Agricultural and Forest Meteorology 372 (2025) 110686

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Network of networks: Time series clustering of AmeriFlux sites

David E Reed [a,b,c,*], Housen Chu [d,*], Brad G. Peter [e], Jiquan Chen [c,f], Michael Abraha [f,g], Brian Amiro [h], Ray G. Anderson [i], M. Altaf Arain [j], Paulo Arruda [k], Greg A. Barron-Gafford [l], Carl Bernacchi [m], Daniel P. Beverly [n], Sebastien C. Biraud [d], T. Andrew Black [o], Peter D. Blanken [p], Gil Bohrer [q], Rebecca Bowler [r], David R. Bowling [s], M. Syndonia Bret-Harte [t], Mario Bretfeld [u], Nathaniel A. Brunsell [v], Stephen H. Bullock [w], Gerardo Celis [x], Xingyuan Chen [y], Aimee T. Classen [z,aa], David R. Cook [ab], Alejandro Cueva [ac], Higo J. Dalmagro [ad], Kenneth Davis [ae], Ankur Desai [af], Alison J. Duff [ag], Allison L. Dunn [ah], David Durden [ai], Colin W. Edgar [t], Eugenie Euskirchen [t], Rosvel Bracho [aj], Brent Ewers [ak], Lawrence B. Flanagan [al], Christopher Florian [ai], Vanessa Foord [am], Inke Forbrich [an], Brandon R. Forsythe [ae], John Frank [ao], Jaime Garatuza-Payan [ap], Sarah Goslee [aq], Christopher Gough [ar], Mark Green [as], Timothy Griffis [at], Manuel Helbig [au,av], Andrew C. Hill [aw], Ross Hinkle [ax], Jason Horne [ae], Elyn Humphreys [ay], Hiroki Ikawa [az], Go Iwahana [ba], Rachhpal Jassal [o], Bruce Johnson [bb], Mark Johnson [bc,bd], Steven A. Kannenberg [be], Eric Kelsey [bf], John King [bg], John F. Knowles [bh], Sara Knox [bi,bj], Hideki Kobayashi [bk], Thomas Kolb [bl], Randy Kolka [aw], Ken W. Krauss [bm,bn], Lars Kutzbach [bo], Brian Lamb [bp], Beverly Law [bq], Sung-Ching Lee [br], Xuhui Lee [a], Heping Liu [bp], Henry W. Loescher [ai,bs], Sparkle L. Malone [a], Roser Matamala [ab], Marguerite Mauritz [bt], Stefan Metzger [af,bu], Gesa Meyer [bv], Bhaskar Mitra [bw], J. William Munger [bx], Zoran Nesic [o], Asko Noormets [by], Thomas L. O'Halloran [bz], Patrick T O'Keeffe [bp], Steven F. Oberbauer [ca], Walter Oechel [cb], Patty Oikawa [cc], Paulo C. Olivas [cd], Andrew Ouimette [ce], Gilberto Pastorello [cf], Jorge F. Perez-Quezada [cg], Claire Phillips [ch], Gabriela Posse [ci], Bo Qu [cj], William L. Quinton [ck], Michele L. Reba [cl], Andrew D. Richardson [cm], Valentin Picasso [cn], Adrian V. Rocha [co], Julio C. Rodriguez [cp], Roel Ruzol [cq], Scott Saleska [cr], Russell L. Scott [cs], Adam P. Schreiner-McGraw [ct], Edward A.G. Schuur [cu], Maria Silveira [cv], Oliver Sonnentag [cj], David L. Spittlehouse [cw], Ralf Staebler [cx], Gregory Starr [cy], Christina Staudhammer [cy], Chris Still [bq], Cove Sturtevant [ai], Ryan C. Sullivan [ab], Andy Suyker [cz], David Trejo [cg], Masahito Ueyama [da], Rodrigo Vargas [db], Brian Viner [dc], Enrique R. Vivoni [dd], Dong Wang [de], Eric J. Ward [df,dg], Susanne Wiesner [dh], Lisamarie Windham-Myers [di], David Yannick [cy],

* Corresponding authors.
E-mail addresses: david.edwin.reed@gmail.com (D.E. Reed), hchu@lbl.gov (H. Chu).

https://doi.org/10.1016/j.agrformet.2025.110686
Received 16 January 2025; Received in revised form 21 May 2025; Accepted 5 June 2025
Available online 24 June 2025

Enrico A. Yepez[ap], Terenzio Zenone[dj], Junbin Zhao[dk], Donatella Zona[cb]

[a] Yale School of the Environment, Yale University, New Haven, CT, 06511, United States
[b] University of Science and Arts of Oklahoma, Division of Science & Physical Education, OK, 73018, United States
[c] Department of Geography, Environment & Spatial Sciences, Michigan State University, East Lansing, MI, 48823, United States
[d] Climate and Ecosystem Sciences Division, Lawrence Berkeley National Lab, Berkeley, CA, 94702, United States
[e] Department of Geosciences, University of Arkansas, Fayetteville, AR, 72701, USA
[f] Center for Global Change and Earth Observations, Michigan State University, East Lansing, MI, 48823, USA
[g] LI-COR, 4647 superior street, Lincoln, NE, 68504, USA
[h] Department of Soil Science, University of Manitoba, Winnipeg, MB, R3T2N2, Canada
[i] USDA-Agricultural Research Service, US Salinity Laboratory, Agricultural Water Efficiency and Salinity Research Unit, Riverside, CA 92507-4617 USA
[j] School of Earth, Environment and Society, McMaster University, Hamilton, Ontario, L8S 4K1, Canada
[k] Programa de Pós-Graduaçao em Física Ambiental, Universidade Federal de Mato Grosso (UFMT), Cuiabá, Mato Grosso, Brazil
[l] University of Arizona. School of Geography, Development, & Environment, Tucson, AZ, 85721, USA
[m] USDA Agricultural Research Service, Global Change and Photosynthesis Research Unit, Urbana, IL 61801, United States
[n] Paul O'Neill School of Public and Environmental Affairs, University of Indiana, Bloomington, IN 47405, USA
[o] Faculty of Land and Food Systems, University of British Columbia, Vancouver, BC V6T 1Z4, Canada
[p] Department of Geography, University of Colorado, Boulder, CO 80309-0260, USA
[q] Department of Civil, Environmental & Geodetic Engineering, The Ohio State University, Columbus, OH, 43210, USA
[r] British Columbia Ministry of Forests, 499 George St. Prince George, BC, V2L 1R5, Canada
[s] School Of Biological Sciences, The University of Utah, Salt Lake City, UT 84112, USA
[t] Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK, 99775, USA
[u] Department of Ecology, Evolution, and Organismal Biology, Kennesaw State University, Kennesaw, GA, 30144, USA
[v] Dept of Geography and Atmospheric Science, University of Kansas, 1475 Jayhawk Blvd, Lawrence, KS, 66044, USA
[w] Centro de Investigacion Cientifica y de Educacion Superior de Ensenada, Baja California, Mexico
[x] Department of Anthropology and Environmental Dynamics Program, University of Arkansas, Fayetteville, AR 72701, USA
[y] Earth System Measurements & Data, Pacific Northwest National Laboratory, Richland, WA, USA
[z] The University of Michigan Biological Station, University of Michigan, Ann Arbor, MI, USA
[aa] Ecology and Evolutionary Biology Department, University of Michigan, Ann Arbor, MI, USA
[ab] Argonne National Laboratory, Environmental Science Division, 9700 S. Cass Ave., Lemont, IL 60439, United States
[ac] El Colegio de la Frontera Sur, Unidad Villahermosa, Tabasco, 86000, Mexico
[ad] Programa de Pós-Graduação em Ciências Ambiental, Universidade de Cuiabá (UNIC), Cuiabá, Mato Grosso 78065-900, Brazil
[ae] Earth and Environmental Systems Institute, College of Earth and Mineral Sciences, The Pennsylvania State University, United States
[af] Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, WI 53706, United States
[ag] US Dairy Forage Research Center, USDA Agricultural Research Service, Madison, WI 53706, United States
[ah] Department of Earth, Environment, and Physics, Worcester State University, Worcester, MA, 01602, USA
[ai] NEON Program, Battelle, 1685 38th Street, Boulder, CO 80301, USA
[aj] School of Forest, Fisheries and Geomatics Sciences. University of Florida, Gainesville, FL 32611, USA
[ak] Department of Botany, Program in Ecology and Evolution, University of Wyoming, Laramie, WY 82071, United States
[al] Department of Biological Sciences, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta, T1K 3M4, Canada
[am] British Columbia Ministry of Forests. 499 George St. Prince George, BC, V2L1R5, Canada
[an] Department of Environmental Sciences, University of Toledo, Toledo, OH, United States
[ao] USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO 80526, USA
[ap] 5 de Febrero 818 Sur, Col. Centro Cd. Obregon, Sonora, 85000, Mexico
[aq] USDA Agricultural Research Service Pasture Systems and Watershed Management Research Unit, University Park, PA 16802, United States
[ar] Department of Biology, Virginia Commonwealth University, Richmond, VA 23284, United States
[as] Department of Earth, Environmental, and Planetary Sciences, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106, USA
[at] Department of Soil, Water, and Climate, University of Minnesota, 1991 Upper Buford Circle, Saint Paul, MN 55108, United States
[au] Department of Physics and Atmospheric Science, Dalhousie University, Halifax, Nova Scotia, B3H 4R2, Canada
[av] GFZ Helmholtz Centre for Geosciences Potsdam, Germany
[aw] USDA Forest Service, Northern Research Station, Grand Rapids, MN 55744, United States
[ax] University of Central Florida, Orlando, FL 32816, USA
[ay] Department of Geography and Environmental Studies, Carleton University, Ottawa, Ontario, Canada
[az] Hokkaido Agricultural Research Center, National Agriculture and Food Research Organization, Sapporo, 0628555, Japan
[ba] International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, AK 99775, USA
[bb] Global Institute for Water Security, University of Saskatchewan, Saskatoon, SK, S7N 3H5, Canada
[bc] Institute for Resources, Environment and Sustainability, University of British Columbia, Vancouver, V6T 1Z4, Canada
[bd] Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, Vancouver, V6T 1Z4, Canada
[be] Department of Biology, West Virginia University, Morgantown, WV, 26506, USA
[bf] Judd Gregg Meteorology Institute, Plymouth State University, Plymouth, NH 03264, United States
[bg] Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695, USA
[bh] Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, 59717 USA
[bi] Department of Geography, the University of British Columbia, Vancouver, BC, V6T 1Z2, Canada
[bj] Department of Geography, McGill University, Montreal, QC H3A 0B9, Canada
[bk] Research Institute for Global Change, Japan Agency for Marine-Earth Science and Technology, Yokohama, 236-0001, Japan
[bl] School of Forestry, Northern Arizona University, Flagstaff, AZ 86001 USA
[bm] US. Geological Survey, Lafayette, LA 70506, United States
[bn] Louisiana Universities Marine Consortium (LUMCON), Chauvin, LA, 70344, USA
[bo] Institute of Soil Science, Center for Earth System Research and Sustainability (CEN), Universität Hamburg, Hamburg, 20146, Germany
[bp] Department of Civil & Environmental Engineering, Washington State University, Pullman, WA 99164, United States
[bq] Forest Ecosystems & Society, College of Forestry, Oregon State University, Corvallis, OR, 97331, USA
[br] Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Jena 07745, Germany
[bs] Institute of Alpine and Arctic Research, University of Colorado, Boulder, CO, 80301 USA
[bt] Biological Sciences, University of Texas at El Paso, 500W University Ave, El Paso, TX 79902, United States
[bu] AtmoFacts, Longmont, CO 80503, USA
[bv] Climate Research Division, Environment and Climate Change Canada, Victoria, BC, V8N 1V8, Canada
[bw] Information and Computational Sciences Department, The James Hutton Institute, Aberdeen, Scotland, AB11 6XJ, United Kingdom
[bx] School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

[by] Department of Ecology and Conservation Biology, Texas A&M University, College Station, TX 77843, United States
[bz] Department of Forestry & Environmental Conservation, Clemson University, Clemson, SC, 29634, USA
[ca] Institute of Environment and Department of Biological Sciences, Florida International University, Miami, FL 33199, United States
[cb] Department of Biology, San Diego State University, San Diego, CA 92182, United States
[cc] Department of Earth and Environmental Sciences, California State University East Bay Hayward, CA 94542, United States
[cd] Institute of Environment and Department of Earth and Environment, Florida International University, Miami, FL 33199, United States
[ce] USDA Forest Service, Northern Research Station, Durham, NH 03824, United States
[cf] Scientific Data Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States
[cg] Department of Environmental Sciences and Renewable Natural Resources, University of Chile, Santiago, Chile
[ch] USDA-ARS Northwest Sustainable Agroecosystems Research Unit, Pullman, WA 99163, United States
[ci] Instituto de Clima y Agua, CIRN, INTA, Buenos Aires, Argentina
[cj] Département de géographie, Université de Montréal, Montréal, QC H2V 0B3, Canada
[ck] Cold Regions Research Centre, Wilfrid Laurier University, Waterloo, ON, N2L 3C5, Canada
[cl] USDA, ARS, Delta Water Management Research Unit, Jonesboro, AR. 72401, United States
[cm] Northern Arizona University, School of Informatics, Computing, and Cyber Systems, and Center for Ecosystem Science and Society, Flagstaff AZ 86011, USA
[cn] Department of Plant and Agroecosystem Sciences, University of Wisconsin Madison, Madison, WI 53706, United States
[co] Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 45665, United States
[cp] Departamento de Agricultura y Ganadería, Universidad de Sonora, Hermosillo, 83000, Mexico
[cq] Center for Research on Sustainable Forests, University of Maine, 5755 Nutting Hall, Rm 263, Orono, ME 04469-5755, United States
[cr] Department of Ecology and Evolutionary Biology, The University of Arizona, Tucson, AZ, USA
[cs] Southwest Watershed Research Center, USDA-ARS, Tucson, AZ 85719, United States
[ct] USDA Agricultural Research Service, Cropping Systems & Water Quality Research Unit, Columbia, MO, 65211, United States
[cu] Center for Ecosystem Science and Society, Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, 86011 USA
[cv] Department of Soil, Water, and Ecosystem Science, University of Florida, Ona, FL, 33865 USA
[cw] British Columbia Ministry of Forests, Victoria, BC, V8V 1T7, Canada
[cx] Environment and Climate Change Canada, Downsview, ON, M3H 5T4, Canada
[cy] Department of Biological Sciences, University of Alabama, Tuscaloosa, AL, 35487, USA
[cz] Department of Agronomy and Horticulture, University of Nebraska, Lincoln, NE, USA
[da] Graduate School of Agriculture, Osaka Metropolitan University, Sakai, 599-8531, Japan
[db] School of Life Sciences, Arizona State University, Tempe, AZ, 85287, USA
[dc] Savannah River National Laboratory: Aiken, SC, United States
[dd] School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ 85287-8704, United States
[de] USDA-Agricultural Research Service, San Joaquin Valley Agricultural Sciences Center, Water Management Research Unit, Parlier, CA 93648-9757, USA
[df] Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, 20740, USA
[dg] Biospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, 20771, USA
[dh] Technical University of Denmark, Department of Environmental and Resource Engineering, Kongens Lyngby, Denmark
[di] USGS Water Resources Mission Area, 345 Middlefield Road, Menlo Park, CA 94025, United States
[dj] Institute of Research on Terrestrial Ecosystems (IRET) National Research Council (CNR), Via Castellino 111, Napoli 80131, Italy
[dk] Department of Biogeochemistry and Soil Quality, Norwegian Institute of Bioeconomy Research, ÅS, 1433, Norway

ARTICLE INFO

Keywords:
AmeriFlux network
Eddy covariance
Site uniqueness
Site clustering

ABSTRACT

Environmental observation networks, such as AmeriFlux, are foundational for monitoring ecosystem response to climate change, management practices, and natural disturbances; however, their effectiveness depends on their representativeness for the regions or continents. We proposed an empirical, time series approach to quantify the similarity of ecosystem fluxes across AmeriFlux sites. We extracted the diel and seasonal characteristics (i.e., amplitudes, phases) from carbon dioxide, water vapor, energy, and momentum fluxes, which reflect the effects of climate, plant phenology, and ecophysiology on the observations, and explored the potential aggregations of AmeriFlux sites through hierarchical clustering. While net radiation and temperature showed latitudinal clustering as expected, flux variables revealed a more uneven clustering with many small (number of sites < 5), unique groups and a few large (> 100) to intermediate (15–70) groups, highlighting the significant ecological regulations of ecosystem fluxes. Many identified unique groups were from under-sampled ecoregions and biome types of the International Geosphere-Biosphere Programme (IGBP), with distinct flux dynamics compared to the rest of the network. At the finer spatial scale, local topography, disturbance, management, edaphic, and hydrological regimes further enlarge the difference in flux dynamics within the groups. Nonetheless, our clustering approach is a data-driven method to interpret the AmeriFlux network, informing future cross-site syntheses, upscaling, and model-data benchmarking research. Finally, we highlighted the unique and underrepresented sites in the AmeriFlux network, which were found mainly in Hawaii and Latin America, mountains, and at under-sampled IGBP types (e.g., urban, open water), motivating the incorporation of new/unregistered sites from these groups.

1. Introduction

Environmental observation networks, such as the eddy-covariance flux measurement networks, are foundational for monitoring Earth's response to climate change, management, and natural disturbance (Baldocchi et al., 2001; Jones et al., 2021; Loescher et al., 2022; Novick et al., 2018). These networks are often built in a so-called "bottom-up" fashion, where individual investigators establish research sites based primarily on discrete research objectives, and then later, the sites are combined into cooperative observation networks, such as AmeriFlux (Baldocchi et al., 2024; Novick et al., 2018), MexFlux (Tarin-Terrazas et al., 2020; Vargas et al., 2013), and FLUXNET (Baldocchi, Falge, Gu, et al., 2001). These collaborative bottom-up networks can have various coverage and biased representations for different ecoregions and biome types within the networks (Pallandt et al., 2022; Villarreal et al., 2019). While the representativeness of sites is often not an issue for investigators doing research at individual sites, regional, continental, and global evaluations rely on available sites' data and are sensitive to bias in site locations. The assessments of existing observation networks also guided new site locations and future experimental designs (Malone et al., 2022; Pallandt et al., 2022; Papale et al., 2015; Sulkava et al., 2011). For example, a recent study suggested that redistributing sites could improve network representativeness from ~30–40 % to ~80–85 % across Latin America by adding additional optimally distributed sites across the undersampled environmental space (Villarreal and Vargas, 2021). In another undersampled environment, largely bottom-up emerging efforts are increasingly siting flux towers in urban locations (e.g., Davis et al., 2017) to study human activity. Often, these bottom-up networks consisted of groups of sites co-located within the same ecoregions across gradients of land cover/land use, plant functional types, management (e.g., forestry, agriculture), disturbances (e.g., wildfire), elevation, edaphic or hydrologic regimes (Chu et al., 2023; Stoy et al., 2023).

Advancing our understanding of the Earth's surface beyond the footprint of individual tower sites requires observational sites and their formed network to adequately represent the target domains (e.g., regions, continents, and terrestrial globe). Ecoregions have been proposed as a framework to design environmental observation networks where the Earth's surface is delineated into smaller, quantifiable, self-similar, and non-overlapping units based on biophysical properties (e.g., climate, soil, geology, potential vegetation type, possible land use) (Koenig, 1999; Omernik, 2004). Representative observation sites were then determined and established for each ecoregion. Similarly, the International Geosphere-Biosphere Programme (IGBP) was a global coordination effort to classify the Earth's land-surface into standard ecosystem types, helping to foster global- to regional-scale science through scaling individual sites (Loveland et al., 1999). This top-down approach was adopted for the National Ecological Observatory Network (NEON) and the Long-Term Agroecosystem Research (LTAR) network, enabling research efforts and investments to be distributed in a cost-, labor-, and time-efficient fashion (Bean et al., 2021; Jones et al., 2021). Ultimately, the strength of an observation network lies in both the individual sites, to address site-level questions, and in the representatives of the distribution of sites, in order to address larger-scale science.

Many attempts have been made to evaluate the networks' representativeness, often based on potential driver variables of fluxes (e.g., climate, soil, and plant functional type) (Hargrove and Hoffman, 2004; Hargrove et al., 2003; Sulkava et al., 2011; Sundareshwar et al., 2007; Villarreal et al., 2018). These studies assumed certain functional similarities that a site could represent the flux behaviors at other unmeasured areas if they shared similar function types—an assumption also adopted by most upscaling research (Jung et al., 2020; Xiao et al., 2010; Zheng et al., 2020). With this assumption, flux measurements at sparse locations can be extrapolated to regions, continents, and the globe by controlling for patterns in ecosystem structure and climate sensitivities. However, while ecoregions are often classified as discrete entities, the boundaries may be gradual, such as the gradient from forest to savanna to grassland, and may shift over time. Additionally, some ecoregions may be inherently heterogeneous (Kumar et al., 2023), making it incorrect to assume that all locations within an ecoregion will exhibit the overall (mean) properties of the whole, a phenomenon known as the ecological fallacy problem (Openshaw, 1984). While seemingly straightforward, determining how a single site represents a larger region often requires assumptions or simplifications, thus adding considerations when scaling.

With rich time series data (>107 h) collected at hundreds of AmeriFlux sites, temporal information can be utilized to quantify site similarities and directly test assumptions of representativeness. Different approaches have been proposed to harness the temporal information of the flux dynamics (Baldocchi, Falge, and Wilson, 2001; Falge, Tenhunen, et al., 2002; Hill et al., 2024; Mahecha et al., 2007; Stoy et al., 2009; Wilson et al., 2003). An emerging trend in the literature is that measured fluxes at distinct locations can exhibit common behavior across time and space, suggesting that a site can be representative of spatially proximal locations to some degree (Hollinger et al., 2004; Poe et al., 2020; Post et al., 2015). A significant fraction of this temporal coherence occurs at the diel and seasonal scales, highlighting the influence of climatology, seasonality, and phenology on ecosystem fluxes (Stoy et al., 2009). With ecosystem fluxes strongly connected to environmental conditions, studies have shown that the functional behavior of ecosystem fluxes can be independent of vegetation type and climatic region, with an example of high-latitude ecosystems might behave similarly to tropical ecosystems under similar environmental conditions (Krich et al., 2021), and that spatial coherence can be on the order of 400 km based on the model-data assessment (Hilton et al., 2013). Across the upper Midwestern U.S., Poe et al. (2020) demonstrated that a cluster of sites from multiple ecosystem types has high temporal and spatial coherence, implying that flux information can be scalable among sites and across landscapes (Reed et al., 2021). Although there is currently a large amount of data available, relatively little work has been done on quantifying how similar (or dissimilar) data from sites are to each other at a network scale, and it is largely assumed that sites within ecoregions are similar and data from across ecoregions are dissimilar.

Here, we propose a novel approach to quantitatively examine flux sites' similarity of measured ecosystem fluxes by extracting key time series characteristics—diel and seasonal dynamics. Harmonic analyses will be performed on the extracted time series features to explore their similarity and potential grouping among sites using hierarchical clustering. We will compare the clustering results along climatological gradients and across commonly adopted ecosystem classifications—ecoregions and IGBP vegatation types—to assess the degree of uniqueness of these clusters. This study addresses the following questions: 1) Do sites from the same ecoregion or IGBP types share similar time series characteristics? Here, we hypothesize that sites within the same ecoregion (or IGBP type) have higher similarity than sites across groups. 2) Is the similarity of time series characteristics among sites a function of spatial distance? We expect the extent of time series similarity to be distance-limited. 3) Within the AmeriFlux, which sites, ecoregions, or IGBP types show distinct time series characteristics, or are currently under-represented? Here, we expect ecoregions in Latin America and high latitudes to be relatively poorly represented. Solutions to these questions help disseminate the representativeness of existing sites, identify underrepresented locations, allow for model-data benchmarking of site selections, and thus establish a framework for using existing site locations for cross-site examinations that can help interpret patterns in the response of ecosystems to climate change, disturbance events, and anthropogenic management.

2. Methods and materials

2.1. Flux observations and composite time series

Land-atmosphere exchange (i.e., flux) observations are made using the eddy-covariance methodology (Baldocchi et al., 1988). A coalition of researchers measure carbon, water, energy, and momentum fluxes in ecosystems across the Americas and share these data via AmeriFlux to facilitate large-scale meta-analysis (Novick et al., 2018). Flux and supporting environmental data were retrieved from the AmeriFlux BASE data product (accessed September 2023) (Chu et al., 2023). Starting with all sites with data available, we first down-selected sites based on the data records (i.e., $\geq$ 3 years, allowing for robust extraction of time series features) and availability of the core flux variables of the eddy covariance methodology, including net ecosystem exchange (NEE) of carbon dioxide (CO_2), sensible heat flux (H), latent heat flux (LE), and friction velocity (USTAR, i.e., a measure of momentum flux). The study included 343 sites with at least one flux variable available, with 191 sites having at least 6 years of data. While focusing on the flux variables, we also included in the analyses the main environmental variables used to explain fluxes, including net radiation (NETRAD), air temperature (TA), vapor pressure deficit (VPD), and soil water content (SWC). NETRAD was used instead of photosynthetically active radiation (PAR) or shortwave incoming radiation (SW_IN) since more sites report NETRAD (~88 %) compared to PAR (~77 %) or SW_IN (85 %). Our preliminary tests showed similar results using either of these three radiation variables. The top level of tower height for TA or VPD and of soil column depth for SWC are used for further analyses if a site has multi-level measurements. If not provided, VPD was calculated from TA and relative humidity (RH) (Monteith and Unsworth, 2008). No gap-filled data were used. While our analyses focused on directly measured observations mentioned above, we conducted supplemental analyses using the derived ecosystem respiration (RECO) and gross primary production (GPP) for a subset of sites (186) with the available AmeriFlux-FLUXNET data product (accessed October 2024). This supplemental analysis aimed to explore the sensitivity of using NEE against GPP/RECO. Still, our study mainly focused on the large group sites (>300) with the core flux variables (i.e., NEE, LE, H, USTAR) (Table 1). All variables mentioned above were reported at a half-hourly or hourly resolution. For readers interested in details of the AmeriFlux BASE and FLUXNET data products, please refer to Chu et al. (2023) and Pastorello et al. (2020).

Ancillary data, including sites' geolocations, IGBP land-cover types, and instrument heights, were retrieved from the AmeriFlux BADM (Biological, Ancillary, Disturbance, and Metadata) data product (AmeriFlux Management Project, 2020). The IGBP type was a remote-sensing-based land classification that mainly considered the current land cover/use, plant functional types, and phenology (Loveland et al., 1999), i.e., an atomistic classification. Level I Ecoregions generated by the Commission for Environmental Cooperation were extracted based on the sites' geolocations (Commission for Environmental Cooperation, 1997; Griffith et al., 1998; Omernik, 2004). The ecoregion delineation was based on Omernik's framework (Omernik, 1987), mainly considering the climate zones, land-surface forms, potential natural vegetation, and soil types, i.e., a holistic classification. The study sites span 16 IGBP types and 20 ecoregions (see Tables S1-S2 for a site list and data citation and Table S3 for a breakdown of IGBP and ecoregion groups). Last, we used 30-year mean temperature and precipitation from Climatic Research Unit (CRU) data to represent each site's long-term climatic condition (Harris et al., 2020).

All variables from the BASE data product were first filtered by their respective expected plausible ranges (e.g., −50 °C–50 °C for air temperature). If NEE was absent, we calculated it from turbulent CO_2 flux (FC) and storage flux (SC). Nighttime FC was first filtered using USTAR thresholds on a per-site basis, using the REddyProc library (Reichstein et al., 2005; Wutzler et al., 2018). If not provided, storage flux was assumed to be negligible at short-vegetation sites and calculated from single-level CO_2 concentration at tall-vegetation sites (Papale et al., 2006). No gap-filled data were used to extract information from the observation record unbiasedly (Vekuri et al., 2023). At this point, the data were treated as original time series from each site and ready for aggregation (Fig. 1a). For GPP and RECO, we used GPP_NT_VUT_REF and RECO_NT_VUT_REF variables from the AmeriFlux FLUXNET data product. NT, VUT, and REF refer to nighttime-based partitioning, variable USTAR threshold, and reference selected based on model efficiency (Pastorello et al., 2020). A brief summary of flux partitioning can be found in the Supplementary Text S1.

At each site, each variable was aggregated into a single annual composite time series. This time series consists of data from multiple years, consisting of diel-seasonal frequencies (Figs. 1b and S1), given that diel and seasonal scales contain the most temporal information of our target variables (Chu et al., 2023; Poe et al., 2020; Stoy et al., 2009). All AmeriFlux data were reported in local standard time, and southern hemisphere sites were not shifted temporally to match northern hemisphere norms. We first upscaled the original data from 30-minute to a 2-hour resolution (e.g., 0:00–2:00 h). We then separated data from all available years into 52 non-overlapping periods of 7 days (e.g., Jan 1st-7th), with any remaining days (e.g., December 31st) combined into the last period. Next, we calculated the median diel variation for each 7-day period at 2-hour timesteps, noting that each 7-day period contains data from multiple (≥3) years of data. Last, we constructed a new characteristic composite diel-seasonal time series of each 7-day period over the course of a single year (i.e., 12 timesteps/day × 52 periods = 624 timesteps). Any gaps in the composite time series only exist when a site had missing data larger than a composite window size (e.g., all data from 0:00–2:00 h, from January 1st-7th were missing from all years). These were then filled through interpolation within each window and across the windows. If the interpolation filled >20 % of gaps within each time window or across windows, we dropped a site variable from further analyses. This composite time series ultimately reduced the large time series (see gray points in Fig. 1b) for further clustering analyses (see the red line in Fig. 1b).

For the preliminary tests, we ran similar analyses based on composite time series aggregated at 15-day/1-hour, 7-day/2-hour, and 5-day/3-hour (window length/temporal resolution) scales to assess the sensitivities of time series aggregation. The results suggested the aggregation windows/resolutions changed the site pairs' distances marginally and potentially the clustering at a small scale (e.g., proximate sites) but did not impact the overall results substantially. We also tested the sensitivities of data record lengths by using sites with >10 years of available data (~110 sites). We compared the results against those generated from the middle 3-year and 6-year periods of the original record. We found that the record lengths had only marginal influence on the site pairs' distances'. Finally, we briefly tested the impact of gap-filling on results for NEE, LE, and H fluxes using the AmeriFlux FLUXNET data product. We found generally good linear relationships between site pairs' distances from non-gap-filled and gap-filled data. Details of these preliminary tests were discussed in Supplementary Text S2 (Figures S1, S6-S10). Unless specified, all following results and discussions focused on the 7-day/2-hour aggregation, using the full data record from each site, and using non-gap-filled data from each site.

We interpreted the characteristic composite time series as a multi-year climatological representation of a site's diel-seasonal dynamics. The extracted temporal features, particularly the phases, period, and amplitudes (Falge et al., 2002), reflect an ecosystem's phenology (e.g., growing/dormant seasons, wet/dry periods) and ecophysiological potential (e.g., maximum daytime CO_2 uptake vs. nighttime respiration, daytime vs. nighttime evapotranspiration). While conceptually similar to other frequency-based methods (e.g., Fourier, wavelet), our proposed method does not impose a specific shape function (e.g., sinusoidal wave, mother wavelet) in constructing the composite time series. Thus, our composite time series can capture irregular or asymmetric diel/seasonal patterns (e.g., afternoon water stress (Vickers et al., 2012)) or abrupt transitions (e.g., monsoon onset, Scott et al., 2009). Practically, this climatological composite time series averaged out interannual variability, smoothed out the random noise around high-frequency sampling error (Moncrieff et al., 1996), filled a portion of the data gaps in the

Table 1
A summary of sites used in univariate, multivariate analyses, and the calculations of the Harmonic Uniqueness Parameter. Sites needed ≥ 3 years of data from flux variables (NEE, LE, H, USTAR) to be included.

Variable	Number of Sites Used		
	Univariate Analyses	Multivariate Analyses	Harmonic Uniqueness Parameter
NEE	313	313 (All NEE, LE, H, USTAR required)	313 (All NEE, LE, H, USTAR required)
LE	332		
H	336		
USTAR	338		
NETRAD	287	Not Used	Not Used
TA	328		
VPD	326		
SWC	215		
GPP	186		
RECO	186		

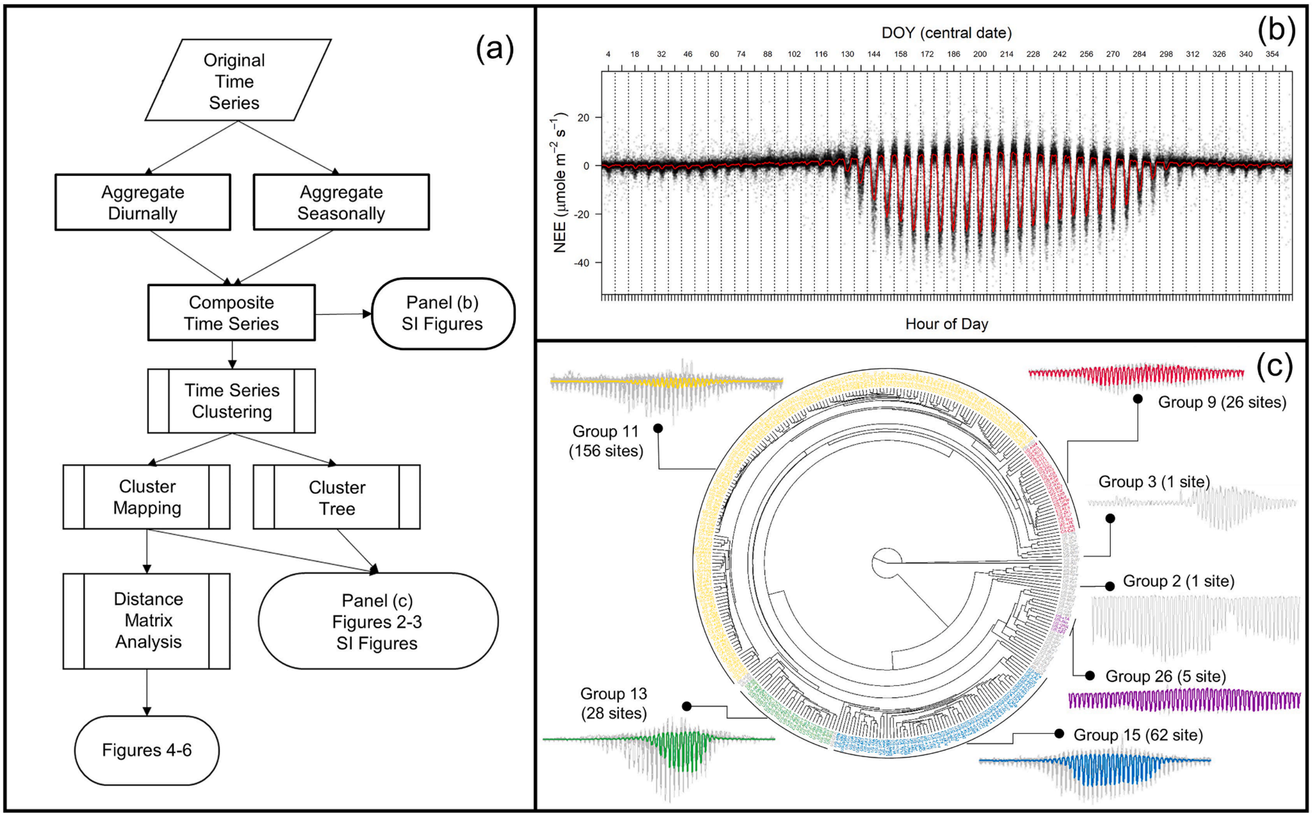

Fig. 1. Conceptual figure of the methodical framework (a). Panel (b) is an example of the composite time series of NEE from the US-Oho site—a deciduous broadleaf forest in the Eastern Temperate Forests (See Figure S1 for an enlarged version). Gray points and red lines denote the original and composite time series. The composite time series captured the dormant/growing seasons and the diel amplitudes of daytime uptake and nighttime respiration. Panel (c) shows an example of a radial tree diagram (i.e., dendrogram) of site clustering, and subplots show composite time series from the selected groups (See Figures S27 and S28 for enlarged versions).

original records (Falge et al., 2001), and reduced the computational time of further clustering analyses (Aghabozorgi et al., 2015; Cheng and Adepeju, 2014).

2.2. *Time series analysis and hierarchical clustering*

To quantify the similarities among the composite time series, we computed the dynamic time warping (DTW) distances for each variable among all available sites (see Figures S2-S5 for example). DTW measures the distances between any two given time series, focusing on the generalized shape of the curve while allowing moderate compression or stretching, following Giorgino (2009):

$$D(x,y) = min\left(d_\phi(x,y)\right) \quad (1)$$

where $D(x,y)$ is the DTW minimum distance (d_ϕ) between the time series x and y of a given variable from two sites, given that DTW methods allow for distortion in the time series. In other words, due to the warping nature of DTW, where a time series can be stretched or compressed while being evaluated against another time series, the DTW distance is the difference between the x and y data with the least amount of deformation applied to the time series. Given that the time series was prepared with the same lengths and resolutions, the DTW distances mainly reflect the difference in diel/seasonal amplitudes and phases. The *proxy* R package was used to extract the DTW distances (Buchta, 2022). We interpreted the DTW distances as a measure of similarity between two sites in the target variable's climatological diel/seasonal dynamics, e.g., any deviation in seasonal phases or differences in amplitudes. To allow comparison between variables, we normalized DTW distances using the minimum and maximum DTW distances of each variable across all site pairs:

$$D_v^*(x,y) = \frac{D_v(x,y) - min\left(D_v(x,y)\right)}{max(D_v(x,y)) - min(D_v(x,y))} \quad (2)$$

where v represents the variable, including NEE, H, LE, USTAR, NETRAD, SWC, TA, and VPD.

We then adopted hierarchical clustering to construct the hierarchy of site clusters (i.e., clustering trees, Fig. 1c) based on the sites' DTW distances. Hierarchical clustering assumes all sites as a single cluster initially and consecutively splits it into separate clusters and, ultimately, the end nodes of each individual site (Aghabozorgi et al., 2015). The radial length of a branch after a split indicates the similarity between the split branches' end nodes (sites), with shorter branch lengths indicating more similarity. We chose this approach because it retains a relative idea of general similarity between clusters/sites, i.e., the longer a radial branch after a split indicates a relatively larger degree of difference or divergence. To aid the interpretation, we trimmed the clustering trees to the optimal number of clusters determined based on the performance of six cluster validity indices (Arbelaitz et al., 2013), searching between 30 and 75 clusters (Supplementary Text S3). Unless specified, we focused on the optimized clustering groups in further sections. Still, the untrimmed clustering trees were presented (e.g., Fig. 1c). For better presentation, clustering groups with fewer than five sites were not color-coded separately in the clustering trees and maps. However, it is worth noting that these small groups represented sites that were unique and relatively different from others (e.g., groups 2–3 in Fig. 1c). The analyses above were carried out for each of the eight target variables (i.e., univariate) for all available sites (i.e., 215 (SWC) to 338 (USTAR)) and then for all flux variables combined (i.e., multivariate, 313 sites).

Additionally, we ran similar univariate analyses for 186 sites with GPP and RECO data (Table 1). The *dtwclust* and *ape* R packages were used in the clustering analyses and tree generation mentioned above (Paradis and Schliep, 2023; Sarda-Espinosa, 2023).

We further examined the dependency of site pairs' DTW distances by IGBP group, ecoregion, and spatial proximity. We compared the DTW distances across sites within and across IGBP groups or ecoregions using the analysis of variance (ANOVA) and Tukey *post hoc* tests (Sokal and Rohlf, 1995). We also examined the dependency of site pairs' DTW distances on the sites' spatial proximity using a moving average lowpass filter. The above tests explored whether sites in the same IGBP groups, ecoregions, or spatial proximity had lower DTW distances than those across groups or far apart. Lastly, to provide a multivariate similarity, we calculated the Harmonic Uniqueness Parameter (HUP) for each site as the mean of the normalized DTW distances to all other sites:

$$HUP(x) = \frac{1}{N_V}\left(\sum\nolimits_{v=1}^{N_V}\frac{1}{N_Y}\left(\sum\nolimits_{y=1}^{N_Y} D_v^*(x,y)\right)\right) \tag{3}$$

where N_Y and N_V denote all other sites and all target variables, i.e., NEE, LE, H, and USTAR, respectively (Table 1). In theory, HUP can range from 0 to 1, but in practice, most values fall between 0.15 and 0.35, with a higher HUP value indicating a site's relatively higher uniqueness.

3. Results

3.1. *Site clustering*

Environmental variables (e.g., TA, NETRAD) showed site clustering mainly reflecting the expected climate gradients, primarily but not exclusively evident along the latitudinal gradient (see Figs. 2a and S9a for maps, Figures S11 and S15 for composite time series, and Figures S12b and S16b for clustering trees). VPD also showed changes with latitude, along with small clusters in separate geographic regions. For example, there was a cluster of 17 sites mostly in Florida (light purple, Figures S19-S20) and a cluster of 22 sites primarily from the Northwestern Forested Mountains (dark purple). Several clusters emerged for about 28 sites in the North American Deserts and Southern Semi-Arid Highlands, featuring high spring VPD declining after the summer monsoon (Groups 1–15, Figures S19-S21). For SWC, there was

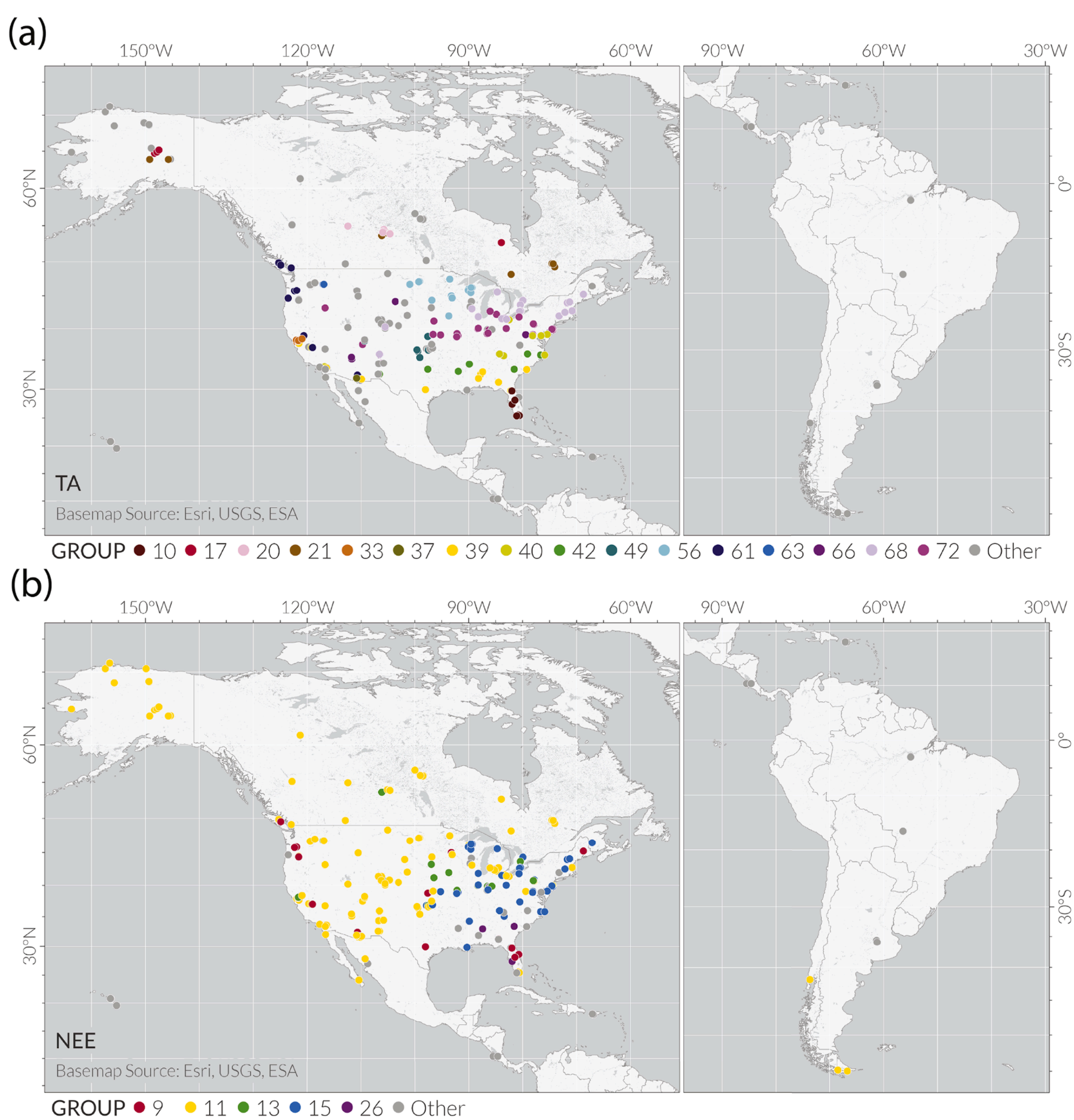

Fig. 2. Geographic locations of AmeriFlux sites clustered based on TA (a) and NEE time series (b). Color codes indicate the sites' clustering groups. For simplicity, all groups with fewer than five sites are shown in gray. See Figures S16a, S20a, S24a, S32a, S36a, S40a, S44a, and S48a for similar maps using the NETRAD, VPD, SWC, H, LE, USTAR, GPP, and RECO time series and an interactive map at https://cartoscience.users.earthengine.app/view/flux-networks.

no clear large-scale geographic or ecoregion-specific pattern (Figure S24a), reflecting the cross-site heterogeneity (e.g., ecohydrology, soil properties, local topography, measurement depth, etc.). In sum, clustering based on environmental variables generally reflected the sites' climate and ecoregions, demonstrating the general robustness of the clustering methods. Latin America had a sparse, relatively low site density. Hence, sites from those regions were often clustered into small groups.

The univariate clustering based on flux variables (i.e., NEE, LE, H, USTAR) showed more uneven distributions in the clusters compared to the environmental variables. Often, around 30–50 sites were grouped into 20–40 small groups with only one or a few sites, while the rest were grouped into 1–2 large groups (>100 sites) and a few intermediate-size groups (15–70). To some extent, the uneven clustering reflected the unevenly distributed nature of the AmeriFlux network, and thus, sites from underrepresented IGBP groups (e.g., urban, open water), ecoregions (e.g., Hawaii, ecoregions in Latin America), geolocations (e.g., mountainous areas), or disturbance/management regimes (e.g., alfalfa (periodic harvest), rice paddy (seasonal flooding)) had distinct diel-seasonal dynamics in the fluxes and were clustered into unique, small groups. The rest of the sites, while varying marginally or moderately among each other, were grouped into a few large or intermediate clusters. For example, NEE had one large cluster of 156 sites across multiple ecoregions and IGBP groups (Group 11, Figs. 2b, S21, Tables S6-S7). Group 15 (62 sites) was primarily deciduous broadleaf forests (41.9 %) and grasslands (17.7 %) in the Eastern Temperate and Northern Forests. Group 9 (28 sites) was nearly all croplands (85.7 %). In contrast, 32 groups contained only 1–3 sites, mostly from underrepresented IGBP groups and ecoregions. About 186 out of the 313 sites with NEE had partitioned GPP and RECO for our analyses. The GPP and RECO clustering showed a larger proportion of sites within the network were classified as unique amid a smaller number of sites within each group (Figures S44 and S48). RECO has two groups with 14 % of sites, while GPP has groups of 28 % and 12 % of sites. In general, GPP and RECO showed a relatively even cluster distribution compared to using NEE. Also, GPP and RECO showed more distinct annual mean fluxes among groups than NEE (Figure S53).

For LE, ~67 % of the sites were clustered into two large groups of 114 and 110 sites across North America, except for the southeastern region (Figure S32, Tables S6-S7). Around 50 sites, mostly from the Eastern Temperate Forests and Great Plains, were clustered into 3 intermediate-size groups. For H, ~71 % of the sites were clustered into two large groups of 121 and 119 sites across North America, except for the southwestern region (Figure S36, Tables S6-S7). Eighteen sites, mainly from the Northern American Deserts and Southern Semi-Arid Highlands, were grouped together (dark purple, Figure S36), and 14 sites were grouped together largely from the Northwestern Forests Mountains and Mediterranean California (blue, Figure S36). For USTAR, ~51 % of sites were clustered into one large group of 174 sites across multiple ecoregions (Figure S40). The LE, H, and USTAR clusters generally reflected the differences in the annual mean fluxes (Figure S53).

The multivariate clustering based on all flux variables showed a relatively even cluster distribution featuring multiple intermediate-size groups (Figs. 3–4, Table S8-S9). The largest Group 59 (70 sites) spanned multiple ecoregions and IGBP classifications (dark purple, Fig. 3). This group is characterized by low annual mean LE, H, and NEE (dark purple, Fig. 4). Group 60 (62 sites) included mostly grasslands (35.5 %) and croplands (41.9 %) in the Eastern Temperate Forests and Great Plains, featured by moderate LE, low H, and various NEE (light purple Figs. 3–4). Group 55 (32 sites) was located in North American Deserts, Mediterranean California, and Southern Semi-Arid Highlands, with high H but low NEE and LE (light blue Figs. 3–4). Groups 29 and 31 (19 and 16 sites) were dominated by deciduous broadleaf and evergreen needleleaf forests across Eastern Temperate and Northern Forests, with high USTAR and moderate to high NEE (orange and yellow, Figs. 3–4). Notably, all of the groups above had at least one NEON or AmeriFlux Core site in each group (Fig. 3b), suggesting these long-term sites generally represented the (current) major clusters of the network. It is also worth mentioning that several known adjacent sites over similar ecosystems (e.g., US-Ha1/US-xHA, US-Bar/US-xBR, US-Kon/US-xKZ, US-EML/US-xHE) were clustered into the same groups, demonstrating the robustness of the site clustering. Last, 84 sites were clustered into 58 small groups (gray, Fig. 3). These sites were mainly from currently underrepresented IGBP groups (e.g., urban, open water, snow/ice), ecoregions (e.g., Hawaii, and ecoregions in Latin America), geographical features (e.g., mountain, floodplain), and disturbance/management regimes (e.g., periodic harvest, wildfire, prescribed burning, forest plantations, restored wetlands).

3.2. Site similarity

The connection between clustering results and climate is explored using mean air temperature and precipitation at the site level (Fig. 4k and S51). Through this lens, sites are more likely to be included in the largest NEE cluster (Group 11) if they are drier, with annual precipitation of <1000 mm, while sites more likely to be included in smaller, more unique clusters have typically hotter and wetter climates (Figure S54a). A similar clustering is seen in LE (Figure S54b) with a threshold of 1000 mm precipitation and of <10 °C for TA, while there were larger amounts of overlap between clusters in climate space for H and USTAR (Figures S54c-S54d). When sites are clustered based on all flux variables (NEE, LE, H, USTAR, Fig. 4k), there are large clusters in the dry and cold (<1000 mm, and <10 °C) and also dry and hot (<1000 mm, and >10 °C) climate spaces, while small and unique clusters again are more likely in hot and wet climates.

We further examined the normalized DTW distance—a measure of similarity in the temporal dynamics between site pairs—across sites within and among the selected ecoregions and IGBP groups (Figs. 5 and S52). Environmental variables generally showed lower DTW distances (i.e., similar) within the same ecoregions, except for NETRAD in Northwestern Forested Mountains, VPD in North American Deserts and Mediterranean California, and SWC in Eastern Temperate Forests and Mediterranean California (Figure S55e-h). For IGBP groups, shrublands and wetlands tended to have higher inter-site DTW distances than all others (Figure S55a-d). It should be noted that SWC was not always measured at wetland sites, particularly at the wetter/inundated locations, potentially biasing the current results. Flux variables generally showed lower DTW distances within the same ecoregions and IGBP groups (Fig. 5). However, several groups had high DTW distances closer to or even higher than the cross-group DTW distances. For example, sites in Mediterranean California showed much higher inter-site variability for NEE, LE, and H than other groups, as did sites in Eastern Temperate Forests for NEE and LE, Northern Forests for USTAR, and Northwestern Forested Mountains for H and USTAR. For IGBP groups, croplands had higher DTW distances for NEE and LE than the others. Wetlands had higher DTW distances for LE, shrublands had higher DTW distances for H, and evergreen needleleaf forests had higher DTW distances for USTAR. In sum, while sites in the same IGBP groups and ecoregions generally showed similar flux dynamics (lower DTW distances), the above exceptions highlighted the inter-site variability, potentially as high as that across groups, for certain ecoregions and IGBP groups.

When examining the site pairs' similarity against the spatial distance, most sites showed high similarity in adjacent locations and became distinct with an increased distance as expected (Figs. 6 and S46-S52). Yet, the dependence of similarity on spatial distance differed by variable and IGBP group, and in some cases, bimodal behavior was apparent. One example of bimodal site-pair similarity is NEE for Eastern Temperate Forests (Fig. 6 g), highlighting how most sites are similar within that ecoregion, but certain site pairs show dissimilar NEE, even at relatively close distances across the landscape. Another example of difference among IGBP groups is the larger amount of variability in NEE

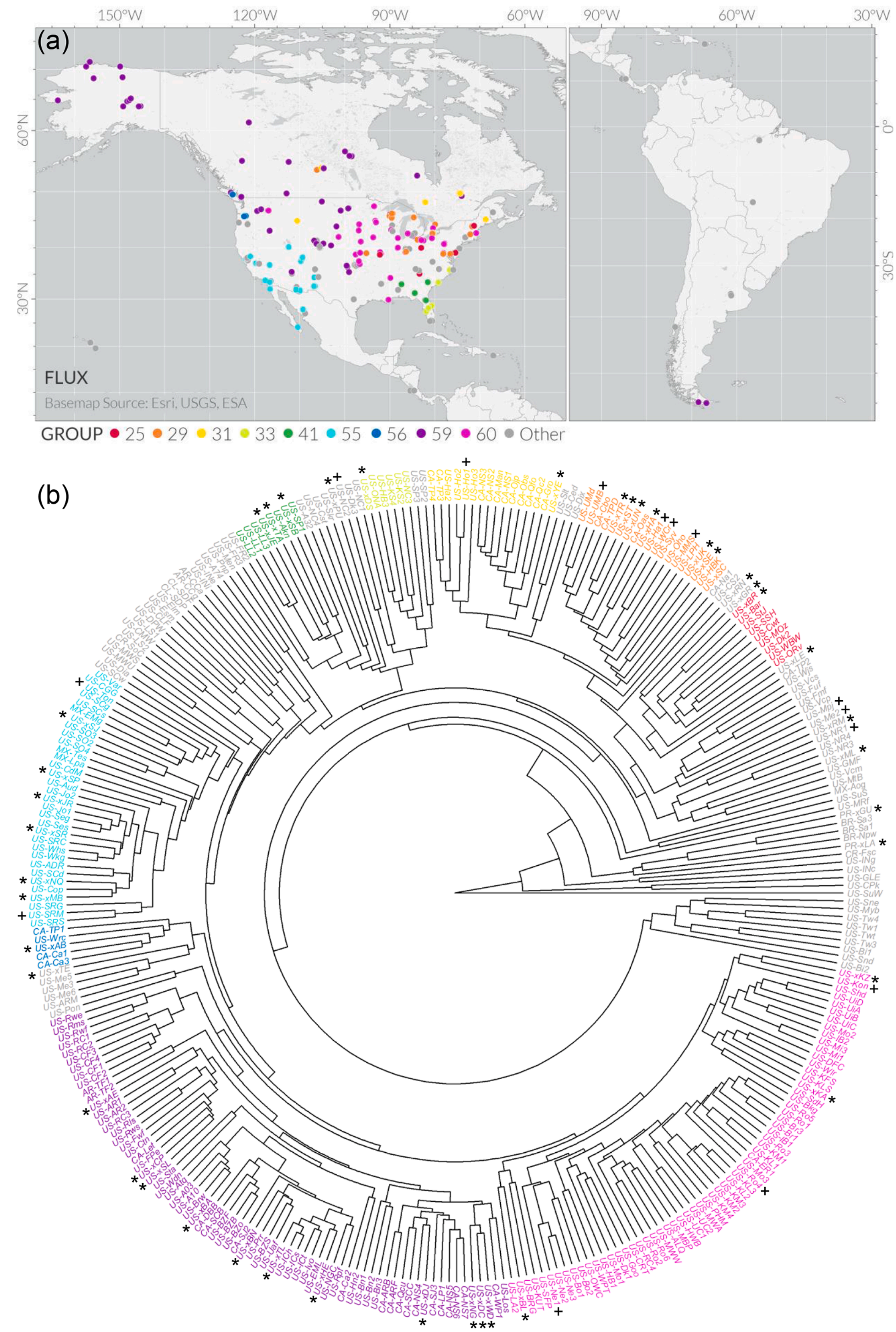

Fig. 3. Site locations (a) and radial tree diagram (b) of AmeriFlux sites clustered using all flux time series (i.e., NEE, LE, H, USTAR). Color codes in both panels indicate the sites' clustering groups. For simplicity, all groups with fewer than five sites are in gray colors. Each end node in panel (b) represents an AmeriFlux site specified by its Site ID, marked with asterisks (*) and plus signs (+) for the NEON and AmeriFlux Core sites. The radial distance at which a branch is split indicates the similarity of the split branches' end nodes (sites). For example, a split closer to the center (origin) suggests that the divided branches (and their end nodes) are very dissimilar. See Figures S43-S44 for similar tree diagrams highlighted by ecoregion and IGBP classification.

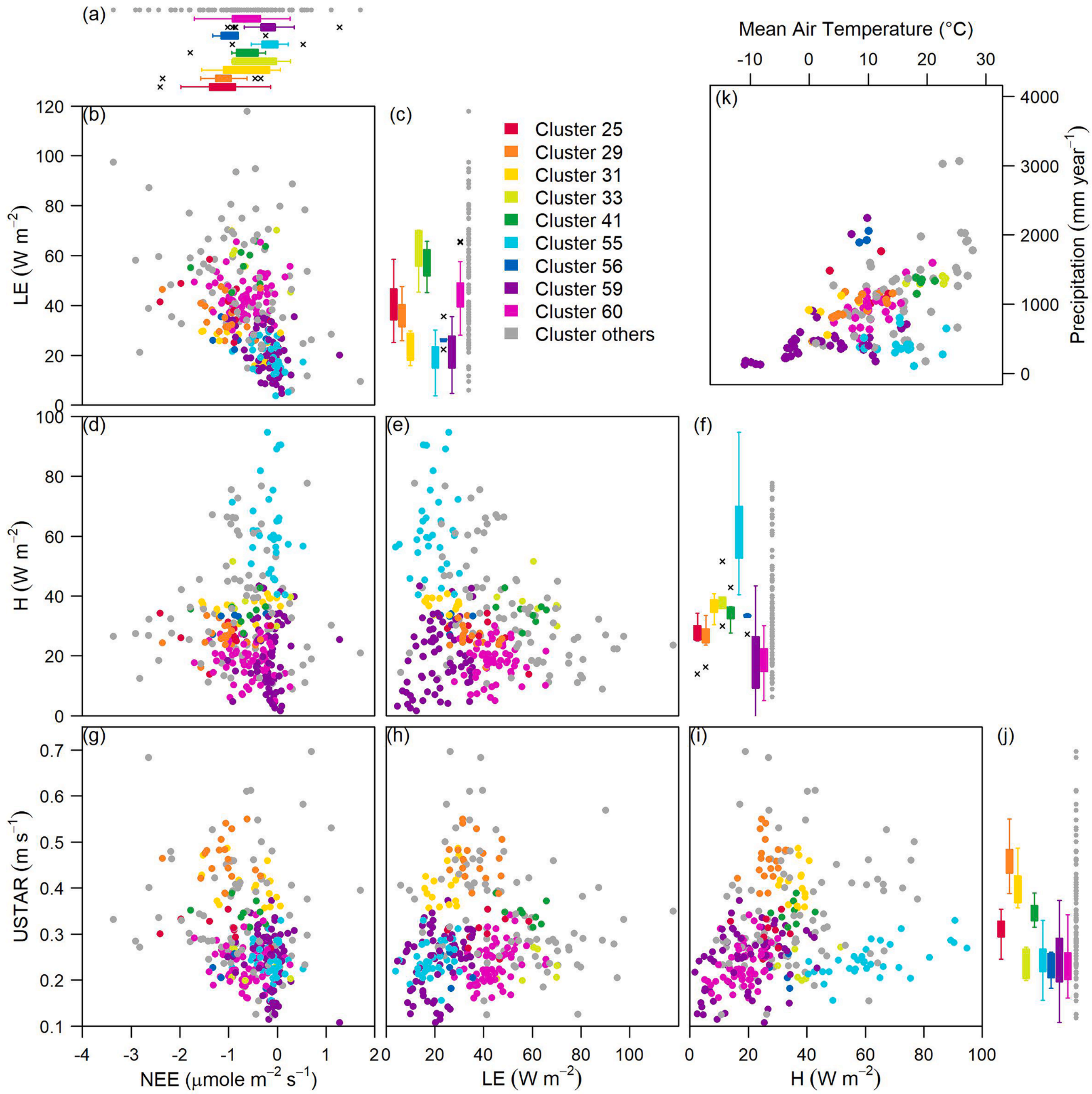

Fig. 4. Distribution of AmeriFlux sites by (a-j) annual mean fluxes and (k) mean air temperature and annual precipitation. Color codes denote clusters using all flux variables in all panels, similar to Fig. 3. Figures b, d, e, g, h, and i are scatterplots between annual mean NEE, LE, H, and USTAR. Each dot represents an AmeriFlux site. Figures a, c, f, and j show boxplots of annual mean fluxes by clusters. Boxes and whiskers showed each cluster's interquartile ranges (IQR, 25th-75th percentiles) ± 1.5 * IQR. Figure (k) uses the Climatic Research Unit (CRU) time series (TS) 4.05 gridded dataset from 1981 to 2020 (0.5 × 0.5°). For better presentation, the figure scales were zoomed in to focus on the main groups. A few sites with extremely large/small NEE (US-Bi2, US-Inc, US-Ing, US-SuS, US-SuW), H (AR-TF1, AR-TF2, US-BMM, US-ICs, US-xNW), and USTAR (US-GLE) were not shown.

in evergreen needleleaf forests (Fig. 6b) compared to deciduous forests (Fig. 6d). In general, LE and NETRAD showed an apparent decrease in similarity with distance (Figures S49 and S52), implying site pairs are more likely to show divergent behavior for those variables across the landscape. In contrast, USTAR, VPD, and SWC showed a weak relationship between the sites' similarity and distance (Figures S58, S61-S62), with a larger variability between sites at close spatial scales. Differences between IGBP classifications exist, particularly for NEE in forests and croplands.

The Harmonic Uniqueness Parameter revealed sites that were relatively unique and distinct from others in the current network (Fig. 7). At the continental scale, sites in Hawaii and most ecoregions in Latin America showed a relatively high HUP, reflecting their current underrepresentation in the AmeriFlux network. However, at a finer scale, several sites from specific IGBP groups, mountains, or management regimes also revealed high HUP (>0.24). For example, around 13 mountainous sites from the Northwestern Forested Mountains (e.g., US-CPk, US-GLE, US-MtB) and Mediterranean California (e.g., US-SO2, US-SO3) showed their relative uniqueness. Urban (e.g., US-INc) and open water (e.g., US-Pnp, US-Men) sites, with distinct flux dynamics, had high HUP. Lastly, several managed forests from the Eastern Temperate Forests (e.g., US-NC2, US-SP2, US-Cst) and croplands from Mediterranean

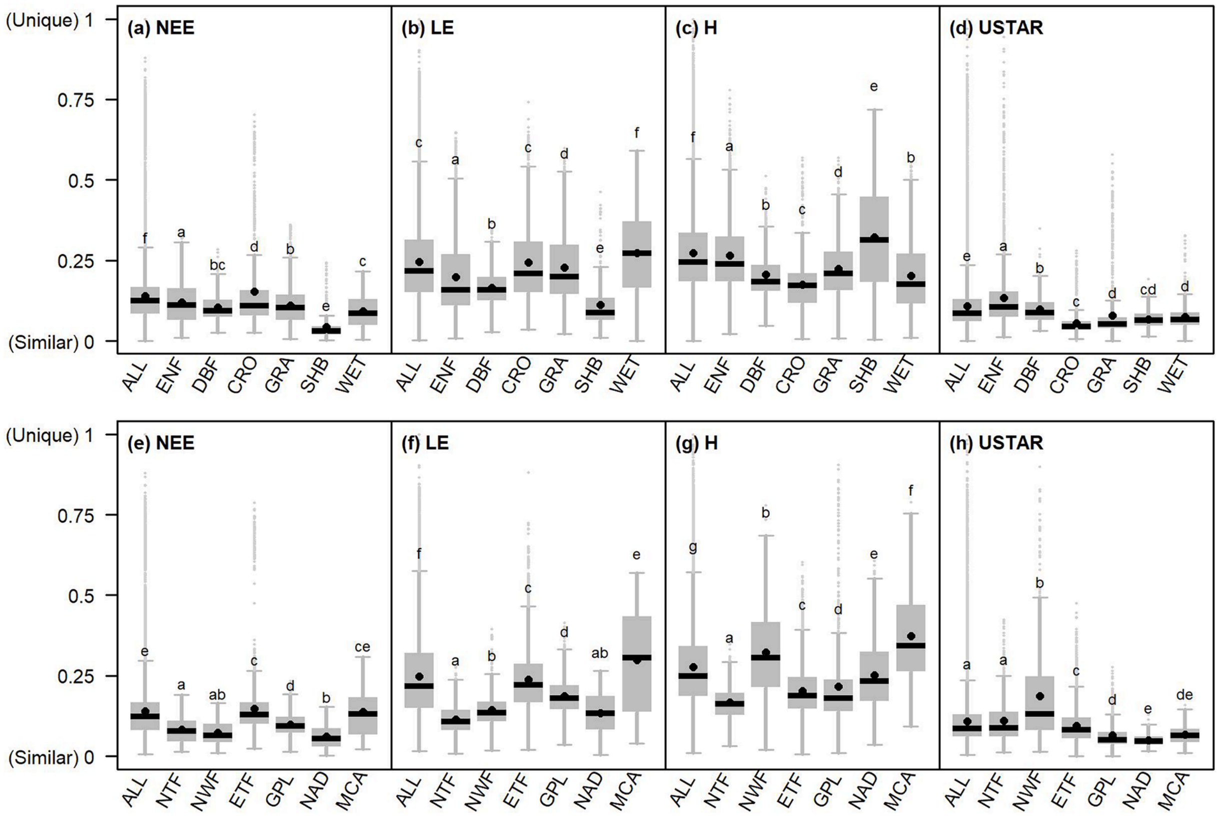

Fig. 5. The box plots of normalized dynamic time warping (DTW) distances for flux variables among selected IGBP groups (a-d) and ecoregions (e-h). Only groups with sufficient samples (>20 sites) were included. Black dashes and circles indicated each group's medians and means. Gray boxes and whiskers showed each group's interquartile ranges (IQR, 25th-75th percentiles) ± 1.5 * IQR. Different annotated letters indicate significant differences among groups. Abbreviations: ALL: all site pairs except those in the same group, ENF: evergreen needleleaf forest, DBF: deciduous broadleaf forest, CRO: cropland, GRA: grassland, SHB: open and closed shrubland, WET: wetland, NTF: Northern Forests, NWF: Northwestern Forested Mountains, ETF: Eastern Temperate Forests, GPL: Great Plains, NAD: North American Deserts, MCA: Mediterranean California. See Figure S55 for a similar figure for environmental variables.

California (e.g., US-Tw3 (alfalfa) and US-Twt (rice)) also showed high HUP.

4. Discussion

4.1. Site similarity and dependency

In this work, we show sites across the AmeriFlux network are generally similar to each other when compared to the same ecoregion or IGBP type, as opposed to sites from other ecoregions or IGBP type. However, in some cases, such as LE fluxes in wetlands and in the Mediterranean ecoregions, sites within those groups are more dissimilar relative to each other than relative to a site from outside of that ecoregion or IGBP type, highlighting where some ecosystem classification schemes could be potentially insufficient. When examining site similarity as a function of spatial distance, we find that sites closer in proximity to each other are generally more similar to each other, but there is a perhaps surprisingly large amount of variation in the degree of spatial similarity between ecoregion and IGBP type. As our approach was solely based on the time series dynamics, without inputting the sites' characteristics, our results provided an independent and quantitative way to evaluate and revisit commonly adopted site classifications, such as IGBP groups and ecoregions.

One goal of flux networks is to provide sufficient, representative information for the regions, continents, or globe, ultimately enabling the upscaling of fluxes from individual sites' footprints to the terrestrial globe (Loescher et al., 2022). Focusing on a few ecoregions with abundant sites, we found Northwestern Forested Mountain, Mediterranean California, and Eastern Temperate Forests had substantial cross-site variability in flux dynamics (Figures S63, S64, S65, S68), even as high as that across ecoregions. This large within-region variability indicates the landscape complexity inherent within the ecoregion concept, and ultimately, there may not be a one-size-fits-all for ecoregion classification. Ecoregion classifications, such as those adopted in this study, were often based on climate, land-surface forms, or potential natural vegetation. Similarly, many studies that evaluated observational network representativeness were often based on driver variables of fluxes or site characteristics (Hargrove et al., 2003; Schimel et al., 2007; Sundareshwar et al., 2007). Our results suggested that those approaches' assumptions may need to be revisited or further refined, as the dynamics in the drivers may not be sufficient to reflect the dynamics in fluxes.

Often, upscaling research is based on quantifying the flux sensitivity to drivers within specific functional classifications such as plant functional type. The implicit assumption of this approach is that a single site provides sufficient detail to capture target processes and with adequate abstraction to allow upscaling and extrapolation to regions or the globe (Pacala and Kinzig, 2002). The functional classification assumes that the flux towers' footprints can represent other unmeasured areas if they share similar plant functional types. Clearly, models trained or parameterized using data at the flux towers can be applied and upscaled to the terrestrial globe (Jung et al., 2020; Zheng et al., 2020). Here, we do not examine the relationship between fluxes and drivers across the landscape but instead quantify how time series within the same functional classifications vary. Our results suggested that the IGBP groups—a commonly adopted classification—may be insufficient to capture the ecosystems' flux dynamics (Page et al., 2024) and justified the need to seek other flux-based classifications, such as ecosystem functional types (Villarreal et al., 2018, 2019). Particularly, wetlands, shrublands,

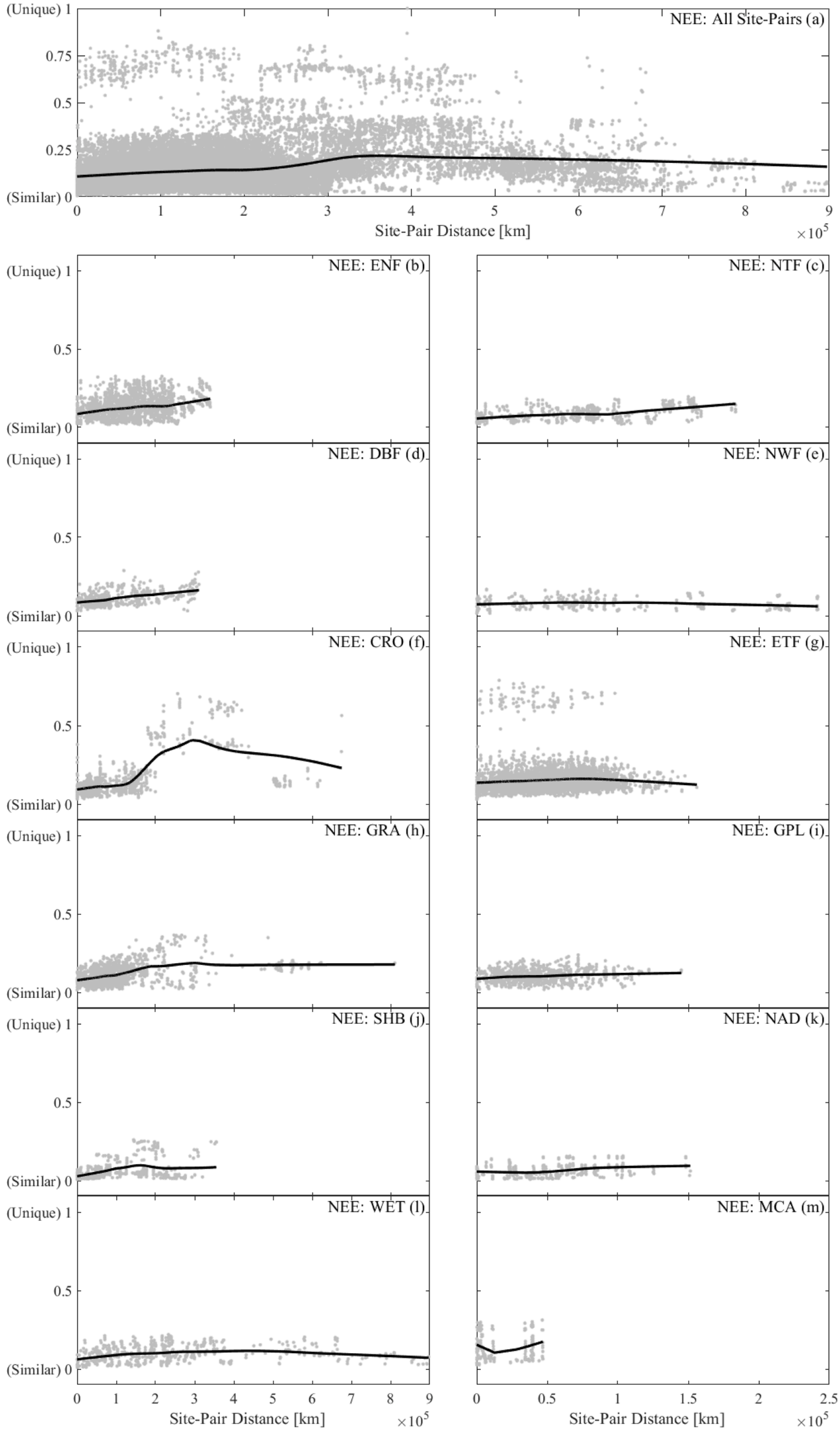

Fig. 6. Normalized site-pair dynamic time warping (DTW) distance as a function of site-pair distance for NEE across all site-pairs (a) and by (b, d, f, h, j, l) IGBP groups and (c, e, g, i, k, m) ecoregions. Only groups with sufficient samples (>20 sites) were included. Black lines represent moving average smoothing. Please refer to Fig. 5 for the IGBP and ecoregion abbreviations. See Figures S56-S62 for similar figures of other variables.

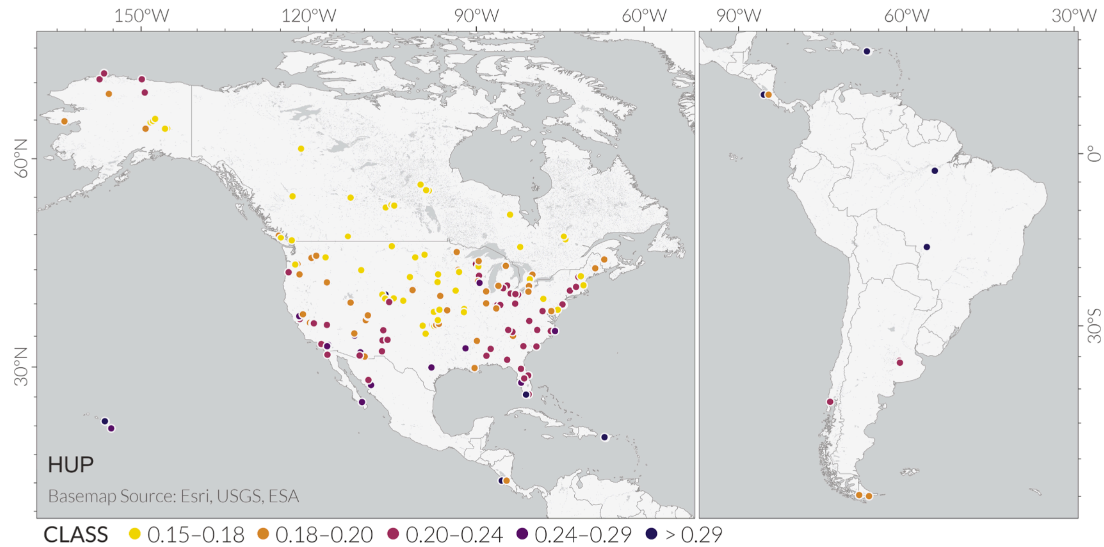

Fig. 7. The AmeriFlux sites highlighted by the sites' Harmonic Uniqueness Parameter (HUP). HUP represents the multivariate similarity using all flux time series (NEE, LE, H, USTAR). A higher HUP value indicates a site's relatively higher uniqueness.

croplands, and evergreen needleleaf forests showed high within-group variability in specific fluxes, highlighting the importance of other unaccounted functional properties such as plant species composition, stand age, salinity, hydrological regimes, and management in driving the observed cross-site flux variability (Hemes et al., 2019; Paul-Limoges et al., 2015; Qu et al., 2023; Vickers et al., 2012; Xiao et al., 2014; Xie et al., 2021).

Upscaling is challenging, particularly when extrapolating fine-scale point observations to a larger region (Newman et al., 2019). Flux towers typically sample an area of around 0.001–10 km^2 (Chu et al., 2021), and the measured fluxes are upscaled using a multitude of models, remote sensing products, and climate data sets to a regional or continental scale, i.e., 10^6–10^9 km^2 (Jung et al., 2020; Zheng et al., 2020). Even in adjacent areas, the validation of flux upscaling beyond the flux tower footprints is often limited (Hollinger et al., 2004; Oren et al., 2006; Post et al., 2015; Schmidt et al., 2012). We show that the degree to which processes change spatially is ecosystem dependent. We found that sites with the same IGBP type likely had similar temporal characteristics in fluxes when spatially proximal. This finding highlights the potential validity of upscaling that extrapolates single-point observations to some distance across the landscape, although management or disturbance history at sites should be considered. The recent findings of Page et al. (2024) suggest that when using site climate to predict flux behavior, plant functional types provide limited information, but combined with the results presented here, there could be a spatial distance across the landscape within which scaling would be valid for a given ecosystem type. Future research should incorporate ancillary datasets such as remote sensing estimates of ecosystem function to help bridge the gap between the footprint scale (Chu et al. 2021) and the regional scale, given that the results here begin to quantify spatial similarity, but only between discrete flux footprints. Connecting these clustering results to landscape-level ecosystem functioning data would help reveal the usefulness of scaling assumptions embedded in IGBP or similar ecosystem classification schemes.

4.2. Site clustering, interpretation, and implications

We advocate that our clustering of site data can serve as an alternative guide to reorganizing and interpreting the AmeriFlux network. This time series clustering approach allows us to quantitatively organize the sites' relationships, identify those unique to the current network, and group sites with similar temporal dynamics. For example, one can interpret our multivariate clustering as a quantitative measure of sites' similarity in their flux dynamics, i.e., a flux-based classification of AmeriFlux sites. The information can inform future syntheses, upscaling studies, and model-data benchmarking on site selection/grouping, balancing between those unique sites and those sharing similar temporal dynamics with others.

One can use our clustering to identify sites with similar/distinct flux dynamics within the same IGBP groups, ecoregions, or georegions, i.e., test the hypotheses of (dis)similarity of flux dynamics. Multiple environmental variables (i.e., NETRAD, TA, VPD) showed an expected agreement between their clustering and ecoregions. The findings justified a typical research design of utilizing multiple sites within the same ecoregion, expecting that climate and weather conditions would be similar. Many AmeriFlux sites were established *ad hoc* as site clusters—multiple co-located sites across gradients of land cover/land use, chronosequence stages, management, disturbance, edaphic and hydrological regimes (Chen et al., 2004; Biederman et al., 2018; Forsythe et al., 2020; Goulden et al., 2006; Knox et al., 2014; Law et al., 2001; Verma et al., 2005; Vickers et al., 2012). Furthermore, potential synthesis opportunities could arise by strategically pairing sites established independently but in the same ecoregions (Chu et al., 2023; Stoy et al., 2023), i.e., *post hoc* site clusters (Bormann, 2012; Butterworth et al., 2021; Chen et al., 2018; Zhang et al., 2020). In fact, several IGBP groups (e.g., evergreen needleleaf forest (76 sites), grassland (57), cropland (51)) and ecoregions (e.g., Eastern Temperate Forests (100), Great Plains (48), Northern Forests (40); Figures S63-S68) had relatively abundant sites, now allowing one to examine their (dis)similarity in flux dynamics and the underlying processes that drive cross-site variation within these groups.

Lastly, the current clusters of site data highlight what would be defined as a typical flux site across the AmeriFlux network and unique and underrepresented sites. Flux-based clustering generally showed an uneven distribution compared to the environmental variables, featuring a few intermediate to large groups and many unique, small groups of sites. Given how clustering algorithms function, with unique sites defining the edges of the clustering distribution, sites show relatively more similar behavior in the center. Hence, the flux variables of NEE, H, and LE show a small number of large groups, and within these large groups, there are site-to-site differences. Still, these differences are small compared to the sites at the edges of the clustering distribution. While large clusters exist, we also show the tree structure for each variable to show subgroup dynamics. The contrast among cluster sizes highlighted the significant ecological modulation by the ecosystems on the exchanges of carbon, water, and energy with the atmosphere (Stoy et al., 2009).

This work also explores the unique and underrepresented sites in the AmeriFlux network, motivating the establishment or incorporation of new/unregistered sites from those groups. First, many unique sites were from currently under-sampled IGBP groups (e.g., urban (US-INc), open water (US-Men, US-Pnp)) and ecoregions (e.g., Marine West Coast Forests, Hudson Plains, Hawaii, and ecoregions in Latin America). One factor contributing to under-sampling here is similar to the mid-domain effect, where geographic patterns are influenced by physiographical boundaries such as mountain ranges or land-sea interfaces, causing a high degree of overlap of species ranges in the center of the shared geographic domain and lower species richness (and hence a unique assemblage of species) near the boundaries (Colwell and Lees, 2000). Here, we see sites closer to spatial boundaries and, therefore, further isolated, having a higher degree of uniqueness, e.g., based on HUP, eight out of the eighteen most unique sites are located in Hawaii and Latin America (US-SuW, US-SuS, CR-Fsc, BR-Sa3, BR-Npw, BR-Sa1, PR-xLA, and PR-xLA). Our results provided a quantitative measure of these sites' uniqueness in flux dynamics to the rest of the network, reiterating the importance of recruiting sites from those groups to increase the network's representativeness (Villarreal and Vargas, 2021). Second, several unique sites were known for geographical features (e.g., mountain, floodplain), disturbances (e.g., wildfire, insect outbreak), or management regimes (e.g., periodic harvest, prescribed burning, forest plantation, restored wetlands). For example, several mountainous sites in the Northwestern Forested Mountains (e.g., US-GLE, US-CPk, US-NR1, US-NR3, US-Vcp, US-Vcm, US-xRM, Figure S67) were clustered into small, separated groups. Several restored wetlands (US-Myb, US-Tw1, US-Tw4), croplands (US-Twt (rice), US-Tw3 (alfalfa), US-Bi2 (corn), and a pasture (US-Snd) co-located within a ~10-km radius in California were also separated into different groups (Figure S68). These examples highlight the strong influence of topography, land cover/use, and management in driving distinct flux dynamics (Chen et al., 2004; Duman et al., 2020; Flerchinger et al., 2019; Goulden et al., 2012; Hemes et al., 2019; Xie et al., 2021), which may not be captured by the IGBP or ecoregion groups alone.

4.3. Justifications and limitations

We advise the readers to interpret our clustering in the context of the target variables, scales, features (e.g., periods and amplitudes), and underlying processes that drove the cross-site variation. Multiple environmental variables (i.e., NETRAD, TA, VPD) generally agree between the clustering and ecoregions. Also, several known co-located sites (e.g., US-Ha1/US-xHA, US-Bar/US-xBR, US-Kon/US-xKZ, US-EML/US-xHE, US-xBN/US-Prr) with similar underlying ecosystems were clustered into the same groups in multivariate flux clustering. Both results suggest the robustness of our approach. We discuss the potential limitations of our clustering analyses in the following paragraphs.

First, one issue to contend with now and into the future is the Modifiable Temporal Unit Problem (MTUP) (Cheng and Adepeju, 2014; Çöltekin et al., 2011). Shape-based time series clustering will depend upon the attribute extraction adopted by the analyses. We strategically focused on the diel and seasonal scales, which typically dominate the temporal dynamics of fluxes across sites (Chu et al., 2023; Stoy et al., 2009). The flux dynamics at these two scales were mainly driven by dynamics in environmental conditions (e.g., radiation, temperature, VPD), vegetation physiology (e.g., assimilation, metabolism, water use), and vegetation phenology (Baldocchi, Falge, and Wilson, 2001; Stoy et al., 2005). Our clustering of sites generally reflected the influences of these abiotic and biotic processes. Admittedly, the aggregation process smoothed out the temporal information (and relevant processes) at other scales, such as the hourly (e.g., rain pulse), weekly (e.g., synoptic weather), or interannual scales (e.g., succession, disturbance, climate change). However, information such as management practices is apparent in the composite time series if the management actions were performed on roughly the same calendar dates each year. For example, US-Bi1 and US-Tw3 are alfalfa fields with regular mowing during the growing season that clearly impacts the flux dynamics. If one focuses on aggregating at other scales, we anticipate that the clustering results may change, shifting weight between the temporal processes that guide clustering. For example, most previous studies focused solely on seasonal dynamics, which typically highlight phase shifts in growing seasons (e.g., peak, duration) across ecosystem types (Falge et al., 2002; Wilson et al., 2003; Yang and Noormets, 2021) and focused more on the peak flux magnitudes and durations of growing seasons and less on the amplitudes of growing/non-growing seasons and daytime/nighttime. Lastly, the site clusters may vary from using one flux variable to another, highlighting the importance of considering multifaceted ecosystem responses. Potentially, multivariate clustering provided a more comprehensive view of the sites' similarities and relationships.

Previous clustering work has often focused on climate and environmental data, but here we cluster sites based on the direct flux observations themselves. One of the pioneer studies by Hargrove et al. (2003) used climatic, physiographic, and edaphic factors to delineate ecoregions, but suggested future work to incorporate and weight ecoregions by their carbon flux magnitude. A similar analysis by Kumar et al. (2023) for the LTAR Network found similar spatial representativeness across the conterminous United States, where larger regions have more generalized environments, and smaller regions are more specialized, while they noted that management practices were not considered. With this work focused on clustering ecosystem flux observations and not on clustering climate drivers of fluxes, we cannot estimate spatial locations of the network that are poorly represented, as previous network clustering work has done (Villarreal and Vargas, 2021). Instead, we highlight unique sites within the current network, and the results presented here should be taken as a parallel method to understand network representation.

Future research should explore the site similarity and clustering when large-scale standardized GPP and RECO become available. We strategically chose to focus on NEE rather than GPP and RECO for the following reasons. First, NEE is directly measured, while GPP and RECO are derived from NEE based on various models and assumptions. The flux partitioning process introduced additional model assumptions and potential biases (Vickers et al. 2009; Baldocchi et al. 2015). While varied in assumptions and implementations, flux partitioning generally involves models for RECO and/or GPP based on NEE data. Studies based on independent observations, such as stable C isotopes (Lee et al., 2020) and solar-induced chlorophyll fluorescence (Kira et al., 2021), suggested a potential overestimate of RECO from the common partitioning methods. Second, current partitioning algorithms can be problematic in nontypical cases (e.g., tropics, arctics, wetlands, open water). Partitioning NEE at these sites remains challenging because many theoretical or methodological assumptions are unlikely to be valid. Third, many efforts have attempted to generate a standardized, multi-approach ensemble of partitioned products across the networks (e.g., FLUXNET2015, Pastorello et al., 2020). At the time of writing, the FLUXNET Data System Initiative and the AmeriFlux ONEFlux processing project have about 56 % of our studied sites have standardized GPP and RECO available, limiting our analyses to only a subset of AmeriFlux sites. While additional sites are added frequently, based on the current site availability, GPP and RECO yield more evenly distributed clusters and distinct annual fluxes among groups. Future investigations using a larger number of sites with partitioned fluxes and different partitioning methodologies are worth approaching.

Our clustering analyses were also constrained by the current site locations and their distribution within the network, reflecting the unevenly partitioned clusters in flux variables. Several IGBP groups (e.g., snow/ice, water, urban, barren land, deciduous needleleaf forest) and ecoregions (e.g., Arctic Cordillera, Hudson Plains, Marine West Coast

Forests, Temperate Sierras, and most ecoregions in Latin America) had no or only a few flux sites currently in AmeriFlux (Perez-Quezada et al., 2023). The few available sites from these groups were often clustered into unique, small groups distinct from other clusters (i.e., large DTW distance). On the other hand, several IGBP groups (e.g., evergreen needleleaf forest, grassland, cropland) and ecoregions (e.g., Eastern Temperate Forests, Great Plains, Northern Forests) had relatively abundant sites, further revealing finer-granular (i.e., low/moderate DTW distance) sub-clusters within each group (Figures S63-S65). The contrast between the small and large groups implied that more undiscovered clusters in undersampled groups/ecoregions may emerge as the AmeriFlux network grows and expands.

5. Conclusion

We used an empirically based, harmonic approach to cluster the AmeriFlux sites based on their time series features independent from the commonly adopted IGBP classifications or ecoregions. Such an approach quantified sites' (dis)similarity in diel/seasonal amplitudes and phases of fluxes, reflecting the effects of climate, plant phenology, and ecophysiological potentials. While most environmental variables (NETRAD, TA, and VPD) showed an evident latitudinal/regional clustering as expected, flux variables revealed an uneven clustering with many small, unique groups and a few large to intermediate groups. The contrast highlighted the strong biological modulation of ecosystem CO_2, water, and energy fluxes, and this modulation varied spatially across the landscape. Many unique sites were from currently under-sampled IGBP types and ecoregions, with very distinct flux dynamics compared to the rest of the network. Yet, at a finer-granular level, local topography, disturbance, management, edaphic, and hydrological regimes further drove the flux dynamics difference within the same IGBP groups and ecoregions. We suggest that our clustering can be used to better interpret the AmeriFlux network, thereby identifying sites with similar/distinct flux dynamics for future cross-site syntheses. The information generated from the clusters can also inform future upscaling studies and model-data benchmarking on site selection and grouping. Finally, our clustering analysis highlighted unique and underrepresented sites in the AmeriFlux network, motivating the establishment or incorporation of new/unregistered sites from underrepresented groups, such as IGBP types of snow/ice, open water, urban, barren land, and deciduous needleleaf forest, and ecoregions of Arctic Cordillera, Hudson Plains, Marine West Coast Forests, Temperate Sierras, and Latin America.

Data statement

This study utilized publicly accessible data from the AmeriFlux Network, where individual PIs established research sites and shared their data with the network. Site-level funding and collaboration were acknowledged in Supplement Table S2. This study followed the AmeriFlux data use policy and contacted all sites' PIs during the manuscript preparation phase. All site-level personnel involved with data collection were invited to participate and contribute to this project.

All AmeriFlux data discussed in this paper are publicly available at AmeriFlux (https://ameriflux.lbl.gov/), accessed in September 2023. Ecoregion maps are accessed through the Ecological Regions website (https://www.epa.gov/eco-research/ecoregions). All studied sites' general information, clustering, and Harmonic Uniqueness Parameter results are provided in the Supporting Information (Table S1) of Reed et al. (2024). The R codes for data processing and clustering are available at https://github.com/chuhousen/flux_timeseries_cluster (DOI: https://doi.org/10.5281/zenodo.12585997). The accompanying interactive web map showing the site clusters and time series is available at https://cartoscience.users.earthengine.app/view/flux-networks.

CRediT authorship contribution statement

David E Reed: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Housen Chu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Brad G. Peter:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology. **Jiquan Chen:** Writing – review & editing, Writing – original draft, Validation, Resources, Methodology. **Michael Abraha:** Writing – review & editing, Validation, Resources, Methodology. **Brian Amiro:** Writing – review & editing, Validation, Resources, Methodology. **Ray G. Anderson:** Writing – review & editing, Validation, Resources, Methodology. **M. Altaf Arain:** Writing – review & editing, Validation, Resources, Methodology. **Paulo Arruda:** Writing – review & editing, Validation, Resources, Methodology. **Greg A. Barron-Gafford:** Writing – review & editing, Validation, Resources, Methodology. **Carl Bernacchi:** Writing – review & editing, Validation, Resources, Methodology. **Daniel P. Beverly:** Writing – review & editing, Validation, Resources, Methodology. **Sebastien C. Biraud:** Writing – review & editing, Validation, Resources, Methodology. **T. Andrew Black:** Writing – review & editing, Validation, Resources, Methodology. **Peter D. Blanken:** Writing – review & editing, Validation, Resources, Methodology. **Gil Bohrer:** Writing – review & editing, Validation, Resources, Methodology. **Rebecca Bowler:** Writing – review & editing, Validation, Resources, Methodology. **David R. Bowling:** Writing – review & editing, Validation, Resources, Methodology. **M. Syndonia Bret-Harte:** Writing – review & editing, Validation, Resources, Methodology. **Mario Bretfeld:** Writing – review & editing, Validation, Resources, Methodology. **Nathaniel A. Brunsell:** Writing – review & editing, Validation, Resources, Methodology. **Stephen H. Bullock:** Writing – review & editing, Validation, Resources, Methodology. **Gerardo Celis:** Writing – review & editing, Validation, Resources, Methodology. **Xingyuan Chen:** Writing – review & editing, Validation, Resources, Methodology. **Aimee T. Classen:** Writing – review & editing, Validation, Resources, Methodology. **David R. Cook:** Writing – review & editing, Validation, Resources, Methodology. **Alejandro Cueva:** Writing – review & editing, Validation, Resources, Methodology. **Higo J. Dalmagro:** Writing – review & editing, Validation, Resources, Methodology. **Kenneth Davis:** Writing – review & editing, Validation, Resources, Methodology. **Ankur Desai:** Writing – review & editing, Validation, Resources, Methodology. **Alison J. Duff:** Writing – review & editing, Validation, Resources, Methodology. **Allison L. Dunn:** Writing – review & editing, Validation, Resources, Methodology. **David Durden:** Writing – review & editing, Validation, Resources, Methodology. **Colin W. Edgar:** Writing – review & editing, Validation, Resources, Methodology. **Eugenie Euskirchen:** Writing – review & editing, Validation, Resources, Methodology. **Rosvel Bracho:** Writing – review & editing, Validation, Resources, Methodology. **Brent Ewers:** Writing – review & editing, Validation, Resources, Methodology. **Lawrence B. Flanagan:** Writing – review & editing, Validation, Resources, Methodology. **Christopher Florian:** Writing – review & editing, Validation, Resources, Methodology. **Vanessa Foord:** Writing – review & editing, Validation, Resources, Methodology. **Inke Forbrich:** Writing – review & editing, Validation, Resources, Methodology. **Brandon R. Forsythe:** Writing – review & editing, Validation, Resources, Methodology. **John Frank:** Writing – review & editing, Validation, Resources, Methodology. **Jaime Garatuza-Payan:** Writing – review & editing, Validation, Resources, Methodology. **Sarah Goslee:** Writing – review & editing, Validation, Resources, Methodology. **Christopher Gough:** Writing – review & editing, Validation, Resources, Methodology. **Mark Green:** Writing – review & editing, Validation, Resources, Methodology. **Timothy Griffis:** Writing – review & editing, Validation, Resources, Methodology. **Manuel Helbig:** Writing – review & editing, Validation, Resources, Methodology. **Andrew C. Hill:** Writing – review & editing, Validation, Resources, Methodology. **Ross Hinkle:** Writing – review &

editing, Validation, Resources, Methodology. **Jason Horne:** Writing – review & editing, Validation, Resources, Methodology. **Elyn Humphreys:** Writing – review & editing, Validation, Resources, Methodology. **Hiroki Ikawa:** Writing – review & editing, Validation, Resources, Methodology. **Go Iwahana:** Writing – review & editing, Validation, Resources, Methodology. **Rachhpal Jassal:** Writing – review & editing, Validation, Resources, Methodology. **Bruce Johnson:** Writing – review & editing, Validation, Resources, Methodology. **Mark Johnson:** Writing – review & editing, Validation, Resources, Methodology. **Steven A. Kannenberg:** Writing – review & editing, Validation, Resources, Methodology. **Eric Kelsey:** Writing – review & editing, Validation, Resources, Methodology. **John King:** Writing – review & editing, Validation, Resources, Methodology. **John F. Knowles:** Writing – review & editing, Validation, Resources, Methodology. **Sara Knox:** Writing – review & editing, Validation, Resources, Methodology. **Hideki Kobayashi:** Writing – review & editing, Validation, Resources, Methodology. **Thomas Kolb:** Writing – review & editing, Validation, Resources, Methodology. **Randy Kolka:** Writing – review & editing, Validation, Resources, Methodology. **Ken W. Krauss:** Writing – review & editing, Validation, Resources, Methodology. **Lars Kutzbach:** Writing – review & editing, Validation, Resources, Methodology. **Brian Lamb:** Writing – review & editing, Validation, Resources, Methodology. **Beverly Law:** Writing – review & editing, Validation, Resources, Methodology. **Sung-Ching Lee:** Writing – review & editing, Validation, Resources, Methodology. **Xuhui Lee:** Writing – review & editing, Validation, Resources, Methodology. **Heping Liu:** Writing – review & editing, Validation, Resources, Methodology. **Henry W. Loescher:** Writing – review & editing, Validation, Resources, Methodology. **Sparkle L. Malone:** Writing – review & editing, Validation, Resources, Methodology. **Roser Matamala:** Writing – review & editing, Validation, Resources, Methodology. **Marguerite Mauritz:** Writing – review & editing, Validation, Resources, Methodology. **Stefan Metzger:** Writing – review & editing, Validation, Resources, Methodology. **Gesa Meyer:** Writing – review & editing, Validation, Resources, Methodology. **Bhaskar Mitra:** Writing – review & editing, Validation, Resources, Methodology. **J. William Munger:** Writing – review & editing, Validation, Resources, Methodology. **Zoran Nesic:** Writing – review & editing, Validation, Resources, Methodology. **Asko Noormets:** Writing – review & editing, Validation, Resources, Methodology. **Thomas L. O'Halloran:** Writing – review & editing, Validation, Resources, Methodology. **Patrick T O'Keeffe:** Writing – review & editing, Validation, Resources, Methodology. **Steven F. Oberbauer:** Writing – review & editing, Validation, Resources, Methodology. **Walter Oechel:** Writing – review & editing, Validation, Resources, Methodology. **Patty Oikawa:** Writing – review & editing, Validation, Resources, Methodology. **Paulo C. Olivas:** Writing – review & editing, Validation, Resources, Methodology. **Andrew Ouimette:** Writing – review & editing, Validation, Resources, Methodology. **Gilberto Pastorello:** Writing – review & editing, Validation, Resources, Methodology. **Jorge F. Perez-Quezada:** Writing – review & editing, Validation, Resources, Methodology. **Claire Phillips:** Writing – review & editing, Validation, Resources, Methodology. **Gabriela Posse:** Writing – review & editing, Validation, Resources, Methodology. **Bo Qu:** Writing – review & editing, Validation, Resources, Methodology. **William L. Quinton:** Writing – review & editing, Validation, Resources, Methodology. **Michele L. Reba:** Writing – review & editing, Validation, Resources, Methodology. **Andrew D. Richardson:** Writing – review & editing, Validation, Resources, Methodology. **Valentin Picasso:** Writing – review & editing, Validation, Resources, Methodology. **Adrian V. Rocha:** Writing – review & editing, Validation, Resources, Methodology. **Julio C. Rodriguez:** Writing – review & editing, Validation, Resources, Methodology. **Roel Ruzol:** Writing – review & editing, Validation, Resources, Methodology. **Scott Saleska:** Writing – review & editing, Validation, Resources, Methodology. **Russell L. Scott:** Writing – review & editing, Validation, Resources, Methodology. **Adam P. Schreiner-McGraw:** Writing – review & editing, Validation, Resources, Methodology. **Edward A.G. Schuur:** Writing – review & editing, Validation, Resources, Methodology. **Maria Silveira:** Writing – review & editing, Validation, Resources, Methodology. **Oliver Sonnentag:** Writing – review & editing, Validation, Resources, Methodology. **David L. Spittlehouse:** Writing – review & editing, Validation, Resources, Methodology. **Ralf Staebler:** Writing – review & editing, Validation, Resources, Methodology. **Gregory Starr:** Writing – review & editing, Validation, Resources, Methodology. **Christina Staudhammer:** Writing – review & editing, Validation, Resources, Methodology. **Chris Still:** Writing – review & editing, Validation, Resources, Methodology. **Cove Sturtevant:** Writing – review & editing, Validation, Resources, Methodology. **Ryan C. Sullivan:** Writing – review & editing, Validation, Resources, Methodology. **Andy Suyker:** Writing – review & editing, Validation, Resources, Methodology. **David Trejo:** Writing – review & editing, Validation, Resources, Methodology. **Masahito Ueyama:** Writing – review & editing, Validation, Resources, Methodology. **Rodrigo Vargas:** Writing – review & editing, Validation, Resources, Methodology. **Brian Viner:** Writing – review & editing, Validation, Resources, Methodology. **Enrique R. Vivoni:** Writing – review & editing, Validation, Resources, Methodology. **Dong Wang:** Writing – review & editing, Validation, Resources, Methodology. **Eric J. Ward:** Writing – review & editing, Validation, Resources, Methodology. **Susanne Wiesner:** Writing – review & editing, Validation, Resources, Methodology. **Lisamarie Windham-Myers:** Writing – review & editing, Validation, Resources, Methodology. **David Yannick:** Writing – review & editing, Validation, Resources, Methodology. **Enrico A. Yepez:** Writing – review & editing, Validation, Resources, Methodology. **Terenzio Zenone:** Writing – review & editing, Validation, Resources, Methodology. **Junbin Zhao:** Writing – review & editing, Validation, Resources, Methodology. **Donatella Zona:** Writing – review & editing, Validation, Resources, Methodology.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

David Reed reports financial support was provided by DOE Visiting Faculty Program. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the AmeriFlux site teams for sharing their data and metadata with the network. This work would not have been possible without the collegial interactions of many of our friends, peers, and mentors who have passed on in the course of collecting these datasets. Funding for these flux sites is acknowledged in the site data DOI, shown in Table S2. AmeriFlux data were made available through the data portal and processing maintained by the AmeriFlux Management Project, supported by the U.S. Department of Energy (DOE) Office of Science, Office of Biological and Environmental Research, under contract DE-AC02-05CH11231. NEON is a project sponsored by the U.S. National Science Foundation (NSF) and managed under the cooperative support agreement DBI-2217817 to Battelle. D. Reed and H. Chu acknowledge the support from the DOE Visiting Faculty Program for this work. ADR acknowledges support from NSF award 2105828 for continuity analysis of collocated NEON and AmeriFlux sites. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of our stakeholders or sponsors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2025.110686.

Data availability

AmeriFlux data discussed in this paper are publicly available at AmeriFlux (https://ameriflux.lbl.gov/), R codes for clustering are available at https://github.com/chuhousen/flux_timeseries_cluster.

References

Aghabozorgi, S., Seyed Shirkhorshidi, A., Ying Wah, T., 2015. Time-series clustering – a decade review. Inf. Syst. 53, 16–38.

AmeriFlux Management Project. (2020, June 25). BADM: biological, ancillary, disturbance, and metadata. Retrieved March 7, 2024, from https://ameriflux.lbl.gov/data/badm/.

Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J.M., Perona, I., 2013. An extensive comparative study of cluster validity indices. Pattern. Recognit. 46 (1), 243–256.

Baldocchi, D., Sturtevant, C., Contributors, Fluxnet, 2015. Does day and night sampling reduce spurious correlation between canopy photosynthesis and ecosystem respiration? Agric. For. Meteorol. 207, 117–126.

Baldocchi, D., Novick, K., Keenan, T., Torn, M., 2024. AmeriFlux: its impact on our understanding of the "breathing of the biosphere", after 25 years. Agric. For. Meteorol. 348 (109929), 109929.

Baldocchi, D.D., Hicks, B.B., Meyers, T.P., 1988. Measuring biosphere-atmosphere exchanges of biologically related gases with micrometeorological methods. Ecology. 69 (5), 1331–1340.

Baldocchi, D.D., Falge, E., Wilson, K., 2001a. A spectral analysis of biosphere-atmosphere trace gas flux densities and meteorological variables across hour to multi-year time scales. Agric. For. Meteorol. 107 (1), 1–27.

Baldocchi, D.D., Falge, E., Gu, L.H., Olson, R., Hollinger, D., Running, S., et al., 2001b. 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 (11), 2415–2434.

Bean, A.R., Coffin, A.W., Arthur, D.K., Baffaut, C., Holifield Collins, C., Goslee, S.C., et al., 2021. Regional frameworks for the USDA long-term agroecosystem research network. Front. Sustain. Food Syst. 4. https://doi.org/10.3389/fsufs.2020.612785.

Biederman, J.A., Scott, R.L., Arnone, J.A., Iii, Jasoni, R.L., Litvak, M.E., Moreo, M.T., et al., 2018. Shrubland carbon sink depends upon winter water availability in the warm deserts of North America. Agric. For. Meteorol. 249 (Supplement C), 407–419.

Bormann, H., 2012. Assessing the soil texture-specific sensitivity of simulated soil moisture to projected climate change by SVAT modelling. Geoderma 185-186, 73–83.

Buchta C., M. D. (2022). proxy: distance and similarity measures (Version version 0.4-27). Comprehensive R archive network (CRAN). Retrieved from https://CRAN.R-project.org/package=proxy.

Butterworth, B.J., Desai, A.R., Townsend, P.A., Petty, G.W., Andresen, C.G., Bertram, T. H., et al., 2021. Connecting land–Atmosphere interactions to surface heterogeneity in CHEESEHEAD19. Bull. Am. Meteorol. Soc. 102 (2), E421–E445.

Chen, J., Paw, U., K. T, Ustin, S.L., Suchanek, T.H., Bond, B.J., Brosofske, K.D., Falk, M., 2004. Net ecosystem exchanges of carbon, water, and energy in young and old-growth Douglas-fir forests. Ecosystems. 7, 534–544.

Cheng, T., Adepeju, M., 2014. Modifiable temporal unit problem (MTUP) and its effect on space-time cluster detection. PLoS. One 9 (6), e100465.

Chen, L., Dirmeyer, P.A., Guo, Z., Schultz, N.M., 2018. Pairing FLUXNET sites to validate model representations of land-use/land-cover change. Hydrol. Earth. Syst. Sci. 22 (1), 111.

Chu, H., Luo, X., Ouyang, Z., Chan, W.S., Dengel, S., Biraud, S.C., et al., 2021. Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites. Agric. For. Meteorol. 301-302, 108350.

Chu, H., Christianson, D.S., Cheah, Y.-W., Pastorello, G., O'Brien, F., Geden, J., et al., 2023. AmeriFlux BASE data pipeline to support network growth and data sharing. Sci. Data 10 (1), 614.

Cöltekin, A., De Sabbata, S., Willi, C., Vontobel, I., Pfister, S., Kuhn, M., Lacayo, M., 2011. Modifiable temporal unit problem. Persistent Problems in Geographic Visualization. https://doi.org/10.5167/UZH-54263.

Colwell, R.K., Lees, D.C., 2000. The mid-domain effect: geometric constraints on the geography of species richness. Trends Ecol. Evol. (Amst.) 15 (2), 70–76.

Commission for Environmental Cooperation. (1997). Ecological regions of North America: toward a common perspective. Commission for Environmental Cooperation.

Davis, K.J., Deng, A., Lauvaux, T., Miles, N.L., Richardson, S.J., Sarmiento, D.P., et al., 2017. The Indianapolis Flux Experiment (INFLUX): a test-bed for developing urban greenhouse gas emission measurements. Elementa (Washington, D.C.) 5. https://doi.org/10.1525/elementa.188.

Duman, T., Huang, C.-W., Litvak, M.E., 2020. Recent land cover changes in the Southwestern US lead to an increase in surface temperature. Agric. For. Meteorol., 108246

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., et al., 2001. Gap filling strategies for long term energy flux data sets. Agric. For. Meteorol. 107 (1), 71–77.

Falge, E., Baldocchi, D., Tenhunen, J., Aubinet, M., Bakwin, P., Berbigier, P., et al., 2002. Seasonality of ecosystem respiration and gross primary production as derived from FLUXNET measurements. Agric. For. Meteorol. 113 (1–4), 53–74.

Flerchinger, G.N., Fellows, A.W., Seyfried, M.S., Clark, P.E., Lohse, K.A., 2019. Water and carbon fluxes along an elevational gradient in a sagebrush ecosystem. Ecosystems. https://doi.org/10.1007/s10021-019-00400-x.

Forsythe, J.D., O'Halloran, T.L., Kline, M.A., 2020. An Eddy covariance mesonet for measuring greenhouse gas fluxes in Coastal South Carolina. Brown Univ. Digest Addict. Theory Appl.: DATA 5 (4), 97.

Giorgino, T., 2009. Computing and visualizing dynamic time warping alignments in R: the dtw package. J. Stat. Softw. 31, 1–24.

Goulden, M.L., Winston, G.C., McMillan, A.M.S., Litvak, M.E., Read, E.L., Rocha, A.V., Rob Elliot, J., 2006. An eddy covariance mesonet to measure the effect of forest age on land–atmosphere exchange. Glob. Chang. Biol. 12 (11), 2146–2162.

Goulden, M.L., Anderson, R.G., Bales, R.C., 2012. Evapotranspiration along an elevation gradient in California's Sierra Nevada. J. Geophys. Res.: Biogeosci. 117, G03028.

Griffith, G.E., Omernik, J.M., Azevedo, S.H., 1998. Ecological Classification of the Western Hemisphere. U.S. Environmental Protection Agency.

Hargrove, W., Hoffman, F., 2004. Potential of multivariate quantitative methods for delineation and visualization of ecoregions. Environ. Manage 34, S39–S60.

Hargrove, W.W., Hoffman, F.M., Law, B.E., 2003. New analysis reveals representativeness of the AmeriFlux network. Eos. (Washington. DC) 84 (48), 529–535.

Harris, I., Osborn, T.J., Jones, P., Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci. Data 7, 109.

Hemes, K.S., Chamberlain, S.D., Eichelmann, E., Anthony, T., Valach, A., Kasak, K., et al., 2019. Assessing the carbon and climate benefit of restoring degraded agricultural peat soils to managed wetlands. Agric. For. Meteorol. 268, 202–214.

Hill, A., Forbrich, I., Schäfer, K.V.R., Vargas, R., 2024. Empirical dynamic modeling reveals complexity of methane fluxes in a temperate salt marsh. J. Geophys. Res.: Biogeosci. 129 e2023JG007630.

Hilton, T.W., Davis, K.J., Keller, K., Urban, N.M., 2013. Improving North American terrestrial CO_2 flux diagnosis using spatial structure in land surface model residuals. Biogeosciences. 10 (7), 4607–4625.

Hollinger, D.Y., Aber, J., Dail, B., Davidson, E.A., Goltz, S.M., Hughes, H., et al., 2004. Spatial and temporal variability in forest-atmosphere CO_2 exchange. Glob. Chang. Biol. 10 (10), 1689–1706.

Jones, J.A., Groffman, P.M., Blair, J., Davis, F.W., Dugan, H., Euskirchen, E.E., et al., 2021. Synergies among environmental science research and monitoring networks: a research agenda. Earths. Future 9 (3). https://doi.org/10.1029/2020ef001631.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., et al., 2020. Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. Biogeosciences. 17 (5), 1343–1365.

Kira, O., Y-Y. Chang, C., Gu, L., Wen, J., Hong, Z., Sun, Y., 2021. Partitioning net ecosystem exchange (NEE) of CO_2 using solar-induced chlorophyll fluorescence (SIF). Geophys. Res. Lett. 48 (4). https://doi.org/10.1029/2020gl091247.

Knox, S.H., Sturtevant, C., Matthes, J.H., Koteen, L., Verfaillie, J., Baldocchi, D., 2014. Agricultural peatland restoration: effects of land-use change on greenhouse gas (CO2 and CH4) fluxes in the Sacramento-San Joaquin Delta. Glob. Chang. Biol. https://doi.org/10.1111/gcb.12745.

Koenig, W.D., 1999. Spatial autocorrelation of ecological phenomena. Trends Ecol. Evol. (Amst.) 14 (1), 22–26.

Krich, C., Migliavacca, M., Miralles, D.G., Kraemer, G., El-Madany, T.S., Reichstein, M., et al., 2021. Functional convergence of biosphere–atmosphere interactions in response to meteorological conditions. Biogeosciences. 18 (7), 2379–2404.

Kumar, J., Coffin, A.W., Baffaut, C., Ponce-Campos, G.E., Witthaus, L., Hargrove, W.W., 2023. Quantitative representativeness and constituency of the long-term agroecosystem research network and analysis of complementarity with existing ecological networks. Environ. Manage 72 (4), 705–726.

Law, B.E., Thornton, P.E., Irvine, J., Anthoni, P.M., Van Tuyl, S., 2001. Carbon storage and fluxes in ponderosa pine forests at different developmental stages. Glob. Chang. Biol. 7 (7), 755–777.

Lee, S.-C., Christen, A., Black, T.A., Jassal, R.S., Ketler, R., Nesic, Z., 2020. 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.

Loescher, H.W., Vargas, R., Mirtl, M., Morris, B., Pauw, J., Yu, X., et al., 2022. Building a global ecosystem research infrastructure to address global grand challenges for macrosystem ecology. Earths. Future 10 (5). https://doi.org/10.1029/2020ef001696.

Loveland, T.R., Zhu, Z., Ohlen, D.O., Brown, J.F., Reed, B.C., Yang, L., 1999. An analysis of the IGBP global land-cover characterization process. Photogramm. Eng. Remote Sensing. 65, 1021–1032.

Mahecha, M.D., Reichstein, M., Lange, H., Carvalhais, N., Bernhofer, C., Grünwald, T., et al., 2007. Characterizing ecosystem-atmosphere interactions from short to interannual time scales. Biogeosciences. 4 (5), 743–758.

Malone, S.L., Oh, Y., Arndt, K.A., Burba, G., Commane, R., Contosta, A.R., et al., 2022. Gaps in network infrastructure limit our understanding of biogenic methane emissions for the United States. Biogeosciences. 19 (9), 2507–2522.

Moncrieff, J.B., Malhi, Y., Leuning, R., 1996. The propagation of errors in long-term measurements of land-atmosphere fluxes of carbon and water. Glob. Chang. Biol. 2 (3), 231–240.

Monteith, J.L., Unsworth, M.H., 2008. Principles of Environmental Physics, 3rd ed. Elsevier, Burlinton, USA.

Newman, E.A., Kennedy, M.C., Falk, D.A., McKenzie, D., 2019. Scaling and complexity in landscape ecology. Front. Ecol. Evol. 7. https://doi.org/10.3389/fevo.2019.00293.

Novick, K.A., Biederman, J.A., Desai, A.R., Litvak, M.E., Moore, D.J.P., Scott, R.L., Torn, M.S., 2018. The AmeriFlux network: a coalition of the willing. Agric. For. Meteorol. 249, 444–456.

Omernik, J.M., 1987. Ecoregions of the conterminous United States. Ann. Assoc. Am. Geogr. 77 (1), 118–125.

Omernik, J.M., 2004. Perspectives on the nature and definition of ecological regions. Environ. Manage 34 (Suppl 1), S27–S38.
Openshaw, S., 1984. Ecological fallacies and the analysis of areal census data. Environ. Plan. a 16 (1), 17–31.
Oren, R.A.M., Hsieh, C., Stoy, P., Albertson, J., Mccarthy, H.R., Harrell, P., Katul, G.G., 2006. Estimating the uncertainty in annual net ecosystem carbon exchange: spatial variation in turbulent fluxes and sampling errors in eddy-covariance measurements. Glob. Chang. Biol. 12 (5), 883–896.
Pacala, S., Kinzig, A.P., 2002. Introduction to theory and the common ecosystem model. Funct. Conseq. Biodivers.: Empiric. Prog. Theoret. Extens. 169–174.
Page, C.J., Abramowitz, G., De Kauwe, M.G., Pitman, A.J., 2024. Are plant functional types fit for purpose? Geophys. Res. Lett. 51 (1). https://doi.org/10.1029/2023gl104962.
Pallandt, M.M., Kumar, J., Mauritz, M., Schuur, E.A.G., Virkkala, A.-M., Celis, G., et al., 2022. Representativeness assessment of the pan-Arctic eddy covariance site network and optimized future enhancements. Biogeosciences. 19 (3), 559–583.
Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., et al., 2006. Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences. 3 (4), 571–583.
Papale, D., Black, T.A., Carvalhais, N., Cescatti, A., Chen, J., Jung, M., et al., 2015. Effect of spatial sampling from European flux towers for estimating carbon and water fluxes with artificial neural networks. J. Geophys. Res.: Biogeosci. 120 (10), 1941–1957.
Paradis, E., & Schliep, K. (2023). ape: analyses of Phylogenetics and Evolution (Version 5.7-1). Comprehensive R Archive Network (CRAN). Retrieved from https://cran.r-project.org/web/packages/ape/index.html.
Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., et al., 2020. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. Sci. Data 7, 225.
Paul-Limoges, E., Black, T.A., Christen, A., Nesic, Z., Jassal, R.S., 2015. Effect of clearcut harvesting on the carbon balance of a Douglas-fir forest. Agric. For. Meteorol. 203, 30–42.
Perez-Quezada, J.F., Meijide, A., Leitner, S., 2023. Greenhouse gas measurements in underrepresented areas of the world. Front. Soil. Sci. 3. https://doi.org/10.3389/fsoil.2023.1240930.
Poe, J., Reed, D.E., Abraha, M., Chen, J., Dahlin, K.M., Desai, A.R., 2020. Geospatial coherence of surface-atmosphere fluxes in the upper Great Lakes region. Agric. For. Meteorol. 295, 108188.
Post, H., Hendricks Franssen, H.J., Graf, A., Schmidt, M., Vereecken, H., 2015. Uncertainty analysis of eddy covariance CO2 flux measurements for different EC tower distances using an extended two-tower approach. Biogeosciences. 12 (4), 1205–1221.
Qu, B., Roy, A., Melton, J.R., Andrew Black, T., Amiro, B., Euskirchen, E.S., et al., 2023. A boreal forest model benchmarking dataset for North America: a case study with the Canadian Land Surface Scheme including biogeochemical cycles (CLASSIC). Environ. Res. Lett.: ERL [Web Site] 18 (8), 085002.
Reed, D.E., Poe, J., Abraha, M., Dahlin, K.M., Chen, J., 2021. Modeled surface-atmosphere fluxes from paired sites in the upper great lakes region using neural networks. J. Geophys. Res. Biogeosci. 126 (8). https://doi.org/10.1029/2021jg006363.
Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., et al., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob. Chang. Biol. 11 (9), 1424–1439.
Sarda-Espinosa, A. (2023). Dtwclust: time series clustering along with optimizations for the dynamic time warping distance (Version 5.5.12). Retrieved from https://cran.r-project.org/web/packages/dtwclust/index.html.
Schimel, D., Hargrove, W., Hoffman, F., MacMahon, J., 2007. NEON: a hierarchically designed national ecological network. Front. Ecol. Environ. 5 (2), 59–59.
Schmidt, A., Hanson, C., Chan, W.S., Law, B.E., 2012. Empirical assessment of uncertainties of meteorological parameters and turbulent fluxes in the AmeriFlux network. J. Geophys. Res.: Biogeosci. 117 (G4), G04014.
Scott, R.L., Jenerette, G.D., Potts, D.L., Huxman, T.E., 2009. Effects of seasonal drought on net carbon dioxide exchange from a woody-plant-encroached semiarid grassland. J. Geophys. Res.: Biogeosci. 114 (G4).
Sokal, R.R., James Rohlf, F., 1995. Biometry. Macmillan.
Stoy, P.C., Katul, G.G., Siqueira, M.B.S., Juang, J.-Y., McCarthy, H.R., Kim, H.-S., et al., 2005. Variability in net ecosystem exchange from hourly to inter-annual time scales at adjacent pine and hardwood forests: a wavelet analysis. Tree Physiol. 25 (7), 887–902.
Stoy, P.C., Richardson, A.D., Baldocchi, D.D., Katul, G.G., Stanovick, J., Mahecha, et al., 2009. Biosphere-atmosphere exchange of CO 2 in relation to climate: a cross-biome analysis across multiple time scales. Biogeosciences. 6, 2297–2312.
Stoy, P.C., Chu, H., Dahl, E., Cala, D.S., Shveytser, V., Wiesner, S., et al., 2023. The global distribution of paired eddy covariance towers. bioRxiv. https://doi.org/10.1101/2023.03.03.530958.
Sulkava, M., Luyssaert, S., Zaehle, S., Papale, D., 2011. Assessing and improving the representativeness of monitoring networks: the European flux tower network example. J. Geophys. Res.: Biogeosci. 116 (G3), 2156–2202.
Sundareshwar, P.V., Murtugudde, R., Srinivasan, G., Singh, S., Ramesh, K.J., Ramesh, R., et al., 2007. Environmental monitoring network for India. Science (1979) 316 (5822), 204–205.
Tarin-Terrazas, T., Alvarado-Barrientos, S., Cueva-Rodríguez, A., Hinojo-Hinojo, C., del Castillo, E.G., Sánchez-Mejía, Z., et al., 2020. MexFlux: sinergias para diseñar, evaluar e informar soluciones climáticas naturales. Element. Para Polític. Públ. 4 (2), 99–118.
Vargas, R., Yépez, E.A., Andrade, J.L., Ángeles, G., Arredondo, T., Castellanos, A.E., et al., 2013. Progress and opportunities for monitoring greenhouse gases fluxes in Mexican ecosystems: the MexFlux network. Atmósfera 26 (3), 325–336.
Vekuri, H., Tuovinen, J.-P., Kulmala, L., Papale, D., Kolari, P., Aurela, M., Laurila, T., Liski, J., Lohila, A., 2023. A widely-used eddy covariance gap-filling method creates systematic bias in carbon balance estimates. Sci. Rep. 13, 1720.
Verma, S.B., Dobermann, A., Cassman, K.G., Walters, D.T., Knops, J.M., Arkebauer, T.J., et al., 2005. Annual carbon dioxide exchange in irrigated and rainfed maize-based agroecosystems. Agric. For. Meteorol. 131 (1), 77–96.
Vickers, D., Thomas, C.K., Martin, J.G., Law, B., 2009. Self-correlation between assimilation and respiration resulting from flux partitioning of eddy-covariance CO2 fluxes. Agric. For. Meteorol. 149 (9), 1552–1555.
Vickers, D., Thomas, C., Pettijohn, C., Martin, J.G., Law, B., 2012. Five years of carbon fluxes and inherent water-use efficiency at two semi-arid pine forests with different disturbance histories. Tellus. B Chem. Phys. Meteorol. 64 (1), 17159.
Villarreal, S., Vargas, R., 2021. Representativeness of FLUXNET sites across Latin America. J. Geophys. Res.: Biogeosci. 126 (3) e2020JG006090.
Villarreal, S., Guevara, M., Alcaraz-Segura, D., Brunsell, N.A., Hayes, D., Loescher, H.W., Vargas, R., 2018. Ecosystem functional diversity and the representativeness of environmental networks across the conterminous United States. Agric. For. Meteorol. 262, 423–433.
Villarreal, S., Guevara, M., Alcaraz-Segura, D., Vargas, R., 2019. Optimizing an environmental observatory network design using publicly available data. J. Geophys. Res.: Biogeosci. 124 (7), 1812–1826.
Wilson, K.B., Baldocchi, D., Falge, E., 2003. Diurnal centroid of ecosystem energy and carbon fluxes at FLUXNET sites. J. Geophys. Res. https://doi.org/10.1029/2001JD001349.
Wutzler, T., Lucas-Moffat, A., Migliavacca, M., Knauer, J., Sickel, K., Šigut, L., et al., 2018. Basic and extensible post-processing of eddy covariance flux data with REddyProc. Biogeosciences. 15, 5015–5030.
Xiao, J., Zhuang, Q., Law, B.E., Chen, J., Baldocchi, D.D., Cook, D.R., et al., 2010. A continuous measure of gross primary production for the conterminous United States derived from MODIS and AmeriFlux data. Remote Sens. Environ. 114 (3), 576–591.
Xiao, J., Davis, K.J., Urban, N.M., Keller, K., 2014. Uncertainty in model parameters and regional carbon fluxes: a model-data fusion approach. Agric. For. Meteorol. 189–190, 175–186, 0.
Xie, X., Chen, J.M., Gong, P., Li, A., 2021. Spatial scaling of gross primary productivity over sixteen mountainous watersheds using vegetation heterogeneity and surface topography. J. Geophys. Res. Biogeosci. 126 (5). https://doi.org/10.1029/2020jg005848.
Yang, L., Noormets, A., 2021. Standardized flux seasonality metrics: a companion dataset for FLUXNET annual product. Earth. Syst. Sci. Data 13 (4), 1461–1475.
Zhang, Q., Barnes, M., Benson, M., Burakowski, E., Oishi, A.C., Ouimette, A., et al., 2020. Reforestation and surface cooling in temperate zones: mechanisms and implications. Glob. Chang. Biol. 26 (6), 3384–3401.
Zheng, Y., Shen, R., Wang, Y., Li, X., Liu, S., Liang, S., et al., 2020. Improved estimate of global gross primary production for reproducing its long-term variation, 1982–2017. Earth. Syst. Sci. Data 12 (4), 2725–2746.