

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Papers in Natural Resources

Natural Resources, School of

10-15-2019

Robust observations of land-to-atmosphere feedbacks using the information flows of FLUXNET

Tobias Gerken

Benjamin L. Ruddell

Rong Yu

Paul C. Stoy

Darren T. Drewry

Follow this and additional works at: <https://digitalcommons.unl.edu/natrespapers>



Part of the [Natural Resources and Conservation Commons](#), [Natural Resources Management and Policy Commons](#), and the [Other Environmental Sciences Commons](#)

This Article is brought to you for free and open access by the Natural Resources, School of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Papers in Natural Resources by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

ARTICLE OPEN

Robust observations of land-to-atmosphere feedbacks using the information flows of FLUXNET

Tobias Gerken^{1,2,6*}, Benjamin L. Ruddell², Rong Yu³, Paul C. Stoy^{1,7} and Darren T. Drewry^{4,5}

Feedbacks between atmospheric processes like precipitation and land surface fluxes including evapotranspiration are difficult to observe, but critical for understanding the role of the land surface in the Earth System. To quantify global surface-atmosphere feedbacks we use results of a process network (PN) applied to 251 eddy covariance sites from the LaThuile database to train a neural network across the global terrestrial surface. There is a strong land–atmosphere coupling between latent (LE) and sensible heat flux (H) and precipitation (P) during summer months in temperate regions, and between H and P during winter, whereas tropical rainforests show little coupling seasonality. Savanna, shrubland, and other semi-arid ecosystems exhibit strong responses in their coupling behavior based on water availability. Feedback couplings from surface fluxes to P peaks at aridity ($P/\text{potential evapotranspiration } ET_p$) values near unity, whereas coupling with respect to clouds, inferred from reduced global radiation, increases as P/ET_p approaches zero. Spatial patterns in feedback coupling strength are related to climatic zone and biome type. Information flow statistics highlight hotspots of (1) persistent land–atmosphere coupling in sub-Saharan Africa, (2) boreal summer coupling in the central and southwestern US, Brazil, and the Congo basin and (3) in the southern Andes, South Africa and Australia during austral summer. Our data-driven approach to quantifying land atmosphere coupling strength that leverages the global FLUXNET database and information flow statistics provides a basis for verification of feedback interactions in general circulation models and for predicting locations where land cover change will feedback to climate or weather.

npj Climate and Atmospheric Science (2019)2:37

; <https://doi.org/10.1038/s41612-019-0094-4>

INTRODUCTION

The terrestrial land surface and atmosphere are coupled through a complex set of interactions and feedbacks that determine the fluxes of mass and energy between the two systems. Weather and climate are well known to determine the productivity of terrestrial ecosystems, but the functioning of the land surface can likewise modify weather and climate patterns.^{1–4} In some cases, the influence of one of these systems can propagate through the other to influence itself, establishing a *feedback*. An understanding of land–atmosphere feedbacks is essential for determining the regional impacts of climate variability and change on the ecosystem services humanity has come to depend, but remains a major challenge as analytical tools to quantify feedbacks have only recently been developed.^{4–10} Feedback processes in nature are difficult to directly observe and to infer, as cause and effect relationships may become obscured or break down when a process influences itself through an intermediary,¹¹ as is often the case in the Earth System. Feedback processes amplify or buffer inputs, resulting in exaggerated or muted responses to perturbations, the latter of which can be difficult to identify. The most severe uncertainties in our climate models are believed to feature feedback.^{12–14} These challenges necessitate the development and application of novel methods to quantify feedback processes in the Earth System. This study presents direct observations of global land to atmosphere information flow through the use of a global network of surface energy flux and meteorological observations,

introducing a statistical approach to characterize temporal and spatial variability in land–atmosphere coupling strength.

Energy exchange between the land surface and atmosphere provides a primary method of interaction — and thereby feedback — between the two systems. Downwelling solar energy absorbed by the land surface warms the soil and vegetation and drives fluxes of sensible and latent heat between the land surface and atmosphere. These energy fluxes modify the composition of the atmospheric boundary-layer (ABL) and drive convective processes that deepen the ABL and result in entrainment of air from the free troposphere.^{15–19} These changes then impact near-surface temperature and humidity as well as precipitation processes, resulting in potential feedbacks through ecosystem physiological response to favor subsequent latent or sensible heat fluxes that in turn impact ABL processes.²⁰ Such feedback processes may intensify with future climate changes,²¹ with the potential to impact critical functions such as water availability²² and ecosystem resilience,²³ and to intensify phenomena such as heat waves,^{24–26} drought,^{27–29} and local convective precipitation.^{30,31} Likewise, the impact of canopy photosynthesis and evapotranspiration on cloud development — and the potential for future climate change to further these effects^{32–34} — make understanding land–atmosphere feedback processes central to predictions of the future availability of ecosystem services.

These examples illustrate the complexity of the coupled land–atmosphere system. Disentangling the presence and strength of positive and negative feedbacks is an ongoing

¹Montana State University, Department of Land Resources and Environmental Sciences, 334 Leon Johnson Hall, Bozeman, MT 59717, USA. ²Northern Arizona University, School of Informatics, Computing, and Cyber Systems, 1295 N Knoles Dr., Flagstaff, AZ 86011, USA. ³University of Nebraska-Lincoln, School of Natural Resources, 101 Hardin Hall, Lincoln, NE 68583, USA. ⁴NASA Jet Propulsion Laboratory, Carbon Cycle and Ecosystems Group, 4800 Oak Grove Drive, Pasadena, CA 91109, USA. ⁵The Ohio State University, Department of Food, Agricultural and Biological Engineering, 590 Woody Hayes Drive, Columbus, OH 43210, USA; ⁶Present address: The Pennsylvania State University, Department of Meteorology and Atmospheric Science, 503 Walker Building, University Park, PA 16802, USA. ⁷Present address: University of Wisconsin–Madison, Department of Biological Systems Engineering, 460 Henry Mall, Madison, WI 53706, USA. *email: tobias.gerken@montana.edu

challenge in understanding how ecosystems and their management impact Earth System processes. Feedback is inherently nonlinear, and its study therefore calls for methods free from assumptions of linear proportionality, simple correlation, or isolated causes and effects.⁴ Likewise, and despite the sophistication of global climate models, our current model-based assessments almost certainly miss key feedback processes or scales of interaction, because they represent hypotheses about poorly understood processes, and not real-world observations of those processes. To robustly measure feedback and critique these hypotheses, we therefore require a method for direct, in situ, and relatively assumption-free (nonlinear and empirical) observation of directional functional couplings.

Process Networks (PNs) characterize the state of a system as a pattern of flows of mass, energy, and/or information that correspond to key system functions.¹¹ Information flow statistics are a robust and mature method for delineating PNs, and have been previously applied to the direct and explicit measurement of feedback between the land surface and atmosphere using flux tower observations.^{35–39} PNs have been shown to accurately diagnose interactions between turbulent fluxes and the atmosphere in ecohydrological systems,^{35–37,40} and have accurately described functional differences between starkly diverse land surface ecosystems at continental scales.^{35,41,42} This paper's choice of Transfer Entropy to delineate PNs⁴³ is ideal to measure directional, scale-specific, and nonlinear couplings that characterize land-to-atmosphere feedbacks.

To investigate feedback between land and atmosphere, we focus on relationships among land surface turbulent energy fluxes of sensible (H) and latent heat (LE) and three atmospheric variables: downward global shortwave radiation (R_g) as an indicator of cloud cover, air temperature (T_a), and precipitation (P). Analysis of these terms provides a core set of surface flux and atmospheric variables that link land and atmosphere through turbulent flux exchange. The strength of the process coupling is quantified through the information flow as given by the normalized transfer entropy (T'). $T'(X \rightarrow Y, \tau)$ is a measure of the predictability of the time series Y from time series X at a time lag of τ . While characteristic τ for significant $T'(X \rightarrow Y, \tau)$ depend on the variable pair, sub-daily timescales are believed to be the primary timescales for flux-based land–atmosphere feedbacks and the average T' for τ from 0.5 to 18 h (T'_{Avg}) is used here to capture the functional relationships for land–atmosphere interactions.

Statistically significant values of information flow and feedback ($p < 0.05$) are established using the method of shuffled surrogates where surrogate T' values are calculated using randomly shuffled time series of X_t and Y_t to remove any correlation between variables. These surrogates are then compared to the observed T' .^{11,44} The fraction of instances during which significant process coupling is observed at any τ (F_{Sig}) is calculated as

$$F_{Sig}(X \rightarrow Y) = \frac{N_{Sig}(X \rightarrow Y)}{N_{Tot}(X \rightarrow Y)}, \quad (1)$$

where N_{Sig} represents the total number of observations during which $T'(X \rightarrow Y, \tau)$ is significant at any τ while N_{Tot} is the total number of observations taken into account. When F_{Sig} approaches 1, a coupling process is robustly significant, but when it approaches zero, the process is weak or absent.

The recent development of regional networks of co-located meteorological and carbon dioxide, water, and energy flux measurements provides a new opportunity to assess land–atmosphere coupling across terrestrial biomes and climate space globally. Here we leverage the LaThuile FLUXNET database,⁴⁵ which provides globally distributed, to a degree standardized observations of land–atmosphere fluxes of energy and water using the eddy covariance technique spanning 251 sites (Supplementary Fig. 1) and representing 11 major IGBP

(International Geosphere-Biosphere Programme) vegetation classes (Supplementary Table 1) across a large range of aridities, and over 10000 site months. We use these data to calculate information flows from land surface fluxes to atmosphere (i.e. coupling strength) and then train an artificial neural network (ANN) of land–atmosphere coupling strength across the terrestrial surface. The use of observational “big data” represents a unique approach to characterizing temporal and spatial variability in land–atmosphere coupling and feedback without a priori assumptions about underlying processes that is capable of directly observing and resolving critical processes at the interface between surface, vegetation, and convective ABL. Our approach therefore complements large-scale climate models and reanalysis products, for which these processes and feedbacks currently remain parameterized due to their comparatively coarse resolution. Given that our information flow PN methodology quantifies the presence, strength, direction, and significance of land-to-atmosphere coupling using a large sample of in situ flux tower observations, it can be used as an independent control for existing models and theory, in addition to providing unique insights into these difficult-to-observe processes. In this work, we focus on the land-to-atmosphere portion of land–atmosphere coupling, by investigating the directional information flow from H and LE to future states of atmospheric variables.

Several studies have identified global land–atmosphere coupling strength and associated “coupling hotspots” using climate models^{5,21,46} and reanalysis data,^{47,48} primarily focused on soil moisture–precipitation feedbacks. Here we examine these hypothesized feedback hotspots in the context of an empirical data-driven analysis, broadening coupling mechanisms to the specific surface fluxes (latent and sensible heat) that are directly measured through FLUXNET, and which are directly responsible for convection and changes in the near-surface atmosphere that impact ecosystem function. Given that PNs are a methodological distinct tool for the analysis of environmental data, they can serve as a validation tool for process-based models and more conventional observational analysis.

RESULTS

There is pronounced seasonality in the magnitude of land-to-atmosphere coupling (Figs 1, 2) and seasonal patterns differ between the six coupling pairs considered in this work. Pairs (LE $\rightarrow T_a$), (LE $\rightarrow P$), and ($H \rightarrow T_a$) exhibit low F_{Sig} during winter months in temperate regions compared to high coupling strength during summer, whereas the seasonality is reversed for (LE $\rightarrow R_g$) and ($H \rightarrow R_g$). The coupling process ($H \rightarrow P$) shows significant coupling during both winter and summer. Tropical regions exhibit little seasonality, except for ($H \rightarrow R_g$) and (LE $\rightarrow R_g$). These are much stronger during June through August, which also broadly coincides with the dry season in Amazonia. As expected, the increase in feedback coupling between winter and summer within temperate regions also coincides with a strong increase in vegetation density and greenness, represented here as the Normalized Difference Vegetation Index (NDVI).

Spatial patterns in feedback coupling strength are related to IGBP biome type (Fig. 3). The lack of a clear annual cycle for ($H \rightarrow P$), visible in Fig. 2, stems mainly from forest ecosystems (mixed forest, MF; deciduous broad leaf forest, DBF; evergreen needle leaf forest, ENF; evergreen broad leaf forest, EBF), which are abundant in FLUXNET, and closed shrubland (Fig. 3, see Supplementary Table 1). EBF, which encompasses tropical rainforests, exhibits strong coupling with little seasonality except for (LE/ $H \rightarrow R_g$), where it follows the general trend of low coupling during boreal summer. Savanna and shrubland type systems (savanna, SAV; woody savanna, WSA, open shrubland, OSH; closed shrubland, CSH) also show pronounced land–atmosphere feedback dynamics that deviate from seasonal cycles found in other biomes. Strongly

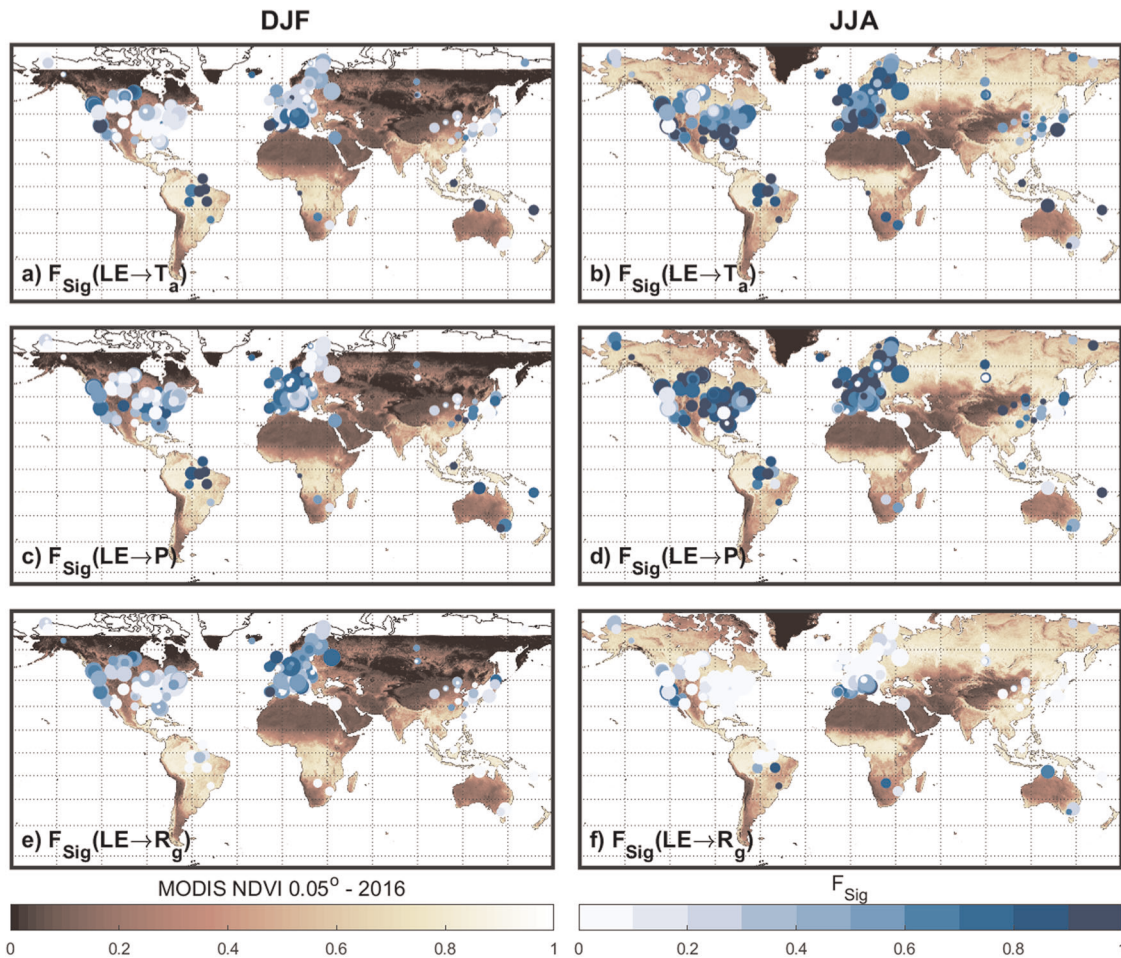


Fig. 1 The fraction of significant surface to atmosphere interaction as determined by a process network (PN) with respect to latent heat flux (LE), F_{Sig} (blue colorscale), for 251 sites of the FLUXNET LaThuile dataset and calculated as average across DJF (northern hemisphere winter) (a, c, e) and JJA (northern hemisphere summer) (b, d, f). Larger circles indicate more years of tower data considered for calculation of F_{Sig} . The background image is DJF and JJA NDVI from MODIS for 2016 (dark colorscale). When F_{Sig} approaches 1, a significant coupling is present, and when it approaches 0 the coupling is weak or absent

increased ($\text{LE} \rightarrow R_g$) and reduced ($\text{LE} \rightarrow P$) during summer, and strong ($\text{LE} \rightarrow T_a$) throughout the year for these generally semi-arid to arid ecosystems, highlight the importance of interactions between biome and prevailing climate in governing land to atmosphere coupling behavior. The amplitude in coupling strength for annual cycles approaches 0.8–1.0 for all couplings except ($\text{LE} \rightarrow R_g$) for which the amplitude tends to be less than 0.5.

There is a great deal of variability in land-to-atmosphere feedback coupling strength between biomes, climate zones, and seasons. To aid interpretation, we plot the feedback coupling strengths described above against monthly values of T_a and monthly aridity (P/ET_p) (Figs 4, 5). There is an increasing trend of increasing feedback strength (T'_{Avg}) with increasing monthly T_a across all biomes. Land–atmosphere coupling is largely absent at $T_a < 0^\circ\text{C}$, which can be expected given that energy inputs and ecosystem activity are generally minimized under these conditions. Savanna and shrubland type ecosystems exhibit the lowest coupling originating from LE at high monthly temperatures ($T_a > 20^\circ\text{C}$), whereas the situation is reversed with high coupling originating from H when $T_a > 20^\circ\text{C}$.

The behavior of T'_{Avg} is more complex with respect to P/ET_p . There is little relationship between $T'_{\text{Avg}}(\text{LE} \rightarrow T_a)$ and aridity, but a clear relationship emerges for $T'_{\text{Avg}}(\text{H} \rightarrow T_a)$, in which feedback originating from H increases with aridity for all vegetation types. The feedback coupling from surface fluxes to P peaks at P/ET_p values near unity and WSA, OSH, and CSH exhibit the strongest

feedbacks of all vegetation types in that range of P/ET_p . For the coupling between surface fluxes and cloud cover as indicated by R_g , we find that there is little feedback for $P/ET_p > 1$. Savanna (SAV and WSA) and shrub (CSH and OSH) vegetation classes generally exhibit the highest feedback for T_a and R_g for low P/ET_p . While we chose to present T'_{Avg} as the coupling metric in this work, there is considerable variation in coupling timescales between variable pairs (Supplementary Figs 4–6), which in itself shows dependence on T and P/ET_p and may be related to the timescales needed to effectively connect land-surface and atmospheric processes. For example, the dominant coupling timescales in the order of 6–12 h between surface fluxes and P or R_g show substantial time-lags in the atmosphere's response to surface fluxes, which are consistent with timescales typically found in convective boundary layers.

The extrapolation of observed feedback strength (T'_{Avg}) from FLUXNET sites to the global map reveals several hotspots of land–atmosphere coupling that stand out from global average feedback strength (Fig. 6; see Supplementary Figs 7–12 for monthly data). The ANN models had R^2 values of 0.69 to 0.92 for ($\text{H} \rightarrow P$) and ($\text{H} \rightarrow R_g$), respectively (Supplementary Table 2), with no evidence of overfitting, so the model's extrapolation is robust when tested against the 251 FLUXNET sites and over 10,000 observed site-months. We find strong land-to-atmosphere feedback in sub-Saharan Africa ($\text{LE} \rightarrow T_a$ and $\text{LE} \rightarrow P$), the central and southwestern US during summer ($\text{H} \rightarrow R_g$ and $\text{H} \rightarrow T_a$), the southern Andes, South Africa and Australia during DJF ($\text{H} \rightarrow R_g$),

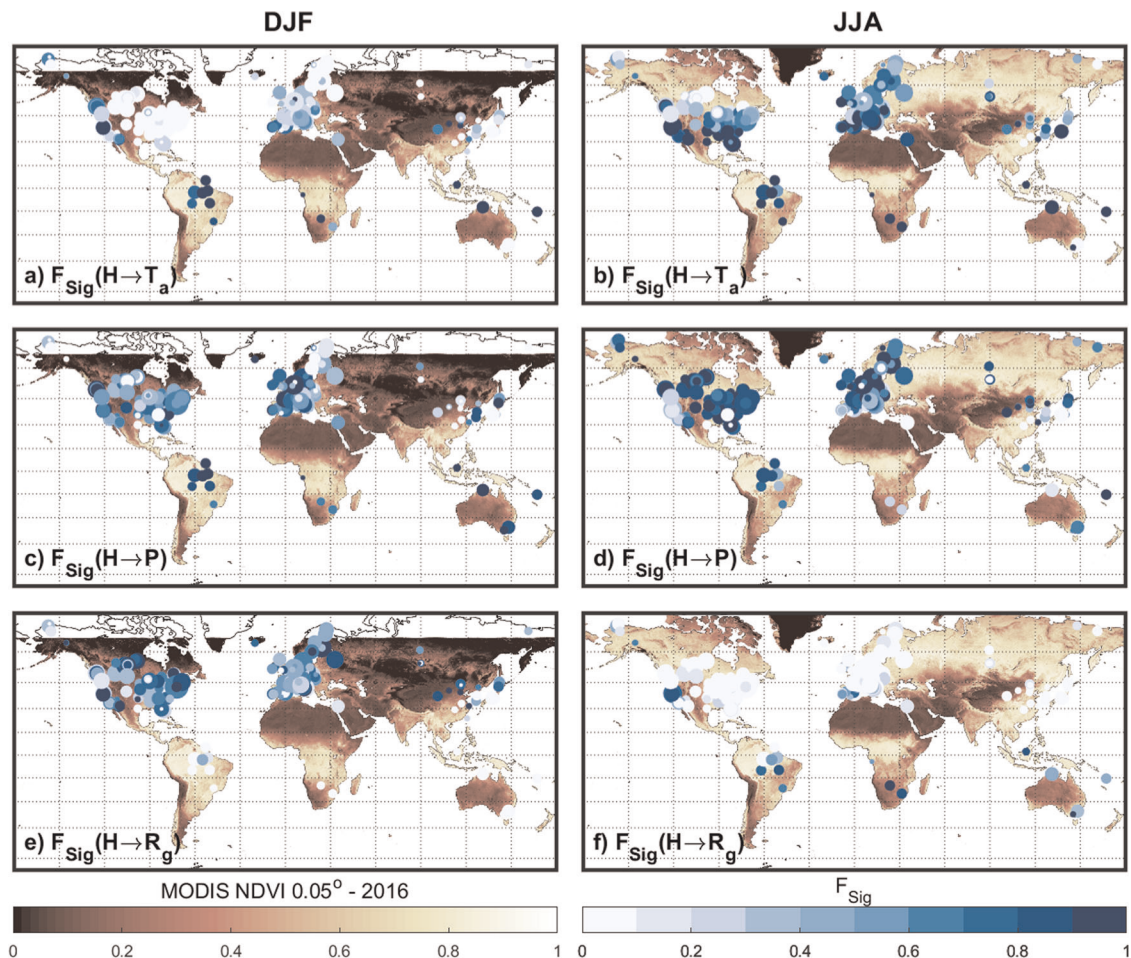


Fig. 2 Same as Fig. 1, but for sensible heat flux (H)

Amazonia ($LE \rightarrow P$ and $LE \rightarrow R_g$), agricultural areas in eastern Brazil ($H \rightarrow T_a$ and $H \rightarrow R_g$), the African Rift Valley ($LE \rightarrow T_a$ and $H \rightarrow R_g$) as well as the Congo, where strong coupling persists throughout the year, but switches from ($LE \rightarrow R_g$) in DJF to ($H \rightarrow R_g$ and $LE \rightarrow T_a$) in JJA. This plot is thematically similar to the soil moisture based results from Koster et al.⁵

DISCUSSION

PNs and other empirical methods based on information theory applied to environmental “big data” provide a wealth of information about land–atmosphere coupling. Specifically, PNs provide information about functional relationships between ecosystem variables that can be used to investigate processes such as land–atmosphere coupling and feedbacks as well as their response to environmental change. Using an ANN to extrapolate these couplings to the global scale, we identified several hotspots of land–atmosphere coupling (Fig. 6). Monthly data are presented in Supplementary Figs 7–12. Unlike previous studies e.g. ^{5,46} which used process-based models, the ANN is based on empirical extrapolation of observations and does not include a priori assumptions about functional relationships to demonstrate the existence of feedbacks. It can therefore be used to complement global models, which require (i) process relationships to be known and (ii) may require parameterizations to include processes that are under-resolved due to their global nature.

We investigated six couplings between turbulent fluxes and atmospheric/near surface properties by taking advantage of databases that incorporate observations of a wide range of

surface meteorology and fluxes. Couplings of H and LE to P and R_g are directly related to the hydrologic cycle, in contrast to the coupling with temperature, which is more related to near surface conditions and cover type. The ANN trained on PN results identifies feedback hotspots in the southwestern and central US similar to⁵ but does not reproduce the hotspot on the Indian subcontinent. However, for the southern African hotspot we find that the coupling signal is strongest for H , LE , and T_a rather than precipitation, and is more pronounced in DJF. For the US hotspot, we find a stronger signal for H , LE , and R_g rather than P . The ANN also detects the hotspots in the Congo Basin, South Africa, Australia and to some extent Brazil (for H to R_g and T_a), in agreement with Notaro and Zeng et al.^{46,48} Similarly, several regional studies highlighted the strong coupling between surface and air temperatures for semi-arid regions in the US and Europe,^{6,49} which is reflected in the PN results for the southwest US and to some extent for the Iberian peninsula. Compared to previous studies, we find a stronger coupling of LE to R_g and P in Amazonia, further highlighting the importance of tropical rain-forest function for cloud development and regional precipitation.⁵⁰ We find R_g to exhibit much clearer land to atmosphere coupling than P , which can be expected given that not all clouds produce precipitation. The reduced coupling could also indicate that models are overly sensitive with respect to their precipitation response or that the PN has problems detecting feedback in P due to the sparseness of precipitation events. While the latter cannot be excluded, global and even regional models rely on cumulus parameterizations for precipitation generation, which have well-known difficulties in producing realistic precipitation.^{51,52}

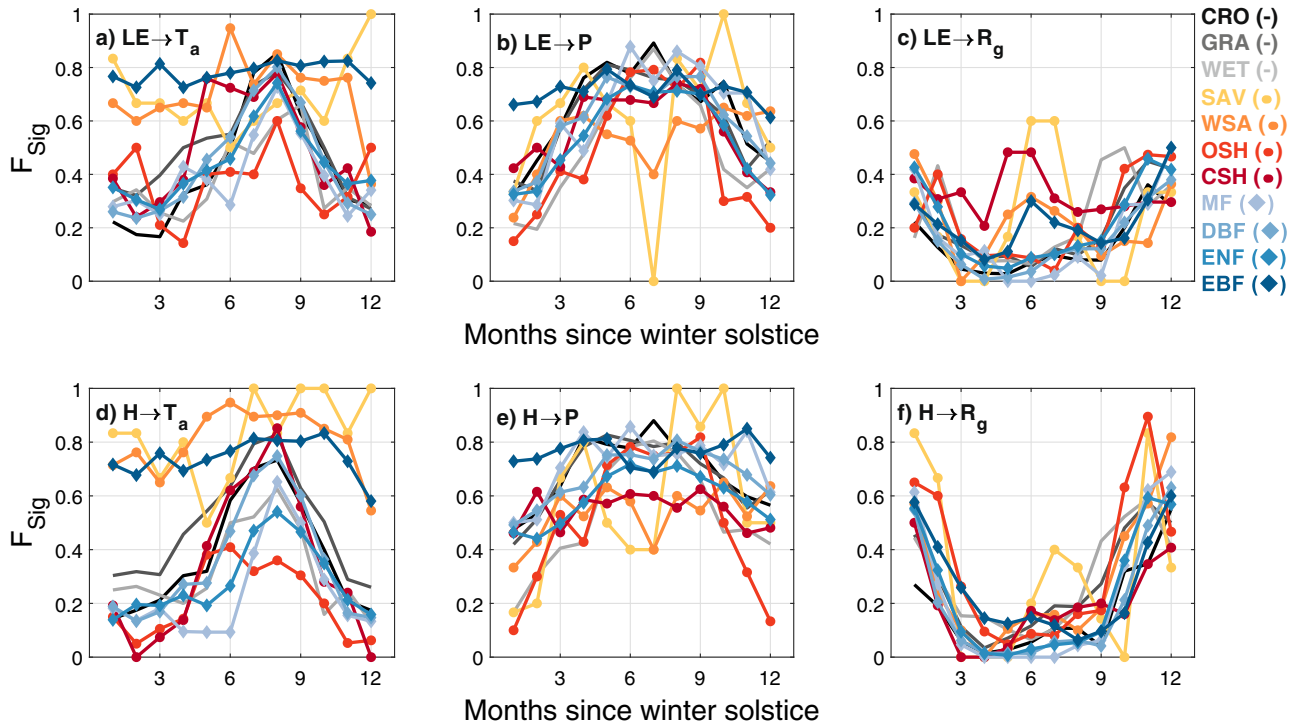


Fig. 3 The annual cycle of significant feedback between land and atmosphere (F_{Sig}) for LE (a–c) and H (d–f), separated by land cover type. The time axis is given in months since winter solstice to align northern and southern hemisphere. When F_{Sig} approaches 1, a significant coupling is present, and when it approaches 0 the coupling is weak or absent

Extrapolation of empirical PN results to the global scale shows two distinct advantages compared to global scale modeling approaches. As a statistical method, global results at a high resolution (e.g. 0.25°) are computationally cheaper than running an Earth System Model, while also providing detailed information on land–atmosphere coupling on spatial and seasonal scales. Also, through considering multiple land–atmosphere feedback pathways, PNs are capable of providing information that can be used to improve process-level understanding of feedbacks not accessible in more complex models. At the same time, data-driven approaches such as PNs and ANNs are not constrained by physically realistic limitations and cannot prove cause–effect relationships. This should not be considered a limitation but as a feature. Combined with domain expertise, data-driven methods can be very useful in guiding research toward regions and processes that merit further scientific attention.

The PN and ANN reveal that dryland ecosystems exhibited the strongest ecosystem–atmosphere feedback due to variability in available water (Figs 3–6). We find the highest couplings between surface fluxes and precipitation at $P/ET_p \sim 1$, highlighting the importance of sufficient water supply and soil moisture in controlling land–atmosphere interactions.^{53,54} Interestingly, for savannas, high monthly mean temperatures ($T_a > 20^\circ\text{C}$) are associated with low $T_{Avg}(LE \rightarrow P)$, indicating the water limited state of these systems during the dry season and the associated absence of coupling. Similarly transition periods between wet and dry seasons and monsoon circulations are important for soil moisture–precipitation coupling.^{47,49} Vegetation response to water limitations occurs on a continuum from isohydric (plants closely regulate transpiration through stomatal conductance in response to atmospheric vapor pressure deficit) to anisohydric (plants have little regulation of stomatal conductance). From these species-level traits, ecosystem-level drought responses emerge.^{28,55} Grasses, which were thought to be mostly anisohydric, often exhibit isohydric behavior in semi-arid environments,^{56–58} supporting the notion that semi-arid grasslands can exhibit

substantial feedbacks with the atmosphere. The resulting interplay between vegetation, surface-energy flux partitioning and atmospheric control also influences the development of local convection, which can be an important ecosystem moisture source.^{31,59–62} Substantial feedbacks between biosphere and precipitation,⁶³ highlighting the need of an accurate representation of the biosphere’s response to temperature, radiation, and water availability for predicting hydrometeorological and climatological feedbacks.

The strong coupling between turbulent fluxes and P for semi-arid systems (i.e. savannas, Figs 5, 6) is particularly interesting in the light of their pronounced seasonality. Given the fact that the analysis covers monthly system state, and precipitation inputs are highly pulsed, intermediate P/ET_p might correspond to rapidly changing moisture supplies at the surface that elicit responses in the land–atmosphere system. The increase in coupling between LE and H to T_a and R_g for small P/ET_p further highlights the importance of convective processes that impact ABL growth and present multiple avenues for feedbacks mediated by the surface–ABL system.^{15–19}

PNs applied across aridity gradients can be used to better understand potential changes to land–atmosphere interactions and ecosystem functioning across temporal and spatial scales. Given that semi-arid ecosystems are critical to the carbon cycle and climate,^{64–66} and are likely to expand⁶⁷ and deteriorate⁶⁸ under climate change, the ability of PNs to quantify their coupling to the atmosphere is of particular importance. Additionally, projected changes in aridity are expected to exhibit complex changes across the globe,^{69,70} increasing the uncertainty for land–atmosphere interactions and feedbacks.

This study is not without limitations related to data availability and uncertainty. This study relies on near surface observations as a proxy for land–atmosphere coupling rather than direct observations of boundary-layer processes that mediate these couplings and feedbacks due to a lack of continuous and spatially distributed ABL observations, which needs to be addressed by

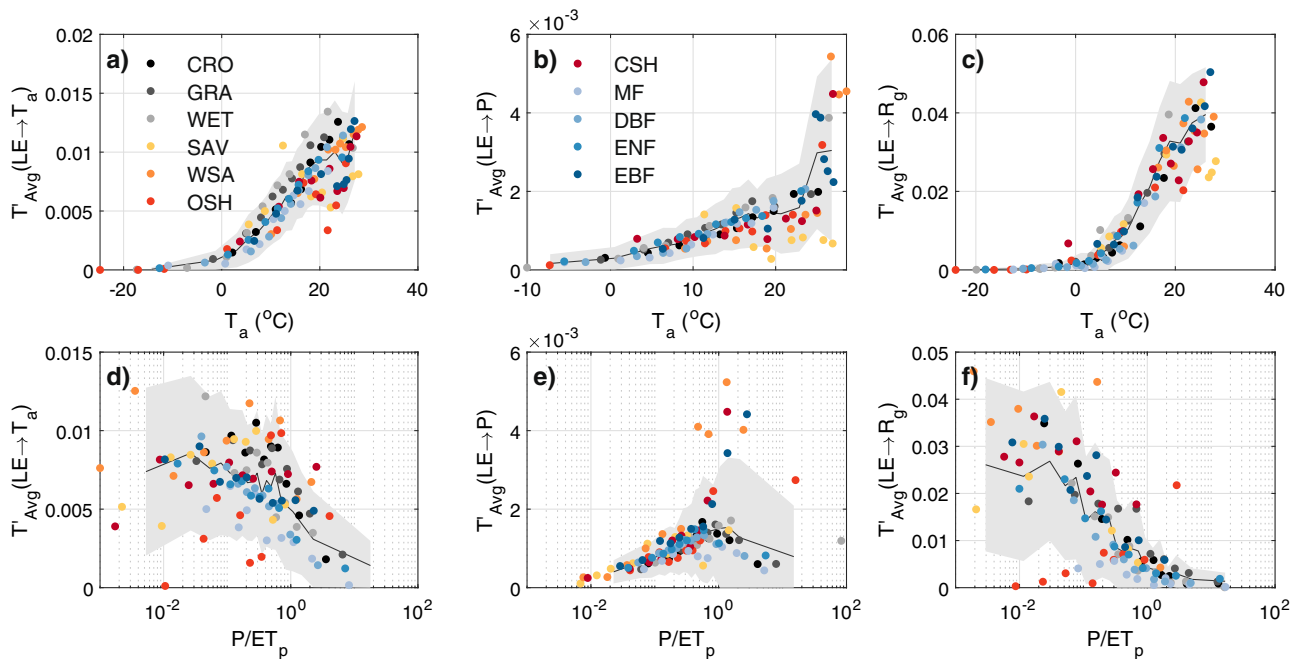


Fig. 4 Average transfer entropy (T_{Avg}) of latent heat (LE) to air temperature (T_a), precipitation (P), and global radiation (R_g) as a function of environmental conditions for temperature (a–c) and aridity index P/ET_p (d–f) (note the logarithmic x-axis) and separated by land cover type. Values are binned into 10 categories with equal number of members for each bin, and then averaged. The black lines and gray shaded areas give the mean and standard deviations across all data. The normalized standard deviation for each bin is presented in Supplementary Fig. 2. Higher transfer entropies indicate stronger land-to-atmosphere information flow

the community.⁴ Also, PNs allow for the detection of coupling relationships irrespective of assumptions of linearity or sign of the relationship. At the same time, this means that PNs do not provide information about exact nature of coupling relationships, which can then be explored with more conventional methods. Similarly, turbulent flux measurements as collected by FLUXNET do not close the surface energy balance^{40,71–74} and it is unclear whether H and LE are similarly affected and to what extent this impacts the results generated by statistical methods applied to this database. Also, FLUXNET does not systematically cover all global biomes, and tends to under-sample remote and harsh environments. The southern hemisphere, northern Africa, and central Asia are particularly under-represented, limiting our ability to assess the systems' responses to global environmental change and implications for surface–atmosphere feedbacks. This has the potential to limit generalizability of our results to the globe as indicated by negative (i.e. non-physical) T_{Avg} values in some remote areas. Similarly, the extrapolation of PN results using an ANN relies on the use of climatological averages. FLUXNET LaThuile contains 251 sites with approximately 1000 site months, which translates to on average 3.5 site years per site and may thus lead to mismatches between climatological states and flux observations, which may result in biases for the extrapolated ANN.

H and LE flux dynamics are closely coupled through the surface energy balance but observed couplings of H and LE diverge (e.g. there is significant coupling between H and T_a , in a given region but not for LE and T_a or vice versa). This behavior of the PN is likely due to the fact that despite their correlation, H and LE are rarely of the same magnitude. The PN is implicitly sensitive to the absolute magnitude and the time rates of change in the time series and thus acts as a low-pass filter on information flow from flux variations (see Supplemental Note 1 for additional details).

Despite its limitations, given the PNs good agreement with previous studies in diagnosing feedback hotspots and its coupling response with respect to T_a and P/ET_p (Figs 4, 5), which is in line with ecohydrological expectations,⁵³ we have confidence that observed information flow from the growing body of

environmental big data — through networks such as FLUXNET — can be used to provide unique insights on land–atmosphere feedbacks from an empirical perspective and can serve as independent empirical verification for process-based climate models, potentially driving progress in the improvement of climate models toward the representation of critical processes for projecting land–atmosphere interactions and feedbacks.

In conclusion, we demonstrate that PN results can be used as independent validation for process-based models based on observed information flows between the land surface and atmosphere. As hypothesized by prior models and research, savanna, shrubland, and other semi-arid ecosystems exhibit a strong response in their atmospheric feedback behavior based on seasonal water availability and aridity. Information flow from surface to atmosphere for other variables exhibited seasonal variability with the exception of tropical rainforests and was a strong function of air temperature. In the light of dryland expansion their vulnerability to climate change, this might strongly impact land–atmosphere coupling including important precipitation processes.

METHODS

Observations

Observed variables were obtained from the FLUXNET LaThuile synthesis dataset, which encompasses data from 251 sites representing nearly 1000 site years at a temporal resolution of 0.5 h. To ensure data quality, we (i) used only original data and gap-filled data of high quality, as indicated by the data-quality flag provided by FLUXNET; (ii) excluded site years with less than 50% available data; (iii) excluded outliers that were identified by exceeding six standard deviations compared to detrended data which had the diurnal cycle removed using a periodic anomaly (except for P); and (iv) excluded site-months with <500 observations (out of ~1400 possible per month). This resulted 10398 site-months being used for this analysis. Monthly T was then calculated across sites that represent 11 major IGBP vegetation classes (Supplementary Table 1). Monthly potential evapotranspiration (ET_p) values, which are calculated from the Penman–Monteith equation using FLUXNET measurements and provided as part of the

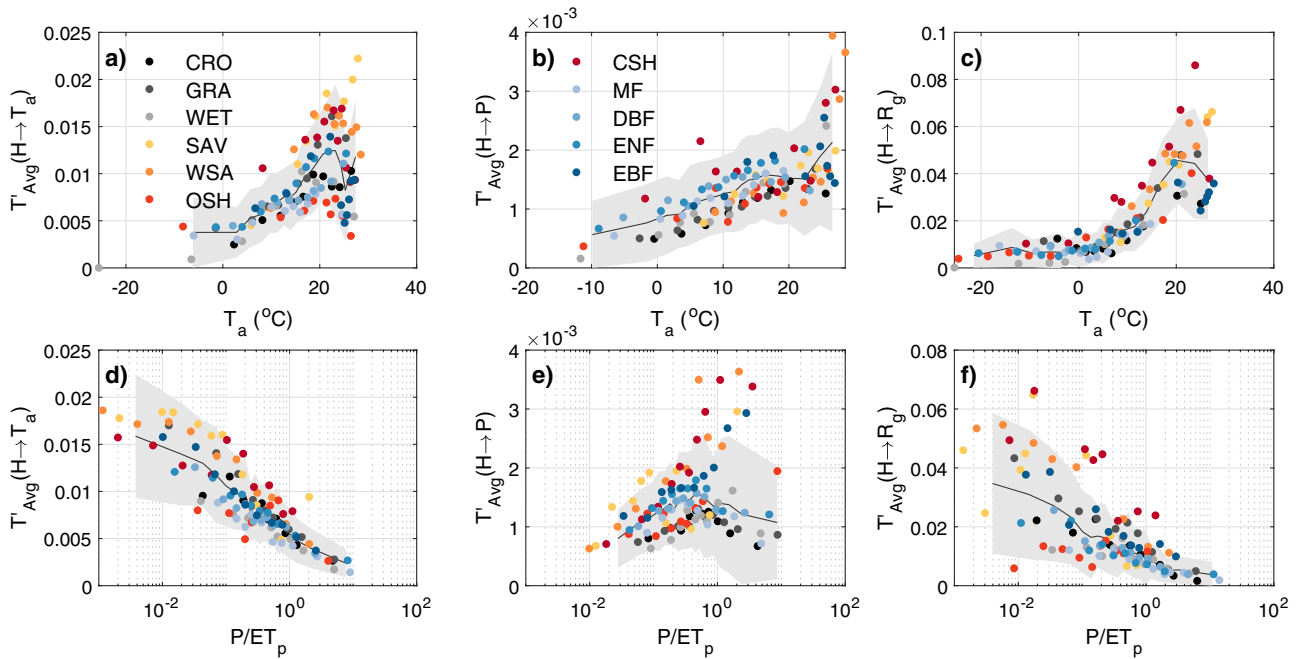


Fig. 5 Same as Figure 5, but for T'_{Avg} of H to T_a , P , and R_g . The normalized standard deviation for each bin is presented in Supplementary Fig. 3

LaThuile dataset, are used to quantify the aridity of the sites through the ratio P/ET_p .

Process network (PN)

Functional relationships between ecosystem and atmospheric variables are calculated using a PN employing the open source package ProcessNetwork version 1.4.⁷⁵

Transfer entropy (T) was calculated as^{11,43}

$$T(X_t \rightarrow Y_t, \tau) = \sum_{y_t, y_{t-1}, x_{t-\tau}} p(y_t, y_{t-1}, x_{t-\tau}) \log \frac{p(y_t | (y_{t-1}, x_{t-\tau}))}{p(y_t | y_{t-1})}, \quad (2)$$

where the predictability of time series Y_t based on knowledge of time series X_t at time lag τ is calculated using y_{t-1} as the immediate history of Y_t and $x_{t-\tau}$ as the history of X_t at τ . The p denotes the corresponding probability density functions. T is bounded between 0 and $\log(m)$, where m is the number of discrete microstates y taken by variable Y_t . We normalize T to a unit-less fraction by division with its upper limit $[\log(m)]$, yielding the normalized transfer entropy (T'). $T'(X \rightarrow Y, \tau)$ was calculated for 0.5 h increments of τ from 0.5 to 18 h and then averaged across all 36 increments to yield T'_{Avg} (Supplementary Figs 4, 5 present additional information on the underlying significant timescales of T'_{Avg}).

In order to achieve a balance between entropy estimation accuracy and limited observations in the numerical estimation of $p(y)$, $m = 20$ was used dividing H and LE into 20 bins (referred to as microstates in information theory nomenclature) of equal width. Note that these microstates do not have a physical significance, but serve as the basis for determining the underlying relationship between X and Y . $T(X \rightarrow Y, \tau)$ measures additional information that is provided by knowledge of X at time-lag τ in addition to information provided by the history of Y itself. It is a statistical index for physically causal and directional coupling (not correlation), albeit with limitations. The reader is also referred to previous works^{11,44} for details on PN calculation methodology.

Artificial neural network (ANN)

To extrapolate results from site level to the entire land surface, artificial neural networks (three-layer feed forward) were trained for each flux coupling. The ANN was trained to extrapolate T'_{Avg} values from FLUXNET using gridded data at 0.25° resolution. ANN training inputs were monthly T_a , R_g , P , ET_p , the enhanced vegetation index (EVI), IGBP class, elevation, and absolute latitude. ANN outputs were monthly T'_{Avg} values for each coupling at 0.25° resolution using the auxiliary datasets described at the end of the methods section.

ANNs have been widely used for climate change and ecosystem research and, given their skill in dealing with noisy and unbalanced datasets,⁷⁶ they are well suited for PN research as they do not require geographically well distributed training sites and are robust to uneven distribution in IGBP classes or climates in the training dataset. To minimize overfitting, which ANNs are sensitive to, we employed Bayesian regularization backpropagation as the training function in a three-layer feed-forward ANN. This improves generalization for small and noisy datasets.⁷⁷ We divided the training dataset randomly into training (70%) and test (30%) datasets and the performance of the ANN was evaluated on the test set using the Pearson's R coefficient (Supplementary Table 2). The ANN analysis was performed with the Neural Network Toolbox in Matlab2014b.

Auxiliary data for ANN extrapolation

In addition to FLUXNET site level data (T_a , R_g , P , and ET_p), the ANN was trained using IGBP (provided by FLUXNET), elevation above sea level, and the EVI. For elevation, we used data provided by FLUXNET, if present. In the absence of site provided data or if only an approximate height was provided, United States Geological Survey (USGS) Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) was substituted using a nearest neighbor approach. EVI from the Moderate Resolution Imaging Spectrometer (MODIS) Monthly L3 Global V006 (Terra: MOD13C2; Aqua: MYD13C2) was assigned to the sites using the following procedure:

1. If both MODIS Terra and Aqua 0.05-degree data was available for the given month and year, their mean was used applying a nearest neighbor approach.
2. If either MODIS Terra or Aqua were available, the available data was used.
3. If neither were present (e.g. for site months before the Terra and Aqua launch dates) the long-term mean for Terra and Aqua for that location and month were used.
4. The value was set to missing, if no acceptable EVI values could be calculated (e.g. snow and ice cover during winter for all years).

Information about the gridded datasets used for the ANN can be found in Supplementary Table 3. The IGBP landcover was assigned using the dominant land cover class (as percent cover within the 0.25° × 0.25° grid cell) from the MODIS MCD12C1 product. Data availability for EVI is shown in Supplementary Fig. 13. Global monthly mean meteorological data (T_a , R_g , P , and ET_p) for ANN extrapolation are obtained from GLDAS (Global Land Surface Data Assimilation System) V2.1 at 0.25° resolution. We use the 30-year climatological average (1981–2010) for ANN extrapolation. On

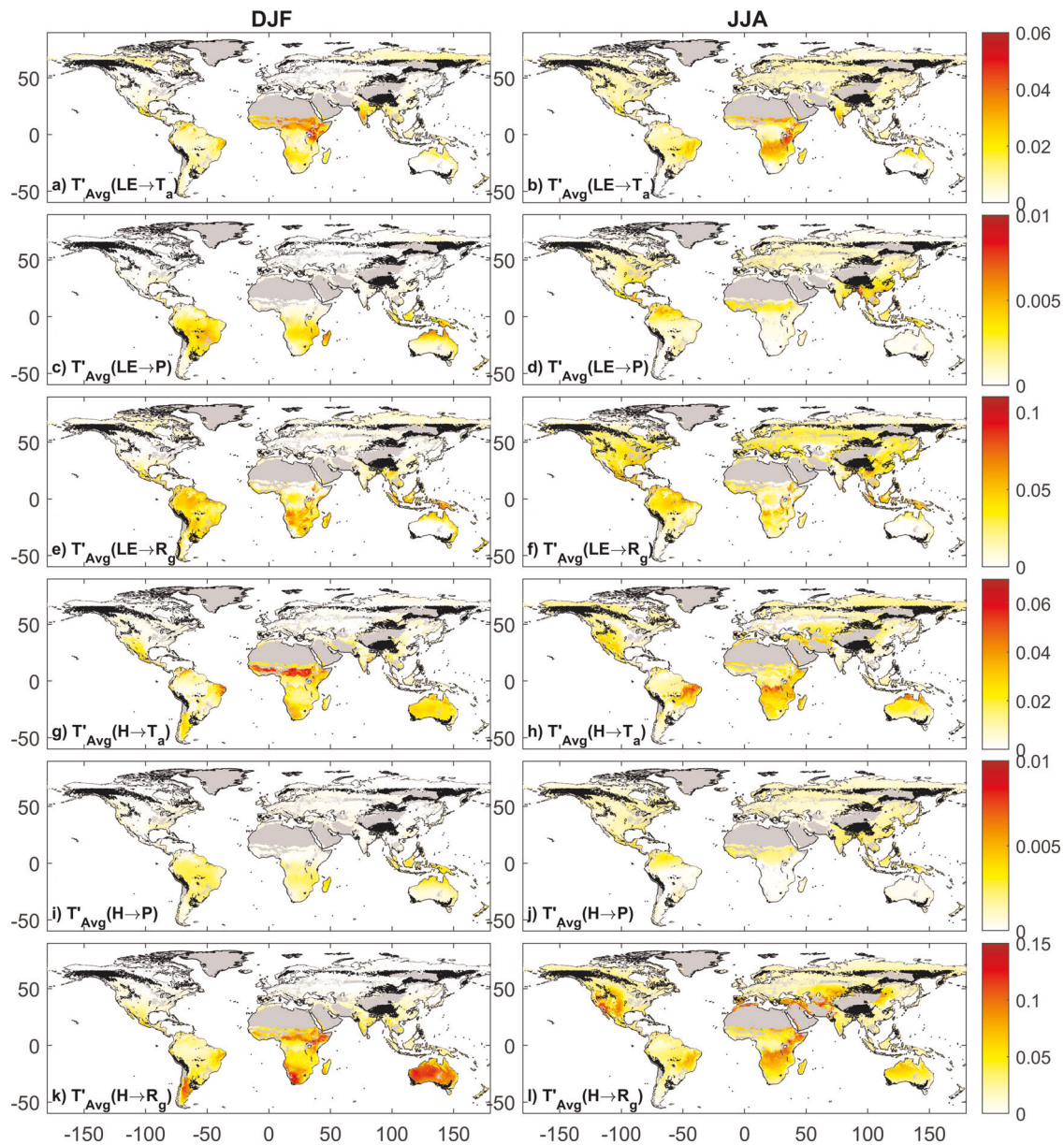


Fig. 6 Average estimated transfer entropy (T'_{Avg} , red colorscale) of latent (LE) and sensible heat (H) to air temperature (T_a), precipitation (P), and global radiation (R_g) for northern hemisphere winter (DJF) (**a**, **c**, **e**, **g**, **j**, **k**) and summer (JJA) (**b**, **d**, **f**, **h**, **j**, **l**) extrapolated from FLUXNET sites to the global map using an artificial neural network (ANN). Gray and black areas indicate lack of input data and negative (non-physical) T'_{Avg} values, respectively. Dark and red areas indicate strong land to atmosphere feedback associated with the indicated land surface flux process

overview about total data availability for training of the ANN is given in Supplementary Fig. 14.

DATA AVAILABILITY

PN and ANN results are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

The Process Network code is published as: *ProcessNetwork/ProcessNetwork_Software* —File Exchange —MATLAB Central Available at: <http://www.mathworks.com/matlabcentral/fileexchange/41515-processnetwork-processnetwork-software>

Received: 15 January 2019; Accepted: 19 September 2019;
Published online: 15 October 2019

REFERENCES

- Baidya Roy, S. & Avissar, R. Impact of land use/land cover change on regional hydrometeorology in Amazonia. *J. Geophys. Res. Atmospheres* **107**, LBA 4-1 (2002).
- Pielke, R. A. et al. The influence of land-use change and landscape dynamics on the climate system: relevance to climate-change policy beyond the radiative effect of greenhouse gases. *Philos. Trans. R. Soc. Lond. Math. Phys. Eng. Sci.* **360**, 1705–1719 (2002).
- Mahmood, R. et al. Land cover changes and their biogeophysical effects on climate. *Int. J. Climatol.* **34**, 929–953 (2014).
- Santanello, J. A. et al. Land–atmosphere interactions: the LoCo perspective. *Bull. Am. Meteorol. Soc.* **99**, 1253–1272 (2017).
- Koster, R. D. et al. Regions of strong coupling between soil moisture and precipitation. *Science* **305**, 1138–1140 (2004).
- Seneviratne, S. I., Lütthi, D., Litschi, M. & Schär, C. Land–atmosphere coupling and climate change in Europe. *Nature* **443**, 205–209 (2006).

7. Findell, K. L., Gentine, P., Lintner, B. R. & Kerr, C. Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation. *Nat. Geosci.* **4**, 434–439 (2011).
8. Berg, A., Findell, K., Lintner, B. R., Gentine, P. & Kerr, C. Precipitation sensitivity to surface heat fluxes over North America in reanalysis and model data. *J. Hydrometeorol.* **14**, 722–743 (2013).
9. Guillod, B. P. et al. Land-surface controls on afternoon precipitation diagnosed from observational data: uncertainties and confounding factors. *Atmos. Chem. Phys.* **14**, 8343–8367 (2014).
10. Knist, S. et al. Land–atmosphere coupling in EURO-CORDEX evaluation experiments. *J. Geophys. Res. Atmospheres* **122**, 2016JD025476 (2017).
11. Ruddell, B. L. & Kumar, P. Ecohydrologic process networks: 1. *Identif. Water Resour. Res.* **45**, W03419 (2009).
12. Hansen, J. et al. in *Geophysical Monograph Series* 29 (eds Hansen, J. E. & Takahashi, T.), 130–163 (American Geophysical Union, Washington, DC, 1984).
13. Cess, R. D. et al. Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *J. Geophys. Res.* **95**, 16601 (1990).
14. Bony, S. et al. How well do we understand and evaluate climate change feedback processes? *J. Clim.* **19**, 3445–3482 (2006).
15. Jacobs, C. M. J. & De Bruin, Ha. R. The sensitivity of regional transpiration to land-surface characteristics: significance of feedback. *J. Clim.* **5**, 683–698 (1992).
16. Santanello, J. A., Friedl, M. A. & Ek, M. B. Convective planetary boundary layer interactions with the land surface at diurnal time scales: diagnostics and feedbacks. *J. Hydrometeorol.* **8**, 1082–1097 (2007).
17. van Heerwaarden, C. C., Vilà-Guerau de Arellano, J., Moene, A. F. & Holtslag, A. A. M. Interactions between dry-air entrainment, surface evaporation and convective boundary-layer development. *Q. J. R. Meteorol. Soc.* **135**, 1277–1291 (2009).
18. van Heerwaarden, C. C., Vilà-Guerau de Arellano, J., Gounou, A., Guichard, F. & Couvreux, F. Understanding the daily cycle of evapotranspiration: a method to quantify the influence of forcings and feedbacks. *J. Hydrometeorol.* **11**, 1405–1422 (2010).
19. Jimenez, P. A., de Arellano, J. V.-G., Navarro, J. & Gonzalez-Rouco, J. F. Understanding land–atmosphere interactions across a range of spatial and temporal scales. *Bull. Am. Meteorol. Soc.* **95**, E514–E517 (2014).
20. Ek, M. B. & Holtslag, Aa. M. Influence of soil moisture on boundary layer cloud development. *J. Hydrometeorol.* **5**, 86–99 (2004).
21. Dirmeyer, P. A. et al. Evidence for enhanced land–atmosphere feedback in a warming climate. *J. Hydrometeorol.* **13**, 981–995 (2012).
22. Berg, A. et al. Land–atmosphere feedbacks amplify aridity increase over land under global warming. *Nat. Clim. Change* **6**, 869–874 (2016).
23. Jung, M. et al. Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature* **467**, 951–954 (2010).
24. Hirschi, M. et al. Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nat. Geosci.* **4**, ngeo1032 (2010).
25. Gonzalez Miralles, D. et al. Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nat. Geosci.* **7**, 345–349 (2014).
26. Donat, M. G., Pitman, A. J. & Seneviratne, S. I. Regional warming of hot extremes accelerated by surface energy fluxes. *Geophys. Res. Lett.* **44**, 2017GL073733 (2017).
27. Roundy, J. K., Ferguson, C. R. & Wood, E. F. Temporal variability of land–atmosphere coupling and its implications for drought over the Southeast United States. *J. Hydrometeorol.* **14**, 622–635 (2012).
28. Novick, K. A. et al. The increasing importance of atmospheric demand for ecosystem water and carbon fluxes. *Nat. Clim. Change* **6**, 1023–1027 (2016).
29. Gerken, T., Bromley, G. T., Ruddell, B. L., Williams, S. & Stoy, P. C. Convective suppression before and during the United States Northern Great Plains Flash Drought of 2017. *Hydrol. Earth Syst. Sci.* **22**, 4155–4163 (2018).
30. Chen, L. & Dirmeyer, P. A. Impacts of land-use/land-cover change on afternoon precipitation over North America. *J. Clim.* **30**, 2121–2140 (2016).
31. Gerken, T., Bromley, G. T. & Stoy, P. C. Surface moistening trends in the northern North American Great Plains increase the likelihood of convective initiation. *J. Hydrometeorol.* **19**, 227–244 (2018).
32. Vilà-Guerau de Arellano, J., van Heerwaarden, C. C. & Lelieveld, J. Modelled suppression of boundary-layer clouds by plants in a CO₂-rich atmosphere. *Nat. Geosci.* **5**, 701–704 (2012).
33. Vilà-Guerau de Arellano, J., Ouwersloot, H. G., Baldocchi, D. & Jacobs, C. M. J. Shallow cumulus rooted in photosynthesis. *Geophys. Res. Lett.* **41**, 1796–1802 (2014).
34. Teuling, A. J. et al. Observational evidence for cloud cover enhancement over western European forests. *Nat. Commun.* **8**, 14065 (2017).
35. Ruddell, B. L., Yu, R., Kang, M. & Childers, D. L. Seasonally varied controls of climate and phenophase on terrestrial carbon dynamics: modeling eco-climate system state using Dynamical Process Networks. *Landsc. Ecol.* **31**, 165–180 (2016).
36. Goodwell, A. E. & Kumar, P. Temporal information partitioning: characterizing synergy, uniqueness, and redundancy in interacting environmental variables. *Water Resour. Res.* **53**, 5920–5942 (2017).
37. Goodwell, A. E. & Kumar, P. Temporal Information Partitioning Networks (TIPNets): a process network approach to infer ecohydrologic shifts. *Water Resour. Res.* **53**, 5899–5919 (2017).
38. Kang, M., Ruddell, B. L., Cho, C., Chun, J. & Kim, J. Identifying CO₂ advection on a hill slope using information flow. *Agric. Meteorol.* **232**, 265–278 (2017).
39. Sturtevant, C. et al. Identifying scale-emergent, nonlinear, asynchronous processes of wetland methane exchange. *J. Geophys. Res. Biogeosciences* **121**, 188–204 (2016).
40. Gerken, T. et al. Investigating the mechanisms responsible for the lack of surface energy balance closure in a central Amazonian tropical rainforest. *Agric. Meteorol.* **255**, 92–103 (2017).
41. Kumar, P. & Ruddell, B. L. Information driven ecohydrologic self-organization. *Entropy* **12**, 2085–2096 (2010).
42. Yu, R., Ruddell, B. L., Kang, M., Kim, J. & Childers, D. L. Anticipating global terrestrial ecosystem state change using FLUXNET. *Glob. Change Biol.* <https://doi.org/10.1111/gcb.14602> (2019).
43. Schreiber, T. Measuring information transfer. *Phys. Rev. Lett.* **85**, 461–464 (2000).
44. Ruddell, B. L. & Kumar, P. Ecohydrologic process networks: 2. *Anal. Charact. Water Resour. Res.* **45**, W03420 (2009).
45. Baldocchi, D. et al. FLUXNET: a New tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull. Am. Meteorol. Soc.* **82**, 2415–2434 (2001).
46. Notaro, M. Statistical identification of global hot spots in soil moisture feedbacks among IPCC AR4 models. *J. Geophys. Res. Atmos.* **113**, D09199 (2008).
47. Dirmeyer, P. A., Schlosser, C. A. & Brubaker, K. L. Precipitation, recycling, and land memory: an integrated analysis. *J. Hydrometeorol.* **10**, 278–288 (2009).
48. Zeng, X., Barlage, M., Castro, C. & Fling, K. Comparison of land-precipitation coupling strength using observations and models. *J. Hydrometeorol.* **11**, 979–994 (2010).
49. Zhang, J., Wang, W.-C. & Leung, L. R. Contribution of land–atmosphere coupling to summer climate variability over the contiguous United States. *J. Geophys. Res. Atmos.* **113**, D22109 (2008).
50. Spracklen, D. V. & Garcia-Carreras, L. The impact of Amazonian deforestation on Amazon basin rainfall. *Geophys. Res. Lett.* **42**, 2015GL066063 (2015).
51. Tawfik, A. B., Dirmeyer, P. A. & Santanello, J. A. The heated condensation framework. Part I: Description and Southern Great Plains case study. *J. Hydrometeorol.* **16**, 1929–1945 (2015).
52. Hohenegger, C. & Stevens, B. Controls on and impacts of the diurnal cycle of deep convective precipitation. *J. Adv. Model. Earth Syst.* **5**, 801–815 (2013).
53. Seneviratne, S. I. et al. Investigating soil moisture–climate interactions in a changing climate: a review. *Earth-Sci. Rev.* **99**, 125–161 (2010).
54. Dirmeyer, P. A., Gentine, P., Ek, M. B. & Balsamo, G. in *Sub-Seasonal to Seasonal Prediction* (eds Robertson, A. W. & Vitart, F.) 165–181 (Elsevier, Amsterdam, 2019).
55. Roman, D. T. et al. The role of isohydric and anisohydric species in determining ecosystem-scale response to severe drought. *Oecologia* **179**, 641–654 (2015).
56. De Kauwe, M. G., Medlyn, B. E., Knauer, J. & Williams, C. A. Ideas and perspectives: how coupled is the vegetation to the boundary layer? *Biogeosciences* **14**, 4435–4453 (2017).
57. Konings, A. G. & Gentine, P. Global variations in ecosystem-scale isohydricity. *Glob. Change Biol.* **23**, 891–905 (2017).
58. Konings, A. G., Williams, A. P. & Gentine, P. Sensitivity of grassland productivity to aridity controlled by stomatal and xylem regulation. *Nat. Geosci.* **10**, 284–288 (2017).
59. Juang, J.-Y. et al. Hydrologic and atmospheric controls on initiation of convective precipitation events. *Water Resour. Res.* **43**, W03421 (2007).
60. Gentine, P., Holtslag, A. A. M., D’Andrea, F. & Ek, M. Surface and atmospheric controls on the onset of moist convection over land. *J. Hydrometeorol.* **14**, 1443–1462 (2013).
61. Manoli, G. et al. Soil–plant–atmosphere conditions regulating convective cloud formation above southeastern US pine plantations. *Glob. Change Biol.* **22**, 2238–2254 (2016).
62. Gerken, T. et al. High-resolution modelling of interactions between soil moisture and convective development in a mountain enclosed Tibetan Basin. *Hydrol. Earth Syst. Sci.* **19**, 4023–4040 (2015).
63. Green, J. K. et al. Regionally strong feedbacks between the atmosphere and terrestrial biosphere. *Nat. Geosci.* **10**, 410–414 (2017).
64. Poulter, B. et al. Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature* **509**, 600–603 (2014).
65. Ahlström, A. et al. The dominant role of semi-arid ecosystems in the trend and variability of the land CO₂ sink. *Science* **348**, 895–899 (2015).

66. Fu, Z., Dong, J., Zhou, Y., Stoy, P. C. & Niu, S. Long term trend and interannual variability of land carbon uptake—the attribution and processes. *Environ. Res. Lett.* **12**, 014018 (2017).
67. Huang, J., Yu, H., Guan, X., Wang, G. & Guo, R. Accelerated dryland expansion under climate change. *Nat. Clim. Change* **6**, 166 (2016).
68. Huang, J., Yu, H., Dai, A., Wei, Y. & Kang, L. Drylands face potential threat under 2 °C global warming target. *Nat. Clim. Change* **7**, 417 (2017).
69. Greve, P. et al. Global assessment of trends in wetting and drying over land. *Nat. Geosci.* **7**, 716–721 (2014).
70. Greve, P. & Seneviratne, S. I. Assessment of future changes in water availability and aridity. *Geophys. Res. Lett.* **42**, 2015GL064127 (2015).
71. Foken, T. The energy balance closure problem: an overview. *Ecol. Appl.* **18**, 1351–1367 (2008).
72. Foken, T. et al. Results of a panel discussion about the energy balance closure correction for trace gases. *Bull. Am. Meteorol. Soc.* **92**, ES13–ES18 (2011).
73. Stoy, P. C. et al. A data-driven analysis of energy balance closure across FLUXNET research sites: the role of landscape scale heterogeneity. *Agric. Meteorol.* **171–172**, 137–152 (2013).
74. Gao, Z., Liu, H., Katul, G. G. & Foken, T. Non-closure of the surface energy balance explained by phase difference between vertical velocity and scalars of large atmospheric eddies. *Environ. Res. Lett.* **12**, 034025 (2017).
75. Ruddell, B. L. ProcessNetwork/ProcessNetwork_Software—File Exchange—MATLAB Central. https://www.mathworks.com/matlabcentral/fileexchange/41515-processnetwork-processnetwork_software (2015).
76. Lek, S. & Guégan, J. F. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecol. Model.* **120**, 65–73 (1999).
77. Burden, F. & Winkler, D. in *Artificial Neural Networks: Methods and Applications* (ed Livingstone, D. J.) 23–42 (Humana Press, Clifton, NJ, 2009).

ACKNOWLEDGEMENTS

This work used eddy covariance data acquired by the FLUXNET community and in particular by the following networks: AmeriFlux (U.S. Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DE-FG02-04ER63917 and DE-FG02-04ER63911)), AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada (supported by CFCAS, NSERC, BIOCAP, Environment Canada, and NRCan), GreenGrass, KoFlux, LBA, NECC, OzFlux, TCOS-Siberia, USCCC. We acknowledge the financial support to the eddy covariance data harmonization provided by CarboEuropeIP, FAO-GTOS-TCO, iLEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Université Laval, Environment Canada and US Department of Energy and the database development and technical support from Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California-Berkeley and the University of Virginia. The GLDAS-2 data are produced by NASA GSFC Hydrological Sciences Laboratory (HSL). The MODIS MCD12C1 data product was retrieved from the online Data Pool, courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota (lpdaac.usgs.gov/data_access/data_pool). PCS acknowledges contributions from the U. S.

National Science Foundation awards OIA-1632810, DEB-1552976, and EF-1702029 and the USDA Hatch project 228396. DTD acknowledges the support of the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The authors thank Paul Dirmeyer and two anonymous reviewers for their valuable comments. This work is supported by the National Science Foundation under Grant Nos. EF-1241960, and BCS-1026865, Central Arizona-Phoenix Long-Term Ecological Research (CAP LTER). The opinions are those of the authors, and not necessarily the funding agencies.

AUTHOR CONTRIBUTIONS

B.L.R. and T.G. designed research; T.G., B.L.R., R.Y., P.C.S., D.T.D performed research; TG and R.Y. analyzed data; R.Y. performed PN and ANN calculations; T.G., B.L.R., P.C.S., D.T.D. wrote the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at <https://doi.org/10.1038/s41612-019-0094-4>.

Correspondence and requests for materials should be addressed to T.G.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019