Postprocessing of covariates

Timeseries of environmental covariates were generated from several NEON data products at 30-minute resolution, including DP4.00200.001(NEON 2022b) (bundled eddy covariance for *NEE* and *LH*), DP1.00001.001(NEON 2022a) (2D windspeed, *u*), DP1.00003.001(NEON2022e) (air temperature, *T*), DP1.00098.001 (NEON 2022c) (relative humidity, *RH*), and DP1.00023.001(NEON 2022d) (global radiation, R_g). *NEE* and *LH* were calculated by summing the storage and turbulent carbon and water fluxes provided in the NEON DP4.00200.001(NEON 2022b) product. This differs from the *NEE* and *LH* values provided in the net surface-atmosphere exchange variables in the NEON data files slightly because of the quality flags used by NEON in their eddy4R processing pipeline (Metzger *et al* 2017). Currently, NEON raises a quality flag when any carbon dioxide or water vapor mixing ratio value is missing along the tower. When calculating the net flux, any value where the storage flux has the quality flag raised is treated as missing, and as a result, there are long periods where there are missing *NEE* and *LH* values. In most cases where storage fluxes were flagged as missing, only one or two mixing ratio measurements were absent. As a result, we decided that the benefit of increasing data coverage and including these storage flux values where the quality flag had been raised outweighed the potential increase in *NEE* and *LH* error introduced by the missing measurements. Since in these cases each tower has at least two measurements still available to calculate the storage flux, we expect the impact of these missing measurements on storage flux calculations to be small (Nicolini *et al* 2018). Moreover, NEON forest sites have more than four measurement levels, and thus, have at least three measurements to calculate the storage flux even if two are missing.

We applied u_*-filtering followed by gap-filling of *NEE* and *LH* using the REddyProc package (Wutzler *et al* 2018). In brief, *NEE* and *LH* data were filtered for periods of low turbulence that are known to bias eddy covariance fluxes(Papale *et al* 2006) and then gap-filled using the marginal distribution sampling method(Reichstein *et al* 2005). We applied a bootstrapping approach to constrain threshold value of the friction velocity u_*, and used the 50th percentile estimate of u_* to filter out data periods with insufficient turbulence. *VPD* was calculated directly from *RH* and *T*, and *VPD*, R_g, and *T* were all gap-filled using the marginal distribution sampling of mentioned above (Reichstein *et al* 2005), with no u_*-filtering necessary. Wind speed was not gap filled. All the 30-minute aggregated variables mentioned above were averaged to a daily scale. The daily values that fall between the range first quantile (Q1) minus of 1.5 times interquartile (IQR) (Q1 – 1.5 * IQR) and third quantile (Q3) plus 1.5 times IQR (Q3 + 1.5 *IQR) were preserved.

Significance test of information quantities

We performed a significance test for all mutual information and partial information components. We shuffled all the data that was involved in computing mutual information and partial information components 50 times. Then, for each information quantity, we performed a one-sided paired t-test between 50 copies of the information quantity series that were computed from unshuffled data (vertically stacked) and 50 shuffled information quantity series (vertically stacked). The information quantity is concluded as significant if the p-value of the test is smaller than 0.05.

References

Metzger S, Durden D, Sturtevant C, Luo H, Pingintha-Durden N, Sachs T, Serafimovich A, Hartmann J, Li J, Xu K and Desai A R 2017 eddy4R 0.2.0: a DevOps model for community-extensible processing and analysis of eddy-covariance data based on R, Git, Docker, and HDF5 *Geosci Model Dev* **10** 3189–206

National Ecological Observatory Network (NEON) 2022a 2D wind speed and direction (DP1.00001.001)

National Ecological Observatory Network (NEON) 2022b Bundled data products - eddy covariance (DP4.00200.001)

National Ecological Observatory Network (NEON) 2022c Relative humidity (DP1.00098.001) Online: https://data.neonscience.org

National Ecological Observatory Network (NEON) 2022d Shortwave and longwave radiation (net radiometer) (DP1.00023.001)

National Ecological Observatory Network (NEON) 2022e Triple aspirated air temperature (DP1.00003.001)

Nicolini G, Aubinet M, Feigenwinter C, Heinesch B, Lindroth A, Mamadou O, Moderow U, Mölder M, Montagnani L, Rebmann C and Papale D 2018 Impact of CO 2 storage flux sampling uncertainty on net ecosystem exchange measured by eddy covariance *Agric For Meteorol* **248** 228–39

Papale D, Reichstein M, Aubinet M, Canfora E, Bernhofer C, Kutsch W, Longdoz B, Rambal S, Valentini R, Vesala T and Yakir D 2006 Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation *Biogeosciences* **3** 571–83

Reichstein M, Falge E, Baldocchi D, Papale D, Aubinet M, Berbigier P, Bernhofer C, Buchmann N, Gilmanov T, Granier A, Grunwald T, Havrankova K, Ilvesniemi H, Janous D, Knohl A, Laurila T, Lohila A, Loustau D, Matteucci G, Meyers T, Miglietta F, Ourcival J-M, Pumpanen J, Rambal S, Rotenberg E, Sanz M, Tenhunen J, Seufert G, Vaccari F, Vesala T, Yakir D and Valentini R 2005 On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm *Glob Chang Biol* **11** 1424–39

Wutzler T, Lucas-Moffat A, Migliavacca M, Knauer J, Sickel K, Šigut L, Menzer O and Reichstein M 2018 Basic and extensible post-processing of eddy covariance flux data with REddyProc *Biogeosciences* **15** 5015–30

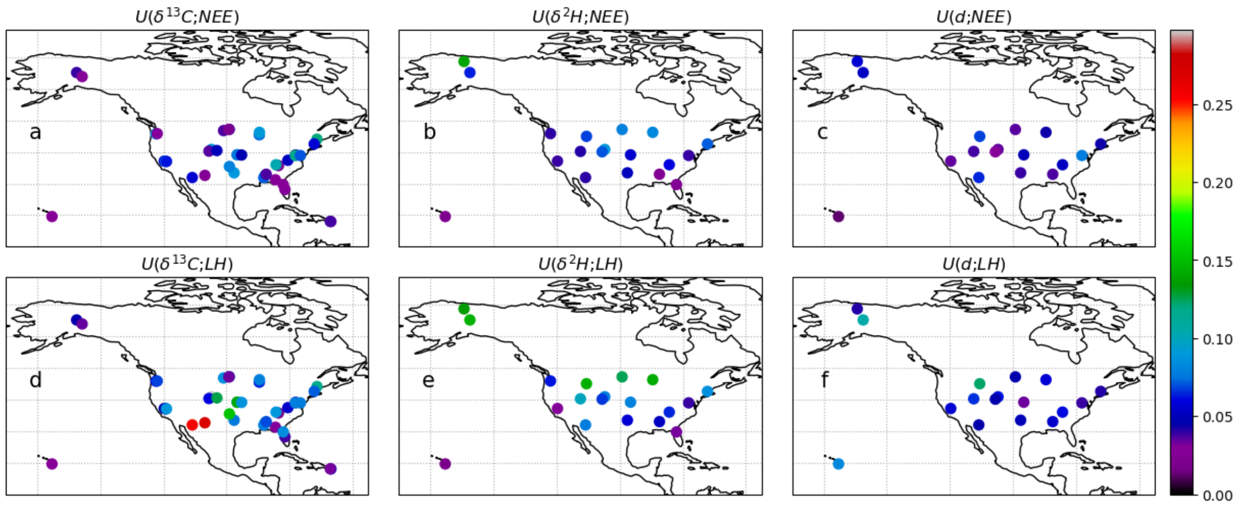

Figure S1 The unique information of (a) $\delta^{13}C$, (b) $\delta^{2}H$, and (c) d isotope data about net ecosystem exchange, *NEE.* The additive information of (d) $\delta^{13}C$, (e) $\delta^{2}H$, and (f) d isotope data about latent heat flux, *LH*.

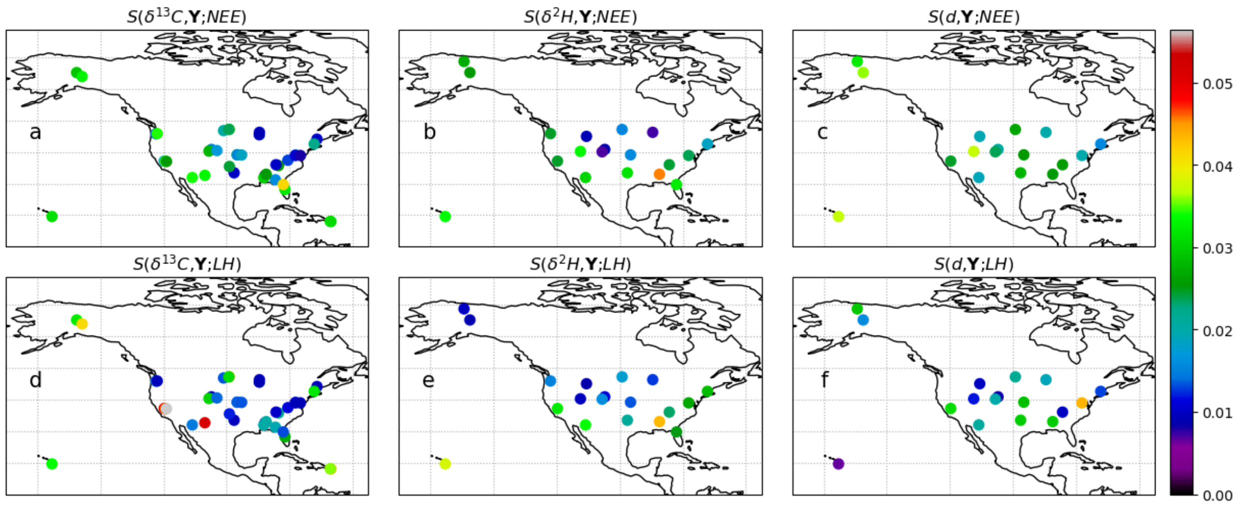

Figure S2 The synergistic information of (a) $\delta^{13}C$, (b) $\delta^{2}H$, and (c) d isotope data about net ecosystem exchange, *NEE. T*he synergistic information of (d) $\delta^{13}C$, (e) $\delta^{2}H$, and (f) d isotope data about latent heat flux, *LH*. The values of S is calculated by averaging across different meteorological variables, indicated by **Y** (e.g., the average over $S(\delta^{2}H,VPD;LH)$, $S(\delta^{2}H,T;LH)$, $S(\delta^{2}H,u;LH)$, and $S(\delta^{2}H,R_g;LH)$).

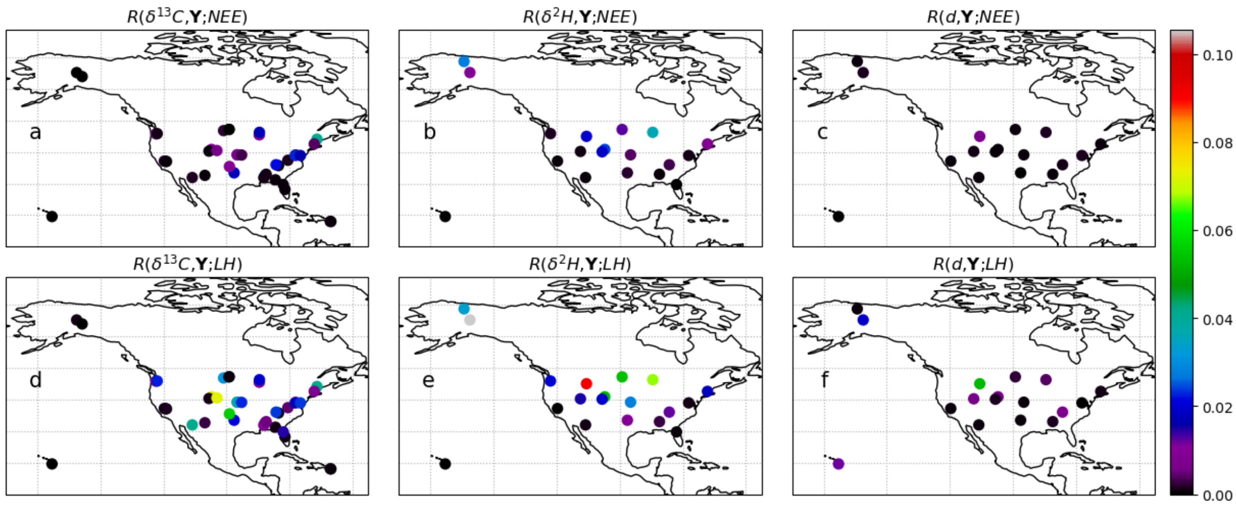

Figure S3 The redundant information of (a) $\delta^{13}C$, (b) $\delta^{2}H$, and (c) d isotope data about net ecosystem exchange, *NEE.* The redundant information of (d) $\delta^{13}C$, (e) $\delta^{2}H$, and (f) d isotope data about latent heat flux, *LH*. The values of R is calculated by averaging across different meteorological variables, indicated by **Y** (e.g., the average over $R(\delta^2H,VPD;LH)$, $R(\delta^2H,T;LH)$, $R(\delta^2H,u;LH)$, and $R(\delta^2H,R_g;LH)$).

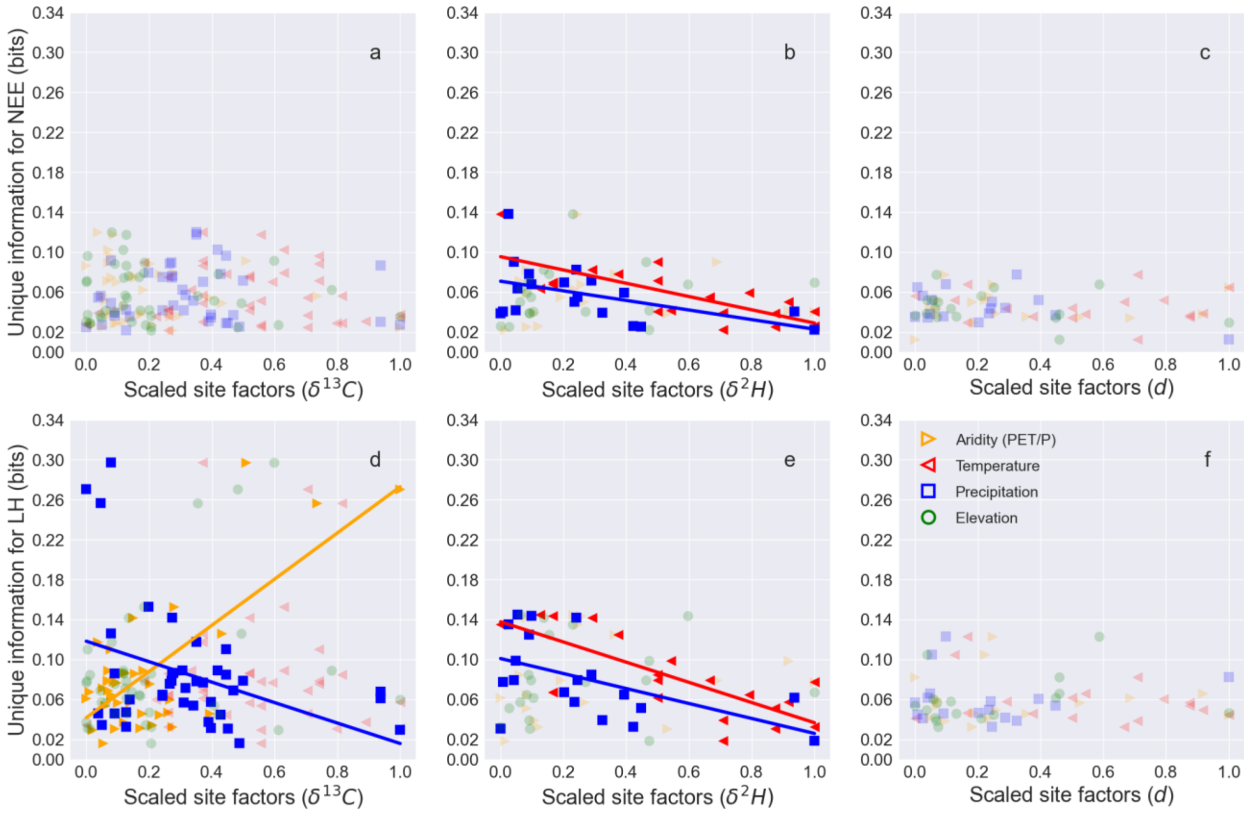

Figure S4 The unique information of (a) $\delta^{13}C$, (b) $\delta^{2}H$, and (c) d isotope data about net ecosystem exchange, *NEE,* against scaled site-specific variables. The total added information of (d) $\delta^{13}C$, (e) $\delta^{2}H$, and (f) d isotope data about latent heat flux, *LH* against scaled site-specific variables. Solid lines indicate a significant p-values (< 0.05) of the slopes.