Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation

Kirsten L. Findell¹*, Pierre Gentine²†, Benjamin R. Lintner³† and Christopher Kerr⁴

Moisture and heat fluxes from the land surface to the atmosphere form a critical nexus between surface hydrology and atmospheric processes, particularly those relevant to precipitation. Although current theory suggests that soil moisture generally has a positive impact on subsequent precipitation, individual studies have shown support both for and against this positive feedback. Broad assessment of the coupling between soil moisture and evapotranspiration, and evapotranspiration and precipitation, has been limited by a lack of large-scale observations. Quantification of the influence of evapotranspiration on precipitation remains particularly uncertain. Here, we develop and apply physically based, objective metrics for quantifying the impacts of surface evaporative and sensible heat fluxes on the frequency and intensity of convective rainfall during summer, using North American reanalysis data. We show that high evaporation enhances the probability of afternoon rainfall east of the Mississippi and in Mexico. Indeed, variations in surface fluxes lead to changes in afternoon rainfall probability of between 10 and 25% in these regions. The intensity of rainfall, by contrast, is largely insensitive to surface fluxes. We suggest that local surface fluxes represent an important trigger for convective rainfall in the eastern United States and Mexico during the summer, leading to a positive evaporation–precipitation feedback.

Observational studies of soil moisture–rainfall interactions generally suffer from insufficient data both spatially and temporally, particularly for soil moisture (SM) and surface turbulent fluxes. In models, the strength of the SM–rainfall feedback depends on model parameterizations and grid resolution. The metrics introduced in the present study constitute powerful diagnostic tools for model intercomparison of simulated land–atmosphere coupling, and for validation of fundamental physical processes incorporated in next-generation earth system models.

Previous work identified three necessary conditions for an initial SM anomaly to impact summertime precipitation: (1) the initial SM anomaly must be large; (2) evaporation must be strongly sensitive to SM; and (3) precipitation must be strongly sensitive to evaporation. The first element emphasizes areas with large SM variability, whereas the latter element emphasizes the connection between the land surface and the atmospheric boundary layer (ABL), and the connection between the ABL and overlying free troposphere. This study emphasizes (3): we demonstrate when and where the North American Regional Reanalysis dataset (NARR; ref. 12) manifests strong precipitation sensitivity to the relative strengths of latent and sensible heat fluxes from the land surface. The link between surface evapotranspiration (ET) and subsequent precipitation is recognized as a critical and highly uncertain step in the SM–precipitation feedback loop.

The NARR dataset comprises dynamically consistent three-hourly gridded fields from 1979 to 2003 at ~30 km grid spacing, providing an extensive testbed for probing mechanistic connections between surface fluxes and subsequent precipitation. NARR is derived from a data assimilation system with precipitation and other near-surface observations assimilated hourly, and atmospheric profiles of temperature, winds, and moisture from rawinsondes and dropsondes assimilated every three hours. Studies documenting the strengths and weaknesses of NARR are discussed in the Supplementary Information. These studies demonstrate that NARR is an improvement over earlier, global reanalysis projects and that it can successfully be used in a wide array of hydrometeorological studies.

The strengths of NARR include its assimilation of precipitation observations and its high spatio-temporal resolution. Precipitation assimilation constrains the diurnal cycle of precipitation, which is poorly captured by current convection schemes; moreover, assimilation of near-surface humidity constrains latent and sensible heat flux partitioning, which is often poorly captured by land surface models. Whereas data density and to a lesser extent frequency are comparable for the continental United States (CONUS) and Mexico, the quality and quantity of observations for Canada are limited; we limit consideration here to data south of 50° N.

Our analysis uses daily gridpoint values of the following quantities: early morning convective triggering potential (CTP) and low-level humidity deficit (Hmd) to assess the large-scale potential for convective development; before-noon evaporative fraction (EF = λE/(H + AE), with λE and H denoting latent and sensible heating, respectively) to assess surface turbulent flux partitioning; and rainfall from noon–6 pm. Temporal offsetting of measurements helps to isolate EF forcing of subsequent precipitation. Further data filtering (described below) removes large-scale influences potentially affecting both EF and precipitation.

Not surprisingly, the likelihood of summertime afternoon rainfall is higher in the humid regions of the NARR domain: Mexico and the southeastern United States show the highest afternoon rainfall probabilities (Supplementary Fig. S1). These...
regions also show a positive EF-precipitation relationship: higher EF, generally linked to wetter soils and increased vegetation coverage\(^1\), is associated with a higher likelihood of afternoon rainfall (Supplementary Fig. S1).

The EF-dependence of rainfall is assessed through two measures: a triggering feedback strength (TFS), reflecting how afternoon rainfall frequency changes with EF, and an amplification feedback strength (AFS), reflecting how accumulated rainfall varies with EF when afternoon rainfall occurs. TFS is given by

\[
\text{TFS} = \sigma_{\text{EF}} \frac{\partial \Gamma(r)}{\partial \text{EF}}
\]

where \(\sigma_{\text{EF}}\) is the standard deviation of EF. We use \(\Gamma(r)\) to denote event probabilities; \(\Gamma(r)\) then denotes the probability of afternoon rain exceeding a 1 mm threshold. AFS is defined analogously, replacing \(\Gamma(r)\) with \(E(r)\), the expected value of afternoon rainfall amount. (Details are provided in equations (2)–(5) of the Methods section.)

To mitigate the impact of large-scale synoptic systems and constrain the analysis to days when local surface turbulent fluxes are most conducive to subsequent convective development, two restrictions are applied to the data included in TFS. First, only days without rainfall between 6 am and noon are retained, limiting the influence from long-duration stratiform rainfall events\(^1\). Second, days with negative CTP are excluded as early morning CTP < 0 conditions have been shown to be typically too stable to support convection\(^1\); afternoon rainfall occurring on such days is assumed to arise from synoptic-scale systems. These restrictions remove 10–30\% of days in the eastern United States and 5–10\% in the western United States and northern Mexico (Supplementary Fig. S2). Although these restrictions do not guarantee removal of all synoptically driven days, they diminish the influence of synoptic systems. The AFS calculation is further restricted to days with afternoon rain (Supplementary Fig. S2), because TFS already accounts for rain-free afternoons. Finally, 50 bootstrap samples (with replacement)\(^1\) are created from the available 2,300 days to assess statistical significance.

The TFS map (Fig. 1a) shows that over the eastern United States and Mexico, higher EF leads to increased afternoon rainfall probabilities. This map is largely consistent with Findell and Eltahir\(^1\). Although the earlier study was limited in scope, using station radiosondes in a 1D boundary layer model, simulations were performed with very dry or wet soils to isolate land surface forcing. Consistency with Findell and Eltahir, in the area they deemed a positive feedback region, supports our finding that surface flux partitioning plays an integral role in afternoon convective triggering in the eastern United States. Independent results from the Atmospheric Radiation Measurement Climate Research Facility Southern Great Plains site (97.5° W, 36.5° N; ref. 20) also show consistency with our near-zero TFS signal at this location.

Figure 1 underscores the importance of land surface and ABL processes for convective triggering in some regions while indicating that rainfall amounts are largely independent of local surface moisture conditions in all but the wettest regions.

Thus, in Mexico and the eastern United States, we argue that the positive EF-precipitation feedback occurs primarily through modification of afternoon rainfall frequency. West of the Mississippi, the land surface exerts little control on local afternoon convection. In other words, where surface moisture is not strictly limited, surface turbulent flux partitioning can shift the local atmosphere from non-convecting to convecting, but other controls, for example, free tropospheric moisture content or large-scale moisture convergence, largely determine how much rainfall occurs, consistent with independent studies\(^21\). The dichotomy between surface controls on convective frequency and intensity constitutes a fundamental advancement in our understanding of land–atmosphere interactions and is consistent with findings from small-scale studies in the US midwest and Europe\(^3,17,22\).

To address whether our results reflect external factors that might impact both morning EF and subsequent precipitation we analysed rainfall for the three subsequent six-hourly periods after the noon–6 pm period considered previously. Figure 2
Using 6 pm–midnight rainfall

Figure 2 | The sensitivity of TFS and AFS to rainfall time period. TFS (top row) and AFS (bottom row) values calculated using rainfall from subsequent time periods: 6 pm–midnight (left column), midnight–6 am (middle column), and 6 am–noon (right column). The middle and right columns use rainfall from the day following the recorded EF value. Shading details as in Fig. 1.

Using midnight–6 am rainfall

Figure 3 | Normalized, non-dimensional versions of sensitivity maps. a, normTFS and b, normAFS. Shading details as in Fig. 1.

shows that the TFS signal persists, but is muted from 6 pm to midnight, and drops off rapidly thereafter, with only small areas exceeding 2.5% after 6 am, indicating that persistent large-scale factors do not drive the Fig. 1 results. AFS remains noisy throughout these subsequent intervals, reflecting the difficulty of distinguishing AFS signal from noise. Additionally, we subdivided the dataset into ten El Niño and seven La Niña years (see Supplementary Information). Results with these subsamples show that TFS and AFS are insensitive to tropical SST forcing (Supplementary Fig. S3).

The ground-breaking analysis of Koster et al. quantified feedback strength in terms of multi-model mean summertime precipitation variability differences of simulations with and without interactive SM. Their central Great Plains ‘hotspot’ clearly differs from the surface flux triggering hotspot east of the Mississippi evident in our Fig. 1a. The different feedback measures considered (EF versus deep SM) partly explain the different locations, as the Koster et al. hotspots include not only the precipitation sensitivity to evapotranspiration considered here, but also evapotranspiration sensitivity to SM. There were tremendous inter-model differences in the assessment of the SM-precipitation feedback strength in the Koster et al. results, tied to model-specific relationships between SM and evapotranspiration and possibly also the details of convection schemes and cloud-radiative feedbacks. A related set of experiments with all prognostic land variables (SM and temperature at all levels, canopy interception reservoir content, various snow-related quantities) prescribed may be a closer analogue to our work because EF is substantially influenced by surface SM and canopy interception; these experiments manifest a larger signal in the eastern United States, albeit with substantial intermodel differences.

To better understand the relative magnitudes of the triggering and amplification metrics, versions of the TFS and AFS with the central derivatives normalized by the ratio of mean EF (EF) to mean rainfall (r) or (Fig. 3) are considered (normTFS and normAFS, see equations (6) and (7) in Methods). Differences between TFS, with derivatives scaled by σ_EF, (Fig. 1a) and normTFS (Fig. 3a) are largest in Mexico, where (T) is highest. The functional
relationship of normTFS to $\overline{EF}$ is nonlinear but relatively smooth (Fig. 4a). Four regimes are evident: for $\overline{EF} < 0.2$, normTFS is highly variable with a near-zero mean; for $0.2 < \overline{EF} < 0.5$, normTFS shows little change with $\overline{EF}$; for $0.5 < \overline{EF} < 0.7$, normTFS shows a slight positive slope with respect to $\overline{EF}$; and for $\overline{EF} > 0.7$, normTFS increases dramatically with $\overline{EF}$. The only instances of negative TFS values are localized to the area north of the Gulf of California (Fig. 3a), reminiscent of a similarly located negative feedback region seen in earlier work. For $\overline{EF}$ in the range of 0.2–0.8, almost the entire envelope of normTFS values in 50 bootstrap samples is positive, suggesting that all but the driest regions consistently exhibit a slight positive triggering feedback between daytime $\overline{EF}$ and subsequent convective rainfall. The EF-dependence is qualitatively consistent with studies based on idealized coupled boundary layer-free troposphere convecting column models. These results further indicate that high-\$\overline{EF}\$ regimes, typically areas with more dense vegetation and wetter soils, the land surface and atmosphere are tightly coupled. Consequently, wet initial surface conditions are more likely to persist by engendering a greater likelihood of afternoon convection. The results for normAFS demonstrate that the triggering effect is far stronger than the amplification effect at all locations (Figs 3b and Fig. 4b). The functional relationships depicted in Fig. 4 represent powerful diagnostics for model assessment and process-based understanding of surface energy flux partitioning controls on precipitation.

The TFS and AFS metrics estimated from the NARR dataset lead to two important conclusions; TFS indicates substantial local turbulent surface flux partitioning control on the frequency of afternoon convection in areas not strictly limited by surface moisture, whereas AFS indicates a small (<1 mm-scale) impact on convective rainfall amounts. The distinction between surface flux impacts on rainfall frequency and intensity is relevant to model development. Models often simulate reasonable rainfall climatologies despite incorrect frequency and intensity statistics: many exhibit persistent low-intensity events rather than more intermittent, higher intensity ones. Improper model representation of land–atmosphere coupling may account for some of these deficiencies over land; the metrics introduced here provide a tool for validating modelled land–atmosphere coupling. This is particularly valuable for improving forecasts of hydroclimatic extremes, comparing land–atmosphere interactions across models, and refining model representations of coupled land–atmosphere processes. Such process-based knowledge is critical for improving future climate prediction capability, for which stationary statistical assumptions based on past or current climate are questionable and areas of strong land–atmosphere coupling are likely to change.

**Methods**

For each grid point, 2,300 summertime days are available for analysis (25 years, 92 days in June–July–August, JJA). For each day, the three-hourly data are locally positioned to determine the data points closest to local 3 am, 6 am...9 pm, and midnight. Each data point contains accumulated rainfall depths over the 3-h period, or average values over the 3-h period for other variables. Early morning atmospheric conditions are assessed through two quantities used in previous work: the convective triggering potential (CTP) and the low-level humidity deficit ($H_{\text{lw}}$). The CTP is a measure of the energy available for convection in the area of the atmosphere 100–300 hPa above the land surface, which is the pressure threshold often used to delineate the daytime boundary layer. $H_{\text{lw}}$ is defined as the sum of the dew-point depressions 50 and 150 hPa above the land surface. CTP and $H_{\text{lw}}$ are determined from the 6 am observation, capturing the state of the low-level atmosphere before sunrise (from 3 to 6 am).

The energy partitioning at the land surface is assessed through the use of the evaporative fraction (EF), defined as the ratio of latent heat flux (evapotranspiration), $l_E$, to sensible ($H$) and latent heat fluxes at the surface: $\text{EF} = l_E/(H + l_E)$. We use EF because it has been shown to be relatively constant during daytime hours, and is little affected by turbulence variability. EF values are calculated for the noontime observation (9 am–noon). Afternoon rainfall is defined over the 6-h period following the noontime EF observation.

The triggering feedback strength (TFS) is a measure of how the probability of afternoon rainfall, $P(r)$, where $r$ is afternoon rain, changes with EF. Rainfall is triggered when afternoon rainfall exceeds a small threshold value, currently set to 1 mm. The results show little qualitative sensitivity to a doubling or halving of this threshold: the location and the relative strength of the high TFS signal does not change, although the quantitative strength of the TFS is slightly impacted by these variations. The amplification feedback strength (AFS) seems to be noisier at lower threshold values.

We treat EF, $H_{\text{lw}}$, and CTP as discrete random variables, with the parameter space of these variables divided into discrete bins. The CTP and $H_{\text{lw}}$ thresholds...
for these bins are pre-defined so that results can be interpreted in the context of the CTP–HI_{HMO} framework. The EF bins, however, are specific to each grid point: they are determined by splitting the observed range of EF data into 10 bins with an equal number of data points in each bin. For each CTP–HI_{HMO} pair, the probability of afternoon rainfall for each EF bin is expressed as:

$$\Gamma(r|\Delta x, \Delta y, EF) \equiv \Gamma(r|\Delta x, \Delta y, EF) \leq (EF < \epsilon + \Delta \epsilon)$$ (2)

We use the more concise notation format on the left of the above equation in subsequent equations. We take advantage of properties of conditional probabilities to calculate the dependency of afternoon rainfall on EF, considering all CTP–HI_{HMO} pairs within each EF bin:

$$\frac{\partial \Gamma(r)}{\partial EF} = \sum_{\Delta x \leq \Delta x, \Delta y \leq \Delta y} \Gamma(r|\Delta x, \Delta y, EF) \times \left[ \Gamma(x,y|\Delta x, \Delta y, EF) \frac{\partial EF}{\partial EF} \right]$$

$$= \sum_{\Delta x \leq \Delta x, \Delta y \leq \Delta y} \left[ \Gamma(x,y|\Delta x, \Delta y, EF) \frac{\partial EF}{\partial EF} \right] + \sum_{\Delta x \leq \Delta x, \Delta y \leq \Delta y} \left[ \Gamma(x,y|\Delta x, \Delta y, EF) \frac{\partial EF}{\partial EF} \right]$$

$$= \sum_{\Delta x \leq \Delta x, \Delta y \leq \Delta y} \left[ \Gamma(x,y|\Delta x, \Delta y, EF) \frac{\partial EF}{\partial EF} \right] + \sum_{\Delta x \leq \Delta x, \Delta y \leq \Delta y} \left[ \Gamma(x,y|\Delta x, \Delta y, EF) \frac{\partial EF}{\partial EF} \right]$$

This allows us to define the amplification feedback strength (AFS) in an analogous manner to the TFS:

$$AFS = \sigma_{EF} \frac{\partial E[r]}{\partial EF}$$ (5)

where $E[r]$ is the expected value of afternoon rainfall amount. The data used in the AFS calculation are further limited to include only days when afternoon rainfall does occur.

Normalized versions of the TFS and AFS are given by scaling the computed derivatives by the ratio of the relevant mean values:

$$\text{normTFS} = \frac{E[r]}{E[r]} \frac{\partial E[r]}{\partial EF} = \frac{E[r]}{\sigma_{EF}}$$ TFS (6)

$$\text{normAFS} = \frac{E[r]}{E[r]} \frac{\partial E[r]}{\partial EF} = \frac{E[r]}{\sigma_{EF}}$$ AFS (7)

References


**Acknowledgements**

We thank R. Stouffer, S. Malyshev, F. D’Andrea, S. Seneviratne and A. Berg for providing comments on the manuscript. We thank A. Wittenberg for help with the determination of ENSO years. P.G. and B.R.L. are supported by an NSF grant and B.R.L. is also supported by a NOAA grant and an NJAES Hatch grant. National Centers for Environmental Prediction (NCEP) North American Regional Reanalysis (NARR) data provided by the National Climatic Data Center from their website: http://nomads.ncdc.noaa.gov/data.php.

**Author contributions**

K.L.F. performed the analysis of the data and wrote the paper. K.L.F., P.G. and B.R.L. jointly designed and refined the study. P.G. and B.R.L. contributed equally to the work. C.K. converted the original data to netCDF format. All authors discussed the results and commented on the manuscript.

**Additional information**

The authors declare no competing financial interests. Supplementary information accompanies this paper on www.nature.com/naturegeoscience. Reprints and permissions information is available online at http://www.nature.com/reprints. Correspondence and requests for materials should be addressed to K.L.F.