



Eddy covariance fluxes over managed ecosystems extrapolated to field scales at fine spatial resolutions

Songyan Zhu^{a,1,*}, Louise Olde^b, Kennedy Lewis^a, Tristan Quaife^{c,d}, Laura Cardenas^b, Nadine Loick^b, Jian Xu^e, Timothy Hill^{a,f}

^a Faculty of Environment, Science and Economy (ESE), University of Exeter, Streatham Campus, Rennes Drive, Exeter EX4 4RJ, UK

^b Rothamsted Research, Sustainable Agriculture Sciences, North Wyke, Devon EX20 2SB, UK

^c Department of Meteorology, National Centre for Earth Observation, University of Reading, UK

^d Department of Meteorology, University of Reading, Reading, UK

^e National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China

^f Quanterra Systems Ltd., Centenary House, Peninsula Park, Exeter EX2 7XE, UK

ARTICLE INFO

Keywords:

Eddy covariance
Remote sensing
Carbon cycles
Livestock grazing

ABSTRACT

To enable an evidence-based management of ecosystems to adapt to the climate crisis, we require fine spatio-temporal resolution estimates of carbon, water, and energy fluxes at the field scale. To overcome the limitations resulting from the coarse spatial resolution of existing flux products, e.g. 500 m (Running et al., 2015), and the challenges in matching eddy covariance (EC) footprints with land use field scales, we for the first time investigate the influence of satellite resolution on flux estimation, which is to support the fine-scale extrapolation of EC fluxes from the tower footprint to the field scale. We validate the extrapolation at 206 FLUXNET2015 tower sites to pave the way for estimating field-scale fluxes extrapolated from three towers in a managed European grazing pasture on a fine-scale, 30 m spatial resolution. The findings suggest that (a) tower-level flux estimates from 30 m satellites were in agreement with fluxes estimated from moderate-resolution satellites, which are extensively employed in literature (R^2 difference $\ll 0.1$); (b) flux estimates were in reasonable agreement with EC measurements (R^2 : 0.7 and annual bias $< 2 \text{ Mg ha}^{-1} \text{ yr}^{-1}$ for carbon fluxes); (c) Sentinel-2 was advantageous in capturing land-use variability over other satellites in European pastures; (d) the machine-learning extrapolation algorithm was resistant to livestock grazing.

1. Introduction

Effective ecosystem management is key in response to the climate crisis (Malhi et al., 2020), and accurate ecosystem-level measurements of ecosystem carbon and water balance are pivotal to evaluating management effectiveness. The rapid development-driven expansion of commodity land uses (Zalles et al., 2021) and the continuing global population growth (Zarocostas, 2022) contribute to the climate crisis we are facing – e.g., the rising concentrations of greenhouse gases [GHGs] (Mauna Loa, 2022), frequent climate extremes (Salomón et al., 2022) and worsened food security (Kotir, 2011). Since the 1990s, 14.6% (ca. 18.5 million km²) of the Earth's land area has been modified (Theobald et al., 2020) and the conversion from natural ecosystems to agriculture – e.g., livestock grazing and crop cultivation – is one major source of

greenhouse gas emissions (FAO, 2022). Accurate and continuous monitoring of carbon, water, and energy exchanges (i.e., fluxes) between the atmosphere and agricultural ecosystems are therefore very important for understanding these ecosystems (Jung et al., 2017) and for improving agricultural effectiveness while minimising emissions (Griscom et al., 2017; Smith et al., 2021).

Eddy covariance (EC) is one of the most effective techniques for providing direct flux measurements at the field scale within the measurement footprint of the installed EC tower, ca. 100 to several thousands of metres radii around the towers (Chu et al., 2021). EC has gained in popularity since the late 1990s (Baldocchi, 2020) and global EC networks (e.g., AmeriFlux, EuroFlux, and AsiaFlux) continue to develop and have collectively accumulated more than 900 towers (Chu et al., 2017). Such global collaboration has successfully led to the release of the

* Corresponding author.

E-mail address: sz394@exeter.ac.uk (S. Zhu).

¹ Present address: School of GeoSciences and National Centre for Earth Observation, University of Edinburgh, Edinburgh EH9 3FF, UK.

FLUXNET2015 database which includes 206 open-access towers and covers data from late 1990s to 2015 (Pastorello et al., 2020). However, the vast majority of Earth's land surface is not monitored because: (1) the EC footprint is relatively small compared to the area of terrestrial ecosystems (Lee et al., 2004), and (2) the majority of existing EC towers (Fig. 1), ca. 85% (Schimel et al., 2015), are clustered in the Northern Temperate ecosystems and the number of EC towers in ecoregions like Amazonia and other tropical systems is relatively few (Pastorello et al., 2020). Furthermore, approximately 38% of the global land surface is agricultural, ca. five billion hectares (FAO, 2022), but the fluxes within this large area are monitored by only 58 FLUXNET towers – 20 for croplands and 38 for grasslands which include pastures (Pastorello et al., 2020).

Extrapolating EC measurements by assimilating satellite observations has been widely recognised as an effective approach to estimating global ecosystem fluxes (Xiao et al., 2012; Jung et al., 2020). However, the literature mainly considers pristine or natural ecosystems (Jung et al., 2019; Joiner and Yoshida, 2020), overlooking management activities which can pose additional challenges to estimating fluxes in managed systems. Further, the majority of EC extrapolation products are derived from moderate-resolution satellites (Schaaf and Wang, 2015) and the coarse spatial resolution (ca. 5–50 km) hinders the estimation of fluxes over managed ecosystems (e.g., grazing pastures) where the land use area can be smaller than one pixel and the internal spatial variability can be strong (Orr et al., 2016; Cardenas et al., 2022). As a result, extrapolating EC fluxes at a fine spatial resolution remains elusive, particularly for managed ecosystems (Fu et al., 2014). Reliable flux estimation from freely available high-spatial-resolution satellite images – e.g., 30 m Landsat-7/-8 (Chander et al., 2009; Vermote et al., 2016) or 10 m/20 m Sentinel-2 (Szantoi and Strobl, 2019) satellites – holds great potential for cost-effective ecosystem management assessment, terrestrial carbon sink preservation, and the pursuit of a carbon-neutral future (Griscom et al., 2017; Novick et al., 2022).

The objective of this study is to extrapolate EC carbon, water, and energy fluxes at a fine spatial resolution (i.e., 30 m) and in a cost-effective manner to better monitor managed ecosystem fluxes with livestock grazing disturbances. This study comprises two parts: In part 1, we first examine our flux estimates at 206 FLUXNET sites, considering relative performance across 11 ecosystem types and using three satellites (open-access for research purposes) with different spatial resolutions varying from 30 m to 5566 m. This knowledge is then used to support the second part of this study, where we consider the impact of management activities when extrapolating EC fluxes to obtain reliable 30 m field-scale estimates over three neighbouring grazed pastures in southwest England (Orr et al., 2016). To do this we use three EC towers located in separate pastures which experience the same climate, and similar land use therefore minimising the impacts of these factors on the extrapolation. As these three pastures have different grazing periods,

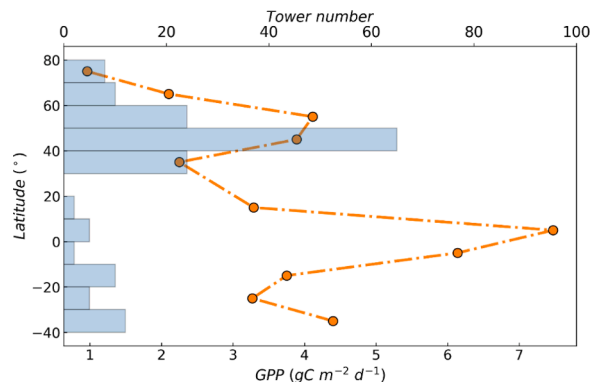


Fig. 1. FLUXNET2015 site numbers (blue bins) and mean gross primary productivity [GPP] (orange dots) at every 10° latitude bins.

they provide a good testbed to extrapolate the EC fluxes and analyse the impacts of grazing activities across sites.

2. Methodology

2.1. Machine learning-based flux extrapolation framework

Inspired by the Unified FLUXes (UFLUX) initiative (<https://sites.google.com/view/uflux>), extrapolating EC fluxes typically makes use of machine-learning algorithms to link EC fluxes and predictors (Jung et al., 2019, 2020) such as (Fig. 2):

$$\text{Flux} = f(\text{Sat.}, \text{Meteo.}, \text{aux.})$$

where, *Sat.* is satellite-based vegetation data products, e.g., (1) normalised difference vegetation index, NDVI (Verhegghen et al., 2014):

$$\text{NDVI} = \frac{\text{NIR} - R}{\text{NIR} + R}$$

where, *NIR* is the surface reflectance of near-infrared band and *R* is the surface reflectance of red band. (2) Enhanced vegetation index, EVI (Jiang et al., 2008):

$$\text{EVI} = 2.5 \times \frac{\text{NIR} - R}{\text{NIR} + 2.4 \times R + 1}$$

And (3) the near-infrared reflectance of vegetation, *NIR_v* (Badgley et al., 2017, 2019):

$$\text{NIR}_v = \text{NDVI} \times \text{NIR}$$

It is noteworthy that we use the two-band EVI (Jiang et al., 2008) here instead of the three-band one which also uses the blue band. This is because the signal-to-noise ratio of blue band is commonly poor (Rocha and Shaver, 2009). The *Meteo.* is reanalysed meteorological forcings including solar radiation and temperature (Joiner and Yoshida, 2020). The *aux.* are auxiliary parameters that typically include site timestamps and plant functional types (PFTs). It is also noteworthy that satellite products usually limit the temporal resolution of extrapolated flux estimates as the meteorological variables are typically resolved more frequently – e.g., the fifth generation atmospheric reanalysis of the European Centre for Medium-Range Weather Forecasts (ERA5) is available on an hourly scale (Hersbach et al., 2020).

We applied this same framework to the flux extrapolation using 206

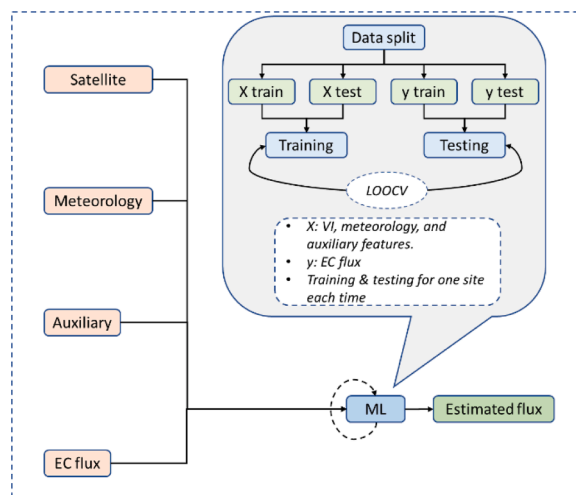


Fig. 2. Schematic overview of the flux extrapolation model and data flow (black arrows). The ML therein refers to machine learning algorithms, EC is eddy covariance, VI is vegetation indices, and LOOCV is the leave-one-out cross validation.

FLUXNET towers and using our three pasture towers (Fig. 2). As the EC tower footprint varies tower by tower and time by time, we interpolated the gridded data (satellite vegetation indices and meteorological forcings) to the tower locations using the inverse distance weighting (IDW) interpolation (Bartier and Keller, 1996) to reduce the uncertainty of only using the nearest pixels. The eXtreme Gradient Boosting (Xgboost) was chosen as the machine-learning algorithm for extrapolation, because it is highly flexible, fast (i.e., supporting parallel processing), and it tolerates a small amount of missing data (Chen et al., 2015). To account for the differences between PFTs, we validated the extrapolation model separately for each PFT represented by the International Geosphere Biosphere Programme (IGBP) by employing the leave-one-out cross validation, LOOCV (Marchetti, 2021) to generate training and test datasets – i.e., we treated data from a single tower ($i \in [1, n]$) as the test dataset and data from all other towers ($1, \dots, i-1, i+1, \dots, n$) within the same PFT as the training dataset. Using this approach, each tower was used once as a test dataset. For example, among the FLUXNET2015 tower network comprising 206 sites, 20 of them were located within cropland areas. The machine learning model was trained on data from 19 towers, and then used to predict fluxes for the remaining tower. This approach was similar to that used by Joiner and Yoshida (2020) and was implemented to avoid issues related to overtraining and incorrect performance assessment that can arise when using conventional k-fold data splitting methods. This method was repeated 206 times for the FLUXNET2015 towers and three times for the pasture sites, with each tower's data being used once as the test dataset. By splitting the data in this way, a comprehensive and reliable performance assessment was achieved.

We utilised statistical metrics in presenting performance of the machine-learning model at the “tower level”. Please note throughout this manuscript, the term “tower level” refers to data, such as fluxes and satellite observations, that were spatially interpolated into EC tower geolocations using the inverse distance weighting (IDW) interpolation method described above. The use of IDW interpolation method matters because it accounts for the difference in spatial resolution between the gridded data used, such as satellite observations (Bartier and Keller, 1996). Unlike other interpolation methods, such as nearest neighbour or bilinear interpolation, which use only a few surrounding pixels, IDW theoretically utilises all pixels for interpolation, giving weights to each pixel that are inversely related to their distance from the target geolocation (Bartier and Keller, 1996). Consequently, closer pixels are given higher weights in the interpolation process, while farther pixels are given lower weights. As the FLUXNET2015 database does not provide open-access footprint information and the spatial resolution of satellite data can be smaller or larger than footprints (Xiao et al., 2012; Tramontana et al., 2015; Ichii et al., 2017; Jung et al., 2019; Joiner and Yoshida, 2020), using IDW interpolation is a viable alternative to extracting pixels according to daily tower footprints that continuously vary (Chu et al., 2017, 2021). Metrics for assessing the model performance were the coefficient of determination (R^2) and mean bias error (MBE), and the uncertainty was measured using the interquartile range (IQR) of the MBE. In addition, we presented the root mean squared error (RMSE) as a point of reference to facilitate the assessment of the model performance for readers.

2.2. Study sites

In this study, we used two sets of EC towers – i.e., global 206 towers from the FLUXNET2015 database and three pasture towers we operated in North Wyke, southwest England. Details regarding the two sets will be given separately as follows.

2.2.1. FLUXNET2015 sites

The FLUXNET2015 database was released in 2020 (Pastorello et al., 2020). It includes 206 open-access towers and another six towers under a relatively restrictive data policy. The 212 towers in total cover 13 IGBP ecosystem types – croplands (CRO, 20 towers), closed shrublands (CSH,

3 towers), deciduous broadleaf forests (DBF, 26 towers), deciduous needleleaf forests (DNF, 1 tower), evergreen broadleaf forests (EBF, 15 towers), evergreen needleleaf forests (ENF, 50 towers), grassland (GRA, 38 towers), mixed forests (MF, 10 towers), open shrublands (OSH, 13 towers), savannahs (SAV, 8 towers), snow and ice (SNO, 1 tower), wetlands (WET, 21 towers), and woody savannahs (WSA, 6 towers). The total site-years are over 1500 years up to 2014 inclusively. In terms of continent, 84 towers are in North America, 71 in Europe, 23 in Australia, 21 in Asia, 7 in South America, and 6 in Africa.

The FLUXNET2015 database provides tower-level net ecosystem exchange (NEE), gross primary productivity (GPP), ecosystem respiration (Reco), sensible heat (H), and latent energy (LE) fluxes as well as meteorological data (incl. solar radiation, temperature, and vapour pressure deficit) at a half-hourly frequency (<https://fluxnet.org/data/fluxnet2015-dataset/fullset-data-product/>). The GPP and Reco data therein are partitioned from NEE using the night-time (Reichstein et al., 2005) and day-time (Lasslop et al., 2010) partitioning methods. Data of all towers are inter-comparable as they were processed and quality controlled using the same data-processing pipeline (Pastorello et al., 2020). In this study, we used all 206 open-access towers to evaluate the influence of satellite resolutions on the flux extrapolation.

2.2.2. North Wyke pasture sites

The three pasture EC towers (Fig. 3) were operated in the North Wyke Farm Platform (NWFP, established in 2010) – a national & global agriculture research facility in southwest England (120–180 m.a.s.l., 50°46'10"N, 3°54'05"W). Annual precipitation was 1032 mm and the average minimum and maximum daily temperatures were 6.8 and 13.5 °C over the past 30 years (Orr et al., 2016; Cardenas et al., 2022). The average wind speed was 10 m s⁻¹ and the main wind direction was westerly (Cardenas et al., 2022). We used three EC towers (data from 2017 to the beginning of 2019) with LI-COR setups [referred to as ‘LI-COR towers/systems’] (LI-COR Inc., 2017). The three pasture farmlets are characterised as high sugar grass (HS, 6.65 ha with the *Lolium perenne* grass variety *AberMagic*), permanent pasture (PP, 6.39 ha predominantly perennial ryegrass, *Lolium perenne*), and high sugar/white clover mix (HSC, 6.49 ha with a combination of *AberMagic* and the white clover variety *AberHerald*) (Fig. 3). The farmlet PP is used as a control treatment for a grassland management system in the region, as it is based on typical pasture and management, and it has not been ploughed for the previous 20 years. In contrast, farmlet HS and HSC were ploughed and re-seeded in 2013 after a baseline measurement period when all the fields were under the same original pasture (Orr et al., 2016). All three farmlets are grazed from April to October with



Fig. 3. Locations three existing LI-COR EC systems. The figure is presented via the Google Earth platform (Lisle, 2006).

cattle (ca. 4 ha⁻¹), lamb (ca. 17 ha⁻¹), and sheep (ca. 10 ha⁻¹).

All three LI-COR systems used the same setup but are mounted at different heights; tower heights and sensors for all the EC systems and environmental variables can be found in Table 1. Gas concentrations and wind velocities were recorded at a 20 Hz frequency.

2.3. Data

2.3.1. FLUXNET2015 EC data

We used 10 years (2005–2014) of daily averaged EC fluxes from the FLUXNET2015 database (Pastorello et al., 2020). These fluxes include (1) net ecosystem exchange using Variable Ustar Threshold (VUT) with reference (NEE_VUT_REF), (2) gross primary production using the night-time (NT) partitioning method (GPP_NT_VUT_REF), ecosystem respiration using the night-time (NT) partitioning method (GPP_NT_VUT_REF), sensible heat using the marginal distribution sampling (MDS) gap-filling method (H_F_MDS), and latent energy using the marginal distribution sampling (MDS) gap-filling method (LE_F_MDS). Further variable explanations can be found at <https://fluxnet.org/data/fluxnet2015-dataset/fullset-data-product/>.

We applied quality control criteria to the fluxes following the literature (Jung et al., 2019; Joiner and Yoshida, 2020). We only allowed daily points where 20% or less came from gap-filled data. For NEE, GPP, and Reco, the uncertainty of NEE should be smaller than 3 g C m⁻² d⁻¹ and the difference between night-time and day-time partitioned GPP & Reco should be also smaller than 3 g C m⁻² d⁻¹.

2.3.2. North Wyke EC data

In addition, we used daily averaged EC data from the three LI-COR towers in the North Wyke Farm Platform, including net ecosystem exchange (NEE), night-time partitioned GPP and ecosystem respiration (Reco) (Reichstein et al., 2005), sensible heat (H), and latent energy (LE) from the three LI-COR systems. The data processing was carried out using the EddyPro (v6.2.2, Li-Cor Inc., Lincoln, NE, USA) software (LI-COR Inc., 2017) and it referred to the FLUXNET2015 standard processing pipeline (Pastorello et al., 2020). The quality control criteria we applied to the North Wyke EC fluxes were the same as the ones for FLUXNET2015 fluxes. As the presence of data gaps poses a challenge to the comprehensiveness of EC flux time series, here the fluxes were gap-filled using the random forest regression (RFR)-based method (Zhu et al., 2022) which exhibited global advantages over other gap-filling methods even in challenging conditions such as those involving extended gaps exceeding one month and managed ecosystems. In addition, this RFR gap-filling method has been validated in North Wyke (Zhu et al., 2023).

Table 1
Technical details on the tower and sensors for the three EC systems.

	LI-COR (closed-path)
Tower height (m)	1.59 (HS), 1.57 (PP), 1.59 (HSC)
Gas analyser	LI-7200 ¹
Sonic anemometer	Windmaster Pro ²
Data logger	LI-7550 ¹
Temperature sensor	HMP155 ³
Humidity sensor	HMP155 ³
Photosynthetic active radiation	LI-190SL ¹
Incoming and reflected radiation	CNR4 ⁴
Soil heat exchange	HFP01 ⁵
Soil moisture and temperature	Hydra Probe II
Precipitation data	Nearby automated weather station ⁶

¹ Li-COR Inc., Lincoln, NE, USA.

² Gill Instruments Ltd., Lymington, UK.

³ Vaisala, Helsinki, Finland.

⁴ Kipp&Zonen, Delft, Netherlands.

⁵ Hukseflux Thermal Sensors B.V., Delft, Netherlands.

⁶ <https://nwfp.rothamsted.ac.uk/>.

2.3.3. Satellite data

We used the atmospherically corrected surface reflectance products from five satellites to calculate the vegetation indices and the data was retrieved from the Google Earth Engine (<https://earthengine.google.com/>) platform (Gorelick et al., 2017). The five satellite products (Table 2) were from (1) Landsat-7/Enhanced Thematic Mapper Plus (ETM+) sensor (Chander et al., 2009), (2) Landsat-8/Operational Land Imager (OLI) sensor (Vermote et al., 2016), (3) Sentinel-2A and 2B harmonised Multi-Spectral Instrument (MSI) sensor (Drusch et al., 2012), (4) Aqua and Terra Moderate Resolution Imaging Spectroradiometer (MODIS) sensor (Schaaf and Wang, 2015), and (5) the Advanced Very High Resolution Radiometer (AVHRR) sensors onboard seven National Oceanic and Atmospheric Administration (NOAA) polar orbiting satellites (Claverie et al., 2014).

As this study includes two parts, in the first part which aims to support the second part, we used one high-resolution (Landsat-7) and two moderate-resolution (MODIS and AVHRR) products for examining if consistent flux estimates can be derived from satellites with different spatial resolutions, as these three satellites cover the date range of FLUXNET2015 (Table 2). In the second part of studying flux extrapolation in North Wyke pastures, we used the three high-resolution satellite products that are open-access for research purposes – i.e., Landsat-7/-8 and Sentinel-2. For Landsat-7/-8, the data were from Collection 2 and Tier 1 as recommended by the data producer United States Geological Survey (USGS) (<https://www.usgs.gov/landsat-missions/landsat-collection-2>). Collection 2 improved the data processing algorithms and data distribution capabilities. The Tier 1 products have the highest data geometric and radiometric quality in comparison to other tiers (Dwyer et al., 2018). The Sentinel-2 product was from the Sentinel-2 satellite constellation (Drusch et al., 2012) which includes two satellites – Sentinel-2A and Sentinel-2B. The Sentinel-2A satellite launched in June 2015 and Sentinel-2B launched in March 2017; the revisit frequency of the combined constellation is five days (Szantoi and Strobl, 2019). The MODIS product was Nadir Bidirectional Reflectance Distribution Function (BRDF)-Adjusted reflectance from two satellites – Terra (launched in 1999 with an overpass time at 10:30 on descending passes) and Aqua (launched in 2002 with an overpass time at 13:30 on ascending passes) depending on which satellite had the best representative pixel from the 16-day period (Schaaf and Wang, 2015). The AVHRR surface reflectance product a NOAA Climate Data Record (CDR) from seven NOAA satellites, i.e., NOAA-7, -9, -11, -14, -16, -17, -18 and -19, and it is noteworthy that the orbital drift of NOAA-19 seriously degraded the data quality since 2014 (Bojanowski and Musial, 2020). Satellite images were interpolated into a daily frequency using the method described in Joiner and Yoshida (2020).

2.3.4. Meteorology forcing data

The meteorological forcing data were from the fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalysis of the global land climate (ERA5-Land) (Hersbach et al., 2020). The ERA5-Land reanalysis provides hourly meteorological data from 1950 to present on a 0.1° x 0.1° spatial resolution. Here, we used 2 m temperature (TA), 2 m dewpoint temperature which was used to calculate the vapour pressure deficit (VPD), surface solar radiation downwards (SW_IN), surface thermal radiation downwards (LW_IN), 10 m u-component of wind (U), 10 m v-component of wind (V), and total precipitation (P). Among these variables, solar radiation, air temperature, and VPD were used because they are commonly acknowledged as primary drivers of fluxes (Reichstein et al., 2005; Moffat et al., 2007) while others were considered as auxiliary meteorological variables that may affect measured and estimated fluxes (Zhu et al., 2022, 2023). As described above, the gridded data were interpolated into EC tower locations using the inverse distance weighting (IDW) interpolation (Bartier and Keller, 1996). Furthermore, due to the high temporal resolution, the meteorological forcing data were interpolated into a daily frequency.

Table 2

Information for used satellite surface reflectance products. The L2 below means Level-2 product and L2A therefore means Level-2 A product. The V5 and V6 separately means Version 5 and Version 6. The integers in R and NIR band columns are the satellite band numbers. Note the Sentinel-2 product was from two satellites (S2A and S2B) with slightly different band wavelength. Therefore two ranges of wavelength were provided for S2A and S2B, respectively.

Satellite	Product	Frequency (days)	Spatial resolution (m)	R band (nm)	NIR band (nm)	Start date
Landsat-7/ETM+	L2	16	30	3 (630–690)	4 (770–900)	28/05/1999
Landsat-8/OLI	L2	16	30	4 (636–673)	5 (851–879)	18/03/2013
Sentinel-2/MSI	L2A	5	10	4 (664.5 / 665)	8 (835.1 / 833)	28/03/2017
Terra & Aqua/MODIS	MCD43A4 (V6)	1	500	1 (620–670)	2 (841–876)	24/02/2000
NOAA/AVHRR	V5	1	5566	1 (640)	2 (860)	24/06/1981

3. Results

3.1. Multi-resolution flux extrapolation comparisons at FLUXNET2015 towers

The scatter plots of machine learning estimates against EC measurements for the same flux were very similar across satellites and/or vegetation indices (Fig. 4). In general, the correlation between EC measured and machine learning estimated fluxes was lower for NEE than the other fluxes, with R^2 averages of 0.42, 0.65, 0.5, 0.61, and 0.57 for NEE, GPP, Reco, H, and LE.

The differences in R^2 when using vegetation indices from same satellite were small, and less than 0.02 for any of fluxes, with NIRv generally performing slightly better than the EVI and NDVI vegetation indices. Comparing between satellites, the largest differences were seen for GPP and NEE, where MODIS exhibited a higher R^2 by 0.04 for GPP and a higher R^2 by 0.03 for NEE over the Landsat-7 and AVHRR satellites (Fig. 4). The R^2 differences between satellites for the fluxes were ≤ 0.02 .

In terms of mean bias error (MBE), the differences between satellites and/or vegetation indices were relatively small too. Similar phenomena were also seen for root mean squared error (RMSE). In most cases, the differences in MBE between vegetation indices from the same satellite were smaller than $0.05 \text{ Mg ha}^{-1} \text{ yr}^{-1}$ for NEE, GPP, and Reco and they were smaller than $5 \text{ MJ m}^{-2} \text{ yr}^{-1}$ for H and LE (Fig. 4). For GPP, Landsat-7 exhibited the smallest absolute MBE of $0.07 \text{ Mg ha}^{-1} \text{ yr}^{-1}$ on average (Fig. 4). For the other two carbon fluxes (i.e., NEE and Reco), the differences between satellites and/or vegetation indices were smaller than $0.1 \text{ Mg ha}^{-1} \text{ yr}^{-1}$ in most cases (Fig. 4). The differences between satellites and/or vegetation indices were smaller than $10 \text{ MJ m}^{-2} \text{ yr}^{-1}$ for H and the differences were smaller than $20 \text{ MJ m}^{-2} \text{ yr}^{-1}$ for LE (Fig. 4). The difference in RMSE between vegetation indices from the same satellite were smaller than $1 \text{ Mg ha}^{-1} \text{ yr}^{-1}$ for NEE, GPP, and Reco and they were smaller than $10 \text{ MJ m}^{-2} \text{ yr}^{-1}$ for H and LE (Fig. 4).

The scatter plots of flux extrapolation estimates varied largely between ecosystem types (Fig. 5). The correlations between EC measured and machine learning estimated fluxes were lower for ecosystem types where the number of existing EC towers was relatively few. Specifically, the correlation was particularly low (< 0.1 for most cases) in close shrublands (CSH) where there were only three EC towers. The correlation was relatively lower in woody savannahs (WSA, 6 towers), savannahs (SAV, 8 towers), and evergreen broadleaf forests (EBF, 15 towers) than in ecosystem type like evergreen needleleaf forests (ENF, 50 towers). In croplands (CRO, 20 towers) and grasslands (GRA, 38 towers), the correlation was relatively high. In terms of R^2 , the average between croplands and grasslands across fluxes was 0.6. The mean bias error (MBE) average was $0.26 \text{ Mg ha}^{-1} \text{ yr}^{-1}$ for NEE, H, and LE and $17.50 \text{ MJ m}^{-2} \text{ yr}^{-1}$ for H and LE (Fig. 5). The mean root mean squared error (RMSE) average was $7.00 \text{ Mg ha}^{-1} \text{ yr}^{-1}$ for NEE, H, and LE and $93.14 \text{ MJ m}^{-2} \text{ yr}^{-1}$ for H and LE (Fig. 5).

3.2. Fine-resolution extrapolation at the three North Wyke pasture sites

The correlations between EC measurements and machine learning estimated fluxes were relatively high at the three pasture sites (Fig. 6).

The averaged R^2 across towers and fluxes was 0.81. The annual estimation error was small, particularly for carbon fluxes ($\ll 1.5 \text{ Mg ha}^{-1} \text{ yr}^{-1}$) (Fig. 6a–c). Regarding ecosystem flux types, the R^2 for ecosystem respiration (Fig. 6c) was greater than for other fluxes. The difference in R^2 between towers was very small except for latent energy (Fig. 6e).

Furthermore, the fitting correlations between measured and estimated fluxes were close to one (Fig. 6). The first quartile, median, and the third quartile of the fitted slopes were 0.77, 0.85, and 0.89, respectively. The fitted slopes at tower HS were slightly lower than the other towers. The low slope, that infers the existence of underestimation by the model, was observed particularly for LE (Fig. 6e).

In general, the cumulative time series of flux estimates were consistent with EC measurements (Fig. 7). The absolute ratio of cumulative error in flux estimation to the overall flux (referred to as R_{err} , it is an absolute value) varied flux by flux and tower by tower (Fig. 7 and Table 3). The R_{err} was smaller for GPP, Reco, and H. The tower-average R_{err} values for these three fluxes were 2.79%, 5.38%, and 7.11%, respectively. Whilst the tower-average R_{err} values for NEE and LE were 11.97% and 17.69%, respectively. No obvious tower-related R_{err} patterns were observed, e.g., the lowest R_{err} for NEE (1.55%) was seen at tower HSC while the lowest R_{err} for GPP (0.75%) and LE (10.50%) were separately seen at tower HS and tower PP.

The four-year averaged (2017–2020) fluxes across all three fields were -0.08 Mg yr^{-1} (NEE), 0.38 (GPP) , 0.41 (Reco) , $103,325.66 \text{ (H)}$, $243,990.57 \text{ MJ yr}^{-1}$ (LE) (Fig. 8). The differences in annual fluxes between the three satellite platforms were generally small – e.g., the differences in GPP and NEE estimation were smaller than 0.01 Mg yr^{-1} and the difference in H was smaller than 5000 MJ yr^{-1} (ca. 5% of the total flux). However, it is noteworthy that the difference in Reco was relatively large, i.e., 0.07 Mg yr^{-1} , which was almost equal to the whole field-scale NEE.

Overall, the extrapolated flux spatial distribution was homogeneous as expected (Cardenas et al., 2022). For example, the GPP spatial variability in any field and for any satellite was smaller than $10^{-3} \text{ Mg yr}^{-1}$. The extrapolated GPP exhibited regional variability which was very small (ca. $0.2 \text{ g C m}^{-2} \text{ d}^{-1}$). Sentinel-2 and Landsat-7 NIRv exhibited strong image texture that may affect satellites capturing the flux spatial variability (Fig. 8a and b). Whilst Landsat-8 NIRv calculated from the surface reflectance product still exhibited cloud-related patterns in North Wyke (Fig. 8c). Compared to Landsat-7, the Sentinel-2 NIRv showed clearer image details. The contrast between land-use types in Sentinel-2 NIRv was more intensive (Fig. 8d).

4. Discussion

4.1. Fine-resolution extrapolation feasibility and implications

4.1.1. Satellite resolution and extrapolation performance

Fine-scale flux extrapolation can be considered viable according to the results from global FLUXNET2015 sites – i.e. the consistent flux estimates using satellites with various spatial resolutions (Fig. 4). The measured EC fluxes were found to closely match the estimates obtained from satellite data, which varied in spatial resolution from 30 m to 5566 m. This suggested that the fine-scale machine learning-based flux

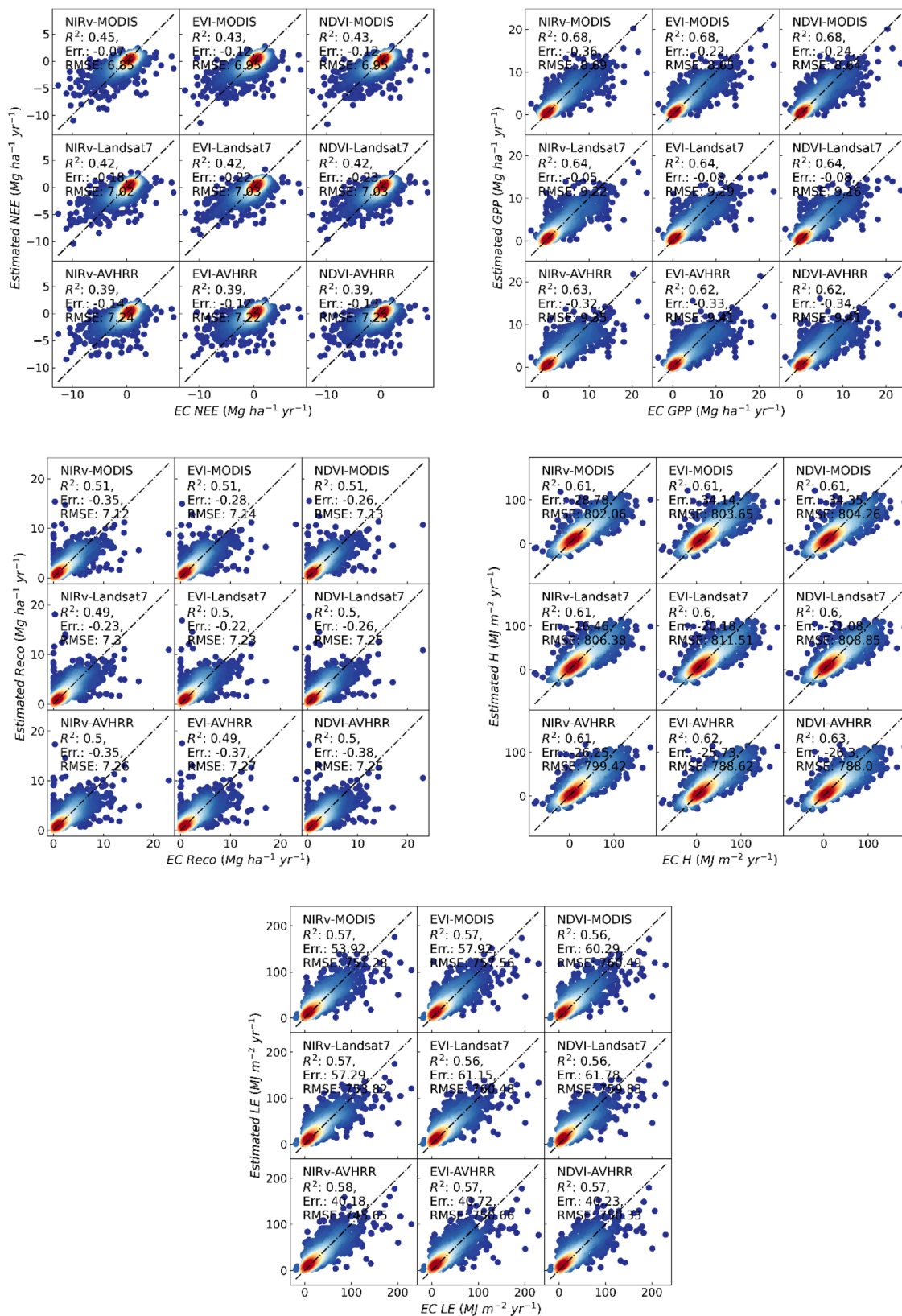


Fig. 4. Scatter plots of tower-level (i.e., using data IDW interpolated from multiple pixels) estimated fluxes against EC measurements for all the FLUXNET2015 towers. The sub-plots in each panel represent the results using different satellites and/or vegetation indices. The R^2 and MBE (i.e., Err.) are provided on the figure. The MBE unit is $\text{Mg ha}^{-1} \text{ yr}^{-1}$ for NEE, GPP, and Reco and it is $\text{MJ m}^{-2} \text{ yr}^{-1}$ for H and LE. The dashed black lines represent the one-to-one relationship.

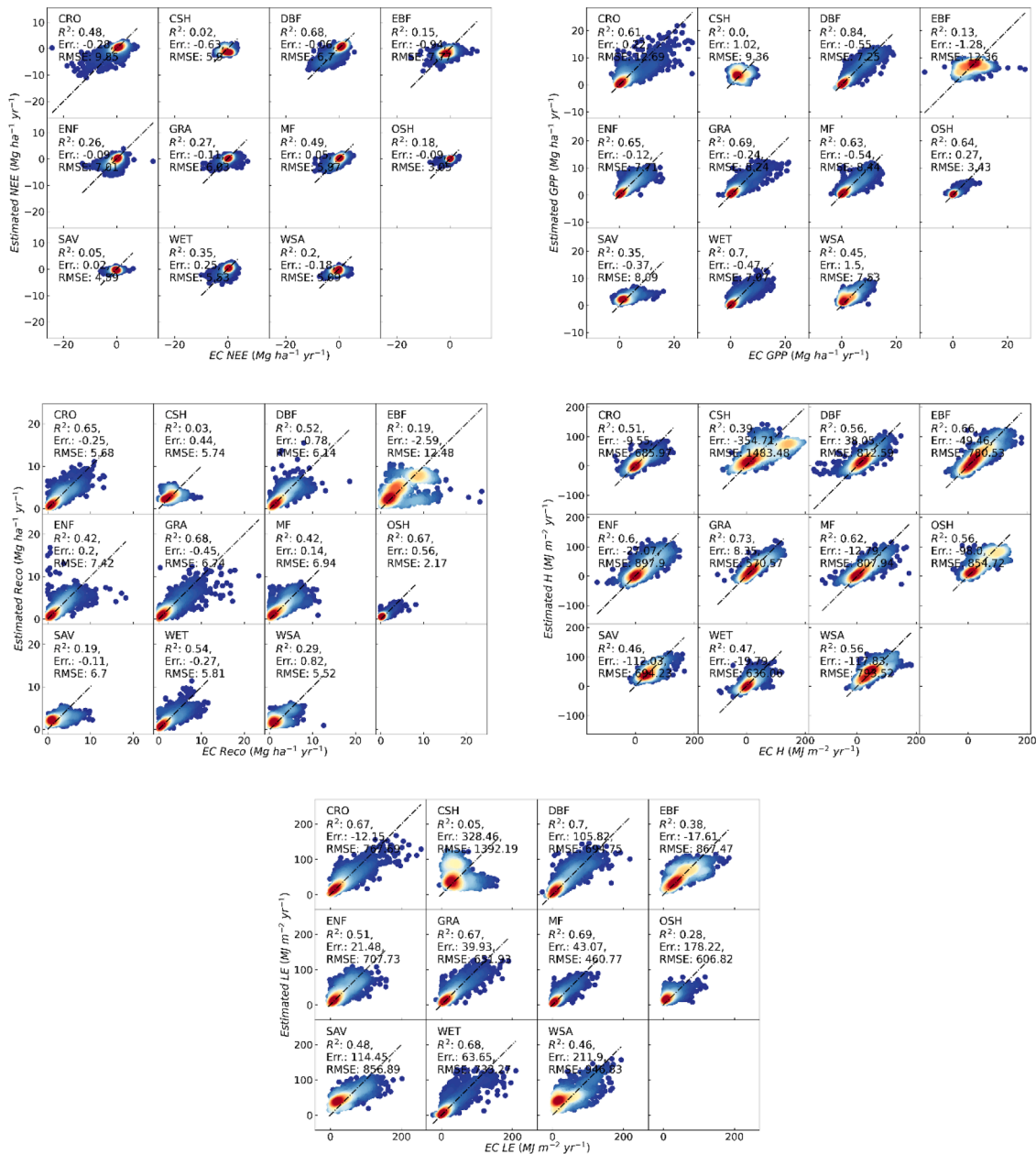


Fig. 5. Scatter plots of estimated fluxes against EC measurements for all the FLUXNET2015 towers across 11 IGBP ecosystem types. The exhibited flux estimates were the averages of using different satellites and vegetation indices. The R^2 and MBE (i.e., Err.) are provided on the figure. The MBE unit is $Mg\ ha^{-1}\ yr^{-1}$ for NEE, GPP, and Reco and it is $MJ\ m^{-2}\ yr^{-1}$ for H and LE. The dashed black lines represent the one-to-one relationship.

estimation or extrapolation was promising given the agreement in estimated fluxes of using satellites with different spatial resolutions. This finding set the fundamentals for the second part of this study at North Wyke. Results from our three pasture towers also supported this argument (Fig. 8). Field-scale flux (i.e., the total extrapolated flux of the whole field) between satellites were in a high-level consistency. However, the selection of satellite is important for capturing the flux spatial variability even if the candidate satellites have the same spatial resolution. We suggest this is because the spatial resolution is not the only factor affecting the effectiveness of satellites in capturing fluxes. The relationship between fluxes and environment is complex and not solely determined by vegetation greenness as indicated by vegetation indices. Instead fluxes depend on multiple factors including the vegetation status, climatic conditions, as well as nutrient and water supply (Chapin III et al., 2011).

The meteorological forcing data (i.e., ERA5-Land) utilised in this

study exhibited a considerably lower spatial resolution (0.1° , ca. 10 km) compared to both the EC tower footprint (with radii of normally hundreds of metres) and the satellites (ranging from 30 to 5566 m). Despite the implementation of spatial interpolation methods in both this study and the literature (Jung et al., 2020; Joiner and Yoshida, 2020), the employment of ERA-Land data, which possessed a higher spatial resolution in comparison to the meteorological forcing data employed in prior research (Jung et al., 2020; Joiner and Yoshida, 2020), could potentially influence the accuracy of flux estimation. It is therefore crucial to acknowledge this limitation. In the forthcoming years, the resolution of local meteorological forcing at high spatial scales or the deployment of high-density *in-situ* networks for meteorological measurements could provide a viable solution.

4.1.2. Ecosystem type and extrapolation performance

The performance of the flux estimation varied greatly with

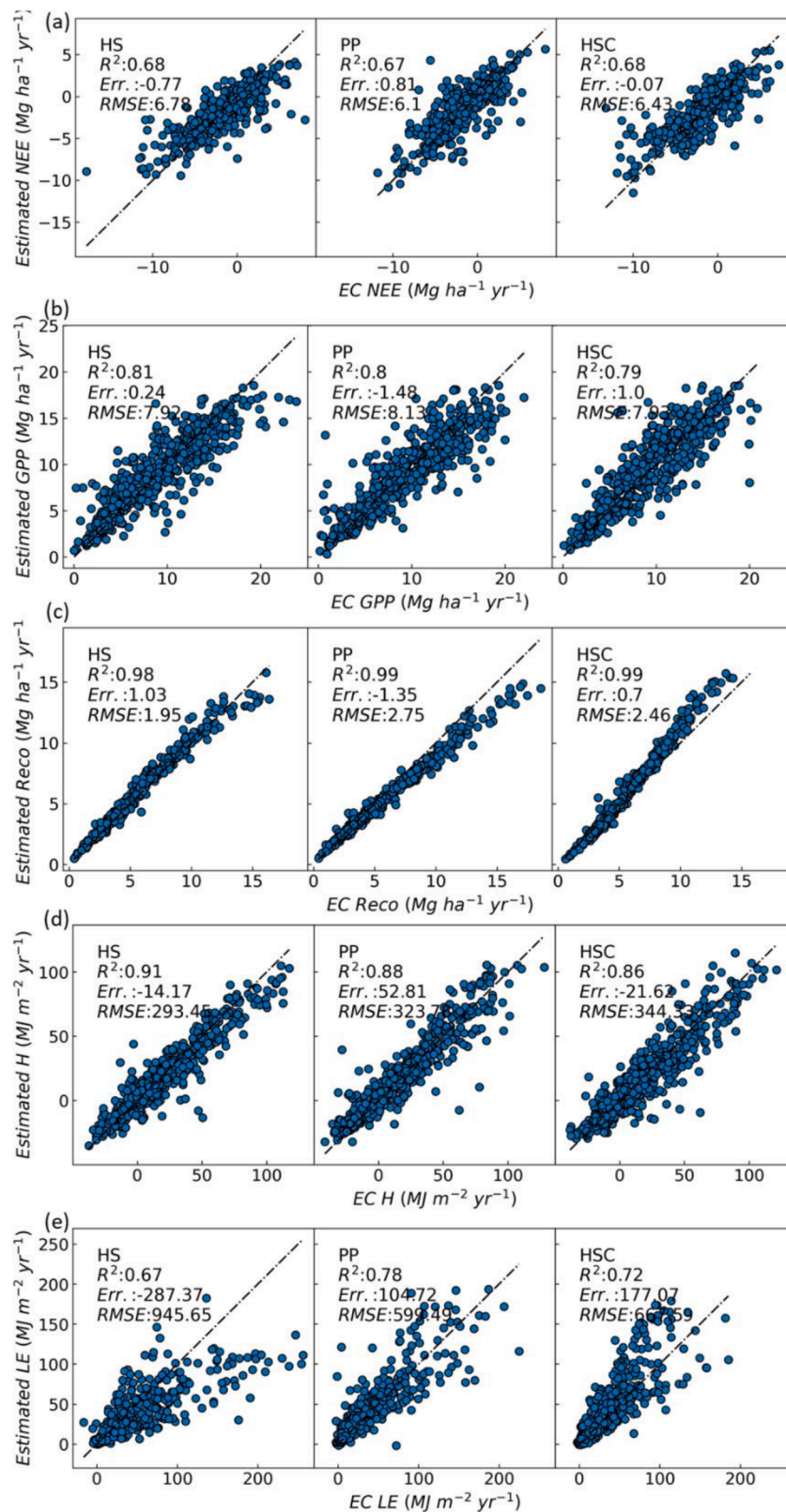


Fig. 6. Tower-level (i.e., using data IDW interpolated from multiple pixels) estimates against EC measurements for net ecosystem exchange (NEE) (a), gross primary productivity (GPP) (b), ecosystem respiration (Reco) (c), sensible heat (H), and latent energy (LE). The R^2 and MBE (i.e., Err.) are provided on the figure. The unit is $\text{Mg ha}^{-1} \text{ yr}^{-1}$ of MBE for (a) – (c) and $\text{MJ m}^{-2} \text{ yr}^{-1}$ for (d) – (e), respectively. The dashed black lines represent the one-to-one relationship.

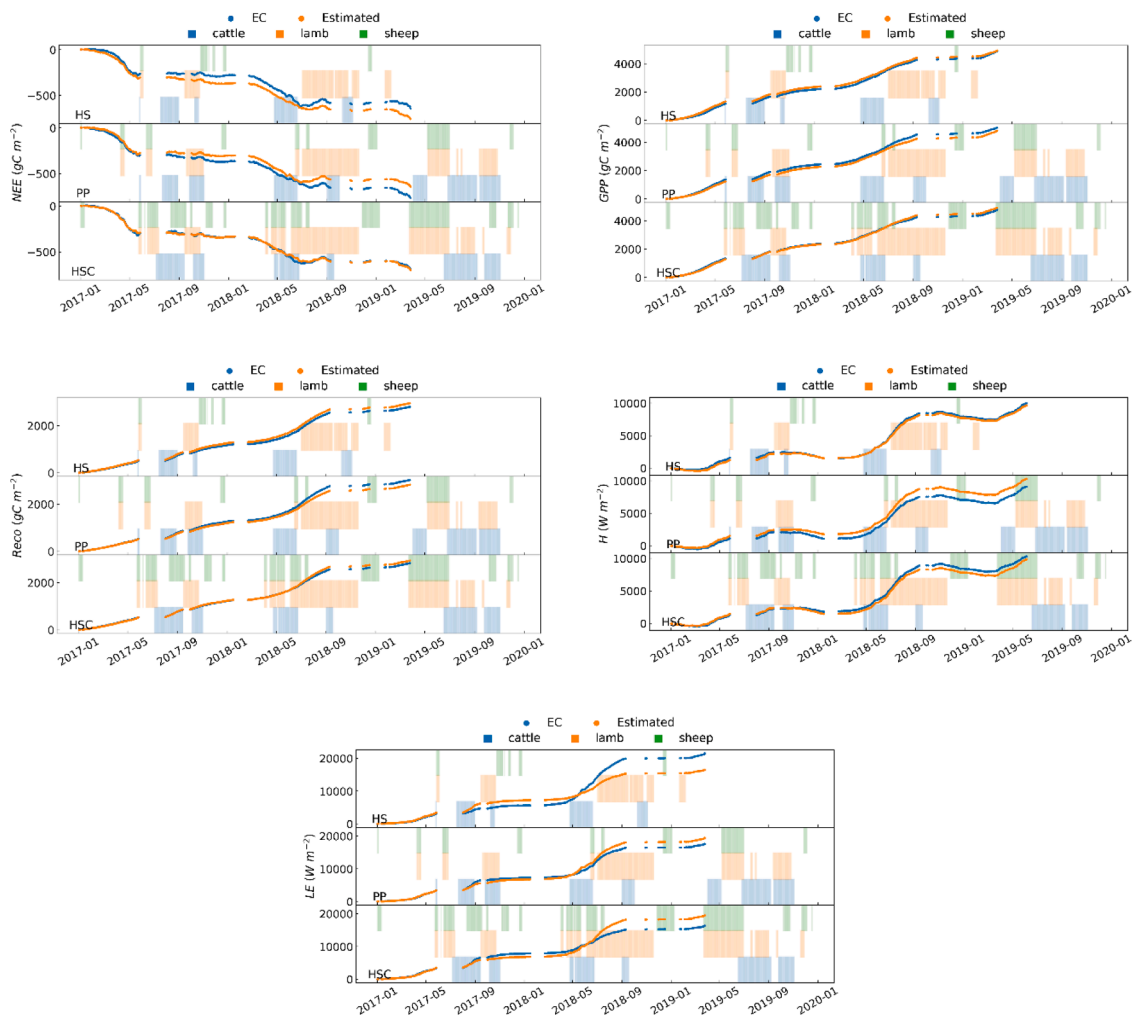


Fig. 7. Cumulative fluxes of EC measurements and the tower-level estimates. The estimates at each site were predicted from the other two sites. The blocks represent the occurrence of grazing events. The results were derived from the Sentinel-2 satellite.

Table 3

The ratio (%) of cumulative error (against the EC measurements) in tower-level flux estimation to the overall flux over two years.

	HS	PP	HSC
NEE	18.30	-16.07	1.55
GPP	0.75	-4.46	3.16
Reco	5.60	-6.78	3.76
H	-3.25	13.29	-4.78
LE	-23.51	10.50	19.07

ecosystem type (Fig. 5). The estimation performed better in grass and forest ecosystems than in shrubland ecosystems. This performance difference can be accounted for by many reasons. First, it was suggested to be relative to the number of existing EC towers. According to Fig. 5, the better performance tended to be seen in ecosystem types with a higher EC sampling rate. This infers that it could be difficult to predict the flux in an ecosystem where no similar samples were included in the training dataset. Second, this performance difference might be related to the ecosystem itself as well. For example, closed shrublands (CSH) had relatively smaller GPP than very productive ecosystems and such weak signals could pose challenges to machine learning in capturing the flux variability (Fig. 5). In addition, the flux seasonality or periodicity of an ecosystem could be very important too. For example, the evergreen broadleaf forests (EBF) had very strong but relatively stationary GPP time series (Jung et al., 2020), therefore the temporal variability can be

difficult to capture and thereby leading to a low explained variance (i.e., R^2) (Fig. 5). Fortunately, the global estimation performance for grasslands which include pastures were good, and this paves the way for our study in the North Wyke pastures.

4.1.3. The values of 30 m flux extrapolation

The 30-m flux extrapolation performance was good (averaged R^2 was 0.55 across fluxes) on the global scale and was satisfactory (averaged R^2 was 0.88 across fluxes) in the North Wyke pastures. It is noteworthy that the averaged R^2 for FLUXNET2015 towers in this study was lower than the literature – e.g., 0.7 in Joiner and Yoshida (2020). This is because Joiner and Yoshida (2020) focused on weekly GPP estimation whilst we also estimated NEE and worked at a daily scale. As NEE is the difference of two greater fluxes (GPP and Reco), estimating NEE is potentially more challenging than GPP (Aubinet et al., 2012). Therefore, the overall performance was good and fine-resolution EC flux extrapolation at a daily scale is feasible. The cumulative errors (Fig. 7) were generally small for GPP ($R_{err} = 2.79\%$). The large error ratio for NEE can be accounted for by the weak signal and the large error for LE was relative to the relatively low energy balance closure here (ca. 60%) due to the measurement height (Cardenas et al., 2022) This low energy balance ratio could reduce the measurement accuracy. More importantly, the estimated flux time series were resistant to the disturbances of livestock grazing (Fig. 7). We did not use grazing periods as a machine learning predictor, but the estimated time series were not strongly affected by grazing. This infers great application potential in other managed

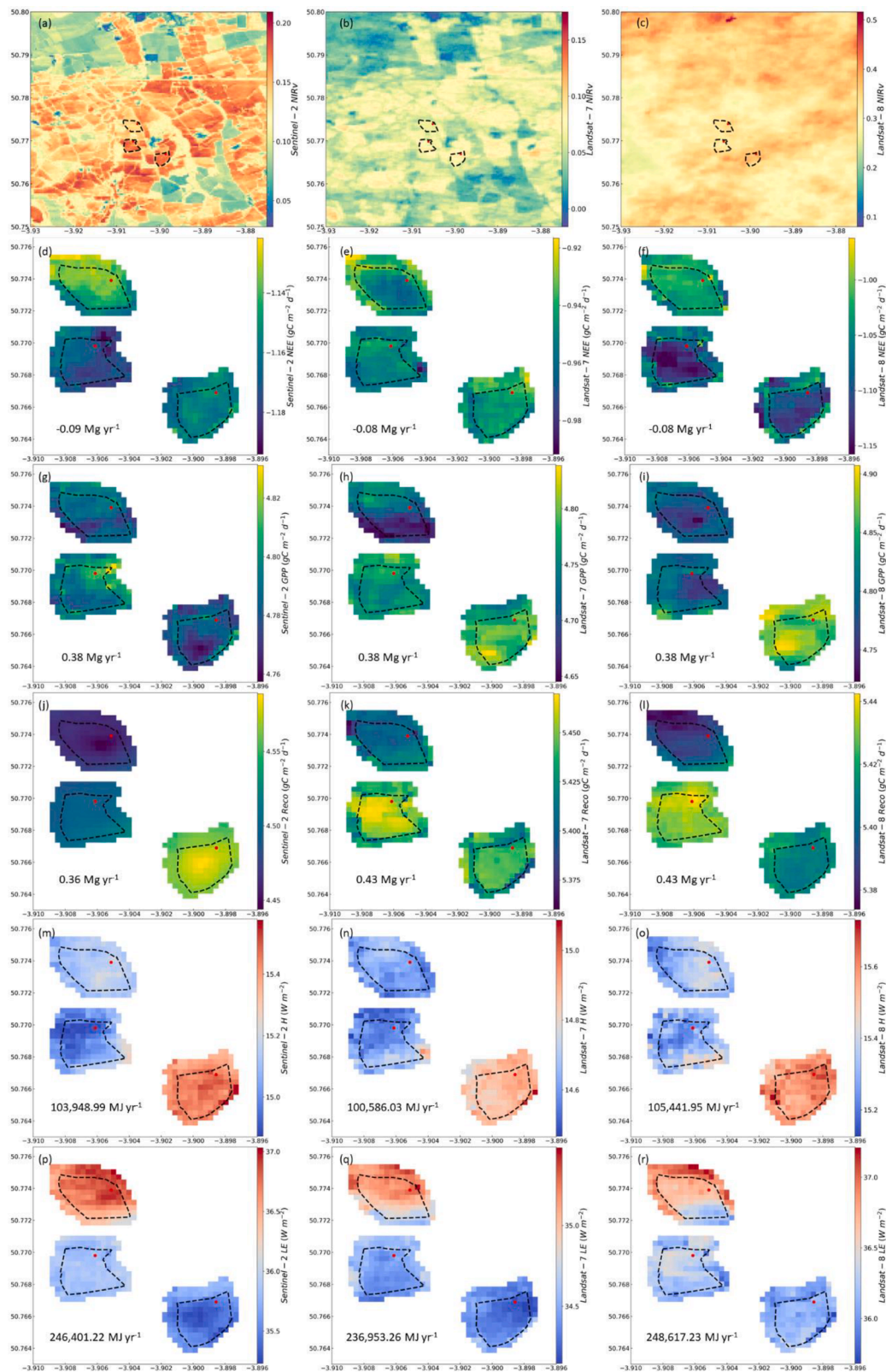


Fig. 8. Multi-year averages (2017–2020) of extrapolated fluxes in the three fields. The flux estimates were extrapolated from EC measurements gap-filled using the random forest regression (RFR) method at the three towers separately located in the three fields. The top row (a)–(c) are multi-year averaged NIRv from the three satellites in order. From top to bottom, the second row to the last rows are multi-year averaged NEE, GPP, Reco, H, and LE fluxes. From left to right, the columns are results using Sentinel-2, Landsat-7, and Landsat-8 NIRv data.

ecosystems with agricultural practices.

Extrapolating EC fluxes at a fine spatial resolution is very informative for ecosystem management. For example, the average area of the three studied fields is 6.5 ha⁻¹, but 80% of the footprint was only 3–70 m upwind from the tower in these farmlets (Cardenas et al., 2022). Hence, the necessity of extrapolating EC towers should be apparent for

quantifying fluxes over the whole ecoregion. The fine spatial scale extrapolation results here were promising, because the current performance of fluxes estimated directly from satellites was reported to be poor (Wang et al., 2017) and the spatial resolution of mainstream extrapolation products was too coarse (> 5 km) for these relative small ecoregions (Jung et al., 2020; Joiner and Yoshida, 2021). The good

validation performance (i.e., high R^2 and small bias, Fig. 6) suggests that the proposed fine-scale extrapolation approach may have high potential if used in other areas.

Despite the consistent flux estimates from satellites with different spatial resolutions (Fig. 4), it is important to recognise that the spatial resolution of satellites is a significant factor in vegetation monitoring (Pax-Lenney and Woodcock, 1997; Ozdogan and Woodcock, 2006; Nelson et al., 2009; Duveiller and Defourny, 2010), particularly for capturing the flux spatial variability across heterogeneous ecosystems (Chu et al., 2021). As an illustration, the performance of our machine learning model demonstrated a significant disparity in estimating fluxes between forests and croplands (which might cover different crop species), with a disparity in R^2 that may reach up to 38% (Fig. 4). In addition, one should be cautious even when selecting satellites with the same spatial resolution for flux estimation. For example, we resampled Sentinel-2 images into the 30 m spatial resolution of Landsat satellites, but Sentinel-2 images exhibited a superior level of spatial detail in capturing features when compared to Landsat in Southwest England (Fig. 8). We used the surface reflectance products from both Sentinel-2 and Landsat satellites to extrapolate EC fluxes. However, it was observed that flux estimates derived from Landsat-7/-8 were more susceptible to the impact of cloud cover in Southwest England. It is plausible that this outcome may have arisen due to the performance of algorithms employed by Landsat-7/-8 for the removal of cloud cover in this region (Masek et al., 2020). As footprint matching between EC and satellite data is beneficial for flux estimation (Kong et al., 2022), it is advisable to pursue such an approach whenever feasible in future studies, despite the fact that using the inverse distance weighted (IDW) spatial interpolation can be a potentially effective alternative.

4.2. Challenges and the future in fine-resolution flux extrapolation

Some challenges remain in fine-resolution flux extrapolation despite the many successful cases in this study.

The first challenge is the disagreement in NEE estimates between extrapolation pathways. We called the NEE estimated directly from EC NEE measurements as direct NEE and NEE estimated by calculating the difference between GPP and Reco estimates as indirect NEE. According to the results in North Wyke (Fig. 8), the difference between direct NEE and indirect NEE was very large even for estimates using Sentinel-2. In this case, independent evidence is in need given the good cross-validation performance for direct NEE, GPP, and Reco (Fig. 6). It remains unknown which pathway in North Wyke was correct and how large its uncertainty will be when applied in other regions.

The second challenge is the justification of field-scale spatial variability. In North Wyke, field-scale flux estimates showed nearly equivalent total flux amounts but slightly different flux spatial variability when using different satellites (Fig. 8). As the fields in North Wyke were homogeneous, the difference in flux spatial variability was relatively small, but the potential uncertainty might be large in heterogeneous managed ecosystems. Sentinel-2 captured the texture information better than Landsat-7/-8. This might be attributed to the higher resolution of Sentinel-2 than Landsat-7/-8 here but more tests in other regions can be necessary before extrapolating EC fluxes.

The third challenge concerns the temporal period of management activities. Pastureland management activities can be divided into (1) continuous activities (e.g., livestock grazing) that occur during a long period of time and (2) discrete ones (e.g., spraying herbicides and cutting the grass) that typically occur within a day (Cardenas et al., 2022). The grazing events seemed not to effect the estimates (Fig. 7) as the flux time series estimated from the other two towers were in accordance with the EC measurements during the grazing periods (Fig. 7). This inferred that the machine learning algorithm can overcome the impacts of gradual but slow loss of aboveground biomass due to livestock grazing. However, large differences between measured EC and estimated flux can be expected around certain discrete management activities (e.g.

harvesting). We currently cannot use management information in the machine learning due to the complexity in quantifying discrete management applications – i.e., many application types, different management magnitude and time of application (Cardenas et al., 2022). In the future, field experiments will help us to better understand the impacts of management on ecosystem fluxes and improve extrapolation performance. In addition, we interpolated satellite imagery (e.g., from Landsat-7) into a daily frequency using the method in Joiner and Yoshida (2021). This is also likely to affect the performance of flux estimation, however, in the future, products with a higher temporal resolution will likely improve flux estimation, and our capacity to capture the impacts of discrete management activities.

5. Conclusion

Considering the requirement to accurately monitor ecosystem fluxes in managed ecosystems, this study (1) examined flux estimation at a global network of sites by extrapolating EC measurements with satellites of various spatial resolutions to support (2) assessments of the successful field-scale flux estimation in three pastures in the southwest of England at a 30 m spatial resolution. The key findings are:

- 1) Satellites with different (spatial) resolutions exhibited agreement in flux estimation.
- 2) Extrapolating EC fluxes on daily 30 m scales is a promising approach.
- 3) Sentinel-2 is recommended as it can capture the image texture better than Landsat-7/-8.
- 4) Strategies to quantify management activities in machine learning algorithms might be a way to better monitor the impact of management on ecosystem fluxes in the future.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the Rothamsted Research Institute Strategic Programme Soil to Nutrition, funded by the Biotechnology and Biological Sciences Research Council (BBS/E/C/00010320). The North Wyke Farm Platform is a UK National Capability, also supported by the Biotechnology and Biological Sciences Research Council (BBS/E/C/000J0100). The Eddy Covariance equipment deployed in this work was funded by CIEL (<https://www.cielivestock.co.uk/>) and the raw data is available on the Farm Platform Portal (<https://nwfp.rothamsted.ac.uk/>). SZ would like to acknowledge a Shell funded PhD studentship, TQ's contribution was supported by the NERC National Centre for Earth Observation (NE/R016518/1).

References

- Aubinet, M., Vesala, T., Papale, D., 2012. Eddy Covariance: A Practical Guide to Measurement and Data Analysis. Springer Science & Business Media.
- Badgley, G., Anderegg, L.D.L., Berry, J.A., Field, C.B., 2019. Terrestrial gross primary production: using NIRV to scale from site to globe. *Glob. Change Biol.* 25, 3731–3740.
- Badgley, G., Field, C.B., Berry, J.A., 2017. Canopy near-infrared reflectance and terrestrial photosynthesis. *Sci. Adv.* 3, e1602244.
- Baldocchi, D.D., 2020. How eddy covariance flux measurements have contributed to our understanding of global change biology. *Glob. Change Biol.* 26, 242–260.

- Bartier, P.M., Keller, C.P., 1996. Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). *Comput. Geosci.* 22, 795–799.
- Bojanowski, J.S., Musial, J.P., 2020. Dissecting effects of orbital drift of polar-orbiting satellites on accuracy and trends of climate data records of cloud fractional cover. *Atmos. Measur. Tech.* 13, 6771–6788.
- Cardenas, L., Olde, L., Loick, N., et al., 2022. CO₂ fluxes from three different temperate grazed pastures using Eddy covariance measurements. *Sci. Total Environ.* 831, 154819.
- Chander, G., Markham, B.L., Helder, D.L., 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sens. Environ.* 113, 893–903.
- Chapin III, F.S., Matson, P.A., Vitousek, P., 2011. *Principles of Terrestrial Ecosystem Ecology*. Springer Science & Business Media.
- Chen T., He T., Benesty M., et al. (2015) Xgboost: extreme gradient boosting. R package version 04-2 1:1–4.
- Chu, H., Baldocchi, D.D., John, R., et al., 2017. Fluxes all of the time? A primer on the temporal representativeness of FLUXNET. *J. Geophys. Res. Biogeosci.* 122, 289–307.
- Chu, H., Luo, X., Ouyang, Z., et al., 2021. Representativeness of Eddy-covariance flux footprints for areas surrounding AmeriFlux sites. *Agric. Forest Meteorol.* 301, 108350.
- Claverie M., Vermote E., others (2014) NOAA climate data record (CDR) of leaf area index (LAI) and fraction of absorbed photosynthetically active radiation (FAPAR) version 4. NOAA National Centers for Environmental Information 10:V5M043BX.
- Drusch, M., Del Bello, U., Carlier, S., et al., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* 120, 25–36.
- Duveiller, G., Defourny, P., 2010. A conceptual framework to define the spatial resolution requirements for agricultural monitoring using remote sensing. *Remote Sens. Environ.* 114, 2637–2650.
- Dwyer, J.L., Roy, D.P., Sauer, B., et al., 2018. Analysis ready data: enabling analysis of the Landsat archive. *Remote Sens.* 10, 1363.
- FAO (2022) Land use in agriculture by the numbers. In: *Food and Agriculture Organization of the United Nations*. <http://www.fao.org/sustainability/news/detail/en/c/1274219/>. Accessed 30 Aug 2022.
- Fu, D., Chen, B., Zhang, H., et al., 2014. Estimating landscape net ecosystem exchange at high spatial-temporal resolution based on Landsat data, an improved upscaling model framework, and eddy covariance flux measurements. *Remote Sens. Environ.* 141, 90–104.
- Gorelick, N., Hancher, M., Dixon, M., et al., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27.
- Griscom, B.W., Adams, J., Ellis, P.W., et al., 2017. Natural climate solutions. *Proc. Natl. Acad. Sci.* 114, 11645–11650.
- Hersbach, H., Bell, B., Berrisford, P., et al., 2020. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* 146, 1999–2049.
- Ichii, K., Ueyama, M., Kondo, M., et al., 2017. New data-driven estimation of terrestrial CO₂ fluxes in Asia using a standardized database of Eddy covariance measurements, remote sensing data, and support vector regression. *J. Geophys. Res. Biogeosci.* 122, 767–795.
- Jiang, Z., Huete, A.R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* 112, 3833–3845.
- Joiner, J., Yoshida Y. (2021) Global MODIS and FLUXNET-derived daily gross primary production, V2. ORNL DAAC, Oak Ridge, Tennessee, USA.
- Joiner, J., Yoshida, Y., 2020. Satellite-based reflectances capture large fraction of variability in global gross primary production (GPP) at weekly time scales. *Agric. Forest Meteorol.* 291, 108092.
- Jung, M., Koirala, S., Weber, U., et al., 2019. The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Sci. Data* 6, 1–14.
- Jung, M., Reichstein, M., Schwalm, C.R., et al., 2017. Compensatory water effects link yearly global land CO₂ sink changes to temperature. *Nature* 541, 516–520.
- Jung, M., Schwalm, C., Migliavacca, M., et al., 2020. Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. *Biogeosciences* 17, 1343–1365.
- Kong, J., Ryu, Y., Liu, J., et al., 2022. Matching high resolution satellite data and flux tower footprints improves their agreement in photosynthesis estimates. *Agric. Forest Meteorol.* 316, 108878.
- Kotir, J.H., 2011. Climate change and variability in Sub-Saharan Africa: a review of current and future trends and impacts on agriculture and food security. *Environ. Dev. Sustain.* 13, 587–605.
- Lasslop, G., Reichstein, M., Papale, D., et al., 2010. Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. *Glob. Change Biol.* 16, 187–208.
- Lee, X., Massman, W., Law, B., 2004. *Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis*. Springer Science & Business Media.
- LI-COR Inc. B (2017) EddyPro software instruction manual. LI-COR Inc, Lincoln, Nebraska, USA.
- Lisle, R.J., 2006. Google Earth: a new geological resource. *Geol. Today* 22, 29–32.
- Malhi, Y., Franklin, J., Seddon, N., et al., 2020. Climate change and ecosystems: threats, opportunities and solutions. *Philos. Trans. R. Soc. B* 375, 20190104.
- Marchetti, F., 2021. The extension of Rippla's algorithm beyond LOOCV. *Appl. Math. Lett.* 120, 107262.
- Masek, J.G., Wulder, M.A., Markham, B., et al., 2020. Landsat 9: Empowering open science and applications through continuity. *Remote Sens. Environ.* 248, 111968.
- Mauna Loa (2022) Carbon dioxide now more than 50% higher than pre-industrial levels. <http://www.noaa.gov/news-release/carbon-dioxide-now-more-than-50-higher-than-pre-industrial-levels>. Accessed 30 Aug 2022.
- Moffat, A.M., Papale, D., Reichstein, M., et al., 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agric. Forest Meteorol.* 147, 209–232.
- Nelson, M., McRoberts, R., Holden, G., Bauer, M., 2009. Effects of satellite image spatial aggregation and resolution on estimates of forest land area. *Int. J. Remote Sens.* 30, 1913–1940.
- Novick, K.A., Metzger, S., Anderegg, W.R., et al., 2022. Informing nature-based climate Solutions for the United States with the best-available science. *Glob. Change Biol.* 28, 3778–3794.
- Orr, R., Murray, P., Eyles, C., et al., 2016. The North Wyke farm platform: effect of temperate grassland farming systems on soil moisture contents, runoff and associated water quality dynamics. *Eur. J. Soil Sci.* 67, 374–385.
- Ozdogan, M., Woodcock, C.E., 2006. Resolution dependent errors in remote sensing of cultivated areas. *Remote Sens. Environ.* 103, 203–217.
- Pastorello, G., Trotta, C., Canfora, E., et al., 2020. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Sci. Data* 7, 1–27.
- Pax-Lenney, M., Woodcock, C.E., 1997. The effect of spatial resolution on the ability to monitor the status of agricultural lands. *Remote Sens. Environ.* 61, 210–220.
- Reichstein, M., Falge, E., Baldocchi, D., et al., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Glob. Change Biol.* 11, 1424–1439.
- Rocha, A.V., Shaver, G.R., 2009. Advantages of a two band EVI calculated from solar and photosynthetically active radiation fluxes. *Agric. Forest Meteorol.* 149, 1560–1563.
- Running, S., Mu, Q., Zhao, M., 2015. MOD17A2H MODIS/terra gross primary productivity 8-day L4 global 500m SIN grid V006. NASA EOSDIS Land Processes DAAC.
- Salomón, R.L., Peters, R.L., Zweifel, R., et al., 2022. The 2018 European heatwave led to stem dehydration but not to consistent growth reductions in forests. *Nat. Commun.* 13, 1–11.
- Schaaf, C., Wang, Z., 2015. MCD43A4 MODIS/Terra+ Aqua BRDF/Albedo Nadir BRDF adjusted RefDaily L3 Global 500 m V006. NASA EOSDIS land processes DAAC.
- Schimel, D., Pavlick, R., Fisher, J.B., et al., 2015. Observing terrestrial ecosystems and the carbon cycle from space. *Glob. Change Biol.* 21, 1762–1776.
- Smith, P., Beaumont, L., Bernacchi, C.J., et al., 2021. Essential outcomes for COP26. *Glob. Change Biol.* 28, 1–3.
- Szantoi, Z., Strobl, P., 2019. Copernicus Sentinel-2 calibration and validation. *Eur. J. Remote Sens.* 52, 253–255.
- Theobald, D.M., Kennedy, C., Chen, B., et al., 2020. Earth transformed: detailed mapping of global human modification from 1990 to 2017. *Earth Syst. Sci. Data* 12, 1953–1972.
- Tramontana, G., Ichii, K., Camps-Valls, G., et al., 2015. Uncertainty analysis of gross primary production upscaling using random forests, remote sensing and Eddy covariance data. *Remote Sens. Environ.* 168, 360–373.
- Verhegghen, A., Bontemps, S., Defourny, P., 2014. A global NDVI and EVI reference data set for land-surface phenology using 13 years of daily SPOT-VEGETATION observations. *Int. J. Remote Sens.* 35, 2440–2471.
- Vermote, E., Justice, C., Claverie, M., Franch, B., 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sens. Environ.* 185, 46–56.
- Wang, L., Zhu, H., Lin, A., et al., 2017. Evaluation of the latest MODIS GPP products across multiple biomes using global eddy covariance flux data. *Remote Sens.* 9, 418.
- Xiao, J., Chen, J., Davis, K.J., Reichstein, M., 2012. Advances in upscaling of eddy covariance measurements of carbon and water fluxes. *J. Geophys. Res. Biogeosci.* 117 <https://doi.org/10.1029/2011JG001889>.
- Zalles, V., Hansen, M.C., Potapov, P.V., et al., 2021. Rapid expansion of human impact on natural land in South America since 1985. *Sci. Adv.* 7, eabg1620.
- Zarocostas, J., 2022. The UN reports global asymmetries in population growth. *Lancet* 400, 148.
- Zhu, S., Clement, R., McCalmont, J., et al., 2022. Stable gap-filling for longer eddy covariance data gaps: a globally validated machine-learning approach for carbon dioxide, water, and energy fluxes. *Agric. Forest Meteorol.* 314, 108777.
- Zhu, S., McCalmont, J., Cardenas, L.M., et al., 2023. Gap-filling carbon dioxide, water, energy, and methane fluxes in challenging ecosystems: comparing between methods, drivers, and gap-lengths. *Agric. Forest Meteorol.* 332, 109365.