Stable isotopes contain substantial additive information about terrestrial carbon and water cycling

Bonan Lia,b,c,d Stephen P. Gooda,b, Richard P. Fiorellae,f,g, Catherine E. Finkenbinera,b, Gabriel J. Bowene,f, David C. Nooneh, Christopher J. Stilli, William R.L. Andereggf,j,h

aDepartement of Biological & Ecological Engineering, Oregon State University

bWater Resources Graduate Program, Oregon State University

cCollege of Earth Ocean and Atmospheric Sciences, Oregon State University

dDepartment of Biological & Agricultural Engineering, University of Arkansas

eDepartment of Geology and Geophysics, University of Utah

fGlobal Change and Sustainability Center, University of Utah

gEarth and Environmental Sciences Division, Los Alamos National Laboratory

hDepartment of Physics, University of Auckland

iDepartment of Forest Ecosystems and Society, Oregon State University

jSchool of Biological Sciences, University of Utah

hWilkes Center for Climate Science and Policy, University of Utah

Corresponding author: Bonan Li

**Abstract**

Stable isotope ratios of H (*δ2H*), O (*δ18O*), and C (*δ13C*) are linked to key biogeochemical process of the water and carbon cycles; however, the degree to which isotope associated processes are reflected in ecosystem flux observations remains unquantified. Here through formal information assessment, new measurements of *δ13C*ofnet ecosystem exchange (*NEE*) as well as *δ2H* and *δ18O* of latent heat (*LH*) fluxes across the United States National Ecological Observation Network are used to determine conditions under which isotope measurements are informative of environmental exchanges. We find all three isotopic datasets individually contain information about *NEE* and *LH* fluxes, with isotope measurements carrying unique information content comparable with wind speed observations. Generally, *δ13C*provides more information about *LH* as aridity increases or mean annual precipitation decreases; *δ2H* provides more information about *LH* as temperatures or mean annual precipitation decreases, and also provides more information about *NEE* as temperatures decrease. These results demonstrate that isotopic variability reflecting biophysical controls on *LH*, and to a lesser extent *NEE*, fluxes is stronger under low precipitation, arid, and cooler conditions. The patterns identified in this study are expected aid in modeling and data interpretation efforts focused on constraining carbon and water cycles mechanisms

**Introduction**

Understanding the interactions and drivers of water and carbon exchanges between terrestrial ecosystems and the atmosphere is crucial to illuminate earth’s current climate as well as forecasting impacts of future change on ecosystems and the climate itself1,2. Over the past several decades, significant efforts have been made to measure and monitor terrestrial carbon and water fluxes, including the widespread development of macroscale eddy covariance (EC) networks to measure ecosystem fluxes3,4. EC flux towers can measure continuous net ecosystem exchange (*NEE*) of CO2 between the land surface and atmosphere at various frequency time domains. Similarly, EC measurements of latent heat flux (*LH*), representing evaporation and transpiration from soils, water bodies, and plant canopies, provides valuable information for understanding regional and global water budgets as well as agricultural applications5,6. Flux measurements have been used for a variety of environmental applications such as calibrating and validating remotely sensed flux estimations7, parameterizing land surface models8, modeling seasonal crop coefficients9, and investigating disturbance impacts such as post-fire carbon balance10. While measurements of *LH* and *NEE* can quantify fluxes themselves, new kinds of data are needed to refine knowledge of the processes driving these fluxes which are central to the carbon and water cycles.

To improve understanding of Earth system processes, the geoscience community has developed a wide array of advanced measurements to complement EC flux data to help constrain environmental processes. These include studies focused on stable isotope fluxes11, Carbonyl Sulfide (COS)12, various radiometric indices such as thermal13 and solar induced fluorescence (SIF)14, and even environmental DNA15. Prominent among these techniques, naturally occurring water and carbon isotopes measurements have been shown to be a powerful tool for understanding a wide array of ecohydrological and biophysical processes because distinct processes are, and are not, often associated with known isotope transformations (i.e. fractionation effects)16. Water isotope ratios (*δ2H* and *δ18O* in water) have been used to partition evapotranspiration into evaporation and transpiration, as evaporated and transpired fluxes from the same ecosystem may have distinct isotope ratios17,18. *δ13C* values of CO2 have been applied to separate *NEE* into its constituent fluxes, as the isotopic composition of photosynthesis can differ from that of ecosystem respiration19. Previous studies of *δ2H*, *δ18O* and *δ13C* examined patterns across distinct ecosystems using cryogenic baths and flask samples, however these approaches are constrained in their ability to provide information about ecosystem-scale processes, which generally requires finer temporal and spatial sampling coverage20,21. The development of automated laser spectroscopy systems mounted on EC towers provides new opportunities to obtain long term spatially and temporally resolved atmosphere profiles of these isotopes22. The recently launched National Ecological Observatory Network (NEON) provides the first standardized measurements of the stable isotope ratios of H2O vapor and CO2 for ecosystems across the USA that can be used to estimate *δ2H* and *δ18O* of *LH* and *δ13C* of *NEE*23.

The development of advanced ecosystem measurements across networks such as NEON presents new scientific possibilities; yet this also raises the fundamental question of how useful new and often expensive data streams are for constraining targeted environmental processes. Many advanced measurements are made at considerable cost and effort, yet their full value as a source of information beyond traditional meteorological observations (i.e., vapor pressure deficit [*VPD*], air temperature [*T*], global radiation [*Rg*], and windspeed [*u*]), is rarely demonstrated in a formal sense, especially within continental-scale networks where variability in environmental conditions occurs across a much wider range than individual sites. Here we capitalize on recent advances in information theory to assess the information content of NEON stable isotope data. These advances allow for the formal quantification of linear and nonlinear interactions between variables (termed mutual information)24, as well as approaches to diagnose how unique the information provided by new data sources is relative to others25,26. This study addresses three related questions: (1) Do new observations (here *δ2H,* *δ18O,* and *δ13C* values) contain useful information about the bulk *NEE* and *LH* fluxes across North America? (2) Can any of the information provided by new (isotope) measurements be obtained from other meteorological variables? and (3) Under which environmental conditions do these new measurements provide the most additional information? In doing so, this study provides a generalizable approach for evaluating the conditions under which novel geoscience data is helpful for understanding the Earth system. It also formally quantifies the conditions under which environmental processes associated with transformations of stable isotope ratios, as measured systematically within continental scale networks*,* are a greater contribution to overall environmental exchanges. This approach thereby provides key process level benchmarks for advancing research into Earth’s integrated carbon and water cycles.

**INFORMATIONAL ASSESSMENT OF NEW ISOTOPE DATASETS**

Informational analysis (see **METHODS** for detailed description) shows that isotope data (*δ13C, δ2H,* and *d*)and traditional meteorological data (*Rg*, *T*, VPD, *u*) each contain significant information about temporal variation in *NEE* and *LH* fluxes (Fig. 1) throughout the NEON sites. While it is expected that all these observations encode information on water and carbon cycling to some degree, our analysis quantifies the relative importance of each dataset. We find that *Rg*, *T*, and *VPD* observations consistently contain more information aboutenvironmental fluxesthan either isotope data or wind speed (*u*)which containcomparable information content (Fig. 1). Though the information provided by *Rg* is almost an order of magnitude larger than the information from *u* and the isotopes, *u* is nevertheless one of the well-established drivers of surface-atmosphere water and carbon exchange and is commonly measured at meteorological stations worldwide27. Overall, there is consistently more inter-site variability in mutual information for *VPD*, *T* and *Rg* than the mutual information contained in isotopes and *u*.

Each individual variable, including the flux isotopic compositions, generally tends to share more information with *LH* than *NEE* (Fig. 1). This indicates that *LH* is generally more easily constrained and predicted based on these environmental observations, possibly because it more strongly captures isotopic differences in the contributing one-way flux compared to *NEE* which is the net sum of two opposing fluxes with less distinct isotope ratios. There is more inter-site variability in the mutual information between each isotope and *LH* than the mutual information between each isotope and *NEE*. Instead of *δ13C* values best constraining *NEE* and *δ2H* or *d* values best constraining *LH,* we find that *δ2H* values on average provide slightly more mutual information than *δ13C* values for both *NEE* and *LH* fluxes*;* however, both these (i.e., *δ2H* and *δ13C*) are more informative than *d.* The amount of information that can be inferred from isotopes (and other variables) about *NEE* and *LH* is highly unlikely to be obtained by random processes (*p* < 0.01).

We decomposed and evaluated the multivariate mutual information betweenenvironmental fluxes, isotope ratios, and other variables (Fig. 2). These results demonstrate that most of the information provided by the isotopes about *NEE* and *LH* is unique to these measurements (*δ13C* and *δ2H*), and thus distinct from what can be learned from traditional meteorological datasets. This unique information provided by *δ13C* and *δ2H* values about *LH* is generally higher than the unique information provided about *NEE.* The unique information provided by *δ13C* and *δ2H* values is slightly higher than that contained within *d* values for both *LH* and *NEE* fluxes. The unique information amounts are also found to vary spatially across the NEON sites (Supplemental Fig. S1), where we find the unique information about *NEE* contained in *δ13C* and *δ2H* values has larger inter-site variability than that contained in *d* values. All the unique information provided by the isotope ratios is statistically significant and highly unlikely to be obtained at random (*p* < 0.01).

In addition to the unique information that *δ13C*, *δ2H*, and *d* values contain about *NEE* and *LH* fluxes, a smaller amount of synergistic and redundant information is also present (Supplemental Fig. S2 and S3). The synergistic components consistent of novel information that is present only when both the isotope data and meteorological data are observed and are quite similar for *δ13C* and *δ2H* with *d* values being marginally larger for *NEE* and *δ13C* is marginally larger for *LH*. The redundant information, which is information already present within the other meteorological observations, generally tends to smaller than the unique and synergistic components. Observations of *δ13C* and *δ2H* contain more redundant information about *LH* than about *NEE*. The unique and redundant information linking isotopes with *NEE* are statistically significant (*p* < 0.01).

The total additional information, represented by the sum of the synergistic information and the unique information, provided by each flux isotope composition to *LH* and *NEE* varies spatially across North American NEON sites (Fig. 3). This additional information can be interpretated as the extra information provided by the isotopes beyond what can be obtained by other variables. In general, *δ2H* tends to provide more additional information for *NEE*, while the *d* provides the least amount of additive information to *NEE*. The fraction of information for isotopes about *NEE* that is additive, i.e. (*U+S*)/(*U+S+R*), is 0.94 for *δ13C*, 0.93 for *δ2H*, and 0.99 for *d*, respectively). For *LH,* *δ2H* and *δ13C* provided more additive information than *d* (Fig. 3A). The fraction of additive information about *LH* is 0.89 for *δ13C*, 0.84 for *δ2H*, and 0.95 for *d*, respectively. The additive information of *δ13C* and *δ2H* relating to *LH* has larger variability among sites than that relating to *NEE* (Fig. 3A-3B and Fig. 3D-3E), and there is less variability in the additive information of *d* about *NEE* (Fig. 3C and Fig. 3F) than in *LH*. All the additive information of these isotopes relating to *NEE* and *LH* is statistically significant (*p* < 0.01).

**ECOHYDROLOGICAL IMPLICATIONS**

Our analysis provides a rigorous evaluation of the quantitative value of isotope ratios to provide useful information about carbon and water fluxes across continental scale gradients. For these bulk fluxes, we showed that the information individually provided by these isotopes was similar to the amount of information provided by wind speed measurements, while providing less information than atmospheric vapor pressure deficit, air temperature, and radiation measurements. The meteorological observations evaluated here are commonly used to drive forecasts of environmental processes28,29. A prior *NEE* simulations showed that radiation was consistently the most sensitive predictor for the simulation of *NEE* at two maize fields with distinct irrigation practices30. Similar sensitivity analysis on global evapotranspiration models indicated that radiation was found to rank in the top three most influential input variables31. Our results are consistent with the fundamental notion that solar radiation is the basis for all ecosystem functions32 (excluding rare energy transformations) and drives most diurnal variation in air temperature and vapor pressure deficit and therefore shares the highest amount of mutual information individually with *LH* and *NEE*, with temperature and moisture levels of secondary importance and isotope metrics and wind speed of tertiary importance. It is not necessarily to yield higher mutual information while adding the micrometeorological variables to the current system than the isotopes as micrometeorological variables [e.g., *I*(*LH*; *Rg*,*T*) vs. *I*(*LH*; *Rg*,*δ2H*)] might constrain ecosystem fluxes in a similar way that yield high redundancy, whereas isotopes might constrain the fluxes in a unique manner.

The meteorological variables evaluated here are known to be inter-related to some extent. For instance, the vapor pressure deficit is strongly dependent on air temperature due to the Clausius-Clapeyron relationship 33 and air temperature is tightly related to the amount of radiation as well as to sensible heat fluxes, which balance *LH* in dissipating net absorbed radiation. Therefore, it is expected that these variables contain some similar information about *NEE* and *LH*. Past studies have highlighted how *NEE* and *LH* respond to changes in vapor pressure deficit, air temperature and radiation across various scales, seasons, and ecosystems34–36. Vapor pressure deficit was found to have direct effect on surface energy partitioning as high vapor pressure deficitrepresents high atmosphere demand and hence high *LH* with constant surface conductance37,38. Yet, high vapor pressure deficit can reduce stomatal conductance and thereby reduce plant photosynthesis39. Wind speed can modulate the rate of evapotranspiration and thereby *LH*40,41. The different effect of vapor pressure deficit and wind speed on *LH* may be underrepresented by other metrics but can be captured if evaluated using information theory-based metrics like those explored here.

Notably, even though individual isotopes are not as informative as other variables for *NEE* and *LH*, the information carried by these isotope ratios was found to be unique and cannot be obtained from the other variables. It is also crucial to understand how different variables interactively provide information to a target of interest because knowledge of the interactive dependencies between the inputs and outputs of a studied system is fundamental for model uncertainty characterizations42–44. In fact, one of the challenges for land surface models is increasing process complexity with the integration of a set of sub-models that dramatically expands the inputs for the modeling parametrization45, which substantially increasing risks of model “equifinality”. Moreover, numerous approaches have been developed to estimate and forecast *NEE* and *LH* using empirically or physically based models46–48. However, these methods are not perfect and often either must make some assumptions or simplifications, which can be subject to significant uncertainty49,50. In general, it is more desirable for most of the inputs in a model to provide unique or synergistic pieces of information51, which can potentially capture different source to target processes52. Therefore, the construction and simplification of ecosystem models should tend towards a direction with maximized sum of unique and synergistic information while minimizing redundant information.

The decomposition of the multivariate mutual information between isotopes, other meteorological variables, and the bulk fluxes offers an opportunity to elucidate how much of the information from isotopes is transferred to the bulk fluxes (*NEE* and *LH*). We showed that isotopes carry unique signals about the ecosystem fluxes that cannot be implied by other variables. The portion of unique information from isotopes measurements for carbon and water isotopes was statistically significant and is general larger than the redundant information. Larger unique information often indicates less inter-dependency of the isotopes on meteorological variables, suggesting that isotope ratios of the fluxes may be capturing an integrated signal that is not well reflected in other meteorological variables via distinct pathways. However, we observed inter-site variations in the unique information provided by the isotopes, indicating that the unique information may be dependent on site-specific conditions (e.g., aridity, precipitation). Similarly, isotopes can provide synergistic information about the bulk fluxes suggesting that the patterns of bulk fluxes can potentially be better characterized and predicted with the isotopes included as an additional constraint.

The additional information provided by isotopes to these bulk fluxes are described by the sum of unique information and synergistic information. The sum of these two components, on average, represents how the remaining uncertainty can be reduced by isotopic observations given other meteorological variables in the system. Our analysis demonstrate that fusing isotope data products can potentially lead to better monitoring and prediction of *NEE* and *LH* in a process modeling framework, as these isotope datasets provide additional information beyond what can be obtained using traditional meteorological variables and are associated with known physical mechanisms, although they can require more involved procedures to collect and process. However, the amount of added information of the isotope datasets is likely to vary across sites, climate, and ecosystems. To assess this, we evaluated the additive information of isotopes based on NEON site conditions via a simple linear regression (Fig. 4). We showed that the additive information that *δ13C* provides about *LH* is influenced by mean annual precipitation, aridity, and site elevation (Fig. 4D), as indicated by a significant slope value from the linear regression. *δ13C* is likely to provide more useful information about *LH* in locations with higher atmospheric evaporative demand relative to precipitation (Fig. 4D, *p* = 8.5 x 10-11) or in locations with less annual precipitation (Fig. 4D, *p* = 0.006) or with higher altitude (Fig. 4D, *p* = 0.024). The additive information *δ2H* provides about *NEE* was shown to be mainly influenced by the site mean annual temperature (Fig. 4B). In particular, *δ2H* tends to be more informative about *NEE* in locations with cooler climates (*p* = 0.007). Similarly, there is more opportunity for *δ2H* to provide additional knowledge about *LH* at locations with cooler climates (*p* = 0.001) or less mean annual precipitation (Fig. 4E, *p* = 0.037). No significant relationship was found between the additional information of the *d* provided to either *NEE* or *LH*.

One of key motivations for measuring stable isotopes of water and carbon fluxes is that they may provide a unique constraint, which has not been formally tested until this study, which allow for the partition of bulk fluxes into their respective constituents53–55. This is because the flux isotope ratios are influenced by distinct biophysical processes, and thus larger amounts of new mutual information between isotopes and environmental fluxes quantifies the conditions under which these processes are more dominate components of overall bulk fluxes. In this light the trends described above (and in Fig 4) are consistent with prior knowledge of isotope geophysics. For instance, equilibrium fractionation factors are sensitive to temperature, particularly at low values16, with broad decreases in vapor *δ2H* observed poleward at continental scales56. Similarly, evaporation is expected to play a larger role in *LH* fluxes under low vegetation, more arid climates57, and this study provides a new way to quantify the relative importance of these isotope processes on bulk fluxes. Finally, it is possible that the isotopes can even be more informative for studying constituents of the bulk fluxes. Indeed, water isotopes have been used to partition evapotranspiration to evaporation and transpiration, but these analyses have yielded diverse results across various ecosystems and variable levels of accuracy 57–59.

It is worth noting that this analysis is based on current available data products and quality control methods. It may be also worthwhile for future studies to investigate if and how the results vary with more available datasets and a wider range of environmental conditions. However, given the power of isotopes for tackling fundamental problems in carbon and water cycling and projecting the future of terrestrial ecosystem function under a rapidly changing climate16,60, our results can be useful to provide guidance for improving model results after the incorporation of isotope flux ratios.

**2 Methods**

**2. 1 Site descripts and data preparation**

This study was conducted at part of terrestrial sites of National Ecological Observatory Network (NEON), which is a continental scale research platform for understanding the fundamental principles that govern ecological responses to climate change, land use change and species invasion61. NEON statistically divided the US into 20 ecoclimatic domains across different biomes and landforms to capture key bioecological aspects of North American ecology62. We used the 30-minute aggregated *NEE*, *LH*, global radiation (*Rg*), tower top air temperature (*T*), and the two-dimensional wind speed (*u*) datasets from the NEON’s eddy covariance bundled datasets63. The vapor pressure deficit (*VPD*) data were derived based on NEON’s relative humidity product 64. These 30-minte variables were gap-filling and further processed to daily scale. More details can be found in **SUPPLEMENTARY MATERIAL**. Daily stable isotope ratios of *NEE* and *LH* were obtained from a recently published datasets23, which was derived based on the surface isotope composition of carbon dioxide and water vapor across NEON sites.

**2.2 Mutual information measures**

Mutual information is a measure of how two random variables are probabilistically dependent on each other in the unit of bits24. Theoretically, the maximum mutual information shared between two discrete random variables is the minimum of the entropy, which is characterized as the uncertainty associated with a random variable24, of these two random variables. Probabilistically, the mutual information can be expressed as:

|  |  |
| --- | --- |
| *I*(X;Y) = | (1) |

where *p*(*x*), *p*(y), and *p*(x,y) are the probability density functions of random variables X, Y, and {X,Y} respectively.

The multivariate mutual information of a single random variable (Z) and a set of random variables {X, Y} characterizes the amount of uncertainty inZ that can be reduced by the knowledge of {X, Y} and can be expressed as:

|  |  |
| --- | --- |
| *I*(X,Y;Z) = | (2) |

where *p*(z), *p*(x,y), and *p*(x,y,z) are the probability density functions of variables Z, {X,Y}, and {X,Y,Z}, respectively and were estimated using a kernel density estimation (KDE) method with a gaussian kernel and Silverman bandwidth selection method65. We rescaled each data point to a common range of [0, 1] before using KDE. We then evaluate the probability density functions from 0 to 1 with a step size of 0.05.

We computed the pairwise mutual information [e.g., *I*(*NEE*;*VPD*), *I*(*LH*;*VPD*) .etc] shared among *VPD*, *T*, *Rg*, *u*, *δ13C*, *δ2H*, and *d* about *NEE* and *LH* iteratively. Due to the limitation of isotope datasets, we computed the mutual information of each variable with the *NEE* and *LH* by subsampling 100 data points without replacement 500 times to ensure constituent data counts in mutual information calculations. Then, the mutual information of the variable of interest and the flux is computed as the average mutual information across 500 resamplings. The mutual information contents computed above are tested for statistical significance (refer to **SUPPLEMENTARY MATERIAL** for details).

**2.3 Partial information decomposition**

The multivariate mutual information can be decomposed into different informational components via a partial information decomposition framework (PID)25,26,52. The PID can decompose *I*(X,Y;Z) into: (1) unique information (U) that is only provided byX or Y solely to the Z, respectively; (2) synergistic information (S) that is the information provide to the Z when X and Y act jointly; (3) redundant information (R) that is the overlapping information provided both by X and Y to the Z25. The PID framework can be formulated as

|  |  |
| --- | --- |
| *I*(X,Y; Z) = + + R + S | (3) |
| *I*(X; Z) = + R | (4) |
| *I*(Y; Z) = + R | (5) |

Whereand are the unique information of X and Y to Z, respectively. R and S are the redundant and synergistic information of X and Y to Z, respectively. All PID components are non-negative real numbers in unit of bits25.

In this study, we quantified the information flow between each flux and each isotope flux ratio by leveraging the PID framework. We first computed the unique information that each isotope flux ratio contributed to each of the bulk flux. The unique information can only be solved when other variables (besides the bulk flux and isotopes) are available in the PID system. Therefore, we defined the decomposed information components that the isotope ratios provided to the bulk fluxes as the averaged unique information across all meteorological variables (*VPD*, *T*, *Rg*, and *u*). As with computing the individual mutual information, we also subsampled 100 data points from each dataset without replacement 500 times. The partial information components of the isotopes were then computed as the averaged information components from 500 iterations. The significance tests were performed similarly to mutual information (refer to **SUPPLEMENTARY MATERIAL** for details).

**Acknowledgements**

The authors want to acknowledge the funding support of the United States National Foundation (EF1802885 and EF 1802880).

**Author contribution**

BL and SPG designed the study. RPF provided flux datasets and gap-filled meteorological datasets and wrote part of the data processing steps in Supplementary material. BL analyzed the data and wrote the manuscript. SPG, RPF, CEF, GJB, DCN, CJS, and WRLA reviewed the manuscript.

**Competing interest statement**

The authors declare no conflicts of interest.

**References**

1. Jung, M. *et al.* Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *J. Geophys. Res.* **116**, G00J07 (2011).

2. Piao, S. *et al.* Interannual variation of terrestrial carbon cycle: Issues and perspectives. *Glob. Chang. Biol.* **26**, 300–318 (2020).

3. Baldocchi, D. Measuring fluxes of trace gases and energy between ecosystems and the atmosphere - the state and future of the eddy covariance method. *Glob. Chang. Biol.* **20**, 3600–3609 (2014).

4. Schimel, D. & Schneider, F. D. Flux towers in the sky: global ecology from space. *New Phytol.* **224**, 570–584 (2019).

5. Zhou, S., Yu, B., Zhang, Y., Huang, Y. & Wang, G. Water use efficiency and evapotranspiration partitioning for three typical ecosystems in the Heihe River Basin, northwestern China. *Agric. For. Meteorol.* **253**–**254**, 261–273 (2018).

6. Zeng, S. *et al.* Integrated land-surface hydrological and biogeochemical processes in simulating water, energy and carbon fluxes over two different ecosystems. *J. Hydrol.* **582**, 124390 (2020).

7. Jia, Z., Liu, S., Xu, Z., Chen, Y. & Zhu, M. Validation of remotely sensed evapotranspiration over the Hai River Basin, China. *J. Geophys. Res. Atmos.* **117**, n/a-n/a (2012).

8. Williams, M. *et al.* Improving land surface models with FLUXNET data. *Biogeosciences* **6**, 1341–1359 (2009).

9. Li, S., Kang, S., Li, F. & Zhang, L. Evapotranspiration and crop coefficient of spring maize with plastic mulch using eddy covariance in northwest China. *Agric. Water Manag.* **95**, 1214–1222 (2008).

10. Lupascu, M., Akhtar, H., Smith, T. E. L. & Sukri, R. S. Post‐fire carbon dynamics in the tropical peat swamp forests of Brunei reveal long‐term elevated CH 4 flux. *Glob. Chang. Biol.* **26**, 5125–5145 (2020).

11. Dubbert, M. & Werner, C. Water fluxes mediated by vegetation: emerging isotopic insights at the soil and atmosphere interfaces. *New Phytol.* **221**, 1754–1763 (2019).

12. Whelan, M. E. *et al.* Reviews and syntheses: Carbonyl sulfide as a multi-scale tracer for carbon and water cycles. *Biogeosciences* **15**, 3625–3657 (2018).

13. Still, C. J. *et al.* Imaging canopy temperature: shedding (thermal) light on ecosystem processes. *New Phytol.* **230**, 1746–1753 (2021).

14. Guan, K. *et al.* Improving the monitoring of crop productivity using spaceborne solar‐induced fluorescence. *Glob. Chang. Biol.* **22**, 716–726 (2016).

15. URycki, D. R., Bassiouni, M., Good, S. P., Crump, B. C. & Li, B. The streamwater microbiome encodes hydrologic data across scales. *Sci. Total Environ.* **849**, 157911 (2022).

16. Bowen, G. J. & Good, S. P. Incorporating water isoscapes in hydrological and water resource investigations. *WIREs Water* **2**, 107–119 (2015).

17. Xiao, W., Wei, Z. & Wen, X. Evapotranspiration partitioning at the ecosystem scale using the stable isotope method—A review. *Agric. For. Meteorol.* **263**, 346–361 (2018).

18. Berkelhammer, M. *et al.* The nocturnal water cycle in an open-canopy forest. *J. Geophys. Res. Atmos.* **118**, 10,225-10,242 (2013).

19. Lee, S.-C. *et al.* Partitioning of net ecosystem exchange into photosynthesis and respiration using continuous stable isotope measurements in a Pacific Northwest Douglas-fir forest ecosystem. *Agric. For. Meteorol.* **292**–**293**, 108109 (2020).

20. Orlowski, N. *et al.* Inter-laboratory comparison of cryogenic water extraction systems for stable isotope analysis of soil water. *Hydrol. Earth Syst. Sci.* **22**, 3619–3637 (2018).

21. Gemery, P. A., Trolier, M. & White, J. W. C. Oxygen isotope exchange between carbon dioxide and water following atmospheric sampling using glass flasks. *J. Geophys. Res. Atmos.* **101**, 14415–14420 (1996).

22. Fiorella, R. P. *et al.* Calibration Strategies for Detecting Macroscale Patterns in NEON Atmospheric Carbon Isotope Observations. *J. Geophys. Res. Biogeosciences* **126**, (2021).

23. Finkenbiner, C. E. *et al.* The NEON Daily Isotopic Composition of Environmental Exchanges Dataset. *Sci. Data* **9**, 353 (2022).

24. Cover, T. M. & Thomas, J. A. *Elements of Information Theory*. (Wiley, 2005). doi:10.1002/047174882X.

25. Goodwell, A. E. & Kumar, P. Temporal information partitioning: Characterizing synergy, uniqueness, and redundancy in interacting environmental variables. *Water Resour. Res.* **53**, 5920–5942 (2017).

26. Williams, P. L. & Beer, R. D. Nonnegative Decomposition of Multivariate Information. (2010).

27. Yusup, Y. & Liu, H. Effects of persistent wind speeds on turbulent fluxes in the water-atmosphere interface. *Theor. Appl. Climatol.* **140**, 313–325 (2020).

28. Cosgrove, B. A. *et al.* Real‐time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *J. Geophys. Res. Atmos.* **108**, 2002JD003118 (2003).

29. Rodell, M. *et al.* The Global Land Data Assimilation System. *Bull. Am. Meteorol. Soc.* **85**, 381–394 (2004).

30. Safa, B., Arkebauer, T. J., Zhu, Q., Suyker, A. & Irmak, S. Net Ecosystem Exchange (NEE) simulation in maize using artificial neural networks. *IFAC J. Syst. Control* **7**, 100036 (2019).

31. Talsma, C. *et al.* Sensitivity of Evapotranspiration Components in Remote Sensing-Based Models. *Remote Sens.* **10**, 1601 (2018).

32. Yetemen, O., Istanbulluoglu, E., Flores-Cervantes, J. H., Vivoni, E. R. & Bras, R. L. Ecohydrologic role of solar radiation on landscape evolution. *Water Resour. Res.* **51**, 1127–1157 (2015).

33. Clausius, R. Ueber die bewegende Kraft der Wärme und die Gesetze, welche sich daraus für die Wärmelehre selbst ableiten lassen. *Ann. der Phys. und Chemie* **155**, 500–524 (1850).

34. Chen, J. *et al.* Assessments of the factors controlling latent heat flux and the coupling degree between an alpine wetland and the atmosphere on the Qinghai-Tibetan Plateau in summer. *Atmos. Res.* **240**, 104937 (2020).

35. Niu, S. *et al.* Thermal optimality of net ecosystem exchange of carbon dioxide and underlying mechanisms. *New Phytol.* **194**, 775–783 (2012).

36. Gu, L. *et al.* Direct and indirect effects of atmospheric conditions and soil moisture on surface energy partitioning revealed by a prolonged drought at a temperate forest site. *J. Geophys. Res.* **111**, D16102 (2006).

37. Tong, B. *et al.* Effects of soil moisture, net radiation, and atmospheric vapor pressure deficit on surface evaporation fraction at a semi-arid grass site. *Sci. Total Environ.* **849**, 157890 (2022).

38. Wang, P., Li, D., Liao, W., Rigden, A. & Wang, W. Contrasting Evaporative Responses of Ecosystems to Heatwaves Traced to the Opposing Roles of Vapor Pressure Deficit and Surface Resistance. *Water Resour. Res.* **55**, 4550–4563 (2019).

39. Grossiord, C. *et al.* Plant responses to rising vapor pressure deficit. *New Phytol.* **226**, 1550–1566 (2020).

40. Yang, Y. *et al.* Short-term forecasting of daily reference evapotranspiration using the reduced-set Penman-Monteith model and public weather forecasts. *Agric. Water Manag.* **211**, 70–80 (2019).

41. Liu, X. & Zhang, D. Trend analysis of reference evapotranspiration in Northwest China: The roles of changing wind speed and surface air temperature. *Hydrol. Process.* **27**, 3941–3948 (2013).

42. Ruddell, B. L., Drewry, D. T. & Nearing, G. S. Information Theory for Model Diagnostics: Structural Error is Indicated by Trade‐Off Between Functional and Predictive Performance. *Water Resour. Res.* **55**, 6534–6554 (2019).

43. Li, B. & Good, S. P. Information-based uncertainty decomposition in dual-channel microwave remote sensing of soil moisture. *Hydrol. Earth Syst. Sci.* **25**, 5029–5045 (2021).

44. Gong, W., Gupta, H. V., Yang, D., Sricharan, K. & Hero, A. O. Estimating epistemic and aleatory uncertainties during hydrologic modeling: An information theoretic approach. *Water Resour. Res.* **49**, 2253–2273 (2013).

45. Fisher, R. A. & Koven, C. D. Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems. *J. Adv. Model. Earth Syst.* **12**, (2020).

46. Wood, D. A. Net ecosystem carbon exchange prediction and insightful data mining with an optimized data-matching algorithm. *Ecol. Indic.* **124**, 107426 (2021).

47. Su, Z. The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrol. Earth Syst. Sci.* **6**, 85–100 (2002).

48. Veroustraete, F., Patyn, J. & Myneni, R. B. Estimating net ecosystem exchange of carbon using the normalized difference vegetation index and an ecosystem model. *Remote Sens. Environ.* **58**, 115–130 (1996).

49. Papale, D. *et al.* Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. *Biogeosciences* **3**, 571–583 (2006).

50. Zhao, W. L. *et al.* Uncertainties Caused by Resistances in Evapotranspiration Estimation Using High-Density Eddy Covariance Measurements. *J. Hydrometeorol.* **21**, 1349–1365 (2020).

51. Wibral, M., Priesemann, V., Kay, J. W., Lizier, J. T. & Phillips, W. A. Partial information decomposition as a unified approach to the specification of neural goal functions. *Brain Cogn.* **112**, 25–38 (2017).

52. Goodwell, A. E., Kumar, P., Fellows, A. W. & Flerchinger, G. N. Dynamic process connectivity explains ecohydrologic responses to rainfall pulses and drought. *Proc. Natl. Acad. Sci.* **115**, (2018).

53. Good, S. P. *et al.* δ 2 H isotopic flux partitioning of evapotranspiration over a grass field following a water pulse and subsequent dry down. *Water Resour. Res.* **50**, 1410–1432 (2014).

54. Conrad, R., Klose, M., Yuan, Q., Lu, Y. & Chidthaisong, A. Stable carbon isotope fractionation, carbon flux partitioning and priming effects in anoxic soils during methanogenic degradation of straw and soil organic matter. *Soil Biol. Biochem.* **49**, 193–199 (2012).

55. Wang, L. *et al.* Partitioning evapotranspiration across gradients of woody plant cover: Assessment of a stable isotope technique. *Geophys. Res. Lett.* **37**, n/a-n/a (2010).

56. Good, S. P., Noone, D., Kurita, N., Benetti, M. & Bowen, G. J. D/H isotope ratios in the global hydrologic cycle. *Geophys. Res. Lett.* **42**, 5042–5050 (2015).

57. Wang, L., Good, S. P. & Caylor, K. K. Global synthesis of vegetation control on evapotranspiration partitioning. *Geophys. Res. Lett.* **41**, 6753–6757 (2014).

58. Lu, X. *et al.* Partitioning of evapotranspiration using a stable isotope technique in an arid and high temperature agricultural production system. *Agric. Water Manag.* **179**, 103–109 (2017).

59. Wen, X., Yang, B., Sun, X. & Lee, X. Evapotranspiration partitioning through in-situ oxygen isotope measurements in an oasis cropland. *Agric. For. Meteorol.* **230**–**231**, 89–96 (2016).

60. Bowling, D. R., Pataki, D. E. & Randerson, J. T. Carbon isotopes in terrestrial ecosystem pools and CO 2 fluxes. *New Phytol.* **178**, 24–40 (2008).

61. Barnett, D. T. *et al.* The plant diversity sampling design for The National Ecological Observatory Network. *Ecosphere* **10**, (2019).

62. Elmendorf, S. C. *et al.* The plant phenology monitoring design for The National Ecological Observatory Network. *Ecosphere* **7**, e01303 (2016).

63. National Ecological Observatory Network (NEON). Bundled data products - eddy covariance (DP4.00200.001). (2022) doi:10.48443/7CQP-3J73.

64. National Ecological Observatory Network (NEON). Relative humidity (DP1.00098.001). (2021) doi:10.48443/S9WY-T644.

65. Silverman, B. W. *Density Estimation for Statistics and Data Analysis*. (Routledge, 2018). doi:10.1201/9781315140919.

**Chart, box and whisker chart

Description automatically generated**

**Figure 1 | Individual mutual information.** Individual mutual information shared between net ecosystem exchange (*NEE*) and each individual meteorologicalvariable (vapor pressure deficit [*VPD*], air temperature [*T*], global radiation [*Rg*], windspeed [*u*]) **(A)**. Individual information shared between latent heat flux (*LH*) and each individual meteorological variable **(B)**. The mean and median values of each boxplot are shown as black triangle and white line, respectively. The double asterisk indicates a significant p-value (<0.01).

**Chart, box and whisker chart

Description automatically generated**

**Figure 2 | The information components of stable isotopes to ecosystem fluxes.** The unique information (*U*), synergistic information (*S*), and redundant information (*R*) of the stable isotope flux ratios on the net ecosystem exchange [*NEE*] **(A)** and latent heat flux (*LH*) **(B)** fluxes. The mean and median values of each boxplot are shown as black triangle and white line, respectively. The double asterisk indicates a significant p-value (<0.01).

**Chart, scatter chart

Description automatically generated**

**Figure 3 | The spatial distribution of total added information, sum of unique information (*U*) and synergistic information (*S*), of different stable isotopes about ecosystem fluxes.** The additive information of *δ13C* about net ecosystem exchange (*NEE*) **(A)** andlatent heat flux (*LH*) **(D)**.The additive information of *δ13H* about *NEE* **(B)** andlatent heat flux (*LH*) **(E)**. The additive information of *d* about *NEE* **(C)** andlatent heat flux (*LH*) **(F)**.

**Graphical user interface, chart, application, scatter chart

Description automatically generated**

**Figure 4 | The relationships between total added information (U + S) of each isotope about the ecosystem fluxes and scaled site-specific variables.** The total added information of *δ*13*C* **(A)**, *δ2H* **(B)**, and *d* **(C)** about net ecosystem exchange (*NEE*) against scaled site-specific variables. The total added information of *δ*13*C* **(D)**, *δ2H* **(E)**, and *d* **(F)** about latent heat flux (*LH*) against scaled site-specific variables. Solid lines indicate a significant p-values (< 0.05) of the slopes.

**Supplementary material**

*Postprocessing of covariates*

Timeseries of environmental covariates were generated from several NEON data products at 30-minute resolution, including DP4.00200.0011 (bundled eddy covariance for *NEE* and *LH*), DP1.00001.0012 (2D windspeed, *u*), DP1.00003.0013 (air temperature, *T*), 4 (relative humidity, *RH*), and DP1.00023.0015 (global radiation, *Rg*). *NEE* and *LH* were calculated by summing the storage and turbulent carbon and water fluxes provided in the NEON DP4.00200.0011 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 pipeline6. 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 of 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 small7. 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 -filtering followed by gap-filling of *NEE* and *LH* using the REddyProc package8. In brief, *NEE* and *LH* data were filtered for periods of low turbulence that are known to bias eddy covariance fluxes9 and then gap-filled using the marginal distribution sampling method10. We applied a bootstrapping approach to constrain threshold value of the friction velocity , and used the 50th percentile estimate of to filter out data periods with insufficient turbulence. *VPD* was calculated directly from *RH* and *T*, and *VPD*, *Rg*, and *T* were all gap-filled using the marginal distribution sampling of mentioned above10, with no -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 of 1.5 times interquartile (IQR) of the first quantile (Q1) and 1.5 times IQR of the third quantile (Q3) were preserved.

*Significant test of information quantities*

We performed the significant 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.

Chart, scatter chart

Description automatically generated

**Figure S1 | The spatial distribution of unique information (*U*).** The unique information of *δ13C* about net ecosystem exchange (*NEE*) **(A)** andlatent heat flux (*LH*) **(D)**.The unique information of *δ13H* about *NEE* **(B)** andlatent heat flux (*LH*) **(E)**. The unique information of *d* about *NEE* **(C)** andlatent heat flux (*LH*) **(F)**.

Chart, scatter chart

Description automatically generated

**Figure S2 | The spatial distribution of synergistic information (*S*).** The synergistic information of *δ13C* about net ecosystem exchange (*NEE*) **(A)** andlatent heat flux (*LH*) **(D)**.The unique information of *δ13H* about *NEE* **(B)** andlatent heat flux (*LH*) **(E)**. The synergistic information of *d* about *NEE* **(C)** andlatent heat flux (*LH*) **(F)**.

Chart

Description automatically generated

**Figure S3 | The spatial distribution of redundant information (*R*).** The redundant information of *δ13C* about net ecosystem exchange (*NEE*) **(A)** andlatent heat flux (*LH*) **(D)**.The redundant information of *δ13H* about *NEE* **(B)** andlatent heat flux (*LH*) **(E)**. The unique information of *d* about *NEE* **(C)** andlatent heat flux (*LH*) **(F)**.

**REFERENCES**

1. National Ecological Observatory Network (NEON). Bundled data products - eddy covariance (DP4.00200.001). (2022) doi:10.48443/7CQP-3J73.
2. National Ecological Observatory Network (NEON). 2D wind speed and direction (DP1.00001.001). (2022) doi:10.48443/77N6-EH42.
3. National Ecological Observatory Network (NEON). Triple aspirated air temperature (DP1.00003.001). (2022) doi:10.48443/Q16J-SN13.
4. National Ecological Observatory Network (NEON). Relative humidity (DP1.00098.001). (2021) doi:10.48443/S9WY-T644.
5. National Ecological Observatory Network (NEON). Shortwave and longwave radiation (net radiometer) (DP1.00023.001). (2022) doi:10.48443/STBF-BH38.
6. Metzger, S. et al. 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, 3189–3206 (2017).
7. Nicolini, G. et al. Impact of CO 2 storage flux sampling uncertainty on net ecosystem exchange measured by eddy covariance. Agric. For. Meteorol. 248, 228–239 (2018).
8. Wutzler, T. et al. Basic and extensible post-processing of eddy covariance flux data with REddyProc. Biogeosciences 15, 5015–5030 (2018).
9. Papale, D. et al. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences 3, 571–583 (2006).
10. Reichstein, M. et al. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob. Chang. Biol. 11, 1424–1439 (2005).