize the sample size. For example, SM data with a time lag $6.01 \leq \Delta t \leq 7.99$ are pooled together in the 7-hour time lag. Figure 1d displays the number of DPR-measured cloud reflectivities n for the months of April-October (2016-2019) as a function of cloud height. Not all heights have the same sample size due to for instance the absence of clouds, radar detection threshold, and path attenuation (Iguchi et al., 2010). The pooled samples of April-October guarantee a sufficiently large sample size at each cloud height. Next, we decompose this final collection of SMAP/L4 - GPM/DPR/L2A samples using HDMR and expand the DPR-measured cloud reflectivities at each separate cloud height as a sum of first- and higher-order structural and correlative contributions of SM and the auxiliary variables.

3.2 High-Dimensional Model Representation

SMCPFs are notoriously challenging to observe and study outside of model environments (Ford et al., 2023), hence innovative analytical approaches are required to study them (Koster et al., 2004; Seneviratne et al., 2006; Findell et al., 2011; Berg et al., 2013; Guillod et al., 2014; Knist et al., 2017). HDMR is particularly appealing in the present context as it expresses all variable interactions in a system in a hierarchical order. This allows us to quantify the individual contribution of SM to the CVP.

Suppose we group all land-surface and atmospheric variables that govern the cloud reflectivity $y = f(\mathbf{x})$ at a given cloud height in a $d \times 1$ vector $\mathbf{x} = (x_1, \ldots, x_d)^\top$. HDMR builds on the finite multivariable function expansion of Sobol′ (1993) and decomposes the output, $y = f(\mathbf{x})$, of the scalar-valued square-integrable function, $f \in L^2(\mathbb{K}^d)$, on the d-dimensional unit cube, $\mathbb{K}^d = \{\mathbf{x} | 0 \leq x_i \leq 1;\ i = 1, \ldots, d\}$, into summands of component functions, $f_i(x_i)$, $f_{ij}(x_i, x_j)$, $\ldots$, $f_{12\ldots d}(x_1, x_2, \ldots, x_d)$, to yield (Li & Rabitz, 2012)

$$y = f_0 + \sum_{i=1}^{n_1} f_i(x_i) + \sum_{1 \le i < j \le d}^{n_2} f_{ij}(x_i, x_j) + \sum_{1 \le i < j < k \le d}^{n_3} f_{ijk}(x_i, x_j, x_k) + \cdots + f_{12\ldots d}(x_1,\, x_2, \ldots, x_d) + \epsilon, \tag{1}$$

where f_0 is the mean output and the residual $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is assumed to be zero-mean normally distributed with a constant variance, σ_ϵ^2. The $n_1 = d$ first-order functions, $f_i(x_i)$, characterize the individual effects of the input variables on the model output. The $n_2 = d(d-1)/2$ second-, $f_{ij}(x_i, x_j)$, $n_3 = d(d-1)(d-2)/6$ third-, $f_{ijk}(x_i, x_j, x_k)$, up to the d^{th}-order component functions, $f_{12\ldots d}(x_1,\, x_2, \ldots, x_d)$, characterize the cooperative contribution of two, three, up to all land-surface variables combined to the cloud reflectivity y. As third- and higher-order independent and cooperative effects are usually negligible in most physical systems (Rabitz & Aliş, 1999; Kucherenko et al., 2011; H. Wang et al., 2017; Falchi et al., 2018; Shereena & Rao, 2019; Gao et al., 2023), our function expansion of the CVP considers only the $n_{12} = n_1 + n_2$ first- and second-order component functions

$$y = f_0 + \sum_{u=1}^{n_{12}} f_u + \epsilon, \tag{2}$$

where subscript u is the index of the component function rather than its order as in equation (1). Thus, $f_1, \ldots, f_d$, signify the first-order component functions and $f_{d+1}, \ldots, f_{d+d(d-1)/2}$ correspond to the second-order component functions. In our implementation, f_0 signifies the mean reflectivity in units of dBZ and the component functions f_u quantify the individual and bivariate contributions of the land-surface and atmospheric variables to the cloud reflectivity.

The component functions must satisfy hierarchical orthogonality to exactly delineate the independent (structural) and cooperative (correlative) contributions of individual and groups of input variables to y (Li & Rabitz, 2012; Gao et al., 2023). This is enforced through a so-called relaxed vanishing condition (Hooker, 2007)

$$\int_0^1 w_u(\mathbf{x}_u) f_u(\mathbf{x}_u) \mathrm{d}x_i = 0 \qquad \text{for all } u \subseteq \{1, \ldots, d\} \text{ and } i \in u, \tag{3}$$

where u is a subset of superset $U = \{1, \ldots, d\}$, $\mathbf{x}_u$ denote the dimensions u of the input vector and

$w_u(\mathbf{x}_u)$ signifies the probability density function (pdf) of $\mathbf{x}_u$. For a second-order component function, the vanishing condition of equation (3) dictates that $f_{ij}(x_i, x_j)$ should be orthogonal to its lower order component functions, $f_i(x_i)$ and $f_j(x_j)$. The component functions are constructed using the extended bases orthonormalized polynomials and associated linear expansion coefficients. D-MORPH regression (Li & Rabitz, 2010) enforces hierarchical orthogonality of the component functions in pursuit of the optimum expansion coefficients. This method is described in Text S2.

The statistical significance of a given component function is readily determined by comparing the performance of the function expansion with and without this component function. Suppose SSR_1 is the sum of squared residuals of the function $y = y_0 + \sum_{i=1}^{d-1} f_i(x_i)$ with $l_1 = (d-1)p$ expansion coefficients and SSR is the same quantity for the same function $y = y_0 + \sum_{i=1}^{d} f_i(x_i)$ expanded with $f_d(x_d)$ and $l = l_1 + p$ coefficients. To reject the null hypothesis, "$\mathcal{H}_0 : f_d(x_d)$ is insignificant", the F-statistic

$$F = \frac{(\mathrm{SSR}_1 - \mathrm{SSR})/(l - l_1)}{\mathrm{SSR}_1/(n - l_1)}, \tag{4}$$

must exceed $F_{\mathrm{crit}} = F_{\mathcal{F}}^{-1}(1 - \alpha | l_1 - l, n - l_1)$ where $F_{\mathcal{F}}^{-1}(p_\alpha | \nu_1, \nu_2)$ is the quantile function of the Fisher-Snedecor distribution with ν_1 and ν_2 degrees of freedom at the critical value $p_\alpha = 1 - \alpha$ and significance level $\alpha \in (0, 1)$. The magnitude of the F-statistic conveys the importance of $f_d(x_d)$ in explaining the CVP and, thus, can be interpreted as a measure of the feedback strength.

Now that we have finished discussing the building blocks of our HDMR data decomposition method, we are left with the selection of land-surface and atmospheric variables $(x_2, \ldots, x_d)$ which complement SM, x_1, in explaining the measured cloud reflectivities, y. We tested many different variables in our analysis and settled on land-surface temperature (LST), leaf area index (LAI), atmospheric temperature (AT), and total precipitable water (TPW) as auxiliary variables. This equates to a 5×1 input vector $\mathbf{x} = (x_1, \ldots, x_5)^\top = (\mathrm{SM}, \mathrm{LST}, \mathrm{LAI}, \mathrm{AT}, \mathrm{TPW})^\top$. LAI and LST modulate evapotranspiration under SM-limited or energy-limited regimes (Seneviratne et al., 2010) and AT and TPW convey information for the SMCPF at synoptic scales about atmospheric preconditioning (Ford, Quiring, et al., 2015; Tuttle

& Salvucci, 2017). This explicit treatment of atmospheric conditions can only raise our confidence in any causal links that are found between SM and CVP. Figure S1 presents a correlogram of the five input variables. Note that we do not consider variables such as the latent heat flux. This derivative product depends on SM, hence would only trouble our inference of the relationships and variables that govern the CVP.

4 Results

4.1 Cloud Height and Temporal Lag of SMCPF

Figure 2 displays the F-statistics of the (a) SM, (b) LST, and (c) LAI component functions as a function of cloud height (1 to 5 km) and time lag ($\Delta t = 7$ and 10 h). The solid line denotes the mean of 1,000 bootstrap trials each with a different selection of $r = 0.75n$ training samples and the light-colored regions portray the associated 95% confidence intervals. The dashed black line in each graph corresponds to the critical F-value at each cloud height using $\alpha = 0.05$. The value of the F-statistic is not constant but altitude dependent. The influence SM, LST and LAI exert on the CVP is dependent on cloud height. In case of SM in panel (a) this equates to a height-dependent SMCPF with a bottom-heavy relationship between SM and CVP. The SMCPF is most pronounced in the lower atmosphere at about 1-3 km above the surface. Above this level, the impact of SM on the CVP decreases rapidly with altitude. As we will shown in Section 4.2, the first-order SM component function $f_1(x_1)$ displays a positive feedback due to a wet soil. A higher SM implies a larger evaporative fraction, promoting moderate PBL growth (see Figure S2) and moisture accumulation (Yin et al., 2015). The CVP at higher altitudes is less dependent on surface SM and controlled more by the upper atmosphere at levels of about 3 km and beyond (Findell & Eltahir, 2003a). Furthermore, a capping inversion layer can inhibit the upward movement of warm, moist air from the surface to the free atmosphere (Findell & Eltahir, 2003b). Indeed, the HDMR-inferred relationship between SM and CVP as articulated by the F-statistic is corroborated by simulation analyses (Findell & Eltahir, 2003a; Koukoula et al., 2019). This physical underpinning inspires confidence in the ability of our methodology to back out SM-cloud feedbacks at different heights. The strong agreement in the results of the two time-lags is a result of SM autocorrelation. The $\Delta t = 7$ hour time lag displays the largest influence on the CVP at all altitudes but the largest two cloud heights near 5 km.

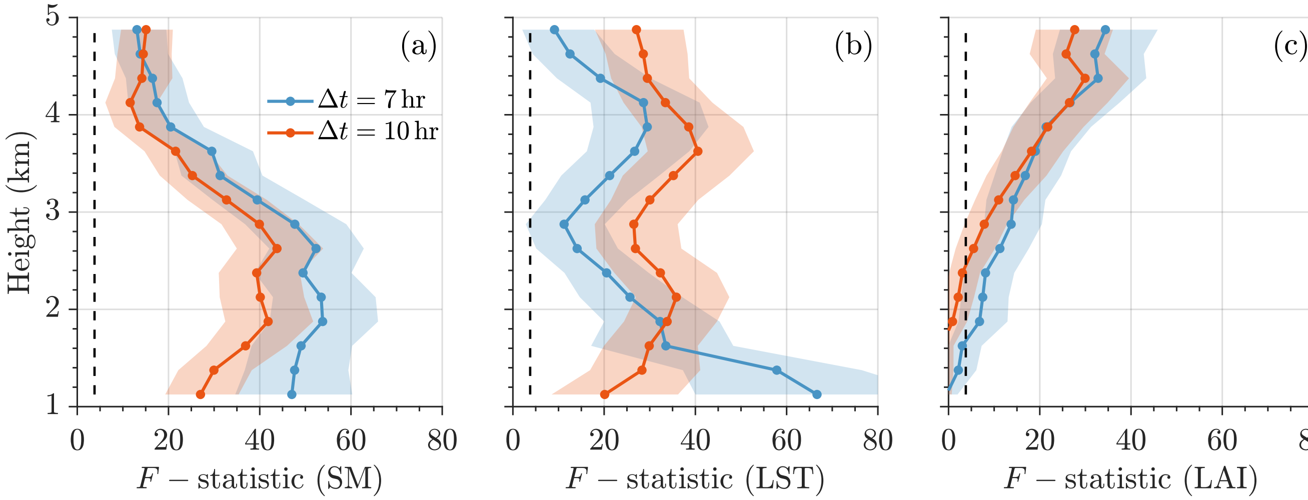

Figure 2: Vertical profiles of the mean F-statistic of the first-order component functions of (a) SM: $f_1(x_1)$, (b) LST: $f_2(x_2)$, and (c) LAI: $f_3(x_3)$ computed from 1,000 bootstrap iterations. Solid blue and red lines differentiate between temporal lags ($\Delta t = 7$ and 10 hours) and black dashed lines represent the critical value at significance level $\alpha = 0.05$, F_{crit}. The light blue and red regions correspond to the 95% bootstrap confidence intervals.

Compared to SM, LST exerts control on CVP across a wider vertical range (in Figure 2b), whose F-statistic shows a bimodal relationship with height, peaking close to the surface with $\Delta t = 7$ hours and at a higher altitude of 3.5-4.0 km with $\Delta t = 10$ hours. As discussed in the next section, $f_2(x_2)$ exhibits a positive correlation with LST, suggesting that positive LST anomalies (or dry soil) play a crucial role in shaping CVP. Therefore, the fact that low-level (1.0-2.5 km) CVP is responsive to LST comes in qualitative agreement with the pathway of negative SMCPF, driven by the effect of positive LST anomalies in catalyzing higher sensible heat flux, convective triggering potential (CTP), and rapid PBL growth. We further support this finding by comparison with the ERA5 reanalysis PBL height in Figure S2. Such observed response of PBL height to wet and dry surface exhibits strong consistency with prior simulation-based and observational studies (Findell & Eltahir, 2003a; Xu et al., 2021; Ford et al., 2023), which indicates two mechanisms for initiating convection: significant moistening of the PBL (over wet soil) and rapid growth of the PBL (over dry soil). In addition, the predictability of LST decreases first at 3.0 km and increases again at 3.5-4.0 km. The reason why LST is significant at a higher altitude may be twofold. On the one hand, the LST anomalies favor strong CTP where air parcels can overcome convective inhibition and reach the level of free convection (Taylor et al., 2012). If we intuitively consider $f_2(x_2)$ the contribution of near-surface air to the cloud reflectivity conditioned on a specific height and time lag, its F-statistic (in Figure 2b) somehow approximates the dynamics of the thermal updraft such that the largest F-statistic value shifts from $\Delta t = 7$ hours to $\Delta t = 10$ hours with height changing from 1.0 km to 5.0 km. On the other hand, local LST may also reflect certain atmospheric conditions such as the melting layer, which typically resides between 3.0-5.0 km above the surface during pre-monsoon and monsoon seasons in the central United States (Song et al., 2021).

The F-statistic of the LAI component, $f_3(x_3)$, informs its poor predictive power in the lower atmosphere, primarily due to the governing effects of SM, LST, and AT (see Figure S3) on initiating convection and the subsequent formation of cloud/precipitation. In contrast, the modest, albeit statistically significant influence of LAI in higher-level CVP can be attributed to its seasonal variations (Savoy & Mackay, 2015) and correlation with the atmospheric conditions (see Figure S1). In Text S3 and Figures S3-S4, we elaborate on our findings in terms of atmospheric controls on CVP which demonstrate a comparable physical underpinning with the land-surface variables.

4.2 The SMCPF across Space

In this section, we focus our attention on the spatial pattern of the SMCPF within the study region. We reiterate that we conduct functional decomposition of the cloud reflectivity using all the samples of April-October (2016-2019) for a specific time lag and cloud height, to guarantee an adequate number of samples and storm events. Our goal here is to present a 4-year averaged spatial distribution of the derived component functions and determine locations of positive and negative SMCPF rather than focusing on interannual and/or cross-season variations.

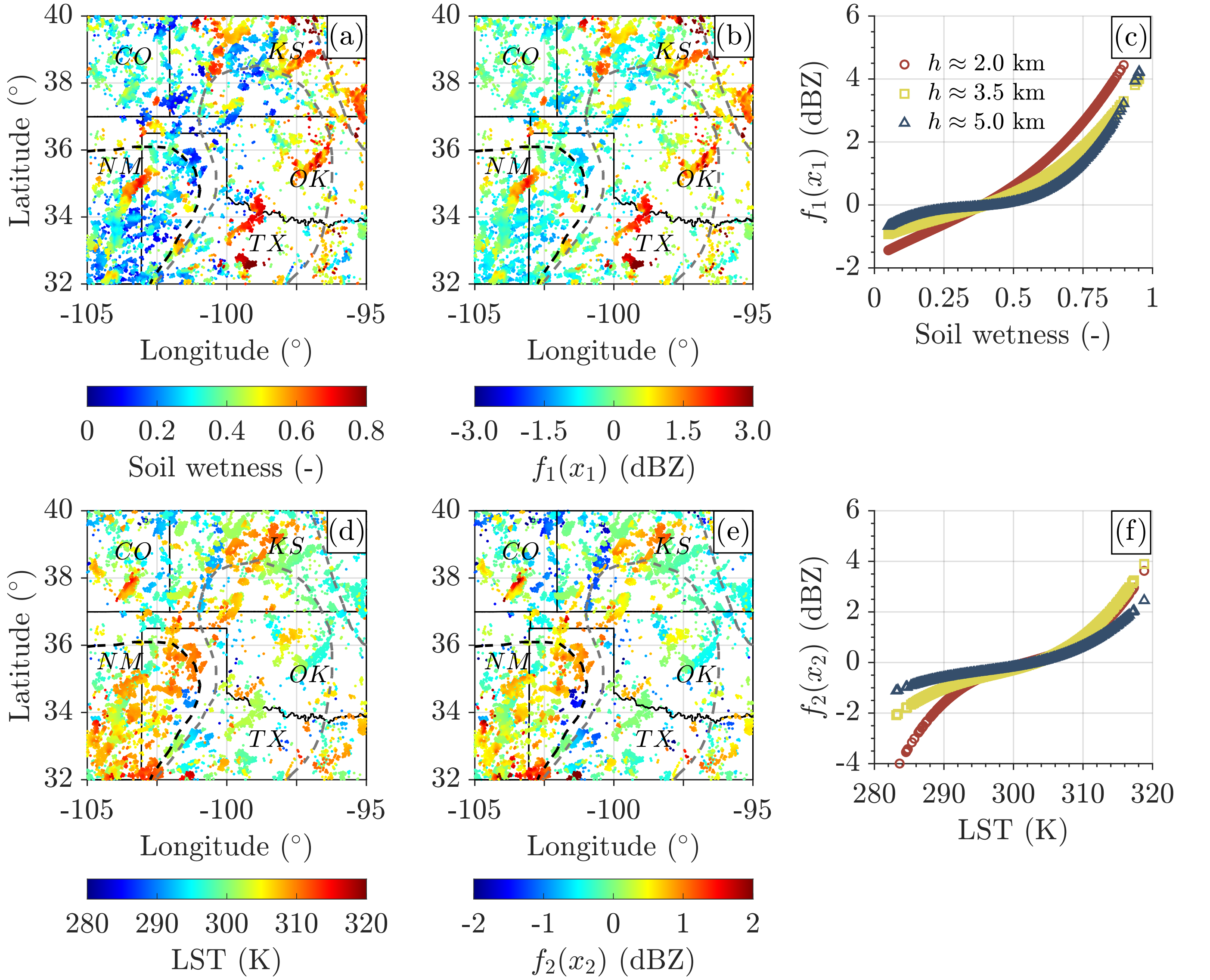

Figure 3: The central United States (95°W-105°W, 32°N-40°N) with (a) antecedent 7-hr SMAP/L4 soil wetness (-) collocated at coordinates of the GPM/DPR/L2A samples and (b) first-order component function of soil wetness, $f_1(x_1)$ (dBZ), evaluated at approximately 2.0 km height. Solid black lines delineate the state borders while dashed black and grey lines depict the negative feedback and transitional regions proposed by Findell and Eltahir (2003b). Panel (c) displays the scatter plots of the samples of antecedent 7-hour SM against the corresponding $f_1(x_1)$ (dBZ), evaluated at three separate heights, 2.0 km (red circles), 3.5 km (yellow squares), and 5.0 km (blue triangles). The bottom row of panels presents the same content as panels (a-c) but for (d) SMAP/L4 LST and (e,f) its associated component function, $f_2(x_2)$.

Figure 3a-b presents the spatial distribution of the antecedent 7-hour SMAP/L4 soil wetness at the top layer (0-5 cm), collocated at the coordinates of the GPM/DPR/L2A samples, alongside the corresponding first-order component function, $f_1(x_1)$ (dBZ), evaluated at 2.0 km. This examination of SM's feedback strength, conditioned on an altitude of 2.0 km and a 7-hour time lag, is of particular interest upon our prior analysis of the F-statistic in Figure 2a. Panels (b-c) reveal the positive feedback from SM represented by $f_1(x_1)$. With a degree of saturation exceeding 0.4, wet soil could increase cloud reflectivity by up to 4 dBZ. The fact that the absolute value of $f_1(x_1)$ decreases with height in Panel (c) again lends support to our inferred height-dependent SMCPF in Section 4.1, underscoring the stronger coupling between SM and CVP in the low-level atmosphere. As a byproduct, we demonstrate in Text S4 and Figure S5 the application of the Marshall-Palmer formula (Marshall & Palmer, 1948) to the transformation of $f_1(x_1)$ (dBZ) into estimates of rainfall rate.

Significant positive feedback of SM is evident in regions such as northern Texas, central Oklahoma, northwestern and southeastern Kansas, and northeastern New Mexico. All these areas, with the exception of northeastern New Mexico, are located inside or close to the 'transitional regions' delineated by dashed grey lines as categorized by Findell and Eltahir (2003b). The middle transitional region, spanning from the semi-arid southwestern to the humid southeastern parts of the central United States, is influenced by both dry and wet soil advantage regimes. Hence, this dual influence explicates the observable positive feedback in the central and eastern sections of the transitional region and negative feedback in the southwestern part (detailed below). These local wet soil anomalies can be attributed to early warm-season mesoscale convective systems (MCSs) and non-MCS rainfall. Typically, the early warm-season MCSs were reported a dominant source of the summer SMCPF (Hu et al., 2021), which are initiated upwind near the Rocky Mountains Foothills and propagate eastward to the central United States (Feng et al., 2019).

Since SM can indirectly exert feedback on cloud and precipitation through heating or cooling the surface (Duerinck et al., 2016), we further delve into examining spatially the samples of antecedent 7-hour LST (K) and their contribution to cloud, $f_2(x_2)$ (dBZ), and rainfall, ΔR (mm/hour), in Figures 3d-f and S6, respectively. $f_2(x_2)$ exhibits a non-linear dependence on LST where LST anomalies exert the most significant influence. From Figure 3d-e, it is suggested that LST above 305 K accounts for an increase of at most 4.0 dBZ in the cloud reflectivity and 2.0 mm/hour in rainfall rate (see Figure S6) at both 2.0 and 3.5 km. On the contrary, the samples with a cooler surface (LST<290 K) seem to foster a more stable atmospheric state, thereby reducing the cloud reflectivity, especially in the near-surface atmosphere ($h \approx 2.0$ km). The underlying LST-driven mechanisms were discussed in the previous section.

Geographically, the most significant effects of these anomalies are evident and clustered in the southwest of the study region, delineated by 101°W-105°W and 32°N-36°N. Within this area, we find a moderate negative correlation ($R = -0.41$, shown in Figure S7a) between surface SM and the LST component function, $f_2(x_2)$. Moreover, we illustrate in Figure S7b that LST contributes to CVP preferentially over dry soil with saturation between 0.1 and 0.4. These findings underscore the presence of the intrinsic SM-LST coupling nested within the SMCPF pathways (Seneviratne et al., 2010), and we can conveniently interpret $f_2(x_2)$ as a proxy for the indirect and negative feedback of SM on CVP. Notably, our identified negative feedback region (101°W-105°W, 32°N-36°N) is consistent with the one proposed by Findell and Eltahir (2003b) (represented by the black dashed line in Figure 3d-e). Several factors can play a role when it comes to the sources of convective clouds and precipitation over the dry soil. For instance, the monsoonal moisture incursion into New Mexico can bring up local humidity and offset the reduced evapotranspiration from the local dry soils (Wallace et al., 1999; Klein & Taylor, 2020). Besides, the Great Plains Low-Level Jet (GPLLJ) can transport abundant moisture southerly from the Gulf of Mexico into the central United States (Ford, Rapp, & Quiring, 2015; Feng et al., 2016).

5 Discussion and Conclusion

This study presents a data-driven approach that uses the functional decomposition of a large database of satellite-measured SM (SMAP/L4) and CVP (GPM/DPR/L2A) for disentangling and quantifying SMCPF in the central United States. Results show that the signs and strengths of the feedback differ among cloud heights and geographical locations. A significant positive feedback is observed in the lower atmosphere, particularly between 1.0 and 3.0 km with a temporal lag of 7 hours. With a degree of saturation over 0.4, wet soil can potentially increase the cloud reflectivity and rainfall rate by up to 4.0 dBZ and 2.0 mm/hour at $h \approx 2.0$ km, evidently in northern Texas, central Oklahoma, northwestern and southeastern Kansas. The negative feedback, indirectly interpreted by the anomalies of LST, is effective with a wider vertical extension from 1.0 km to 4.0 km and a time lag of 7-10 hours. These LST anomalies can explain comparable increments in cloud reflectivity and rainfall rate to SM but in northwestern Texas and southeastern and eastern New Mexico. The identified patterns of SMCPF align qualitatively with previous studies that utilize simulations and observations to investigate the underlying mechanisms and regional categorizations of the feedback (Findell & Eltahir, 2003a, 2003b; Qian et al., 2013; Sathyanadh et al., 2017; Su & Dickinson, 2017; Koukoula et al., 2019; Hu et al., 2021; Ford et al., 2023).

Our approach brings new insights into the observational understanding of the SMCPF characterized by cloud height, time lag, and location and possesses the potential for coupled land-atmosphere model diagnosis. Despite this, certain limitations are highlighted. Even though a decent amount of samples was obtained, they can hardly support extensive analyses over seasonal, interannual, or localized scales due to the substantial downsampling. Another possible limitation is the selection of only five land and atmospheric variables as inputs of the HDMR emulator. We reiterate that this decision is strategically aimed at maximizing the capture of the nonlinear relationship and causal link between cloud and SM. Nonetheless, it concurrently overlooks other pertinent variables that could play a significant role in the SMCPF pathways.

For future work, it is important to conduct a comprehensive analysis employing cloud model simulations and/or reanalysis data sets as inputs of HDMR. This will help diagnose the representativeness of the current-generation coupled land-atmosphere models. We should also build robust HDMR emulators to be integrated with state-of-the-art cloud models for more accurate prediction of convective clouds and precipitation. This necessitates the incorporation of more predictors such as SM gradient (Taylor, 2015; Zhou et al., 2021; Graf et al., 2021; Chug et al., 2023) and evaporative fraction (Taylor et al., 2013; Ford et al., 2023), along with atmospheric variables like wind speed and water vapor mixing ratio (Raymond & Sessions, 2007; Seneviratne et al., 2010). Last but not least, with the advancement of a variety of reanalysis datasets, the methodology can be useful for examining the changes in SMCPF under increasing hydroclimatic extremes at the regional and global scales.

Acknowledgments

The authors would like to acknowledge the support of NASA through the Global Precipitation Measurement Mission program (Grant 80NSSC22K0597) and the Weather and Atmospheric Dynamics program (Grant 80NSSC23K1304), as well as the support of the NSF Division of Information and Intelligent Systems (ExpandAI2ES Grant IIS2324008). The author also wishes to thank the support of the UCI Engineering-Los Alamos National Laboratory (LANL) Fellowship, provided by the UCI Samueli School of Engineering and LANL.

Data and Software Availability

The SMAP/L4 (L4_SM) product is obtained from the National Snow and Ice Data Center at https://nsidc.org/data/spl4smau/versions/7 (Reichle et al., 2022). The GPM/DPR/L2A product (GPM_2ADPR) is obtained from the Goddard Earth Sciences Data and Information Services Center at https://disc.gsfc.nasa.gov/datasets/GPM_2AKaENV_07/summary (Iguchi et al., 2010). MATLAB postprocessing software will be archived in Zenodo along with the final data set of collocated SMAP and DPR samples. A copy of this data set is provided for review in the supporting information.

References

Baker, J. C. A., Castilho de Souza, D., Kubota, P. Y., Buermann, W., Coelho, C. A. S., Andrews, M. B., Gloor, M., Garcia-Carreras, L., Figueroa, S. N., & Spracklen, D. V. (2021). An assessment of land–atmosphere interactions over South America using satellites, reanalysis, and two global climate models. *Journal of Hydrometeorology*, *22*(4), 905–922. https://doi.org/10.1175/JHM-D-20-0132.1

Baker, J. C. A., Garcia-Carreras, L., Buermann, W., De Souza, D. C., Marsham, J. H., Kubota, P. Y., Gloor, M., Coelho, C. A. S., & Spracklen, D. V. (2021). Robust Amazon precipitation projections in climate models that capture realistic land–atmosphere interactions. *Environmental Research Letters*, *16*(7), 074002. https://doi.org/10.1088/1748-9326/abfb2e

Berg, A., Findell, K., Lintner, B. R., Gentine, P., & Kerr, C. (2013). Precipitation sensitivity to surface heat fluxes over north america in reanalysis and model data. *Journal of Hydrometeorology*, *14*(3), 722–743. https://doi.org/10.1175/JHM-D-12-0111.1

Chug, D., Dominguez, F., Taylor, C. M., Klein, C., & Nesbitt, S. W. (2023). Dry-to-Wet Soil Gradients Enhance Convection and Rainfall over Subtropical South America. *arXiv preprint arXiv:2304.04630*. https://doi.org/10.1175/JHM-D-23-0031.1

Cioni, G., & Hohenegger, C. (2017). Effect of soil moisture on diurnal convection and precipitation in large-eddy simulations. *Journal of Hydrometeorology*, *18*(7), 1885–1903. https://doi.org/10.1175/JHM-D-16-0241.1

Deardorff, J. W. (1980). Stratocumulus-capped mixed layers derived from a three-dimensional model. *Boundary-layer meteorology*, *18*, 495–527. https://doi.org/10.1007/BF00119502

Dirmeyer, P. A., Jin, Y., Singh, B., & Yan, X. (2013). Trends in land–atmosphere interactions from CMIP5 simulations. *Journal of Hydrometeorology*, *14*(3), 829–849. https://doi.org/10.1175/JHM-D-12-0107.1

Duerinck, H. M., Van der Ent, R. J., Van de Giesen, N. C., Schoups, G., Babovic, V., & Yeh, P. J. F. (2016). Observed soil moisture–precipitation feedback in Illinois: A systematic analysis over different scales. *Journal of Hydrometeorology*, *17*(6), 1645–1660. https://doi.org/10.1175/JHM-D-15-0032.1

Ek, M. B., & Holtslag, A. A. M. (2004). Influence of soil moisture on boundary layer cloud development. *Journal of hydrometeorology*, *5*(1), 86–99. https://doi.org/10.1175/1525-7541(2004)005⟨0086:IOSMOB⟩2.0.CO;2

Falchi, A., Minisci, E., Kubicek, M., Vasile, M., & Lemmens, S. (2018). HDMR-based sensitivity analysis and uncertainty quantification of GOCE aerodynamics using DSMC. *Stardust Final Conference*, 301–323. https://doi.org/10.1007/978-3-319-69956-1_18

Fast, J. D., Berg, L. K., Feng, Z., Mei, F., Newsom, R., Sakaguchi, K., & Xiao, H. (2019). The impact of variable land-atmosphere coupling on convective cloud populations observed during the 2016 HI-SCALE field campaign. *Journal of Advances in Modeling Earth Systems*, *11*(8), 2629–2654. https://doi.org/10.1029/2019MS001727

Feng, Z., Houze Jr, R. A., Leung, L. R., Song, F., Hardin, J. C., Wang, J., Gustafson Jr, W. I., & Homeyer, C. R. (2019). Spatiotemporal characteristics and large-scale environments of mesoscale convective systems east of the Rocky Mountains. *Journal of Climate*, *32*(21), 7303–7328. https://doi.org/10.1175/JCLI-D-19-0137.1

Feng, Z., Leung, L. R., Hagos, S., Houze, R. A., Burleyson, C. D., & Balaguru, K. (2016). More frequent intense and long-lived storms dominate the springtime trend in central US rainfall. *Nature communications*, *7*(1), 13429. https://doi.org/10.1038/ncomms13429

Ferguson, C. R., & Wood, E. F. (2011). Observed land-atmosphere coupling from satellite remote sensing and reanalysis. *Journal of Hydrometeorology*, *12*(6), 1221–1254. https://doi.org/10.1175/2011JHM1380.1

Findell, K. L., & Eltahir, E. A. B. (2003b). Atmospheric controls on soil moisture–boundary layer interactions. Part II: Feedbacks within the continental United States. *Journal of Hydrometeorology*, *4*(3), 570–583. https://doi.org/10.1175/1525-7541(2003)004⟨0570:ACOSML⟩2.0.CO;2

Findell, K. L., & Eltahir, E. A. B. (2003a). Atmospheric controls on soil moisture–boundary layer interactions. Part I: Framework development. *Journal of Hydrometeorology*, *4*, 552–569. https://doi.org/10.1175/1525-7541(2003)004⟨0552:ACOSML⟩2.0.CO;2

Findell, K. L., Gentine, P., Lintner, B. R., & Kerr, C. (2011). Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation. *Nature Geoscience*, *4*(7), 434–439. https://doi.org/10.1038/ngeo1174

Ford, T. W., Quiring, S. M., Frauenfeld, O. W., & Rapp, A. D. (2015). Synoptic conditions related to soil moisture-atmosphere interactions and unorganized convection in Oklahoma. *Journal of Geophysical Research: Atmospheres*, *120*(22), 11–519.

Ford, T. W., Rapp, A. D., & Quiring, S. M. (2015). Does afternoon precipitation occur preferentially over dry or wet soils in Oklahoma? *Journal of Hydrometeorology*, *16*(2), 874–888. https://doi.org/10.1175/JHM-D-14-0005.1

Ford, T. W., Rapp, A. D., Quiring, S. M., & Blake, J. (2015). Soil moisture–precipitation coupling: Observations from the Oklahoma Mesonet and underlying physical mechanisms. *Hydrology and Earth System. Sciences*, *19*, 3617–3631. https://doi.org/10.5194/hess-19-3617-2015

Ford, T. W., Steiner, J., Mason, B., & Quiring, S. M. (2023). Observation-Driven Characterization of Soil Moisture-Precipitation Interactions in the Central United States. *Journal of Geophysical Research: Atmospheres*, *128*(12), e2022JD037934. https://doi.org/10.1029/2022JD037934

Frye, J. D., & Mote, T. L. (2010). Convection initiation along soil moisture boundaries in the southern Great Plains. *Monthly weather review*, *138*(4), 1140–1151. https://doi.org/10.1175/2009MWR2865.1

Gao, Y., Sahin, A., & Vrugt, J. A. (2023). Probabilistic sensitivity analysis with dependent variables: Covariance-based decomposition of hydrologic models. *Water Resources Research*, *59*(4), e2022WR032834. https://doi.org/10.1029/2022WR032834

Gentine, P., Holtslag, A. A., D'Andrea, F., & Ek, M. (2013). Surface and atmospheric controls on the onset of moist convection over land. *Journal of Hydrometeorology*, *14*, 1443–1462. https://doi.org/10.1175/JHM-D-12-0137.1

Golaz, J. C., Jiang, H., & Cotton, W. R. (2001). A large-eddy simulation study of cumulus clouds over land and sensitivity to soil moisture. *Atmospheric research*, *59*, 373–392. https://doi.org/10.1016/S0169-8095(01)00113-2

Graf, M., Arnault, J., Fersch, B., & Kunstmann, H. (2021). Is the soil moisture precipitation feedback enhanced by heterogeneity and dry soils? A comparative study. *Hydrological Processes*, *35*(9), e14332. https://doi.org/10.1002/hyp.14332

Guillod, B. P., Orlowsky, B., Miralles, D., Teuling, A. J., Blanken, P. D., Buchmann, N., Ciais, P., Ek, M., Findell, K. L., Gentine, P., Lintner, B. R., Scott, R. L., Van den Hurk, B., & I. Seneviratne, S. (2014). Land-surface controls on afternoon precipitation diagnosed from observational data: Uncertainties and confounding factors. *Atmospheric Chemistry and Physics*, *14*(16), 8343–8367. https://doi.org/10.5194/acp-14-8343-2014

Guillod, B. P., Orlowsky, B., Miralles, D. G., Teuling, A. J., & Seneviratne, S. I. (2015). Reconciling spatial and temporal soil moisture effects on afternoon rainfall. *Nature communications*, *6*(1), 6443. https://doi.org/10.1038/ncomms7443

Guo, Z., Dirmeyer, P. A., Koster, R. D., Sud, Y. C., Bonan, G., Oleson, K. W., Chan, E., Verseghy, D., Cox, P., Gordon, C. T., et al. (2006). GLACE: the global land–atmosphere coupling experiment. Part II: analysis. *Journal of Hydrometeorology*, *7*(4), 611–625. https://doi.org/10.1175/JHM511.1

Han, C., Brdar, S., & Kollet, S. (2019). Response of convective boundary layer and shallow cumulus to soil moisture heterogeneity: A large-eddy simulation study. *Journal of advances in modeling earth systems*, *11*(12), 4305–4322. https://doi.org/10.1029/2019MS001772

Hohenegger, C., Brockhaus, P., Bretherton, C. S., & Schär, C. (2009). The soil moisture-precipitation feedback in simulations with explicit and parameterized convection. *Journal of Climate*, *22*(19), 5003–5020. https://doi.org/10.1175/2009JCLI2604.1

Holloway, C. E., & Neelin, J. D. (2010). Temporal relations of column water vapor and tropical precipitation. *Journal of the Atmospheric Sciences*, *67*(4), 1091–1105. https://doi.org/10.1175/2009JAS3284.1

Hooker, G. (2007). Generalized functional ANOVA diagnostics for high-dimensional functions of dependent variables. *Journal of Computational and Graphical Statistics*, *16*(3), 709–732. https://doi.org/10.1198/106186007X237892

Hu, H., Leung, L. R., & Feng, Z. (2021). Early warm-season mesoscale convective systems dominate soil moisture–precipitation feedback for summer rainfall in central United States. *Proceedings of the National Academy of Sciences*, *118*(43), e2105260118. https://doi.org/10.1073/pnas.2105260118

Iguchi, T., Seto, S., Meneghini, R., Yoshida, N., Awaka, J., Le, M., Chandrasekar, V., & Kubota, T. (2010). GPM/DPR level-2 algorithm theoretical basis document. *NASA Goddard Space Flight Center.*

Klein, C., & Taylor, C. M. (2020). Dry soils can intensify mesoscale convective systems. *Proceedings of the National Academy of Sciences*, *117*(35), 21132–21137. https://doi.org/10.1073/pnas.2007998117

Knist, S., Goergen, K., Buonomo, E., Christensen, O. B., Colette, A., Cardoso, R. M., Fealy, R., Fernández, J., García-Díez, M., Jacob, D., Kartsios, S., Katragkou, E., Keuler, K., Mayer, S., van Meijgaard, E., Nikulin, G., Soares, P. M. M., Sobolowski, S., Szepszo, G., ... Simmer, C. (2017). Land-atmosphere coupling in euro-cordex evaluation experiments. *Journal of Geophysical Research: Atmospheres*, *122*(1), 79–103. https://doi.org/10.1002/2016JD025476

Koster, R. D., Liu, Q., Mahanama, S. P. P., & Reichle, R. H. (2018). Improved hydrological simulation using SMAP data: Relative impacts of model calibration and data assimilation. *Journal of Hydrometeorology*, *19*(4), 727–741. https://doi.org/10.1175/JHM-D-17-0228.1

Koster, R. D., Sud, Y. C., Guo, Z., Dirmeyer, P. A., Bonan, G., Oleson, K. W., Chan, E., Verseghy, D., Cox, P., Davies, H., et al. (2006). GLACE: the global land–atmosphere coupling experiment. Part I: overview. *Journal of Hydrometeorology*, *7*(4), 590–610. https://doi.org/10.1175/JHM510.1

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C. T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C.-H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., … Yamada, T. (2004). Regions of strong coupling between soil moisture and precipitation. *Science*, *305*(5687), 1138–1140. https://doi.org/10.1126/science.1100217

Koukoula, M., Nikolopoulos, E. I., Kushta, J., Bartsotas, N. S., Kallos, G., & Anagnostou, E. N. (2019). A numerical sensitivity analysis of soil moisture feedback on convective precipitation. *Journal of Hydrometeorology*, *20*(1), 23–44. https://doi.org/10.1175/JHM-D-18-0134.1

Krakauer, N. Y., Cook, B. I., & Puma, M. J. (2010). Contribution of soil moisture feedback to hydroclimatic variability. *Hydrology and Earth System Sciences*, *14*(3), 505–520. https://doi.org/10.5194/hess-14-505-2010

Kucherenko, S., Feil, B., Shah, N., & Mauntz, W. (2011). The identification of model effective dimensions using global sensitivity analysis. *Reliability Engineering & System Safety*, *96*(4), 440–449. https://doi.org/10.1016/j.ress.2010.11.003

Lasser, M., O, S., & Foelsche, U. (2019). Evaluation of GPM-DPR precipitation estimates with WegenerNet gauge data. *Atmospheric Measurement Techniques*, *12*(9), 5055–5070. https://doi.org/10.5194/amt-12-5055-2019

Li, G., & Rabitz, H. (2010). D-MORPH regression: application to modeling with unknown parameters more than observation data. *Journal of Mathematical Chemistry*, *48*, 1010–1035. https://doi.org/10.1007/s10910-010-9722-2

Li, G., & Rabitz, H. (2012). General formulation of HDMR component functions with independent and correlated variables. *Journal of Mathematical Chemistry*, *50*(1), 99–130. https://doi.org/10.1007/s10910-011-9898-0

Liao, L., & Meneghini, R. (2022). GPM DPR retrievals: Algorithm, evaluation, and validation. *Remote Sensing*, *14*(4), 843. https://doi.org/10.3390/rs14040843

Liu, W., Zhang, Q., Li, C., Xu, L., & Xiao, W. (2022). The influence of soil moisture on convective activity: A review. *Theoretical and Applied Climatology*, *149*(1-2), 221–232. https://doi.org/10.1007/s00704-022-04046-z

Marshall, J., & Palmer, W. M. (1948). The distribution of raindrops with size. *Journal of meteorology*, *5*, 166.

Miralles, A. J., Diego G. Teuling, van Heerwaarden, C. C., & Vilà-Guerau de Arellano, J. (2014). Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature Geoscience*, *7*, 345–349. https://doi.org/10.1038/ngeo2141

Pejcic, V., Saavedra Garfias, P., Mühlbauer, K., Trömel, S., & Simmer, C. (2020). Comparison between precipitation estimates of ground-based weather radar composites and GPM's DPR rainfall product over Germany. *Meteorologische Zeitschrift*, *29*(6), 451–466. https://doi.org/10.1127/metz/2020/1039

Qian, Y., Huang, M., Yang, B., & Berg, L. K. (2013). A modeling study of irrigation effects on surface fluxes and land–air–cloud interactions in the Southern Great Plains. *Journal of Hydrometeorology*, *14*(3), 700–721. https://doi.org/10.1175/JHM-D-12-0134.1

Rabitz, H., & Aliş, Ö. F. (1999). General foundations of high-dimensional model representations. *Journal of Mathematical Chemistry*, *25*(2), 197–233. https://doi.org/10.1023/A:1019188517934

Raymond, D. J., & Sessions, S. L. (2007). Evolution of convection during tropical cyclogenesis. *Geophysical research letters*, *34*(6). https://doi.org/10.1029/2006GL028607

Reichle, R. H., De Lannoy, G., Koster, R. D., Crow, W. T., Kimball, J. S., & Liu, Q. (2022). SMAP L4 global 3-hourly 9 km EASE-Grid surface and root zone soil moisture analysis update, version 7 [Dataset]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/LWJ6TF5SZRG3

Reichle, R. H., De Lannoy, G. J. M., Liu, Q., Koster, R. D., Kimball, J. S., Crow, W. T., Ardizzone, J. V., Chakraborty, P., Collins, D. W., Conaty, A. L., et al. (2017). Global assessment of the SMAP Level-4 surface and root-zone soil moisture product using assimilation diagnostics. *Journal of hydrometeorology*, *18*(12), 3217–3237. https://doi.org/10.1175/JHM-D-17-0130.1

Reichle, R. H., Lucchesi, R. A., Ardizzone, J. V., Kim, G.-K., Smith, E. B., & Weiss, B. H. (2015). *Soil moisture active passive (SMAP) mission level 4 surface and root zone soil moisture (L4_SM) product specification document* (tech. rep.). NASA Goddard Space Flight Center. Greenbelt, MD, United States. https://nsidc.org/sites/default/files/reichle789.pdf

Sakaguchi, K., Berg, L. K., Chen, J., Fast, J., Newsom, R., Tai, S. .-., Yang, Z., Gustafson Jr, W. I., Gaudet, B. J., Huang, M., Pekour, M., Pressel, K., & Xiao, H. (2022). Determining spatial scales of soil moisture—Cloud coupling pathways using semi-idealized simulations. *Journal of Geophysical Research: Atmospheres*, *127*(2), e2021JD035282. https://doi.org/10.1029/2021JD035282

Santanello, J. A., Peters-Lidard, C. D., Kennedy, A., & Kumar, S. V. (2013). Diagnosing the nature of land-atmosphere coupling: A case study of dry/wet extremes in the US southern Great Plains. *Journal of Hydrometeorology*, *14*(1), 3–24. https://doi.org/10.1175/JHM-D-12-023.1

Santanello, J. A., Peters-Lidard, C. D., Kumar, S. V., Alonge, C., & Tao, W.-K. (2009). A modeling and observational framework for diagnosing local land–atmosphere coupling on diurnal time scales. *Journal of Hydrometeorology*, *10*(3), 577–599. https://doi.org/10.1175/2009JHM1066.1

Sathyanadh, A., Prabha, T. V., Balaji, B., Resmi, E. A., & Karipot, A. (2017). Evaluation of WRF PBL parameterization schemes against direct observations during a dry event over the Ganges valley. *Atmospheric Research*, *193*, 125–141. https://doi.org/10.1016/j.atmosres.2017.02.016

Savoy, P., & Mackay, D. S. (2015). Modeling the seasonal dynamics of leaf area index based on environmental constraints to canopy development. *Agricultural and Forest Meteorology*, *200*, 46–56. https://doi.org/10.1016/j.agrformet.2014.09.019

Schär, C., Lüthi, D., Beyerle, U., & Heise, E. (1999). The soil–precipitation feedback: A process study with a regional climate model. *Journal of Climate*, *12*(3), 722–741. https://doi.org/10.1175/1520-0442(1999)012⟨0722:TSPFAP⟩2.0.CO;2

Schlemmer, L., Hohenegger, C., Schmidli, J., & Schär, C. (2012). Diurnal equilibrium convection and land surface–atmosphere interactions in an idealized cloud-resolving model. *Quarterly Journal of the Royal Meteorological Society*, *138*, 1526–1539. https://doi.org/10.1002/qj.1892

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., & Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, *99*(3), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Seneviratne, S. I., Lüthi, D., Litschi, M., & Schär, C. (2006). Land–atmosphere coupling and climate change in europe. *Nature*, *443*, 205–209. https://doi.org/10.1038/nature05095

Seneviratne, S. I., Wilhelm, M., Stanelle, T., van den Hurk, B., Hagemann, S., Berg, A., Cheruy, F., Higgins, M. E., Meier, A., Brovkin, V., Claussen, M., Ducharne, A., Dufresne, J.-L., Findell, K. L., Ghattas, J., Lawrence, D. M., Malyshev, S., Rummukainen, M., & Smith, B. (2013). Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment. *Geophysical Research Letters*, *40*(19), 5212–5217. https://doi.org/10.1002/grl.50956

Shereena, O. A., & Rao, B. N. (2019). HDMR-Based Bayesian Structural System Identification. In *Recent advances in structural engineering, volume 1* (pp. 453–464). Springer. https://doi.org/10.1007/978-981-13-0362-3_36

Sherwood, S. C. (1999). Convective precursors and predictability in the tropical western Pacific. *Monthly Weather Review*, *127*(12), 2977–2991. https://doi.org/10.1175/1520-0493(1999)127⟨2977:CPAPIT⟩2.0.CO;2

Sobol′, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiment*, *1*(4), 407–414.

Song, J. I., Yum, S. S., Park, S. H., Kim, K. H., Park, K. J., & Joo, S. W. (2021). Climatology of melting layer heights estimated from cloud radar observations at various locations. *Journal of Geophysical Research: Atmospheres*, *126*(17), e2021JD034816. https://doi.org/10.1029/2021JD034816

Spennemann, P. C., Salvia, M., Ruscica, R. C., Sörensson, A. A., Grings, F., & Karszenbaum, H. (2018). Land-atmosphere interaction patterns in southeastern South America using satellite products and climate models. *International journal of applied earth observation and geoinformation*, *64*, 96–103. https://doi.org/10.1016/j.jag.2017.08.016

Su, H., & Dickinson, R. E. (2017). On the spatial gradient of soil moisture–precipitation feedback strength in the April 2011 drought in the Southern Great Plains. *Journal of Climate*, *30*(3), 829–848. https://doi.org/10.1175/JCLI-D-13-00185.1

Tavakol, A., Rahmani, V., Quiring, S. M., & Kumar, S. V. (2019). Evaluation analysis of NASA SMAP L3 and L4 and SPoRT-LIS soil moisture data in the United States. *Remote Sensing of Environment*, *229*, 234–246. https://doi.org/10.1016/j.rse.2019.05.006

Tawfik, A. B., Dirmeyer, P. A., & Santanello, J. A. (2015). The heated condensation framework.: Part i: Description and southern great plains case study. *Journal of Hydrometeorology*, *16*(5), 1929–1945. Retrieved July 25, 2023, from http://www.jstor.org/stable/24915510

Taylor, C. M. (2015). Detecting soil moisture impacts on convective initiation in Europe. *Geophysical Research Letters*, *42*(11), 4631–4638. https://doi.org/10.1002/2015GL064030

Taylor, C. M., Birch, C. E., Parker, D. J., Dixon, N., Guichard, F., Nikulin, G., & Lister, G. M. S. (2013). Modeling soil moisture-precipitation feedback in the Sahel: Importance of spatial scale versus convective parameterization. *Geophysical Research Letters*, *40*(23), 6213–6218. https://doi.org/10.1002/2013GL058511

Taylor, C. M., de Jeu, R. A. M., Guichard, ., Harris, P. P., & Dorigo, W. A. (2012). Afternoon rain more likely over drier soils. *Nature*, *489*(7416), 423–426. https://doi.org/10.1038/nature11377

Taylor, C. M., & Ellis, R. J. (2006). Satellite detection of soil moisture impacts on convection at the mesoscale. *Geophysical Research Letters*, *33*(3). https://doi.org/10.1029/2005GL025252

Taylor, C. M., Gounou, A., Guichard, F., Harris, P. P., Ellis, R. J., Couvreux, F., & De Kauwe, M. (2011). Frequency of Sahelian storm initiation enhanced over mesoscale soil-moisture patterns. *Nature Geoscience*, *4*(7), 430–433. https://doi.org/10.1038/ngeo1173

Taylor, C. M., Harris, P. P., & Parker, D. J. (2010). Impact of soil moisture on the development of a Sahelian mesoscale convective system: A case-study from the AMMA special observing period. *Quarterly Journal of the Royal Meteorological Society*, *136*(S1), 456–470. https://doi.org/10.1002/qj.465

Thompson, G., Field, P. R., Rasmussen, R. M., & Hall, W. D. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly weather review*, *136*(12), 5095–5115. https://doi.org/10.1175/2008MWR2387.1

Thompson, G., Rasmussen, R. M., & Manning, K. (2004). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Monthly Weather Review*, *132*(2), 519–542. https://doi.org/10.1175/1520-0493(2004)132⟨0519:EFOWPU⟩2.0.CO;2

Trenberth, K. E. (1999). Atmospheric moisture recycling: Role of advection and local evaporation. *Journal of Climate*, *12*(5), 1368–1381. https://doi.org/10.1175/1520-0442(1999)012⟨1368:AMRROA⟩2.0.CO;2

Tuttle, S. E., & Salvucci, G. D. (2017). Confounding factors in determining causal soil moisture-precipitation feedback. *Water Resources Research*, *53*(7), 5531–5544. https://doi.org/10.1002/2016WR019869

Wallace, C. E., Maddox, R. A., & Howard, K. W. (1999). Summertime convective storm environments in central Arizona: Local observations. *Weather and Forecasting*, *14*(6), 994–1006. https://doi.org/10.1175/1520-0434(1999)014⟨0994:SCSEIC⟩2.0.CO;2

Wang, G., Kim, Y., & Wang, D. (2007). Quantifying the strength of soil moisture–precipitation coupling and its sensitivity to changes in surface water budget. *Journal of Hydrometeorology*, *8*(3), 551–570. https://doi.org/10.1175/JHM573.1

Wang, H., Chen, L., Ye, F., & Chen, L. (2017). Global sensitivity analysis for fiber reinforced composite fiber path based on D-MORPH-HDMR algorithm. *Structural and Multidisciplinary Optimization*, *56*(3), 697–712. https://doi.org/10.1007/s00158-017-1681-9

Williams, I. N. (2019). Evaluating soil moisture feedback on convective triggering: Roles of convective and land-model parameterizations. *Journal of Geophysical Research: Atmospheres*, *124*(1), 317–332. https://doi.org/10.1029/2018JD029326

Xu, Z., Chen, H., Guo, J., & Zhang, W. (2021). Contrasting effect of soil moisture on the daytime boundary layer under different thermodynamic conditions in summer over China. *Geophysical Research Letters*, *48*(3), e2020GL090989. https://doi.org/10.1029/2020GL090989

Yin, J., Albertson, J. D., Rigby, J. R., & Porporato, A. (2015). Land and atmospheric controls on initiation and intensity of moist convection: CAPE dynamics and LCL crossings. *Water Resources Research*, *51*(10), 8476–8493. https://doi.org/10.1002/2015WR017286

Yin, J., Porporato, A., & Albertson, J. (2014). Interplay of climate seasonality and soil moisture-rainfall feedback. *Water Resources Research*, *50*(7), 6053–6066. https://doi.org/10.1002/2013WR014772

Yuan, S., Wang, Y., Quiring, S. M., Ford, T. W., & Houston, A. L. (2020). A sensitivity study on the response of convection initiation to in situ soil moisture in the central United States. *Climate Dynamics*, *54*, 2013–2028. https://doi.org/10.1007/s00382-019-05098-0

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., et al. (2016). Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bulletin of the American Meteorological Society*, *97*(4), 621–638. https://doi.org/10.1175/BAMS-D-14-00174.1

Zhang, L., He, C., & Zhang, M. (2017). Multi-scale evaluation of the SMAP product using sparse in-situ network over a high mountainous watershed, Northwest China. *Remote Sensing*, *9*(11), 1111. https://doi.org/10.3390/rs9111111

Zhang, X., Zhang, T., Zhou, P., Shao, Y., & Gao, S. (2017). Validation analysis of SMAP and AMSR2 soil moisture products over the United States using ground-based measurements. *Remote Sensing*, *9*(2), 104. https://doi.org/10.3390/rs9020104

Zhou, S., Williams, A. P., Lintner, B. R., Berg, A. M., Zhang, Y., Keenan, T. F., Cook, B. I., Hagemann, S., Seneviratne, S. I., & Gentine, P. (2021). Soil moisture–atmosphere feedbacks mitigate declining water availability in drylands. *Nature Climate Change*, *11*(1), 38–44. https://doi.org/10.1038/s41558-020-00945-z

Supporting Information for "Soil Moisture Cloud Precipitation Feedback in the Lower Atmosphere from Functional Decomposition of Satellite Observations"

Yifu Gao[1], Clément Guilloteau[1], Efi Foufoula-Georgiou[1], Chonggang Xu[2],

Xiaoming Sun[2] and Jasper A. Vrugt[1] *

[1] Department of Civil and Environmental Engineering, University of California, Irvine, California, USA.

[2] Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA.

Contents of this file

1. Introduction

2. Texts S1 to S4

3. Figures S1 to S7

Introduction

This supporting information contains the following content: (1) Text S1: A more detailed description of the SMAP/L4 and GPM/DPR/L2A products; (2) Text S2: Component

*Corresponding author: jasper@uci.edu

function construction and D-MORPH regression that enforces hierarchical orthogonality of the component functions in pursuit of the optimum coefficients; (3) Text S3: Investigation into atmospheric controls on cloud vertical profile (CVP); (4) Text S4: Application of Marshall-Palmer formula to converting component function, $f_i(x_i)$, into rainfall rate estimates; (5) Figure S1: Correlogram of the land-surface and atmospheric variables employed as inputs in the High-Dimensional Model Representation (HDMR): soil moisture (SM), land-surface temperature (LST), leaf area index (LAI), atmospheric temperature (AT), and total precipitable water (TPW); (6) Figure S2: Diurnal development of ERA5 reanalysis planetary boundary layer (PBL) height determined for two groups of samples where Groups 1 and 2 highlight positive and negative soil moisture cloud precipitation feedback (SMCPF), respectively; (7) Figure S3: Vertical profiles of the mean F-statistic of the first-order component functions of AT: $f_4(x_4)$ and TPW: $f_5(x_5)$; (8) Figure S4: Scatter plots of the samples of antecedent 7-hour AT (K) and TPW (kg/m^2) against their first-order component functions, $f_4(x_4)$ and $f_5(x_5)$, evaluated at three separate heights; (9) Figure S5: The central United States (95°W-105°W, 32°N-40°N) with antecedent 7-hr SMAP/L4 soil wetness (-) of the top layer (0-5 cm) collocated at coordinates of the GPM/DPR/L2A samples and change in rainfall rate, ΔR (mm/hour), attributed to SM; (10) Figure S6: same as Figure S5 but for LST and its associated impact on rainfall rate; (11) Figure S7: Evidence of negative SMCPF in the southwest (101°W-105°W, 32°N-36°N) of the central United States.

Text S1. Data Description: SMAP/L4 and DPR/L2A Products

The SMAP mission Level 4 SM (L4_SM) product provides 3-hourly estimates of surface and root-zone SM at 9-km spatial resolution with global coverage (Reichle et al., 2015). Despite the malfunction of SMAP's active radar system since July 2015, its passive microwave radiometer has continued to operate and measure brightness temperatures. SMAP L-band (1.4 GHz) brightness temperature data from descending and ascending half-orbit satellite passes (approximately 6:00 AM and 6:00 PM local solar time, respectively) are assimilated into the NASA catchment land-surface model using the Earth-fixed, global, cylindrical 9 km Equal-Area Scalable Earth Grid, Version 2.0 (EASE-Grid 2.0) projection. L4_SM provides surface (see Figure 1a) and root zone SM data in two products. We use the 3-hour time-averaged 9-km geophysical data product (SPL4SMGP) which provides soil wetness (0-1) of the top layer (0-5 cm) and other land-surface variables.

Cloud vertical profiles are derived from the dual-frequency precipitation radar (DPR) aboard the Global Precipitation Measurement (GPM) Core Observatory satellite. Launched in February 2014, the GPM core satellite orbits the Earth about 16 times a day in a non-sun-synchronous orbit with an inclination angle of 65°. The DPR operates at Ku-band (13.6 GHz) and Ka-band (35.5 GHz) frequencies and is an advanced successor to the Tropical Rainfall Measuring Mission precipitation radar. The DPR has the capability of obtaining the raindrop size distribution with improved detection of light rain and precipitating snow due to the addition of the Ka-band radar. This instrument operates in two modes: (1) a higher range resolution, lower sensitivity mode in the inner swath (125 km) and (2) a lower resolution, higher sensitivity mode (Liao & Meneghini, 2022). The KuPR and KaPR sense rain over land and ocean, day and night.

The GPM/DPR/L2A product (GPM_2ADPR) provides a swath of precipitation profiles (see Figure 1b) every 1.5 hours with a spatial resolution of 5 km and vertical increment of 125 m. Each pixel has its own cloud and precipitation profiles such as the cloud reflectivity factor (see Figure 1c), precipitation rate, height of received echos, and so forth. The DPR level-2 algorithm is made up of six different modules named preparation (PRE), vertical profile (VER), classification (CSF), drop size distribution (DSD), surface reference technique (SRT) and solver (SLV) (Iguchi et al., 2010). The SLV module computes the DSD, precipitation rate and related physical quantities by solving the radar equations recursively along range profiles utilizing output received from other modules such as the measured reflectivity profile (PRE), precipitation type (CSF), path integrated attenuation (SRT) and an adjustable $R - D_{\mathrm{m}}$ relationship of precipitation rate R and mass-weighted diameter D_{m} (DSD). We use the major data fields, zFactorFinal (dBZ) and typePrecip (-), which provide vertical profiles of the Ka-band cloud reflectivity factor and an 8-digit ID for precipitation type, respectively. In this study, we take the 250-m average Ka-band cloud reflectivity and exclusively use samples classified as convective precipitation.

Text S2. Component Function Construction and D-MORPH Regression

We construct the component functions using the family of orthogonal polynomial functions (Li & Rabitz, 2012)

$$\begin{array}{ccc} \phi_1(x_i) = a_1 x_i + a_0 & \phi_2(x_i) = b_2 x_i^2 + b_1 x_1 + b_0 & \phi_3(x_i) = c_3 x_i^3 + c_2 x_i^2 + c_1 x_i + c_0 \\ \text{degree } p = 1 & \text{degree } p = 2 & \text{degree } p = 3, \end{array} \tag{S1}$$

where the values of coefficients a, b and c are derived from Gram-Schmidt orthonormalization. This projection operator constructs an orthonormal basis for the polynomial functions on the unit interval of x with respect to an arbitrary weighting function. The component functions are now equal to sums of linear multiples of the orthonormalized polynomial functions of degrees 1 to p

$$f_i(x_i) = \sum_{r=1}^{p} \alpha_r^{(i)i} \phi_r(x_i) \tag{S2a}$$

$$f_{ij}(x_i, x_j) = \sum_{r=1}^{p} \left[\alpha_r^{(ij)i} \phi_r(x_i) + \alpha_r^{(ij)j} \phi_r(x_j)\right] + \sum_{r=1}^{p} \sum_{s=1}^{p} \beta_{rs}^{(ij)ij} \phi_r(x_i) \phi_s(x_j), \tag{S2b}$$

where the extended bases of the second-order component functions will help satisfy the vanishing condition in Equation (3). The use of extended bases has implications for our index notation of the coefficients. Parenthesized symbol(s) in the superscripts of α, β and γ enumerate the component functions. Non-parenthesized superscripts are indices of the input vector, $\mathbf{x}$. If all n_{12} component functions are included in the series expansion of equation (1) then the number of unknown expansion coefficients equals $l = dp + \frac{1}{2}d(d-1)(2p+p^2)$. At the end of Section 3, we introduce the five ($d = 5$) input variables used in our analysis. Thus, with a typical polynomial degree $p = 3$ (Gao et al., 2023) the number of unknown expansion coefficients $l = 165$ is much smaller than the sample size n for each cloud height (Figure 1d). This minimizes the risk of overfitting.

Hierarchical representation of the cloud reflectivity into a finite sum of first- and second-order polynomial component functions offers a significant advantage over function approximation methods such as artificial neural networks. The function expansion delineates marginal and cooperative effects in determining the magnitude and sign of the SMCPF. Furthermore, the expansion coefficients α, β and γ of the component functions of equation (S2) have a closed-form solution for a training record of $(\mathbf{x}, y)$-samples.

We can write equation (2) in matrix form $\mathbf{\Phi c} = \mathbf{b}$ and yield

$$\mathbf{\Phi} = \begin{bmatrix} \boldsymbol{\phi}(\mathbf{x}^{(1)})^\top \\ \vdots \\ \boldsymbol{\phi}(\mathbf{x}^{(N)})^\top \end{bmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{bmatrix} y^{(1)} - y_0 \\ \vdots \\ y^{(N)} - y_0 \end{bmatrix}, \tag{S3a}$$

where $\boldsymbol{\phi}(\mathbf{x})^\top$ is a $1 \times l$ design vector with orthonormalized polynomial functions of equation (S2) (and products thereof) evaluated at their respective entries of $\mathbf{x}$ and arranged in appropriate order, $\mathbf{c}$ is a $l \times 1$ coefficient vector with values of α, β and γ and the $n \times 1$ vector $\mathbf{b}$ stores differences between the measured $y^{(i)}$ and mean y_0 cloud reflectivity for each training sample, $i = (1, \dots, n)$. To offer some protection against underdetermined problems $N < l$ or a rank-deficient design matrix, we remove duplicate entries of the basis functions of the first- and second-order component functions. This reduced system is easier to solve in practice (Li & Rabitz, 2012). First, we determine the least squares values $\widehat{\mathbf{c}}_{\text{ls}}$ of the expansion coefficients

$$\widehat{\mathbf{c}}_{\text{ls}} = (\mathbf{\Phi}^\top \mathbf{\Phi})^\dagger \mathbf{d}, \tag{S4}$$

where the $l \times (l - dp)$ matrix $(\mathbf{\Phi}^\top \mathbf{\Phi})^\dagger$ is the generalized pseudo inverse of the $l \times l$ Gramian matrix, $\mathbf{G} = \mathbf{\Phi}^\top \mathbf{\Phi}$, which satisfies all four Moore-Penrose conditions (Penrose, 1955; Golub & Van Loan, 1996) and whose redundant rows (first dp rows of the first-order basis functions) are removed and $\mathbf{d}$ is the $(l - dp) \times 1$ vector $\mathbf{\Phi}^\top \mathbf{b}$ without the first dp rows. Diffeomorphic modulation under observable response preserving homotopy (D-MORPH) regression (Li & Rabitz, 2010) enforces hierarchical orthogonality of the component functions in pursuit of the optimum coefficients

$$\widehat{\mathbf{c}}_{\mathrm{dm}} = \mathbf{V}_{l-r}(\mathbf{U}_{l-r}^{\top}\mathbf{V}_{l-r})\mathbf{U}_{l-r}^{\top}\widehat{\mathbf{c}}_{\mathrm{ls}}, \tag{S5}$$

where $\mathbf{U}_{l-r}$ and $\mathbf{V}_{l-r}$ equal the last $l-r$ columns of the $l \times l$ matrices $\mathbf{U}$ and $\mathbf{V}$ determined from singular value decomposition $\mathbf{PB} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\top}$ of the product of a $l \times l$ projection matrix $\mathbf{P} = \mathbf{I}_l - \mathbf{G}$ and $l \times l$ constraint matrix $\mathbf{B}$ of inner products of the orthonormalized polynomial functions. This latter matrix $\mathbf{B}$ enforces the relaxed vanishing condition in Equation (3) (Li & Rabitz, 2010), matrix $\mathbf{I}_l$ is the $l \times l$ identity matrix and r is the number of nonzero singular values.

Text S3. Atmospheric Controls on CVP

Figure S3 displays the similar content as Figure 2 but for the two atmospheric conditions: low-level (roughly 1-3 km) AT and TPW, whose component functions are $f_4(x_4)$ and $f_5(x_5)$, respectively. As anticipated, antecedent 7-10 hours low-level AT significantly influences the cloud vertical profile (CVP) within the 1-3 km range, denoting a bottom-heavy relationship. The component function, $f_4(x_4)$, shows a strong negative correlation with AT (as illustrated in Figure S5a) and thus underscores the profound contribution of a cooler early atmosphere to the development of convective clouds/precipitation. The observed sensitivity of CVP to early-stage AT is deemed reasonable since AT is a crucial atmospherically forced synoptic condition for diagnosing the likelihood of deep convection. Conditions of lower AT coupled with higher LST are conducive to higher Convective Available Potential Energy (CAPE) and Convective Triggering Potential (CTP) (Findell & Eltahir, 2003a). Compared to TPW, AT exhibits weaker predictability in the free atmosphere, likely due to TPW's more straightforward connection with the cloud formation (as detailed below). Through integrating the characterized relationships between CVP and $\{\mathrm{SM}, \mathrm{LST}, \mathrm{AT}\}$ (i.e., $f_1(x_1)$, $f_2(x_2)$, and $f_4(x_4)$), we can identify favorable conditions for SM-cloud-precipitation feedback (SMCPF) within the height range of 1-3 km: (i) substantial boundary-layer moistening from wet soil (ii) the existence of a unstable lapse rate facilitated by a warm surface and a cool low-level atmosphere. This finding corroborates the physical mechanisms underlying SMCPF pathways (Wallace & Hobbs, 2006).

Further investigation of antecedent 7-hour TPW shows somehow the opposite pattern against AT. Such dependence of CVP on TPW can be explicated by its reflection of the synoptic scale humidity of the early atmosphere. Intuitively, early TPW can be viewed as a proxy for the amount of water vapor that actually condenses and forms clouds and precipitation later. This is coordinated with the derived positive correlation between TPW and its component function, $f_5(x_5)$ in Figure S5. In addition, TPW can be a precursor to mesoscale convective events. A sharp increase in TPW prior to the convective precipitation is indicative of the deep convection (Sherwood, 1999; Holloway & Neelin, 2010). This possibly explains why the magnitudes of $f_5(x_5)$ and its F-statistics increase with height so that CVP is more sensitive to TPW in the free atmosphere. 7-hour is observed to be the most informative time lag for the TPW-CVP relationship. This comes in excellent agreement with the conclusion of Holloway and Neelin (2010) that, with the involvement of mesoscale convective dynamics, a peak in TPW occurs typically 7-hour prior to the strong precipitation events at Nauru Island.

In summary, atmospheric controls on CVP can be altitude-dependent. The low-level AT, along with SM and LST, exhibits a governing effect on convective clouds/precipitation within the 1-3 km zone. TPW, by contrast, plays a critical role in shaping cloud and precipitation distribution in the free atmosphere.

Text S4. Marshall-Palmer formula

The component function, $f_i(x_i)$ (dBZ), which quantifies the contribution of a variable (e.g., SM) to cloud reflectivity, can be further converted into estimates of rainfall rate through the Marshall-Palmer formula (Marshall & Palmer, 1948)

$$R_0 = \left[\frac{10^{(f_0/10)}}{200}\right]^{5/8} \tag{S6a}$$

$$R_1 = \left\{\frac{10^{[(f_0+f_i(x_i))/10]}}{200}\right\}^{5/8} \tag{S6b}$$

$$\Delta R = R_1 - R_0, \tag{S6c}$$

where R_0 signifies the mean rainfall rate (mm/hour) estimated from the mean cloud reflectivity, f_0 (dBZ), and R_1 is the same quantity but computed using the sum of mean cloud reflectivity and the SM component, $f_0 + f_i(x_i)$ (dBZ). By taking the difference between the two quantities (ΔR), we can readily determine the impact of SM on rainfall rates. As is shown in Figure S4c, the 7-hour wet soil can account for up to a 2 mm/hour increment in rainfall rate at 2.0 km, denoting strong positive feedback.

References

Findell, K. L., & Eltahir, E. A. B. (2003a). Atmospheric controls on soil moisture–boundary layer interactions. Part I: Framework development. *Journal of Hydrometeorology*, *4*, 552-569. doi: 10.1175/1525-7541(2003)004⟨0552:ACOSML⟩2.0.CO;2

Findell, K. L., & Eltahir, E. A. B. (2003b). Atmospheric controls on soil moisture–boundary layer interactions. Part II: Feedbacks within the continental United States. *Journal of Hydrometeorology*, *4*(3), 570–583. doi: 10.1175/1525-7541(2003)004⟨0570:ACOSML⟩2.0.CO;2

Gao, Y., Sahin, A., & Vrugt, J. A. (2023). Probabilistic sensitivity analysis with dependent variables: Covariance-based decomposition of hydrologic models. *Water Resources Research*, *59*(4), e2022WR032834. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022WR032834` doi: 10.1029/2022WR032834

Golub, G. H., & Van Loan, C. F. (1996). *Matrix computations (3rd ed.).* Baltimore: Johns Hopkins.

Holloway, C. E., & Neelin, J. D. (2010). Temporal relations of column water vapor and tropical precipitation. *Journal of the Atmospheric Sciences*, *67*(4), 1091–1105. doi: 10.1175/2009JAS3284.1

Iguchi, T., Seto, S., Meneghini, R., Yoshida, N., Awaka, J., Le, M., ... Kubota, T. (2010). GPM/DPR level-2 algorithm theoretical basis document. *NASA Goddard Space Flight Center*.

Li, G., & Rabitz, H. (2010). D-MORPH regression: application to modeling with unknown parameters more than observation data. *Journal of Mathematical Chemistry*, *48*, 1010–1035. doi: 10.1007/s10910-010-9722-2

Li, G., & Rabitz, H. (2012). General formulation of HDMR component functions with independent and correlated variables. *Journal of Mathematical Chemistry*, *50*(1), 99-130. doi: 10.1007/s10910-011-9898-0

Liao, L., & Meneghini, R. (2022). GPM DPR retrievals: Algorithm, evaluation, and validation. *Remote Sensing*, *14*(4), 843. doi: 10.3390/rs14040843

Marshall, J., & Palmer, W. M. (1948). The distribution of raindrops with size. *Journal of meteorology*, *5*, 166.

Penrose, R. (1955). A generalized inverse for matrices. *Proceedings of the Cambridge Philosophical Society*, *51*(3), 406-413. doi: 10.1017/S0305004100030401

Reichle, R. H., Lucchesi, R. A., Ardizzone, J. V., Kim, G.-K., Smith, E. B., & Weiss, B. H. (2015). *Soil moisture active passive (SMAP) mission level 4 surface and root zone soil moisture (L4_SM) product specification document* (Tech. Rep.). NASA Goddard Space Flight Center. Retrieved from https://nsidc.org/sites/default/files/reichle789.pdf

Sherwood, S. C. (1999). Convective precursors and predictability in the tropical western Pacific. *Monthly Weather Review*, *127*(12), 2977–2991. doi: 10.1175/1520-0493(1999)127⟨2977:CPAPIT⟩2.0.CO;2

Wallace, J. M., & Hobbs, P. V. (2006). *Atmospheric science: an introductory survey* (Vol. 92). Elsevier.

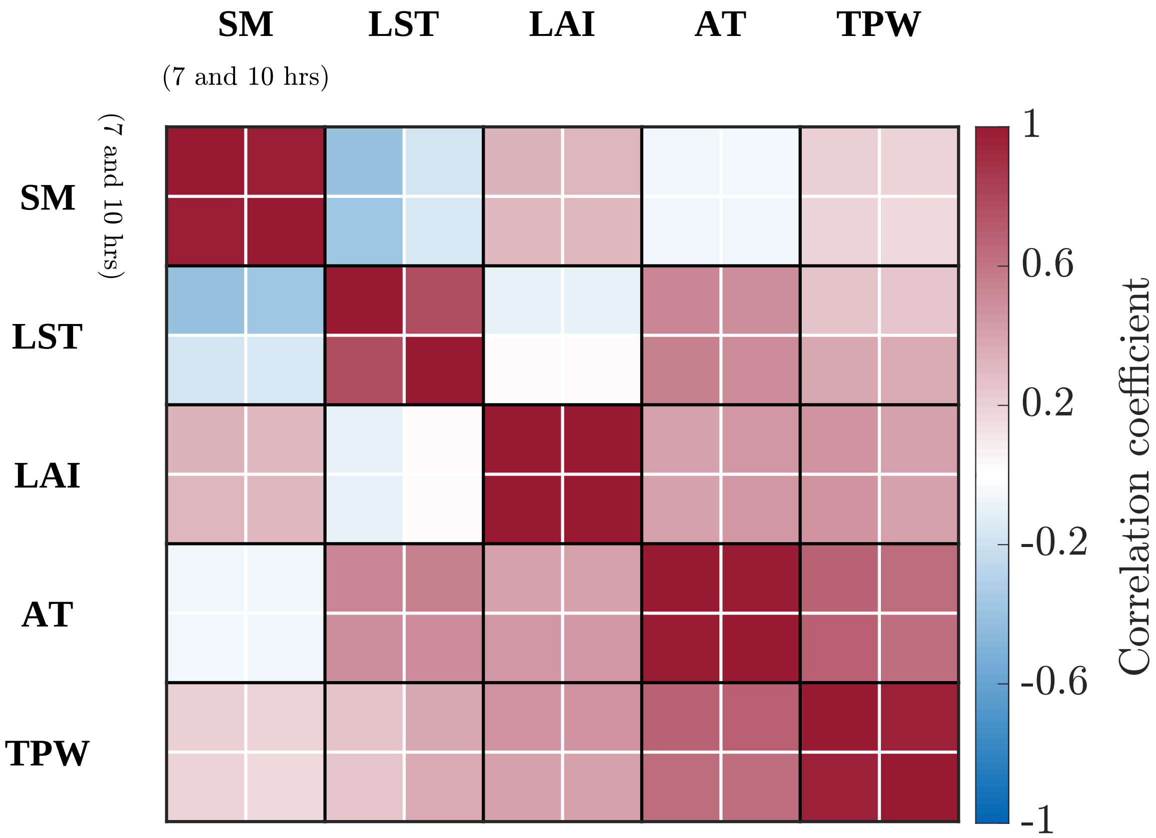

Figure S1. Correlogram of the land-surface and atmospheric variables used as inputs of HDMR. Solid black lines demarcate distinct variables, whereas solid white lines differentiate between time lags ($\Delta t = 7$ and 10 hours) relative to the DPR scanning time.

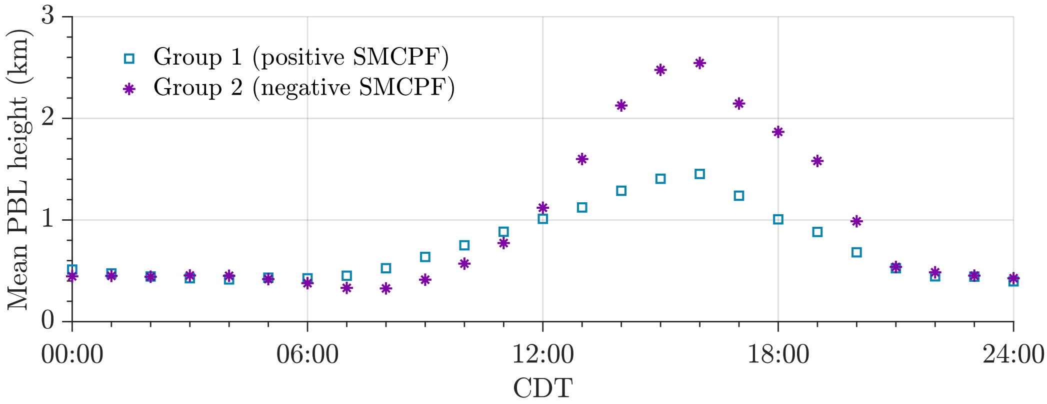

Figure S2. Diurnal development of the mean ERA5 reanalysis PBL height at Central Daylight Time (CDT) determined for two groups of samples. Groups 1 and 2 highlight positive SMCPF (blue squares) with $f_1(x_1) > 1.0$ dBZ and negative SMCPF (violet stars) with $f_2(x_2) > 1.0$ dBZ, respectively, both at a time lag of 7 hours and cloud height of 2.0 km.

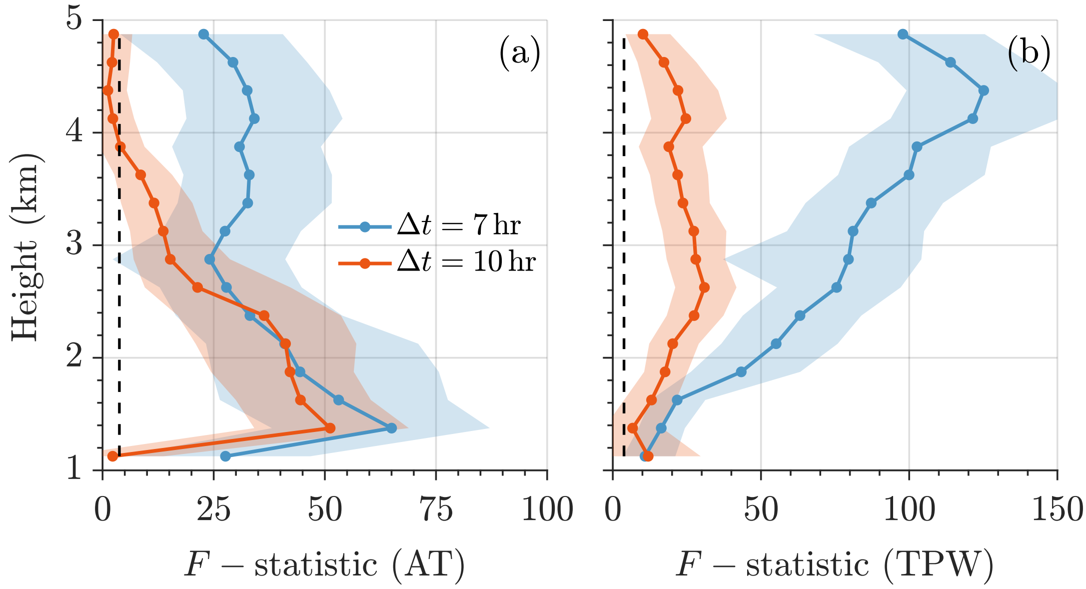

Figure S3. Vertical profiles of the mean F-statistic of the first-order component functions of (a) AT: $f_4(x_4)$ and (b) TPW: $f_5(x_5)$ computed from 1,000 bootstrap iterations. Solid blue and red lines refer to different temporal lags ($\Delta t = 7$ and 10 hours) and black dashed lines represent the critical value at significance level $\alpha = 0.05$, F_{crit}. The light red and blue regions correspond to the 95% bootstrap confidence intervals.

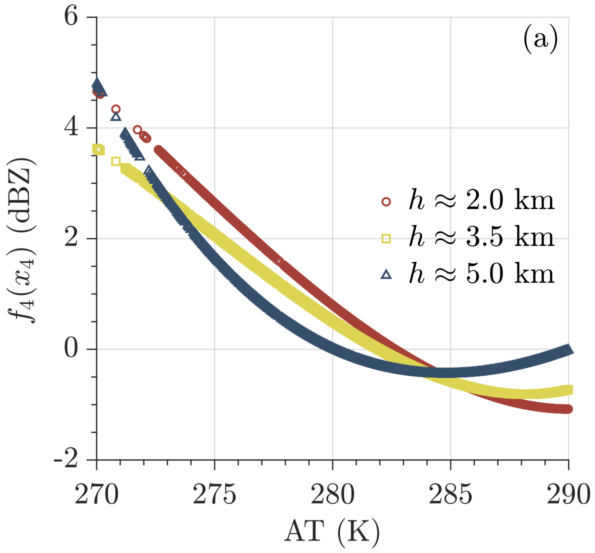

Figure S4. Scatter plots of the samples of antecedent 7-hour (a) AT (K) and (b) TPW (kg/m^2) against their first-order component functions, $f_4(x_4)$ (dBZ) and $f_5(x_5)$ (dBZ), evaluated at three separate heights, 2.0 km (red circles), 3.5 km (yellow squares), and 5.0 km (blue triangles).

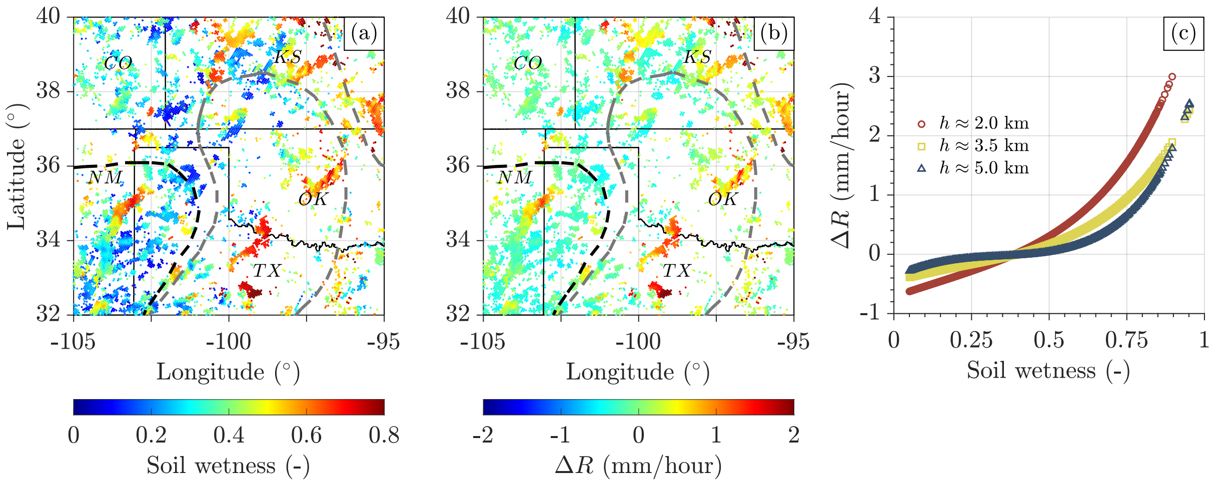

Figure S5. The central United States (95°W-105°W, 32°N-40°N) with (a) antecedent 7-hr SMAP/L4 soil wetness (-) of the top layer (0-5 cm) collocated at coordinates of the GPM/DPR/L2A samples and (b) change in rainfall rate, ΔR (mm/hour), at 2.0 km attributed to SM. Solid black lines delineate the state borders while dashed black and grey lines depict the negative feedback and transitional regions proposed by Findell and Eltahir (2003b). Panel (c) displays the scatter plots of the samples of antecedent 7-hour SM against the corresponding change in rainfall rate, ΔR (mm/hour), evaluated at three separate heights, 2.0 km (red circles), 3.5 km (yellow squares), and 5.0 km (blue triangles).

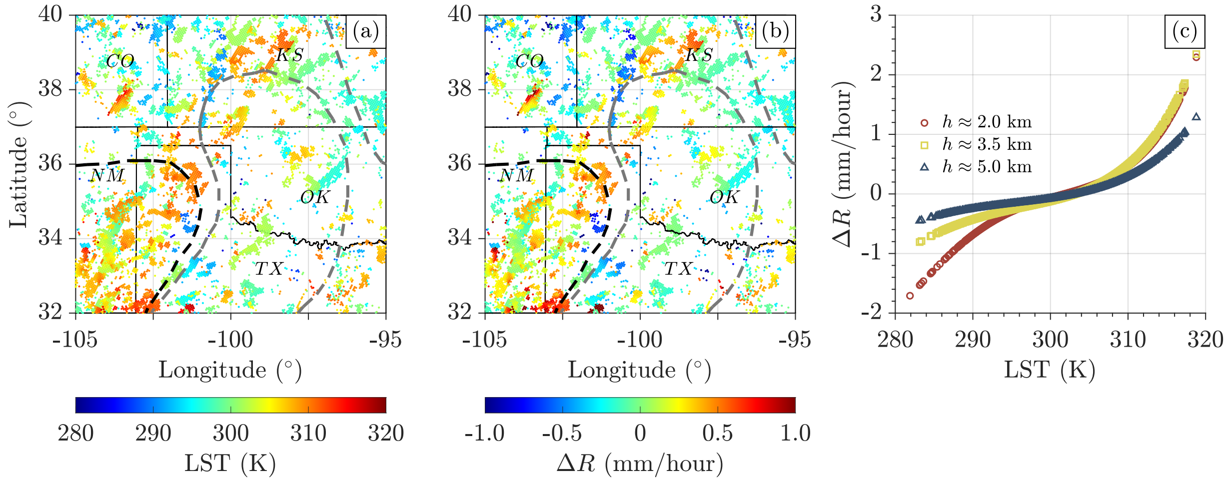

Figure S6. Same as Figure S5 but for SMAP/L4 LST and its associated change in rainfall rate.

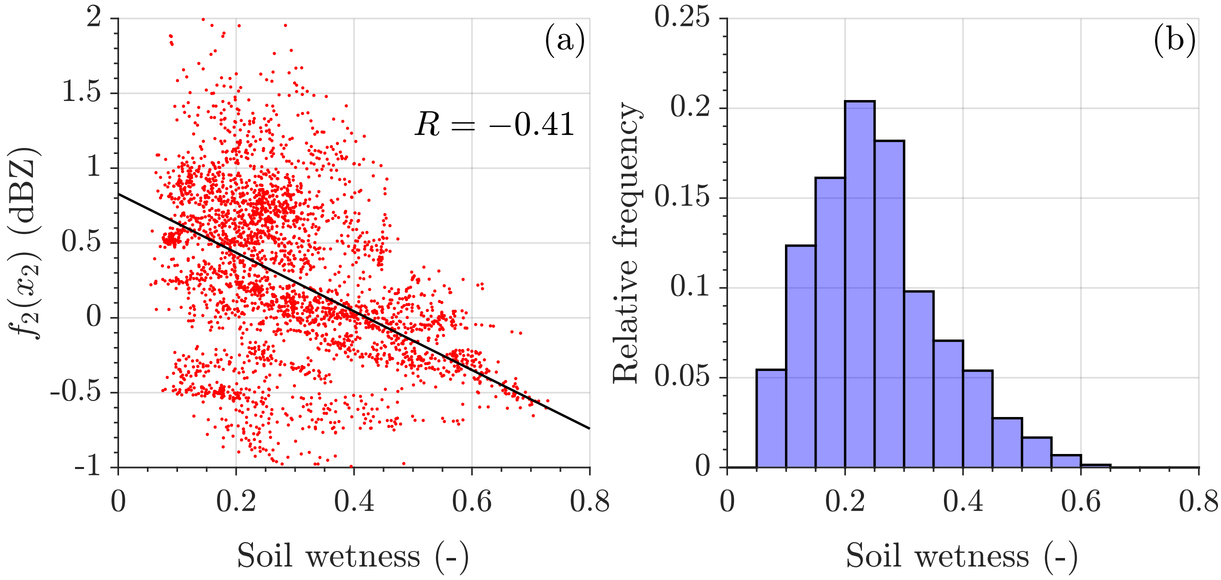

Figure S7. Evidence of negative SMCPF in the southwest (101°W-105°W, 32°N-36°N) of the central United States: (a) scatter plot of the SM samples from this area against the respective LST component function, $f_2(x_2)$; solid black line portrays the least squares fit of a simple regression function to the samples; (b) marginal distribution of SM subsampled from panel (a) with $f_2(x_2) > 0$ (dBZ). The negative correlation ($R = -0.41$) between SM and $f_2(x_2)$ and LST's pronounced contribution to the cloud over dry soils highlight the intrinsic SM-LST coupling. This aligns with the negative SMCPF.