A Deep-Learning Hybrid-Predictive-Modeling Approach for Estimating Evapotranspiration and Ecosystem Respiration

Jiancong Chen[1], Baptiste Dafflon[2], Anh Phuong Tran[2, 3], Nicola Falco[2], and Susan S. Hubbard[2]

[1]Department of Civil and Environmental Engineering, University of California, Berkeley, CA, USA, [2]Earth and Environmental Sciences Area, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [3]Depatment of Water Resources Engineering and Technology, Water Resources Institute, 8, Phao Dai Lang, Dong Da, Hanoi, Vietnam

Abstract: Gradual changes in meteorological forcings (such as temperature and precipitation) are reshaping vulnerable ecosystems, leading to uncertain effects on ecosystem dynamics, including water and carbon fluxes. Estimating evapotranspiration (ET) and ecosystem respiration (R_{ECO}) is essential for analyzing the effect of climate change on ecosystem behavior. To obtain a better understanding of these processes, we need to improve our estimation of water and carbon fluxes over space and time, which is difficult within ecosystems that often have only sparse data. In this study, we developed a hybrid predictive modeling approach (HPM) that integrates eddy covariance measurements, physically-based model simulation results, meteorological forcings, and remote sensing datasets to estimate evapotranspiration (ET) and ecosystem respiration (R_{ECO}) in high space-time resolution. HPM relies on a deep learning algorithm-long short term memory (LSTM)-as well as direct measurements or outputs from physically-based models. We tested and validated HPM estimation results at sites within various sites. We particularly focus on testing HPM in mountainous regions, given their importance for water resources, their vulnerability to climate change, and the recognized difficulties in estimating ET and R_{ECO} in such regions. We benchmarked daily scale estimates of ET and R_{ECO} obtained from the HPM method against measurements made at FLUXNET stations and outputs from the Community Land Model (CLM) at Rocky Mountain SNOTEL stations. At the mountainous East River Watershed site in the Upper Colorado River Basin, we explored how ET and R_{ECO} dynamics estimated from the new HPM approach vary with different vegetation and meteorological forcings. The results of this study indicate that HPM is capable of identifying complicated interactions among meteorological forcings, ET, and R_{ECO} variables, as well as providing reliable estimation of ET and R_{ECO} across relevant spatiotemporal scales, even in challenging mountainous systems. With HPM estimation of ET and R_{ECO} at the East River Watershed, we identified that HPM ET models are sensitive to temperature and radiation inputs whereas NDVI, temperature and radiation all have crucial influences over R_{ECO} dynamics. In general, our study demonstrated that the HPM approach can circumvent the typical lack of spatiotemporally dense data needed to estimate ET and R_{ECO} over space and time, as well as the parametric and structural uncertainty inherent in mechanistic models. While the current limitations of the HPM approach are driven by the temporal and spatial resolution of available datasets (such as meteorological forcing and NDVI data), ongoing advances are expected to further improve accuracy and resolution of ET and R_{ECO} estimation using HPM.

1. Introduction:

Evapotranspiration (ET) and ecosystem respiration (R_{ECO}) are key components of ecosystem water and carbon cycles. ET is an important link between the water and energy cycles: dynamic changes in ET can affect precipitation, soil moisture, and surface temperature, leading to uncertain feedbacks in the environment (Jung et al., 2010; Seneviratne et al., 2006; Teuling et al., 2013). Thus, quantifying ET is particularly essential for improving our understanding of water and energy interactions and watershed response to abrupt and gradual changes in climate, which is critical for water resources management, agriculture, and other societal benefits (Anderson et al., 2012; Jung et al., 2010; Rungee et al., 2019; Viviroli et al., 2007; Viviroli and Weingartner, 2008). R_{ECO}, which represents the sum of autotrophic respiration and respiration by heterotrophic microorganisms in a specific ecosystem, plays a vital role in the response of terrestrial ecosystem to global change (Jung et al., 2017; Reichstein et al., 2005; Xu et al., 2004). As long term exchanges in R_{ECO} have pivotal influences over the climate system (Cox et al., 2000; Gao et al., 2017; IPCC, 2019; Suleau et al., 2011), approaches are needed to estimate and monitor R_{ECO} over relevant spatiotemporal scales. As described below, there are many different strategies for measuring and estimating ET and R_{ECO}, each of which has advantages and limitations. The motivation for this study is the recognition that current methods cannot provide ET and R_{ECO} at space and time scales (e.g., daily) needed to improve prediction of changing terrestrial system behavior, particularly in challenging mountainous watersheds.

Several ground-based approaches have been used to provide *in situ* estimates or measurements of ET and R_{ECO}. Ground based flux chambers capture and measure trace gases emitted from the land surface, which can be used to estimate ET and R_{ECO} (Livingston and Hutchinson, 1995; Pumpanen et al., 2004). However, the microclimate of the environment is affected by the chamber, and the laborious acquisition process and small chamber size typically lead to information with coarse spatiotemporal resolution (Baldocchi, 2014). The eddy covariance method uses a tower with installed instruments to autonomously measure fluxes of trace gases between ecosystem and atmosphere (Baldocchi, 2014; Wilson et al., 2001). The covariance between the vertical velocity and mixing ratios of the target scalar is computed to obtain the fluxes of carbon, water vapor, and other trace gases emitted from the land surface. ET is then calculated from the latent heat flux, and R_{ECO} is calculated from the net carbon fluxes using night-time or daytime partitioning approaches (van Gorsel et al., 2009; Lasslop et al., 2010; Reichstein et al., 2005). The spatial footprint of obtained fluxes is on the order of hundreds of meters, and the temporal resolution of the measurements range from hours to decades (Wilson et al., 2001). Such *in situ* measurements of fluxes have been integrated into the global AmeriFlux (http://ameriflux.lbl.gov/) and FLUXNET (https://FLUXNET.fluxdata.org/) networks, where such data have greatly benefited process investigations and model development undertaken by a wide scientific community. However, given the cost, effort, and power required to install and maintain a flux tower, eddy covariance towers are typically sparse relative to the scale of study sites used to address ecosystem questions. Additionally, the location of a flux tower within a watershed greatly influences measurement representativeness. For example, for logistical reasons, eddy covariance towers are usually installed at valley bottoms of mountainous watersheds (Strachan et al., 2016). However, microclimate caused by complex mountainous terrains (e.g., slope, aspect and elevation) can have different radiation inputs and moisture dynamics compared to flat areas where flux towers are mostly installed. Flux measurements from eddy covariance towers provide a representation of major driver and controls on ET and R_{ECO} in an ecoregion while meteorological forcing variability needs to be accounted to possibly represent various aspects introduced by complex terrain. Thus, though measurements from a single flux tower may not capture heterogeneity in ET and R_{ECO} due to complex terrain, they support the development of statistical or physical-based models integrated with other types of data to provide ET and R_{ECO} estimation in high resolution over space and time.

Physically-based models, which numerically represent land-surface energy and water balance, have also been used to estimate ET and R_{ECO} (Tran et al., 2019; Williams et al., 2009). These physically-based models solve physical equations to simulate the exchanges of energy, heat, water and carbon across atmosphere-canopy-soil compartments. Examples include the Community Land Model (CLM, Oleson et al., 2013). Performance of these models depend on the accuracy of inputs and parameters, such as soil type and leaf area index, which can be difficult to obtain at sufficiently high spatiotemporal resolution. The lack of measurements to infer parameters needed for models often leads to large discrepancies between model-based and flux-tower-based ET and R_{ECO} estimates. Conceptual model uncertainty inherent in mechanistic models can also lead to ET and R_{ECO} estimation uncertainty and errors. For example, Keenan et al. (2019) suggested that current terrestrial carbon cycle models neglect inhibition of leaf respiration that occurs during daytime, which can result in a bias of up to 25%. Chang et al. (2018) used virtual experiments with 3-D terrestrial integrated modeling system to investigate why a lower ratio of transpiration to ET is always produced by large scale land surface models. Their study suggested heterogeneous fluxes caused by uneven hydraulic distribution due to complex terrain are not always considered in process-based models. These conceptual uncertainties, in addition to data sparseness and data uncertainty, further limit the applicability of physically-based models to estimate ET and R_{ECO} at high spatiotemporal scales. Semi-analytical formulations based on combinations of meteorological and empirical parameters provide a reference condition for the water and energy balance. Examples used to estimate potential ET include the Budyko framework and its extensions (Budyko, 1961; Greve et al., 2015; Zhang et al., 2008); the Penman-Monteith's equation (Allen et al., 1998), and the Priestley-Taylor equation (Priestley and Taylor, 1972). Actual ET can then be approximated by multiplying a coefficient associated with water deficit (De Bruin, 1983; Williams & Albertson, 2004). However, even with these empirical formulations many attributes are still difficult to obtain globally at high temporal scales, such as water-vapor deficit, leaf area index, and aerodynamic conductance of different plants.

Remote sensing products, such as Landsat imagery (Irons et al., 2012), Sentinel-2 (Main-Knorn et al., 2017) and the moderate-resolution imaging spectroradiometer (MODIS, NASA. 2008), have also been integrated to estimate ET and R_{ECO} with empirical, statistical, or semi-physical relations (Abatzoglou et al., 2014; Daggers et al., 2018; Mohanty et al., 2017; Paca et al., 2019). Due to the high spatial coverage of remote sensing products, global-scale estimates of ET and R_{ECO} have become feasible. For example, Ryu et al. (2011) proposed the Breathing Earth System Simulator approach, which integrates mechanistic models and MODIS data to quantify ET and GPP with a spatial resolution of 1-5 km and a temporal resolution of 8 days. Ai et al. (2018) extracted enhanced vegetation index, fraction of absorbed photosynthetically active radiation, and leaf area index from the MODIS dataset—and used the rate-temperature curve and strong correlations between terrestrial carbon exchange and temperature to estimate R_{ECO} at 1 km spatial resolution and 8-day temporal resolution. Ma et al. (2018) developed a data fusion scheme that fused Landsat-like-scale datasets and MODIS data to estimate ET and irrigation water efficiency at a spatial scale of ~100 meters. However, even though remote sensing data cover large areas of the earth surface, they typically do not provide information over both high spatial and temporal resolution, and are also subject to cloudy conditions. For example, Landsat has average return periods of 16 days with a spatial resolution of 30 m (visible and near-infrared), whereas MODIS has 1-2 days temporal resolution with a 250 m or 1 km spatial resolution depending on the sensors. These resolutions are typically too coarse to enable exploration of how aspects such as plant phenology, snowmelt, and rainfall impact integrated ecosystem water and energy dynamics.

Combining machine-learning models with remote sensing products and meteorological inputs offers another option for large-scale estimation of ET and R_{ECO}. Remotely sensed data can be good proxies for plant productivity and can be easily implemented into machine-learning models for ET and R_{ECO} estimation, such as for an enhanced vegetation index, land surface water index and normalized differences vegetation index (NDVI) (Gao et al., 2015; Jägermeyr et al., 2014; Migliavacca et al., 2015). Li and Xiao (2019) developed a data-driven model to estimate gross primary production at a spatial and temporal resolution of 0.05° and 8 days, respectively, using MODIS and meterological reanalysis data. Berryman et al. (2018) demonstrated the value of a Random Forest model to predict growing season soil respiration from subalpine forests in the Southern Rocky Mountains ecoregion. Jung et al. (2009) developed a model tree ensemble approach to upscale FLUXNET data, where they successfully estimated ET and GPP. Other methods have used support vector machines, artificial neural networks, random forest, and piecewise regression (Bodesheim et al., 2018; Metzger et al., 2013; Xiao et al., 2014; Xu et al., 2018). These models were trained with ground-measured flux observations and other variables, and then applied to estimate ET over continental or global scales with remote sensing and meteorological inputs. Some of the most important inputs include the enhanced vegetation index, aridity index, temperature, and precipitation. However, the spatiotemporal resolution of these approaches is constrained by the resolution of remote sensing products and meteorological inputs. Additionally, parameters such as leaf area index, cloudiness, and the vegetation types required by those models may not be available at the required resolution, accuracy or location. For example, in systems that have significant elevation gradients, errors may result when valley-based FLUXNET data are used for training and then applied to hillslope or ridge ET and R_{ECO} estimation.

Development of hybrid models that link direct measurements and/or interpretable mechanistic models with data-driven methods can benefit ET and R_{ECO} estimation (Reichstein et al., 2019). While remote sensing data that cover large regions provide promise for informing models, quantitative interpretation of these data needed for input into mechanistic models is still challenging (Reichstein et al., 2019). Physically-based models can provide estimates of ET and R_{ECO}, but the estimate error can be high, owing to parametric, structural, and conceptual uncertainties as described above. Hybrid data-driven frameworks are potentially advantageous because they enable the integration of remote sensing datasets, meteorological forcings, and mechanistic model outputs of ET and R_{ECO} into one model. Machine-learning approaches are then applied to extract the spatiotemporal patterns for ET and R_{ECO} prediction. Hybrid models can utilize the high spatial coverage of remote sensing data (e.g., 30 m of Landsat) and high temporal resolution of direct measurement from flux towers or simulation results from mechanistic models (e.g., daily or hourly scales), thus providing alternative approaches for next-stage, more accurate estimation of ET and R_{ECO} at greater spatial and finer temporal scales—and enhancing our process understanding of water and carbon cycling under climate change.

In this study, we developed a hybrid predictive modeling approach (HPM) to better estimate daily ET and R_{ECO} with easily acquired meteorological data (i.e., air temperature, precipitation and radiation) and remote sensing products (i.e., NDVI). HPM is hybrid as it can use deep learning models to integrate direct measurements from flux towers and physically-based model results (e.g., CLM) with meteorological and remote sensing inputs to capture the complex physical interactions within the watershed ecosystem. After development, we validated HPM performance with the FLUXNET dataset and benchmarked the CLM model at select sites. We then used the HPM for ET and R_{ECO} estimation at the mountainous East River Watershed in Colorado, USA and investigated how ET and R_{ECO} dynamics varies within the East River Watershed.

The remainder of this paper is organized as follows. Section 2 mainly describes the sites considered in this study and how data were acquired and processed. Section 3 presents the methodology of the HPM approach, followed by the results of various use cases presented in Section 4. Discussion and conclusion are provided in Sections 5 and 6, respectively.

2. Site Information, Data Acquisition and Processing

The HPM method was tested using data from a range of different ecosystem types to explore its performance under different conditions. However, we place a particular focus on mountainous sites, given their regional and global importance yet challenges associated with ET and R_{ECO} as described above.

2.1 FLUXNET Stations and Ecoregions

Nine FLUXNET stations were selected for this study (Table 1 and Figure 1), which cover a wide range of climate and elevations. These stations have elevations from 129 m (US-Var) to 3050 m (US-NR1), mean annual air temperature from 0.34℃ (CA-Oas) to 17.92℃ (US-SRM), and mean annual precipitation from 320 mm (US-Whs) to 800 mm (US-NR1). These FLUXNET stations also cover a wide range of vegetation types (i.e., evergreen forest, deciduous forest, and shrublands). As indicated by Hargrove et al. (2003), FLUXNET stations provide a good representation of different ecoregions, which are areas that display recurring patterns of similar combinations of soil, vegetation and landform characteristics (Omernik, 2004). Omernik & Griffith. (2014) delineated the boundaries of ecoregions through pattern analysis that consider the spatial correlation of both physical and biological factors (i.e., soils, physiography, vegetation, land use, geology and hydrology) in a hierarchical level. FLUXNET stations considered in this study mainly locate in 4 unique ecoregions (Table 1). As is described below, we developed local-scale (i.e., point scale) HPM that are representative for different ecoregions using data provided at these FLUXNET stations to estimate ET and R_{ECO}, and validated the HPM estimates with measurements from stations within the same ecoregion.

2.2 SNOTEL Stations

For reasons described below, we performed a deeper exploration of HPM performance within one of the mountainous watershed sites (the East River Watershed of the Upper Colorado River Basin), which is located in the “western cordillera” ecoregion. At this site, we utilized meteorological forcings data from three snow telemetry (SNOTEL) stations. These sites include the Butte (ER-BT, id: 380), Porphyry Creek (ER-PK, id: 701) and Schofield Pass (ER-SP, id: 737) sites. A CLM model was developed at these SNOTEL stations that provides physically-model-based ET estimation (Tran et al., 2019). Table 1 summarizes the SNOTEL stations used in this study and the corresponding climate characteristics. Figure 1 shows the geographical locations of FLUXNET and SNOTEL stations selected in this study.

Table 1. Summary of FLUXNET stations and SNOTEL stations information. * denotes SNOTEL stations and all others are FLUXNET stations. Dfc, Bsk, Csa represent subarctic or boreal climates, semi-arid climate, Mediterranean hot summer climates, respectively. ENF, DBF, WSA, GRA, and OSH represent evergreen needleleaf forest, deciduous broadleaf forests, woody savannas, grasslands, open shrubland, respectively. FLUXNET data were obtained from the FLUXNET2015 database.

Site ID	Site Name	Latitude, Longitude	Elevation (m)	Mean Annual temperature (℃)	Mean Annual Precipitation (m)	Climate Koeppen	Vegetation IGBP	Ecoregions (Level II)	Period of Record
US-NR1	Niwot Ridge	(40.0329, -105.5464)	3050	1.5	800	Dfc	ENF	Western Cordillera	2000-2014
CA-Oas	Saskatchewan-Aspen	(53.6289, -106.1978)	530	0.34	428.53	Dfc	DBF	Boreal Plain	1997-2010
CA-Obs	Saskatchewan-Black Spruce	(53.9872, -105.1178)	628.94	0.79	405.6	Dfc	ENF	Boreal Plain	1999-2010
US-SRM	Santa Rita Mesquite	(31.8214, -110.8661)	1120	17.92	380	Bsk	WSA	Western Sierra Madre Piedmont	2005-2015
US-Ton	Tonzi Ranch	(38.4316, -120.9660)	177	15.8	559	Csa	WSA	Mediterranean California	2002-2015
US-Var	Vaira Ranch-lone	(38.4133, -120.9507)	129	15.8	559	Csa	GRA	Mediterranean California	2002-2015
US-Whs	Walnut Gulch Lucky Hills Shrub	(31.7438, -110.0522)	1370	17.6	320	Bsk	OSH	Western Sierra Madre Piedmont	2008-2015
US-Wkg	Walnut Gulch Kendall Grasslands	(31.7365, -109.9419)	1531	15.64	407	Bsk	GRA	Western Sierra Madre Piedmont	2005-2015
US-Me2	Metolius mature ponderosa pine	(44.4523, -121.5574)	1253	6.28	523	Csb	ENF	Western Cordillera	2012-2015
ER-BT*	East River-Butte	(38.894, -106.945)	3096	2.38	821	Dfc	N/A	Western Cordillera	1995-2017
ER-SP*	East River-Schofield Pass	(39.02, -107.05)	3261	2.46	1064	Dfc	N/A	Western Cordillera	1995-2017
ER-PK*	East River-Porphyry Creek	(38.49, -106.34)	3280	1.97	574	Dfc	N/A	Western Cordillera	1995-2017

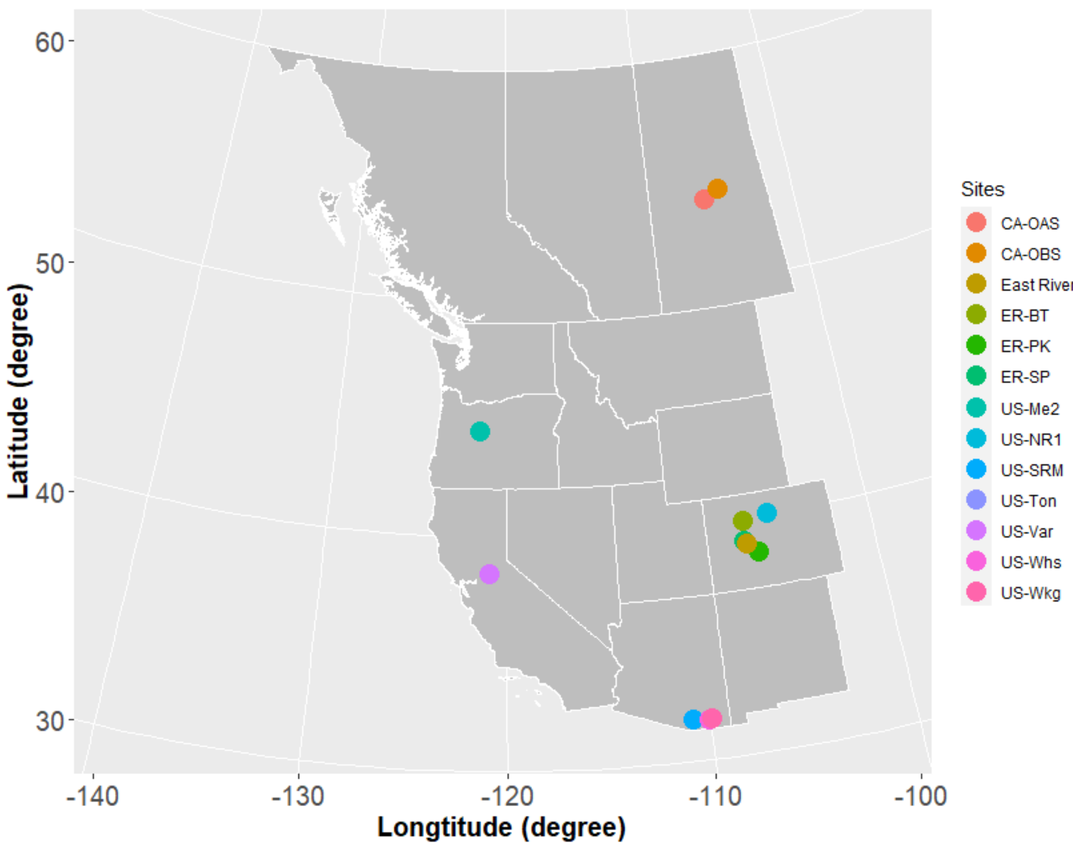

Figure 1. Location of sites considered in this study. Note: US-Ton and US-Var; US-Whs and US-Wkg are at the same locations. East River Watershed is located next to ER-BT. The white lines delineate Western US states and Canadian provinces.

2.3 East River Watershed and Previous Analyses

Data from the East River Watershed were used to explore how ET and R_{ECO} dynamics estimated from the developed HPM vary with different vegetation and meteorological forcings. The East River Watershed is located northeast of the town of Crested Butte, Colorado. This watershed has an average elevation of 3266 m, with significant gradients in topography, hydrology, geomorphology, vegetation, and weather. The watershed has a mean annual temperature around 0℃, with an average of 1200 mm yr^{-1} total precipitation (Hubbard et al., 2018). Consisting of montane, subalpine, and alpine life zones, each with distinctive vegetation biodiversity, the East River Watershed is a testbed for the US Department of Energy Watershed Function Scientific Focus Area Project, led by the Lawrence Berkeley National Laboratory (Hubbard et al., 2018). The project has acquired a range of datasets, including hydrological, biogeochemical, remote sensing, and geophysical datasets.

Recently completed studies at the East River Watershed were used in this study to inform HPM and to assess the results. For example, physically-model-based estimations of ET at this site (Tran et al., 2019) were used herein for HPM development and validation. Falco et al. (2019) used machine-learning-based remote sensing methods to characterize the spatial distribution of vegetation types, slopes, and aspects within a hillslope at the East River

Watershed, which were used with obtained HPM estimates to explore how vegetation heterogeneity influences ET and R_{ECO} dynamics. To perform this assessment, we computed the spatial distribution of vegetation types at watershed scale based on Falco et al. (2019). We evaluated manually and selected 16 locations within the East River Watershed having different vegetation types and slope aspects. These 16 locations were chosen to be at the center of vegetation patched and covered by one vegetation type. A summary of the locations is presented in Table 2; the spatial distribution of the locations is shown in Figure 2.

Table 2: Location and vegetation types of East River Watershed sampling points (Figure 2)

Easting (m)	Northing (m)	Vegetation Type	Aspect	Elevation (m)
327085	4309878	Deciduous Forest	South	2983
326288	4312504	Deciduous Forest	South	3177
330012	4313132	Deciduous Forest	North	3108
326854	4313192	Deciduous Forest	South	3098
328246	4312832	Meadow	South	3095
327010	4315059	Meadow	South	2790
328738	4306139	Meadow	North	2890
334270	4309465	Meadow	North	2929
333406.5	4308340	Riparian Shrubland	South	2760
327846	4312497	Riparian Shrubland	South	2723
334641	4305632	Riparian Shrubland	North	2740
330760	4310097	Riparian Shrubland	South	2855
329573	4314569	Evergreen Forest	South	3026
333106	4307313	Evergreen Forest	North	3102
325056	4310456	Evergreen Forest	South	2961
335141	4309614	Evergreen Forest	North	3131

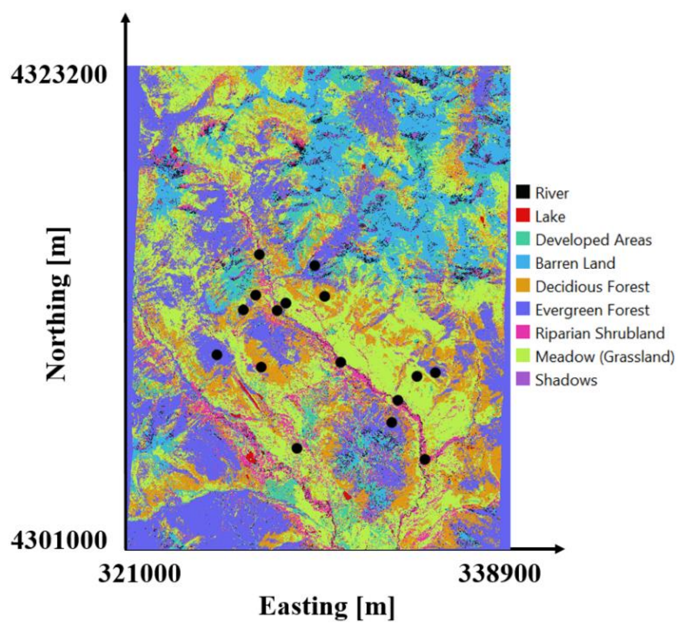

Figure 2: Vegetation classification of the East River, CO Watershed from Falco et al. (2019). East River sites selected in this study are denoted by black circles.

2.4 Data Collection and Processing

To enhance transferability of the developed HPM strategy to less intensively characterized watersheds, we selected only “easy to measure” or “widely available” attributes, such as precipitation, temperature, radiation and

NDVI, as inputs to the HTM model. The data sources used for these inputs include FLUXNET data (https://fluxnet.fluxdata.org/), SNOTEL data (https://www.wcc.nrcs.usda.gov/snow/) and developed CLM model (Tran et al., 2019) at SNOTEL stations, DAYMET meteorological inputs (Thornton et al., 2017) and remote sensing data from Landsat imageries (Irons et al., 2012).

A variety of measured data and model outputs were used to train and validate HPM. We obtained daily meteorological data, including air temperature, precipitation, radiation, ET, and R_{ECO} data, from the FLUXNET database at the selected FLUXNET sites. The pipeline of data processing for FLUXNET dataset is provided at https://FLUXNET.fluxdata.org/. We identified some data gaps and erroneous data (especially during winter seasons) for the ET estimates at US-NR1, which were cleaned following the procedures presented in Rungee et al. (2019). The meteorological data were used as inputs for HPM development, and ET and R_{ECO} data from these sites were used for HPM validation. At the three selected SNOTEL stations, we obtained air temperature, precipitation, and snow-water-equivalent data from the SNOTEL database. Air temperature data at these three SNOTEL stations were processed following Oyler et al. (2015), given potential systematic artifacts. Snow-water-equivalent data are not easily acquired, and thus were not considered as inputs for HPM. However, a categorical variable was constructed to assimilate information regarding snow (Section 3.2.1). CLM models were generated following Tran et al. (2019) for the SNOTEL stations and US-NR1 to assess the spatiotemporal variability of ET at the East River Watershed and for training and validating HPM (Section 4.3). The DAYMET dataset (Thornton et al., 2017) provided gridded daily weather-forcings-attribute estimates at a 1 km spatial resolution. We obtained the incident radiation data from DAYMET at the SNOTEL stations as inputs for HPM. For the East River Watershed sites, meteorological forcings data, including air temperature, precipitation and radiation, were also obtained from DAYMET. The low spatial resolution of DAYMET data introduces uncertainty in HPM estimation of ET and R_{ECO}, which will be discussed in the following sections. We calculated the NDVI time series from the red band (RED) and near-infrared band (NIR) from Landsat 5, 7, and 8 images at all selected FLUXNET sites, SNOTEL stations, and East River Watershed sites at a spatial scale of 30 m.

Since cloud conditions can severely decrease data quality, we used the cloud-scoring algorithm provided in the Google Earth Engine to mask clouds in all retrieved data, only selecting the ones that had a simple cloud score below 20 to ensure data quality. Given the different calibration sensors used in Landsat 5, 7, and 8, we also followed the processes described in Homer et al. (2015) and Vogelmann et al. (2001) to keep NDVI computations consistent over time. Landsat satellites have a return period of 16 days, and thus we performed a reconstruction of NDVI time series to obtain daily scale time data (Section 3.2.2).

3. Hybrid Predictive Modeling Framework

In this section, we illustrate the steps for building an HPM model for ET and R_{ECO} estimation over time and space. Figure 3 presents the general framework of HPM, which includes modules for data preprocessing, model development, model validation, and predictive modeling.

3.1 Model Framework

HPM establishes relationships among meteorological forcings attributes, NDVI, ET, and R_{ECO} (Figure 3). Both input data (e.g., meteorological forcings) and output data (ET and R_{ECO}) used for training and validation are preprocessed for gap filling, smoothing, and data updating. HPM "learns" the complex space-time relationship among meteorological forcings, NDVI, ET, and R_{ECO} using a deep-learning-based module (deeply connected neural networks and a long short-term memory recurrent neural network). HPM then can be used for ET and R_{ECO} estimation at sparsely monitored watersheds. Individual HPM models can be trained in two different ways using ET and R_{ECO} information: with data obtained from flux towers ("data-driven HPM") or with outputs from 1-D physically-based models ("mechanistic HPM"). In both cases, the models obtained with local data are then used to estimate ET and R_{ECO} at other sites in the same ecoregion (see Section 2.1). For ecoregions not represented by FLUXNET sites, it is necessary to develop mechanistic HPM that enables ET and R_{ECO} estimation over space and time.

HPM has several additional modules, including model development, model validation, and model prediction modules. In the HPM model development module, deep-learning algorithms are trained with input features and response data until a pre-defined "stopping criteria" (e.g., root mean squared error, RMSE) is met, indicating subsequent training would lead to minimal improvement. In the validation module, estimation outputs from the "trained HPM models" are compared with other ET and R_{ECO} data obtained from other independent sites or mechanistic models within the same ecoregion. Statistical measures, including adjusted R^2 and mean absolute error (MAE), are computed to evaluate the performance of HPM models. In the predictive model module, meteorological forcings data and remote sensing data are processed at target sites of interest, and the validated HPM model is used to estimate ET and R_{ECO} at these sites. ET and R_{ECO} outputs estimated from HPM at sparsely monitored watersheds then provide alternative datasets for process understanding within the target watersheds.

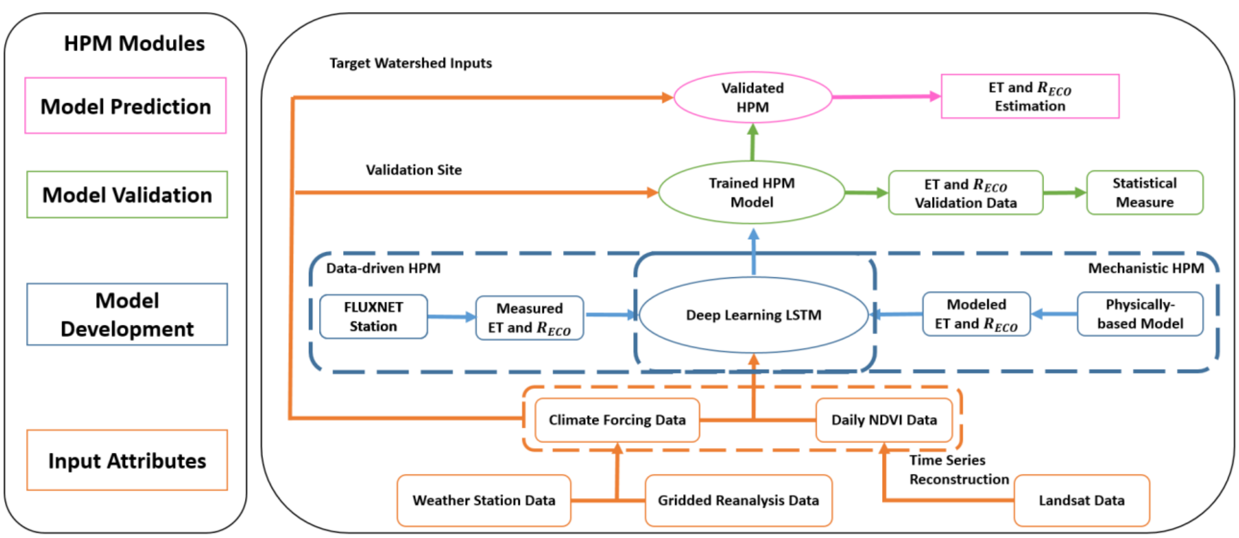

Figure 3: Hybrid Predictive Model (HPM) Framework. The HPM model mainly consists of four modules: Input Attributes, Model Development, Model Validation and Model Prediction, represented by rectangles with colors. Arrows represent the linkages among different modules. Choices of data-driven HPM or mechanistic HPM depend on the ecoregion of target watershed and data availability.

Long short-term memory (LSTM, Hochreiter & Schmidhuber, 1997) is capable of identifying long-term dependencies between climate and environmental data. For example, Kratzert et al. (2018) successfully used LSTM to learn the long-term dependencies in hydrological data (e.g., storage effects within catchments, time lags between precipitation inputs and runoff generation) for rainfall-runoff modeling. LSTM has also been used for gap filling in hydrological monitoring networks in the spatiotemporal domain (Ren et al., 2019). In this study, the outputs (ET or R_{ECO}) denoted as y are predicted from the input $x = [x_1, x_2, \ldots, x_T]$, consisting of the last Tconsecutive time steps of attributes, such as meteorological forcings attributes (e.g., air temperature and precipitation) and remote sensing attributes (i.e., NDVI). In a recurrent neural network (RNN), h_t represents the internal state at every time step t that takes in current input value x_t and previous internal state h_{t-1}, and is recomputed along the time axis using the following equation:

$$h_t = g(Wx_t + Uh_{t-1} + b), \tag{1}$$

where g represents the hyperbolic tangent activation function, W and U are trainable weight metrices of the hidden state h, and b is a bias vector. W, U and b are all trainable through optimization. LSTM introduces the cell state c_t, which makes LSTM powerful in identifying long-term dependencies in a statistical manner. The cell state c_t has three gates structures, including "forget gates" (which determine what information from previous cell states will be forgotten), "input gates" (which determine what information will be conveyed from the forget gate) and "output gates" (which return information from cell state c_t to a new state h_t). With these gate structures, the cell state c_t controls what information will be forgotten, conveyed, and updated over time. The forget gate is formulated as follows:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right), \tag{2}$$

where f_t results in a value between 0 and 1 indicating the degree of information to be forgotten; σ is the logistic sigmoid function, $and\ W_f, U_f$ and b_f are trainable parameters. Next, the input gate decides which values will be updated in the current cell state, and creates a vector of candidate values $\widetilde{c_t}$ in the range of (-1, 1) through a $tanh$ layer, which will be used to update the current state. With the candidate values calculated from the current state, and the information conveyed from the forget gate, we can calculate the current cell state as follows:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i), \tag{3}$$

$$\widetilde{c_t} = tan\,h(W_{\tilde{c}} x_t + U_{\tilde{c}} h_{t-1} + b_{\tilde{c}}), \tag{4}$$

$$c_t = f_t * c_{t-1} + i_t * \widetilde{c_t}, \tag{5}$$

where i_t is the input gate that defines which information of $\widetilde{c_t}$ will be used to update the current cell state and is in the range of (0, 1); c_t represents the current cell state; and $W_{\tilde{c}}, U_{\tilde{c}}, b_{\tilde{c}}, W_i, U_i,\ and\ b_i$ are trainable parameters. Finally, the output gate o_t controls the information of cell state c_t to a new hidden state h_t, which is computed using the following equation:

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o), \tag{6}$$

$$h_t = \tanh(c_t) * o_t, \tag{7}$$

With the new hidden state calculated, ET and R_{ECO} can be calculated using a one unit dense layer:

$$y_t = W_d h_t + b_d, \tag{8}$$

where W_d and b_d are additional trainable parameters. In summary, the LSTM unit calculates the internal state using current meteorological forcings and remote sensing data at every time step. The forget gate, input gate, and output gate decide what information from previous time steps will be kept, updated, and conveyed to the new hidden state. Finally, with a single dense layer, the algorithm will output ET and R_{ECO} estimation from the trained model.

A 70%-30% split between training and validation time series data was applied here, where the first 70% of the data were used for HPM development as a learning process, and 30% of the data were used as validation sets at individual sites. At the East River Watershed, HPM results were also validated with benchmark CLM outputs from Tran et al. (2019) and FLUXNET measurements. We used the mean absolute error (MAE), and adjusted R^2 as the statistical measure to determine model performance. In most models, the configuration of the neural networks includes a first LSTM layer with 50 units, a second LSTM layer with 25 units, and a dense layer with 8 units having L2 regularizers and a final output dense layer. Dropout layers are also embedded in the model to prevent overfitting. There are 11600 and 7600 parameters for the first and second LSTM layers; 208 and 9 for the first and second dense layers and no parameters for the dropout layers. Other configurations of networks may provide better estimation results; however, they are not assessed in this study as the proposed configuration already provide reasonable results. More information about the LSTM-RNN method is provided by Olah. (2015).

3.2 Feature Selection

Key properties influencing ET and R_{ECO} dynamics are linked to snow processes, plant dynamics, moisture stresses, radiation inputs and other relevant processes. However, at sparsely monitored watersheds, only weather reanalysis data and remote sensing data are commonly available. Thus we mainly considered temperature, radiation, precipitation, vegetation indices (e.g., NDVI) and variables inferred from these data as inputs for HPM. Other key attributes that depend on depth and site specific characteristics such as soil moisture and snow depth are not used in current HPM models due to data availability.

3.2.1 Snow information

In mountainous watersheds, snow dynamics significantly influence water and carbon fluxes. Because of the difficulties in measuring snow time series over space, we did not directly use attributes such as snow water equivalent as input to HPM. Instead, we separated precipitation data into snow precipitation (air temperature < 0) and rainfall precipitation (air temperature > 0). This is in line with what has been used in hydrological models such as CLM (Oleson et al., 2013). Note that for certain sites in this study, snow is not present (e.g., US-Ton). In order to capture the dynamics of snow processes, such as accumulation and melting, we constructed a categorical variable (sn), as follows:

$$\boldsymbol{sn} = \begin{cases} 0, \textit{during snow accumlation};\ \ SWE > 0 \textit{ and } SWE < \textit{peak } SWE \\ 1, \textit{during snow melting}; SWE > 0 \textit{ and } \ SWE \leq \textit{peak } SWE \\ 2, \textit{no snow}; SWE = \ 0 \end{cases}, \quad (9)$$

Since data on peak SWE are rarely available because of the difficulties in measuring snow, we also define a proxy categorical variable, $\boldsymbol{sn}$. When no SWE measurements were available, we estimated $\boldsymbol{sn}$ using air and soil temperature data following Knowles et al. (2016), who found significant correlations between the day of peak snow accumulation and first day of air temperature above 0 degrees Celsius, as follows:

$$\boldsymbol{sn} = \begin{cases} 0, \textit{during snow accumlation};\ \ \textit{Air Temperature} < 0 \\ 1, \textit{during snow melting}; \textit{Air Temperature} > 0 \textit{ while Soil Temperature} \leq 0 \\ 2, \textit{no snow}; \textit{Air Temperature and Soil Temperature} > 0 \end{cases}, \quad (10)$$

3.2.2 Vegetation information

To mitigate the long return periods of satellites and the presence of clouds, we reconstructed daily NDVI values based on meteorological forcings data (e.g., air temperature, precipitation, radiation) using deep-learning recurrent neural networks, leading to estimates of NDVI at daily temporal resolution. For example, Figure 4 represents Landsat-derived NDVI and reconstructed NDVI values for two sites at the East River, CO watershed: Butte (ER-BT), and Schofield Pass (ER-SP). Figure 4 reveals that based on meteorological forcings data only, the reconstructions achieved an adjusted R^2 of 0.65. Though not ideal, as satellites continue to advance and more training data becomes available, the accuracy of NDVI temporal reconstruction is expected to increase.

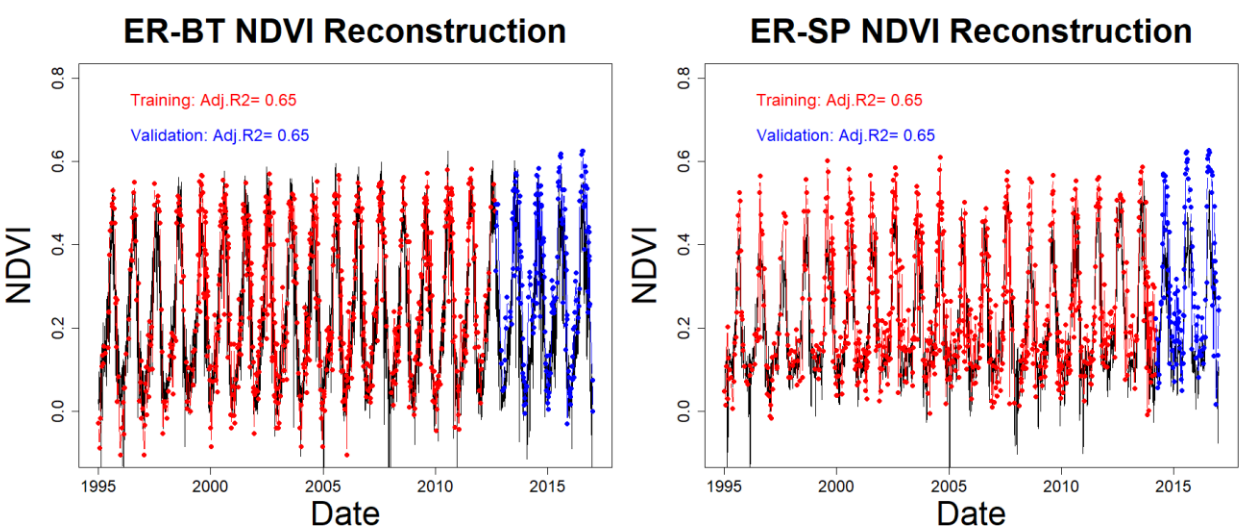

Figure 4: Temporal reconstruction of NDVI at ER-BT (left) and ER-SP (right). Black lines represent reconstructed daily NDVI. Red points are used for training and blue points are used for validation

4. Results

We tested HPM's capabilities using four different use cases to explore different conditions. First, we tested the capability of HPM to estimate long-term temporal dependency among meteorological forcings, ET, and R_{ECO} (Use Case 1; presented in Section 4.1). Second, we validated HPM's capability to estimate the spatial distribution of ET and R_{ECO} over space in selected watersheds, where we developed HPM using existing FLUXNET data (Use Case 2; data-driven HPM, Section 4.2) or outputs from a mechanistic model (Use Case 3; physical-model-based HPM, Section 4.3). In Use Case 4, HPM was used to estimate ET and R_{ECO} at selected sites within the East River Watershed and to distinguish how ET and R_{ECO} dynamics varies in the East River Watershed (Section 4.4). Temporal resolution of HPM models for all Use Cases are at daily scale and the spatial resolution depends on the use of meteorological forcing data. These four use cases illustrate and demonstrate how HPM can be developed and applied at target watersheds where data are sparse.

4.1 Use Case 1: ET and R_{ECO} Time Series Estimation with HPM Developed at FLUXNET Sites

Local HPMs were developed to estimate ET and R_{ECO} using flux tower data obtained from FLUXNET sites listed in Table 1. At all FLUXNET sites, air temperature, precipitation, net radiation, NDVI and soil temperature were used. For US-NR1, CA-Oas and CA-Obs, sn is also included. The results, which are shown in Figure 5 and Table 3, reveal that the HPM approach was effective for estimating ET. Adjusted R^2 between the HPM estimates and flux tower measurements are above 0.85 for all sites, and mean absolute errors are small at a level of ~0.2 mm/d. Figure 5 displays the daily scale estimation of ET from HPM US-NR1 and CA-OAS (other sites provided in supplementary material), and presents monthly mean ET values of measurements, HPM estimations, and differences. The long-term trends in ET are well captured by HPM. At larger temporal scales (monthly or yearly), HPM provides reasonable estimation of ET at these sites. However, short-term fluctuations during the summer are also not well captured by ET, specifically at California sites during the periods when plant transpiration and soil evaporation are constrained by soil moisture (Figure A2 panel a).

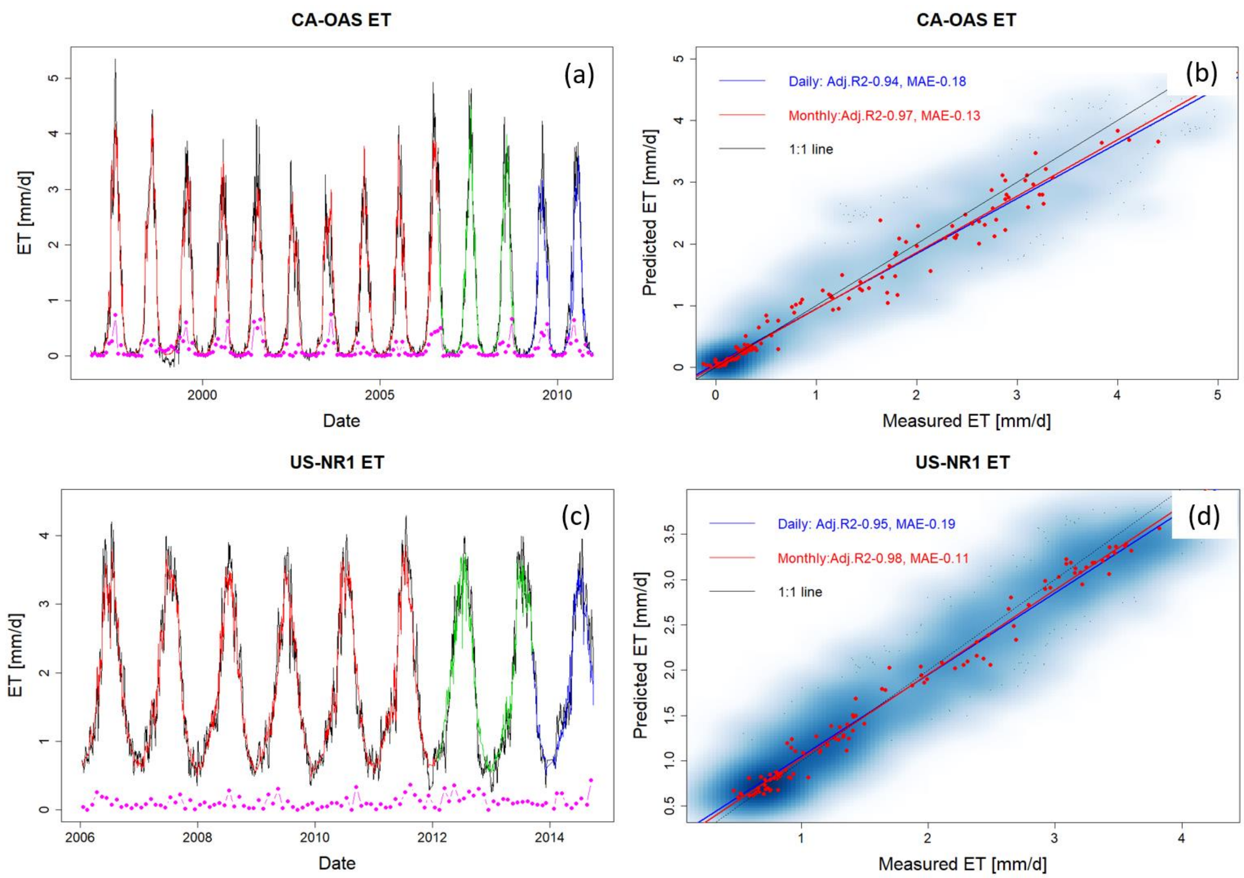

Figure 5: ET estimation with data from FLUXNET sites at CA-OAS and US-NR1. Panels (a) and (c) illustrate the daily estimation of ET with red, green, and blue lines representing data used for training, validation, and prediction, respectively, and the black line showing the eddy covariance measurements. Pink points describe monthly mean difference between HPM estimation and measured data. Panels (b) and (d) show the scatter plots of daily (blue) and monthly (red) ET. Darker blue clouds represent greater density of data points. Results for other sites are included in supplementary materials below (Figures A1 and A2).

Similarly, Table 3 and Figure 6 reveal that HPM was also effective in estimating R_{ECO}, leading to small MAE and adjusted R^2 of 0.8 between estimated and measured R_{ECO} except for US-Ton and US-Var. Figure 6 presents HPM-estimated R_{ECO} at US-NR1 and CA-OAS, with other sites presented in Figures A3 and A4. Long-term dynamics of R_{ECO} are also successfully captured by HPM; however, HPM does not accurately capture R_{ECO} during peak growing seasons. For example, we observed an over estimation of R_{ECO} during 2012 summer at US-Whs, whereas at US-NR1 HPM-estimation during peak growing season are smaller than measured values. While soil moisture is important for R_{ECO} during peak growing season (Ng et al., 2014; Wang et al., 2014), the developed HPM currently does not include soil moisture as a key attribute. HPM R_{ECO} estimation at US-Ton and US-Var show higher uncertainties (i.e., $MAE > 0.4$ and Adj. $R^2 < 0.8$). At these sites limited by water conditions (e.g., US-Ton) and sites with seasonally dry periods (e.g., US-Whs), it is necessary to include variables that could provide information regarding moisture stresses in the subsurface. Soil moisture that directly quantify water stress can be helpful to increase R_{ECO} prediction accuracy

(Noormets et al., 2008). Underestimation of peak growing season R_{ECO} can also come from biases within LSTM training, which is strong in capturing long-term temporal trends but less effective in obtaining peak values, and thus lead to increasing prediction errors during growing season compared to other periods of time.

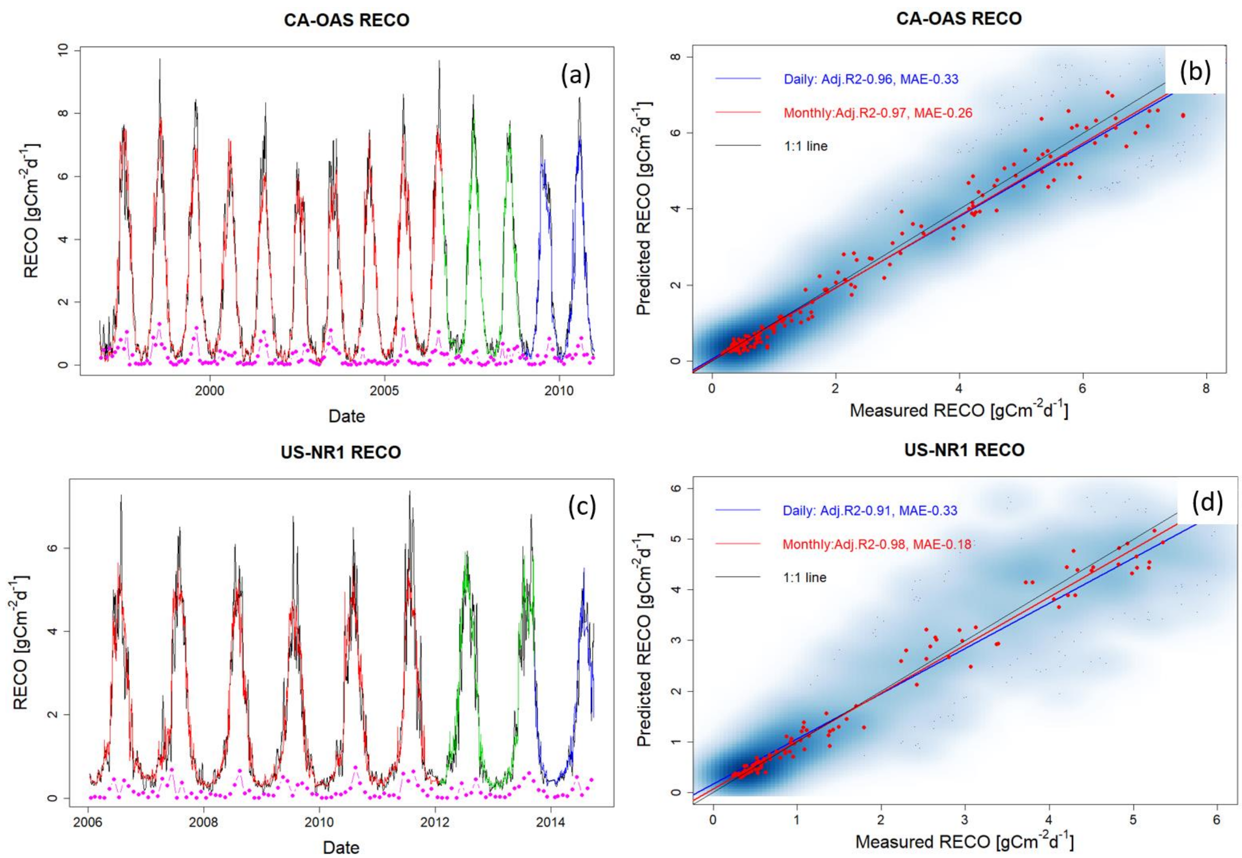

Figure 6: R_{ECO} estimation with data from FLUXNET sites at CA-OAS and US-NR1. Panels (a) and (c) present daily estimation of R_{ECO} with red, green, and blue lines representing data used for training, validation, and prediction, and the black line shows the eddy covariance measurements. Pink points describe monthly mean difference between HPM estimation and measured data. Panels (b) and (d) show the scatter plots of daily (blue) and monthly (red) R_{ECO}. Darker blue clouds represent greater density of data points. Results for other sites are included in supplementary materials below (Figures A3 and A4).

Table 3: Statistical measures of HPM estimation of ET and R_{ECO}

Site ID	Train MAE -ET $[mm/d]$	Test MAE - ET $[mm/d]$	Train Adj. R^2 - ET	Test Adj. R^2 - ET	Train MAE $-R_{ECO}$ $[gCm^{-2}d^{-1}]$	Test MAE $-R_{ECO}$ $[gCm^{-2}d^{-1}]$	Train Adj. R^2 $-R_{ECO}$	Test Adj. R^2 $-R_{ECO}$
US-NR1	0.19	0.11	0.95	0.98	0.33	0.18	0.91	0.98
CA-Oas	0.18	0.13	0.94	0.97	0.33	0.26	0.96	0.97
CA-Obs	0.12	0.09	0.95	0.96	0.29	0.25	0.96	0.97
US-SRM	0.22	0.17	0.92	0.94	0.24	0.19	0.80	0.87
US-Ton	0.22	0.17	0.92	0.94	0.43	0.36	0.76	0.82
US-Var	0.15	0.12	0.92	0.95	0.49	0.38	0.81	0.88
US-Whs	0.13	0.09	0.93	0.96	0.12	0.09	0.84	0.89
US-Wkg	0.19	0.15	0.87	0.91	0.18	0.15	0.85	0.91

4.2 Use Case 2: Ecoregion-Based, Data-Driven HPM Model for ET and R_{ECO} Estimation

While the effort and cost involved in establishing flux towers naturally limit the spatial coverage of obtained measurements, point scale measurements from one FLUXNET station provides representative information about ecosystem dynamics at other locations within the same ecoregion. In this section, we explored the use of a data-driven HPM trained with one FLUXNET station to estimate ET and R_{ECO} at other locations within the same ecoregion. To test this approach, we first trained HPM at a selected FLUXNET stations and validated these HPM models at other FLUXNET stations (ET and R_{ECO} data at testing sites were only used for comparison with HPM prediction) within the same ecoregion. Specifically, we developed HPM models at US-Ton, CA-Oas and US-Wkg, and provided ET and R_{ECO} estimations at US-Var, CA-Obs and US-Whs at three ecoregions, respectively.

Table 4 summarizes how we developed the data-driven HPM models for spatially distributed estimation of ET and R_{ECO} as well as the corresponding statistical summaries. Figures 7 and 8 present the time series of HPM-estimated ET and R_{ECO} compared to measurements from flux towers. HPM estimation at US-Obs, US-Whs and US-Var achieved an adjusted R^2 of 0.87, 0.88 and 0.91 for ET and 0.95, 0.70 and 0.78 for R_{ECO}, respectively. These results show that HPM captures the seasonal and long-term dynamics of ET and R_{ECO}. However at sites that experience seasonally dry periods (e.g., US-Whs), prediction accuracy decreases during the peak growing season. For example, we observed large errors in HPM-based estimations compared to measurements during peak growing seasons (e.g., a 0.5 mm discrepancy in June mean ET). We interpret this discrepancy as the result that current HPM models did not capture water stress conditions, and it is necessary to include other key attributes (e.g., soil moisture) to improve prediction accuracy, especially at these sites with seasonally dry periods. Although the prediction accuracy is not as high as Use Case 1 (Section 4.1), this use case demonstrates that HPM can learn the complicated relationships between responses and features successfully, and that a local data-driven HPM can be used to fuse with data from other subsites for long-term estimation of ET and R_{ECO} within the same ecoregions.

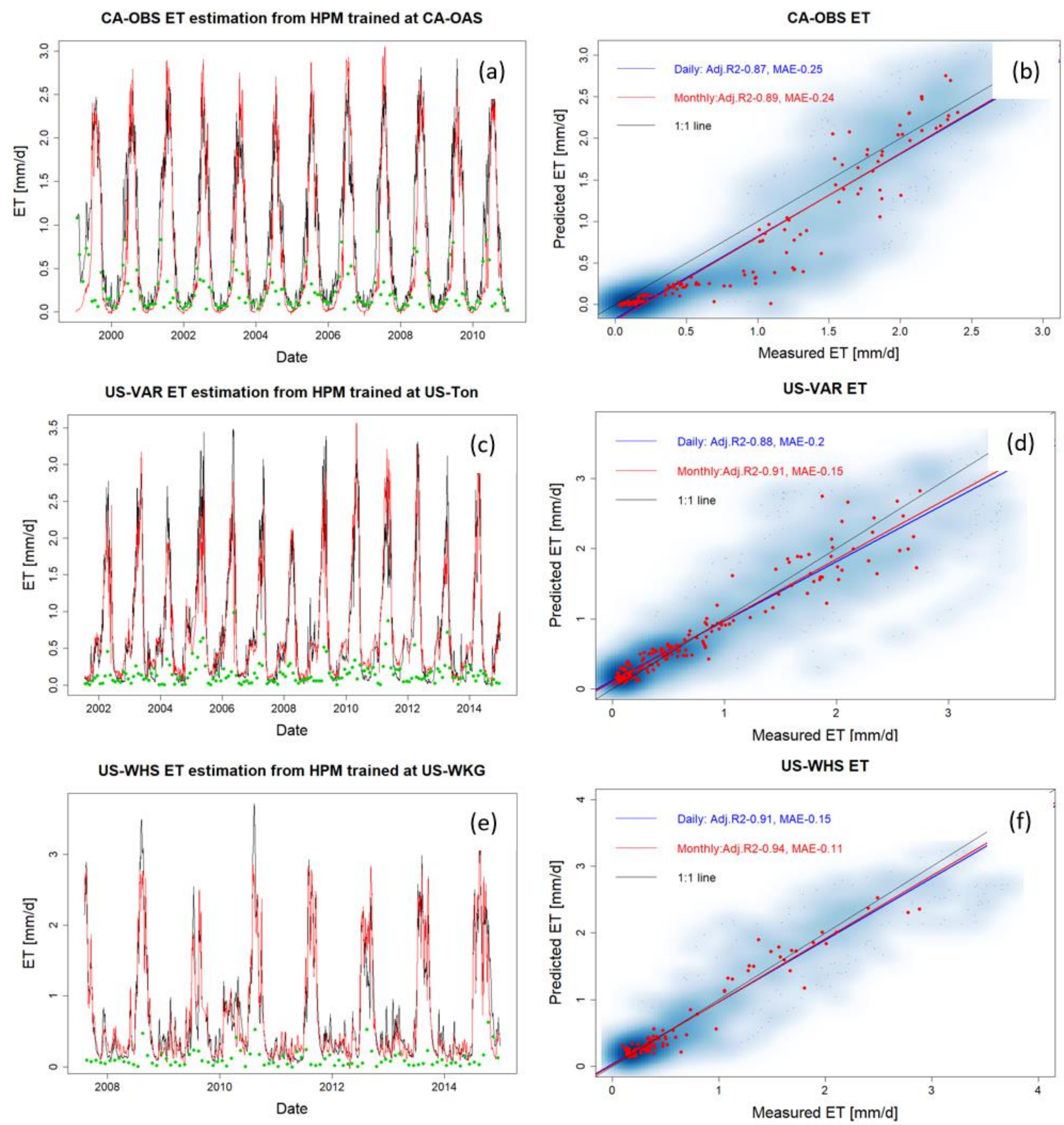

Figure 7. ET estimation at CA-Oas (a), US-Var (c), and US-Whs (e) with HPM trained at US-Ton, US-Wkg, and CA-Oas, respectively. Red and black lines represent HPM estimation and real measurements, with green points denoting the monthly mean difference between HPM estimationss and measurements. Panels (b), (d), and (f) show the scatter plots of daily (blue) and monthly (red) ET at these three sites. Darker blue clouds represent greater density of data points.

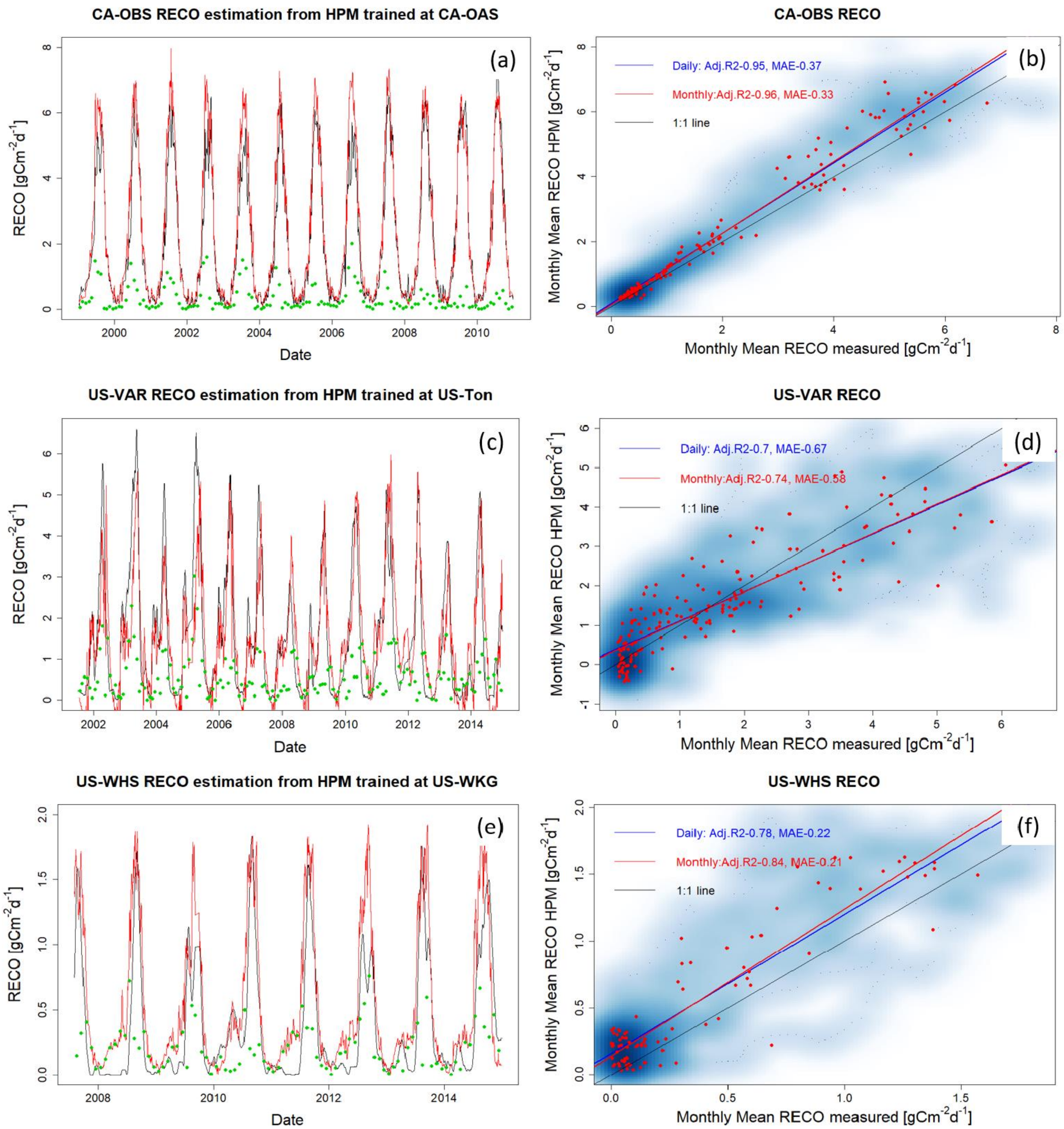

Figure 8. R_{ECO} estimation at CA-Oas (a), US-Var (c), and US-Whs (e) with HPM trained at US-Ton, US-Wkg, and CA-Oas, respectively. Red and black lines represent HPM estimations and real measurements; green points denote the monthly mean difference between HPM estimation and measurements. Panels (b), (d), and (f) show the scatter plots of daily (blue) and monthly (red) R_{ECO} at these three sites. Darker blue clouds represent greater density of data points.

4.3 Use Case 3: Ecoregion-Based, Mechanistic HPM Estimation of ET

Mechanistic HPM, which is trained with ET estimates from 1-D physically-based-model simulations, provides an avenue for estimating ET in ecoregions where direct measurements from eddy covariance tower are not available. In order to test the effectiveness of the mechanistic HPM, we focused on the three SNOTEL stations and US-NR1, which locates in the "Western Cordillera" ecoregion. Mechanistic HPM is coupled with CLM simulations at these sites (Tran et al., 2019). To ensure the CLM physically-based-model simulations can provide alternative datasets to develop mechanistic HPMs, we compared CLM estimation and direct measurements of ET at US-NR1 (Figure S2). The consistent results between measured ET and CLM-estimated ET (adjusted $R^2 = 0.88$; $k = 0.95$) indicate independent CLM simulations can be effectively used to develop the mechanistic HPM.

We applied mechanistic HPM trained with 1-D CLM developed at ER-BT (Tran et al., 2019) to estimate ET at sites classified as part of the same ecoregion (i.e., ER-SP, ER-PK and US-NR1). We then compared ET estimation from HPM to independent CLM-based ET estimations at ER-SP and ER-PK and to direct measurements at US-NR1. Figure 9 shows a high consistency between HPM estimation and the validation data. For all scenarios, an adjusted R^2 of 0.8 or greater is observed (Table 4), which strongly indicates that mechanistic HPM can provide accurate ET estimation at sites of similar ecoregions. These results suggest the broad applicability of mechanistic HPM to estimate ET based on ecoregion characteristics. This approach is expected to be particularly useful for regions where flux towers are difficult to install or where measured fluxes are not representative of the landscape, such as in mountainous watersheds.

Table 4. Statistical summary of HPM estimation over space with FLUXNET sites and SNOTEL stations with CLM

Target Site	**Training Site**	**Level II Ecoregion**	**ET MSE (monthly)**$[mm/d]$	**ET Adj.** R^2	R_{ECO} **MSE(monthly)**$[gCm^{-2}d^{-1}]$	R_{ECO} **Adj.** R^2
CA-Obs	CA-Oas	Boreal Plain	0.39	0.88	0.36	0.97
US-Var	US-Ton	Mediterrean California	0.34	0.70	0.67	0.70
US-Whs	US-Wkg	Western Serra Madre Pidemont	0.13	0.94	0.17	0.85
ER-SP	ER-BT	Western Cordillera	0.20	0.92	-	-
ER-PK	ER-BT	Western Cordillera	0.24	0.90	-	-
US-NR1	ER-BT	Western Cordillera	0.23	0.90		

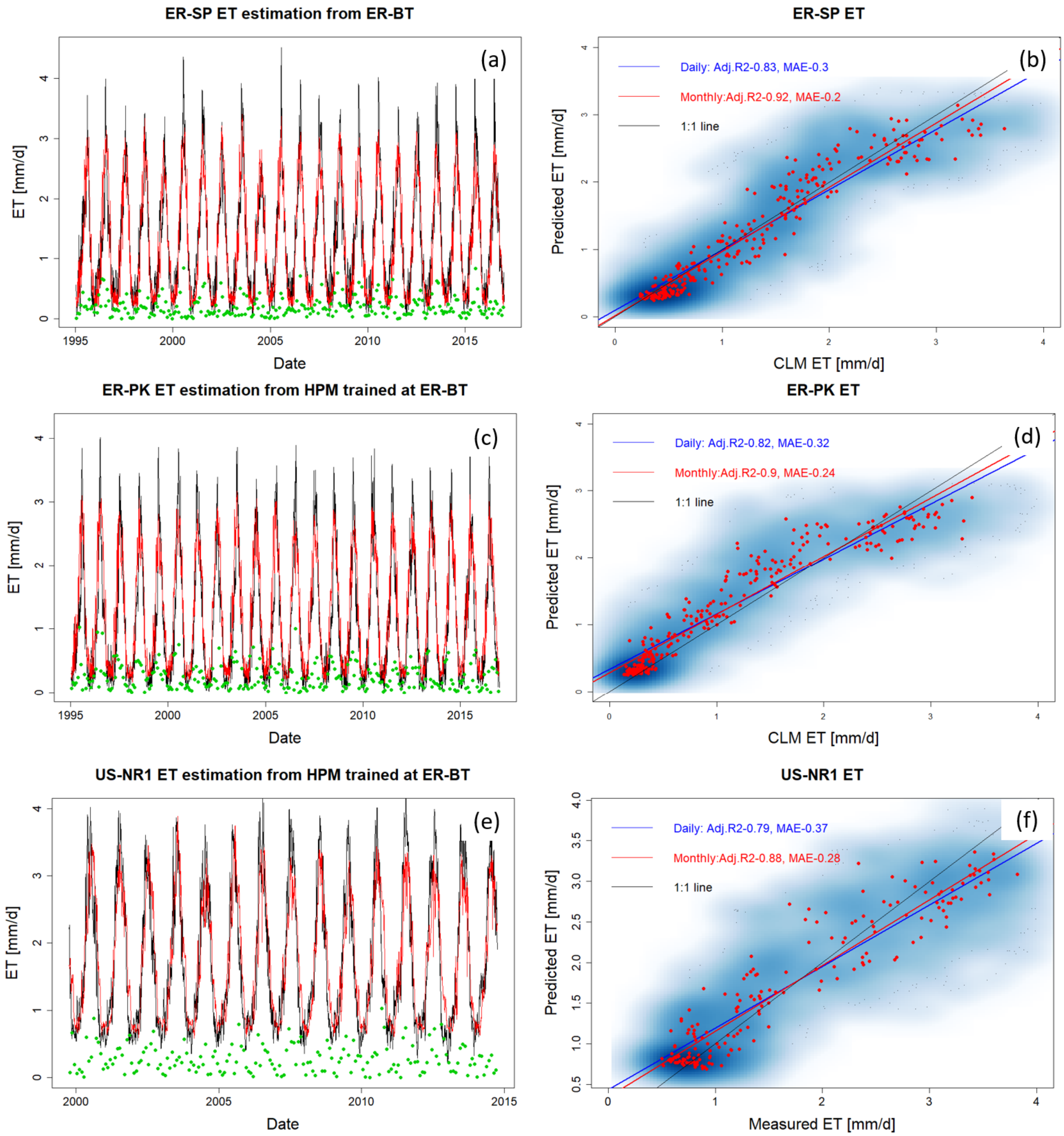

Figure 9. HPMs trained with CLM simulation at ER-BT are used to estimate ET at ER-SP, ER-PK, and US-NR1. Panels (a), (c), and (e) display HPM estimation of ET (red lines), as well as independent CLM estimation at ER-SP, ER-PK, and eddy covariance measurements at US-NR1 (black lines). Panels (b), (d), and (f) show the scatter plots of daily (blue) and monthly (red) ET at these three sites. Darker blue clouds represent greater density of data points.

4.4 Exploration of How ET and R_{ECO} Varies with Meteorological forcings and Vegetation Heterogeneity at the East River Watershed

ET and R_{ECO} estimated from the HPM model at the mountainous East River Watershed in CO enabled us to analyze how vegetation heterogeneity and meteorological forcings heterogeneity influence estimated ET and R_{ECO}

dynamics, and to identify limitations in the developed approach for estimating ET and R_{ECO} across mountainous and heterogeneous watersheds.

NDVI time-series data provide high-resolution (30m scale) information about vegetation variability across the East River Watershed. The spatial distribution of vegetation cover presented in Figure 2 (from Falco et al. 2019) enables us to distinguish different patches of deciduous forests, evergreen forests, meadow grassland and riparian shrublands and retrieve corresponding NDVI time-series. NDVI time series is related with snowmelt processes, whereas earlier snowmelt triggers earlier vegetation growth and result in earlier rise NDVI values (Pedersen et al., 2018). Figure 10 shows Landsat-derived and reconstructed NDVI values for the four different vegetation types within the East River Watershed. March, April and May mean NDVI values in 2012 for site DF1 are 0.07, 0.22 and 0.37 respectively compared to 0.06, 0.15 and 0.33 in 2015. The early rise of NDVI values observed in April 2012 is consistent with the fact that snowmelt occurred much earlier in 2012 than in 2015, as recorded by the SNOTEL Butte station. Earlier increase of NDVI in earlier snowmelt year (2012) was also observed for other vegetation types. In addition, evergreen forests have an extended growing season compared to the other vegetation types. For example, March-mean NDVI for EF1, RS1 and MS1 in 2012 are 0.30, 0.13, 0.11 compared to 0.28, 0.11, 0.08 in 2015, respectively whereas May-mean NDVI for EF1, RS1 and MS1 in 2012 are 0.38, 0.33, 0.35 compared to 0.34, 0.29 and 0.31 in 2015, respectively. Though earlier snowmelt triggers earlier increase in vegetation growth, significant faster greenness was observed for deciduous forests, meadow grasslands and shrublands compared to evergreen forests, where NDVI increased by 0.08, 0.20, 0.24 and 0.30 for evergreen forests, shrublands, grasslands and deciduous forests in 2012, respectively. In addition, peak NDVI is generally smaller in evergreen forests compared to deciduous forests, meadow grasslands and riparian shrublands. NDVI ranges from 0.2 to 0.6 for evergreen forests, whereas larger fluctuations in NDVI are observed for deciduous forests, shrublands and grasslands. The NDVI values during the winter are likely sensing both snow and forest density, due to pixel spatial averaging from Landsat images. Similar to Qiao et al. (2016), we also found that the NDVI of deciduous forests exhibits a significant increase during the growing season, followed by a sharp decline (likely caused by defoliation), and that evergreen forests had a more stable NDVI. Similar sharp decreases in the NDVI of riparian shrublands and meadow grasslands are observed.

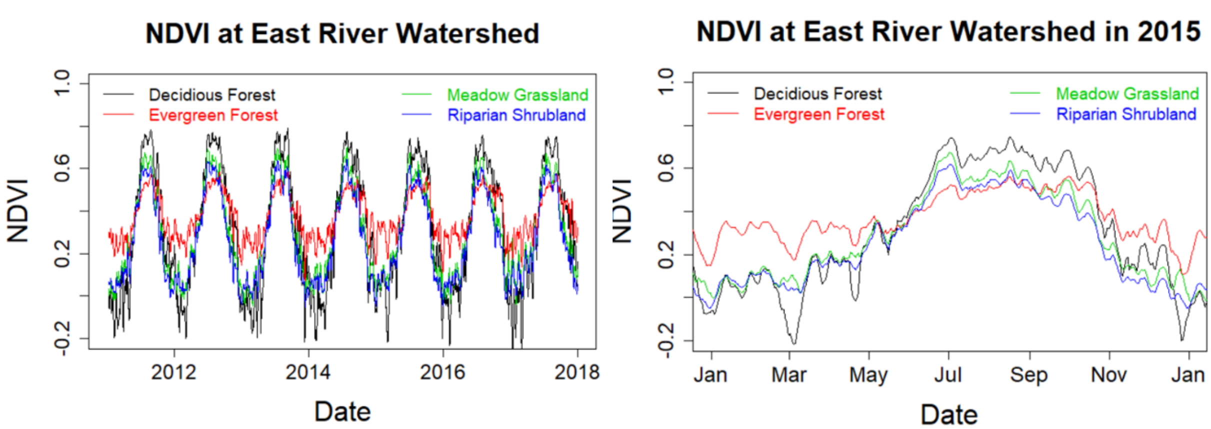

Figure 10: Reconstructed NDVI time series at selected locations in the East River Watershed for 2011 to 2018 (panel a) and for 2015 (panel b, normal water year). Black, red, green, and blue lines represent the time series of NDVI for deciduous forests, meadow grasslands, evergreen forests and riparian shrubland, respectively.

HPM-estimated ET and R_{ECO} also show different dynamics with different vegetation types as a result of differences in snowmelt timing, meteorological forcing and vegetation heterogeneity. Figure 11a and 11b present the time series of estimated ET and R_{ECO} associated with deciduous forests, respectively. Figure 11c and d present the ET and R_{ECO} differences between deciduous forests sites and evergreen forests, shrublands and grasslands. Before peak growing season, evergreen forests have the greatest ET and R_{ECO} compared to the other vegetation types. ET of evergreen forests is about 10% greater than deciduous forests, whereas ET of deciduous forests during peak growing season is greater than evergreen forests, shrublands and meadows. After growing season, the NDVI of deciduous forests is less than 0.2 (loss of leaves) compared to the NDVI of evergreen forests. Before peak growing season, R_{ECO} of evergreen forests is slightly greater than deciduous forests, meadow grasslands and shrublands. During peak growing season, we observed largest R_{ECO} for deciduous forests sites (~ 6 $gCm^{-2}d^{-1}$) followed by meadows, shrublands and evergreen forests. R_{ECO} of deciduous forests is around 17% greater than $R_{ECO}$of evergreen forests. However, we did not observe significant differences in annual ET among these four vegetation types (e.g., DF1: 535 to 573 mm, MS1: 534 to 570 mm, RS1: 532 to 567 mm and EF1: 532 to 569 mm across 7 years in this study). Total annual R_{ECO} of deciduous forests is greater than the other vegetation types (DF1: 642 to 698 gCm^{-2}, MS1: 588 to 636 gCm^{-2}, RS1: 589 to 636 gCm^{-2} and EF1: 592 to 639gCm^{-2}).

Considering the inter-annual variability in snow dynamics, we observed annual ET at $569\ mm$ and $532\ mm$ and annual R_{ECO} at $639\ gCm^{-2}$ and $602\ gCm^{-2}$ at EF1 for 2012 and 2015, respectively. We observed an earlier increase in ET and R_{ECO} in 2012 with March-mean ET and R_{ECO} at 0.69 mm/day and $0.51\ gCm^{-2}d^{-1}$ compared to 0.60 mm/day and 0.47 $gCm^{-2}d^{-1}$ in 2015. During peak growing season, we observed July-mean ET at 3.43 and 3.33 mm/day and R_{ECO} at 4.73 and 4.47 0.47 $gCm^{-2}d^{-1}$ for 2012 and 2015, respectively. Though earlier snowmelt usually triggers summer drought conditions, we observed a significantly greater amount of monsoon precipitation in 2012 (3.06mmd^{-1}) compared to 2015 (1.87mmd^{-1}). Water stress situation caused by earlier snowmelt was largely compensated by earlier monsoon in 2012, and thus we observed higher March, July and annual ET and R_{ECO} compared to 2015. Similar trends have also been observed for deciduous forests, shrublands and meadows in 2012 and 2015.

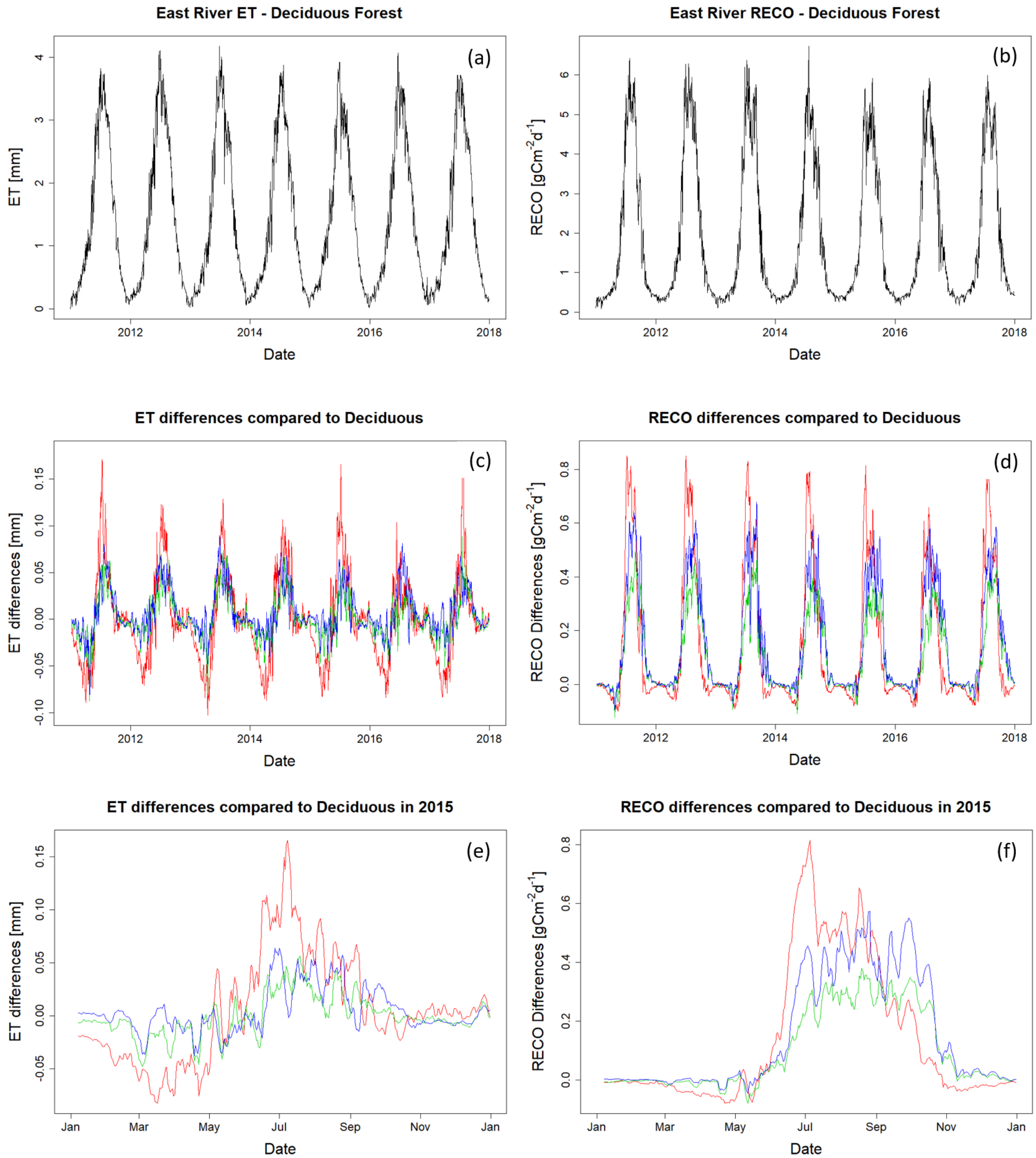

Figure 11: ET (a) and R_{ECO} (b) estimation for the deciduous forest site DF1 at the East River Watershed. Panels (c) and (d) show the differences in ET and R_{ECO} among various vegetation types and DF1. Red, green, and blue lines represent the differences in evergreen forest, meadow, and riparian shrubland compared to DF1. Panels (e) and (f) zoom into 2015 to better display seasonal variations.

ET and R_{ECO} estimation at the East River Watershed from the HPM model further enabled us to assess the role of input attributes and investigate limitations of the HPM approach. Figure 12 shows the absolute value of monthly mean difference in ET (Fig. 12a and Fig. 12b) and R_{ECO} (Fig. 12c and Fig. 12d) across SNOTEL stations (ER-BT, ER-SP and ER-PK) and within selected East River locations. Landsat data enabled us to capture NDVI differences at these sites (Figure 10), but we have identified the insufficient resolution of input meteorological forcing data at the East River sites. We observed a greater differences in temperature and radiation at the SNOTEL sites whereas there's very small differences at the East River sites (Figure S3). Summer temperature differences among SNOTEL sites can be over 3 ℃ but there's a barely 0.2 ℃ differences in DAYMET data used for the East River sites. In addition, a ~80 W/m^2 of radiation differences is observed with SNOTEL data whereas radiation differences stays around 30 W/m^2 for East River sites. Correspondingly, we observed 2.5 times greater differences in ET across SNOTEL stations compared to the sites within the East River watershed. We observed similar level of differences (around 0.8 gCm^{-2}) in R_{ECO} within East River Watershed and across SNOTEL stations. These results indicate HPM ET model is sensitive to temperature and radiation inputs whereas NDVI, temperature and radiation are all influential for HPM R_{ECO} models. Differences in ET and R_{ECO} among SNOTEL sites and East River sites are resulted from the differences in input meteorological forcing data. If high resolution meteorological data becomes available for the East River watershed, we believe the HPM approach can better capture heterogeneities in ET and R_{ECO} at the East River watershed and better distinguish the roles of meteorological forcing and vegetation heterogeneity on ET and R_{ECO} distribution.

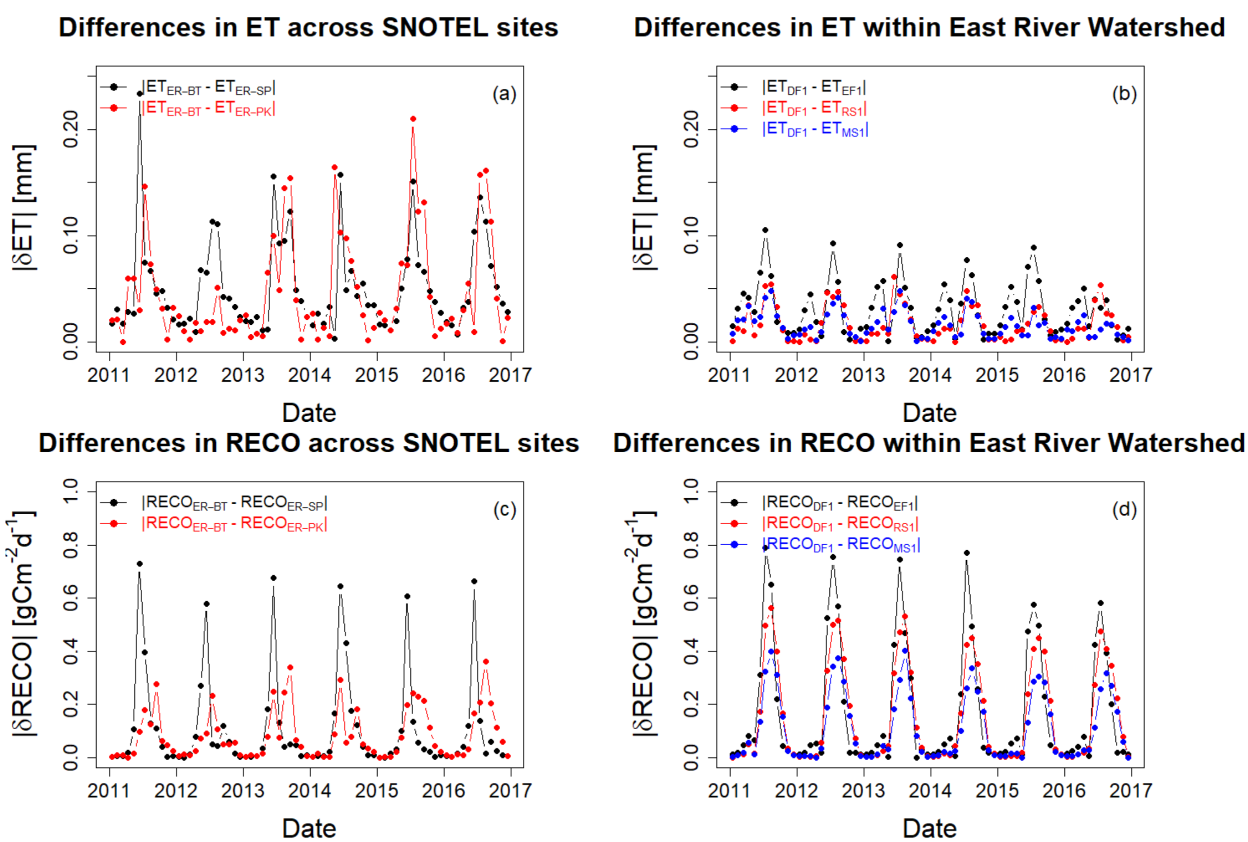

Figure 12. Absolute differences in monthly mean ET and R_{ECO} across SNOTEL stations and within East River Watershed. Panels (a) and (c) describe the absolute differences in monthly mean ET and R_{ECO} between ER-BT, ER-SP, and ER-PK.

Panels (b) and (d) describe the absolute differences in monthly mean ET and R_{ECO} within East River Watershed between deciduous forest (DF1), evergreen forest (EF1), meadow (MS1), and riparian shrubland (RS1).

5. Discussion

Our study demonstrates that HPM provides reliable estimations of ET and R_{ECO} under various climate and vegetation conditions, including data-based HPMs that are trained with FLUXNET data as well as physical-model-based HPMs that are coupled with simulations results from mechanistic models (i.e., CLM in this study). With 70% of the data used for training (model development), ET and R_{ECO} estimation from HPM achieves an adjusted R^2 of 0.9 compared to eddy covariance measurements. With this high estimation accuracy, we demonstrated that this approach could be used for predicting ET and R_{ECO} over time. HPM is capable of "learning" the complex interactions among meteorological forcings, vegetation dynamics, and water and carbon fluxes. The underlying relationships acquired by HPM can serve as a local ecohydrological model for long-term monitoring of ET and R_{ECO} with the aid of remote sensing data, and can fill in gap data during occasional equipment failure. HPM was also successful at estimating the spatial distribution of ET and R_{ECO} through exploiting an ecoregion concept. Using the representative FLUXNET sites in different ecoregions, HPM provided estimates of ET and R_{ECO} at locations using learned relationships from other sites having the same ecoregion classification. For conditions where no FLUXNET sites are within the same ecoregion, our study showed that physically-based models that utilize weather forcings data can provide alternatives for developing mechanistic HPM to estimate ET and R_{ECO}.

With the proposed HPM approach, we investigated the variability in ET and R_{ECO} estimations across different proportions of the East River Watersheds. While we currently do not have continuous measurements of ET and R_{ECO} at the East River Watershed for validation, our results are comparable to other studies that focus on sites within the same ecoregion. HPM-based ET estimation at East River Watershed is comparable to Mu et al. (2013), where ET is computed based upon the logic of the Penman-Monteith equation and MODIS remote sensing data (Figure S1), and the HPM-based R_{ECO} estimation is comparable to what Berryman et al. (2018) discovered, with growing season R_{ECO} ranging between 555 to 607 gCm^{-2} and mean growing season R_{ECO} ranging between $3.01\ to\ 3.30\ gCm^{-2}$. Annual ET between deciduous forests and evergreen forests are not statistically different, which is similar to Mu et al. (2013). Annual R_{ECO} differences between evergreen forests and deciduous forests are around $50\ gCm^{-2}$, which is comparable to Berryman et al. 2018.

We confirmed the important role of vegetation heterogeneity in modeling ET and R_{ECO} dynamics, which further enabled us to better understand ecosystem dynamics at the East River Watershed. As indicated from NDVI time series (Fig 10), evergreen forests have a longer growing season compared to other vegetation types; however, deciduous forests have higher peak NDVI values. Correspondingly, we also observed an earlier increase in ET and R_{ECO} for evergreen forests (before May), but larger ET and R_{ECO} for deciduous forests during peak growing season (around June and July). Similar dynamics were also observed at regions that have different climate conditions. Through assessing the differential mechanisms of deciduous forests and evergreen forests at various sites under Mediterranean climates, Baldocchi et al. (2010) found that deciduous forests had a shorter growing season, but showed a greater capacity for assimilating carbon during the growing season. Evergreen forests, on the other hand, had an extended growing season but with a smaller capacity for gaining carbon. These results were identified through analyzing the relationships among leaf ages, leaf nitrogen level, leaf area, and water use efficiencies of these tree species at the selected Mediterranean sites. They found older leaves tend to have smaller leaf nitrogen and stomata conductance that lead to smaller ET and R_{ECO} during peak growing seasons. Though our approach were not able to quantify the physiology differences among vegetation types, HPM estimation indicated evergreen forests that maintain leaves throughout the year have smaller ET and R_{ECO} during peak growing season compared to other vegetation types.

Dynamic changes in the inter-annual variability of meteorological conditions result in varying growing season length and spatiotemporal variability in ET and R_{ECO}. Earlier snowmelt triggers earlier growth of vegetation, causing earlier rise in ET and R_{ECO}. However, earlier growth in vegetation and increasing demand for water results in drought conditions (Sloat et al., 2015; Wainwright et al., 2020) that decrease ET and R_{ECO}. Timing and amount of monsoon precipitation are also important monsoons can relieve water stress and lead to increases in ET and R_{ECO}. Combination of these events jointly determine the magnitude of annual ET and R_{ECO}. Hu et al. (2010) analyzed flux data at US-NR1 to determine the relationships between growing season lengths and carbon sequestration, and found that extended growing season length resulted in less annual CO_2 uptake. They found that the duration of growing seasons substantially decreases snow water storage, which significantly decreases forest carbon uptake. Wieder et al. (2017) used point-scale CLM to better understand how complex terrain controls landscape-level variation of water, carbon and energy fluxes in the Niwot Ridge mountain ecosystems. With synthetic scenarios (e.g., different snow accumulation dynamics, fluctuations in air temperature), their simulation indicated earlier snowmelt and warmer summertime temperatures might drive divergent plant responses across the landscape. In our study, the combination of early snowmelt and early vegetation growth resulted in higher March ET and R_{ECO} in 2012 compared to 2015. The earlier start of growing season led to occurrences of fore-summer drought that decreases ET and R_{ECO}. However, the substantial earlier monsoon precipitation in 2012 relieved subsurface water stress whereas we observed higher July ET and R_{ECO} compared to other years. In addition, we observed smaller annual ET and R_{ECO} for evergreen forests that have longer growing season compared to other vegetation types. These results suggested HPM is capable of translating these variabilities in meteorological forcing and vegetation variables to ET and R_{ECO} dynamics.

Through comparing the HPM estimation results at different ecoregions, we also identified and assessed the limitations of current selection of input parameters. In the current study, we only used meteorological forcing and remote sensing based variables as inputs for HPM models, because these data are generally acquirable from weather reanalysis datasets and remote sensing products. HPM models with these variables provided reasonable estimates of ET and R_{ECO} for ecoregions limited by energy conditions, however we observed a decreasing prediction accuracy for ecoregions that experience seasonally dry periods. For example, HPM estimates at US-NR1 and CA-OAS achieved very high R^2 and small MAE; but prediction accuracy decreases especially during peak growing season at US-Ton and other water-limiting sites. These results indicate other key variables are necessary in order to capture dynamics during the seasonally dry periods, such as soil moisture measurement. The current HPM models did not use soil moisture as an input variable due to data availability reasons, but we believe and recommend adding soil moisture as well as other key variables to HPMs to further improve model performance at these seasaonlly dry ecoregions when such data becomes available.

Parameterization and spatiotemporal resolution of meteorological forcing data still remain a challenge for improving ET and R_{ECO} estimation at sparsely monitred watersheds. Microcliamte and heterogeneities in meteorological forcing attributes control the mangnitude and timing of ET and R_{ECO} dynamics. Other field observations along the Rocky Mountain ranges have shown that south-facing hillslopes have significantly earlier snowmelt compared to north-facing hillslopes (Kampf et al., 2015; Webb et al., 2018), which are hypothesized to result in significant differences in ET and R_{ECO} dynamics. We compared ET and R_{ECO} differences among SNOTEL sites and East River sites and identified ET differences among SNOTEL sites are greater than the differences among East River sites but R_{ECO} differences are similar between the two groups. Data from weather stations (SNOTEL sites) captured the spatiotemporal heterogeneity in radiation and temperature, however DAYMET data suggested very small differences in radiation and temperature (Figure S3 and S4). The insufficient spatial resolution of input meteorological forcing data limits HPM performance at the East River Watershed. Uncertainties in meteorological inputs can result in large errors (i.e., >20% MAE) and reduce accuracy by 10-30% in ET and R_{ECO} estimations as suggested by Mu et al. (2013) and Zhang et al. (2019). Thus, there is still a significant need for high-spatial-resolution meteorological-forcing data products to enable better estimates of ET and R_{ECO} and assess the governing factors that regulate their spatiotemporal variability.

In addition to the quality of meteorological data, HPM is also influenced by remote sensing inputs accuracy. Incorrectly calculated or pixel-averaged NDVI values from Landsat images can greatly alter HPM outputs for ET and R_{ECO}. Satellite images with different cloud cover have a slight influence over the NDVI values calculated, which do not represent real-time vegetation conditions. Algorithms used to reconstruct daily NDVI time series are also subject to uncertainties. But with recent advances in remote sensing and satellite technologies (McCabe et al., 2017) and harmonized Landsat-Sentinel datasets (Claverie et al., 2018), the spatial and temporal resolution should greatly increase in the future (i.e., 3 m resolution and daily). These advances will lead to more accurate classification of vegetation types and NDVI calculations, which are expected to decrease uncertainty associated with flux estimation

Another source of uncertainty in HPM arises from the choice of hybrid approaches and any parameter uncertainties in mechanistic models. Since HPM relies on accurate ET and R_{ECO} inputs from flux towers or mechanistic models, any uncertainties in measuring or modeling ET and R_{ECO} will propagate to HPM. If HPM is developed with a mechanistic model that has such missing components, these biases will be passed on to HPM estimation of ET and R_{ECO}. Parameter and conceptual model uncertainties in mechanistic models also restrict HPM's ability to "learn" the ecosystem dynamics. In order to reduce potential biasedness, we trained data-based HPM and physical-model-based HPM upon long time series (e.g., > 5 years) with quality assessed data or simulation results, which also enables HPM to better memorize long time dependencies of ecosystem dynamics. Though the quantification of uncertainties remains challenging, efforts have been made to lower these uncertainties using the technical advances described here.

6. Conclusion

In this study, we developed and tested a Hybrid Predictive Modeling (HPM) approach for ET and R_{ECO} estimation, with a focus on mountainous watersheds in the Rocky Mountains. We developed individual HPM models at various FLUXNET sites and at sites where data can supports the proper development of a mechanistic model (e.g., CLM). These models were validated against eddy covariance measurements and CLM outputs. We further used these models for ET and R_{ECO} estimation at watersheds within the same ecoregion to test HPM's capability of providing estimation over space, where only meteorological forcings data and remote sensing data were available. Lastly, we applied the HPM to provide long-term estimation of ET and R_{ECO} and test the sensitivity of HPM to various vegetation types at various sites within the East River Watershed of CO.

Given the promising results of HPM, this work offers an avenue for estimating ET and R_{ECO} using easy-to-acquire or commonly available datasets. This study also suggests that the spatial heterogeneity of meteorological forcings and vegetation dynamics have significant impacts on ET and R_{ECO} dynamics, which may be currently underestimated due to typically coarse spatial resolution of data inputs. Parameters related to energy and soil moisture conditions can be implemented into HPM to increase HPM's accuracy, especially for sites in ecoregions limited by soil moisture conditions. Lastly, it should be pointed out that HPM is not restricted to estimation of ET and R_{ECO} only. We focused here on developing HPM for ET and R_{ECO}, but HPM also has great potential for estimating other parameters important for water and carbon cycles given the right choice of input variables. Indeed, other attributes, such as net ecosystem exchange (Figure A6) and sensible heat flux, might also be accurately captured and represented with HPM, given the right choice of features.

Data availability. The data used in this study are from publicly available datasets. FLUXNET measurements can be accessed at https://FLUXNET.fluxdata.org. SNOTEL data are available at https://www.wcc.nrcs.usda.gov/snow/. DAYMET data can be found at (Thornton et al., 2017) or via Google Earth Engine. Landsat data are available on Google Earth Engine. All data and simulated results and model parameters associated with this article can be found at https://data.ess-dive.lbl.gov/view/doi:10.15485/1633810.

Acknowledgement. This material is based upon work supported as part of the Watershed Function Scientific Focus Area funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-AC02-05CH11231. We thank Haruko Wainwright and Bhavna Arora for providing comments on East River estimations. We also greatly appreciate all the guidance provided by Professor Yoram Rubin and Professor Dennis Baldocchi at UC Berkeley to the first author. We also acknowledge the Jane Lewis Fellowship Committee of the UC Berkeley for providing fellowship support to the first author.

References

Abatzoglou, J. T., Barbero, R., Wolf, J. W. and Holden, Z. A.: Tracking Interannual Streamflow Variability with Drought Indices in the U.S. Pacific Northwest, J. Hydrometeorol., doi:10.1175/jhm-d-13-0167.1, 2014.

Ai, J., Jia, G., Epstein, H. E., Wang, H., Zhang, A. and Hu, Y.: MODIS-Based Estimates of Global Terrestrial Ecosystem Respiration, J. Geophys. Res. Biogeosciences, 123(2), 326–352, doi:10.1002/2017JG004107, 2018.

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: Crop evapotranspiration: Guidelines for computing crop requirements., 1998.

Anderson, M. C., Allen, R. G., Morse, A. and Kustas, W. P.: Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources, Remote Sens. Environ., doi:10.1016/j.rse.2011.08.025, 2012.

Baldocchi, D.: Measuring fluxes of trace gases and energy between ecosystems and the atmosphere - the state and future of the eddy covariance method, Glob. Chang. Biol., doi:10.1111/gcb.12649, 2014.

Baldocchi, D. D., Ma, S., Rambal, S., Misson, L., Ourcival, J. M., Limousin, J. M., Pereira, J. and Papale, D.: On the differential advantages of evergreenness and deciduousness in mediterranean oak woodlands: A flux perspective, Ecol. Appl., 20(6), 1583–1597, doi:10.1890/08-2047.1, 2010.

Berryman, E. M., Vanderhoof, M. K., Bradford, J. B., Hawbaker, T. J., Henne, P. D., Burns, S. P., Frank, J. M., Birdsey, R. A. and Ryan, M. G.: Estimating Soil Respiration in a Subalpine Landscape Using Point, Terrain, Climate, and Greenness Data, J. Geophys. Res. Biogeosciences, 123(10), 3231–3249, doi:10.1029/2018JG004613, 2018.

Bodesheim, P., Jung, M., Gans, F., Mahecha, M. D. and Reichstein, M.: Upscaled diurnal cycles of land-Atmosphere fluxes: A new global half-hourly data product, Earth Syst. Sci. Data, 10(3), 1327–1365, doi:10.5194/essd-10-1327-2018, 2018.

De Bruin, H. A. R.: A model for the Priestley-Taylor parameter alpha., J. Clim. Appl. Meteorol., doi:10.1175/1520-0450(1983)0222.0.CO;2, 1983.

Budyko, M. I.: The Heat Balance of the Earth's Surface, Sov. Geogr., 2(4), 3–13, doi:10.1080/00385417.1961.10770761, 1961.

Chang, L. L., Dwivedi, R., Knowles, J. F., Fang, Y. H., Niu, G. Y., Pelletier, J. D., Rasmussen, C., Durcik, M., Barron-Gafford, G. A. and Meixner, T.: Why Do Large-Scale Land Surface Models Produce a Low Ratio of Transpiration to Evapotranspiration?, J. Geophys. Res. Atmos., doi:10.1029/2018JD029159, 2018.

Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J. C., Skakun, S. V. and Justice, C.: The Harmonized Landsat and Sentinel-2 surface reflectance data set, Remote Sens. Environ., doi:10.1016/j.rse.2018.09.002, 2018.

Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A. and Totterdell, I. J.: Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model, Nature, doi:10.1038/35041539, 2000.

Daggers, T. D., Kromkamp, J. C., Herman, P. M. J. and van der Wal, D.: A model to assess microphytobenthic primary production in tidal systems using satellite remote sensing, Remote Sens. Environ., 211(April), 129–145, doi:10.1016/j.rse.2018.03.037, 2018.

Falco, N., Wainwright, H., Dafflon, B., Léger, E., Peterson, J., Steltzer, H., Wilmer, C., Rowland, J. C., Williams, K. H. and Hubbard, S. S.: Investigating Microtopographic and Soil Controls on a Mountainous Meadow Plant Community Using High-Resolution Remote Sensing and Surface Geophysical Data, J. Geophys. Res. Biogeosciences, doi:10.1029/2018JG004394, 2019.

Gao, X., Mei, X., Gu, F., Hao, W., Li, H. and Gong, D.: Ecosystem respiration and its components in a rainfed spring maize cropland in the Loess Plateau, China, Sci. Rep., doi:10.1038/s41598-017-17866-1, 2017.

Gao, Y., Yu, G., Li, S., Yan, H., Zhu, X., Wang, Q., Shi, P., Zhao, L., Li, Y., Zhang, F., Wang, Y. and Zhang, J.: A remote sensing model to estimate ecosystem respiration in Northern China and the Tibetan Plateau, Ecol. Modell., doi:10.1016/j.ecolmodel.2015.03.001, 2015.

van Gorsel, E., Delpierre, N., Leuning, R., Black, A., Munger, J. W., Wofsy, S., Aubinet, M., Feigenwinter, C., Beringer, J., Bonal, D., Chen, B., Chen, J., Clement, R., Davis, K. J., Desai, A. R., Dragoni, D., Etzold, S., Grünwald, T., Gu, L., Heinesch, B., Hutyra, L. R., Jans, W. W. P., Kutsch, W., Law, B. E., Leclerc, M. Y., Mammarella, I., Montagnani, L., Noormets, A., Rebmann, C. and Wharton, S.: Estimating nocturnal ecosystem respiration from the vertical turbulent flux and change in storage of CO 2, Agric. For. Meteorol., 149(11), 1919–

1930, doi:10.1016/j.agrformet.2009.06.020, 2009.

Greve, P., Gudmundsson, L., Orlowsky, B. and Seneviratne, S. I.: Introducing a probabilistic Budyko framework, Geophys. Res. Lett., 42(7), 2261–2269, doi:10.1002/2015GL063449, 2015.

Hargrove, W. W. and Hoffman, F. M.: Using multivariate clustering to characterize ecoregion borders, Comput. Sci. Eng., 1(4), 18–25, doi:10.1109/5992.774837, 1999.

Hargrove, W. W., Hoffman, F. M. and Law, B. E.: New analysis reveals representativeness of the amerflux network, Eos (Washington. DC)., 84(48), doi:10.1029/2003EO480001, 2003.

Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, Neural Comput., doi:10.1162/neco.1997.9.8.1735, 1997.

Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J. and Megown, K.: Completion of the 2011 national land cover database for the conterminous United States – Representing a decade of land cover change information, Photogramm. Eng. Remote Sensing, doi:10.1016/S0099-1112(15)30100-2, 2015.

Hu, J., Moore, D. J. P., Burns, S. P. and Monson, R.: Longer growing seasons lead to less carbon sequestration by a subalpine forest, Glob. Chang. Biol., 16(2), 771–783, doi:10.1111/j.1365-2486.2009.01967.x, 2010.

Hubbard, S. S., Williams, K. H., Agarwal, D., Banfield, J., Beller, H., Bouskill, N., Brodie, E., Carroll, R., Dafflon, B., Dwivedi, D., Falco, N., Faybishenko, B., Maxwell, R., Nico, P., Steefel, C., Steltzer, H., Tokunaga, T., Tran, P. A., Wainwright, H. and Varadharajan, C.: The East River, Colorado, Watershed: A Mountainous Community Testbed for Improving Predictive Understanding of Multiscale Hydrological–Biogeochemical Dynamics, Vadose Zo. J., 17(1), 0, doi:10.2136/vzj2018.03.0061, 2018.

IPCC: IPCC 2019- Special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystem, Res. Handb. Clim. Chang. Agric. Law, doi:10.4337/9781784710644, 2019.

Irons, J. R., Dwyer, J. L. and Barsi, J. A.: The next Landsat satellite: The Landsat Data Continuity Mission, Remote Sens. Environ., doi:10.1016/j.rse.2011.08.026, 2012.

Jägermeyr, J., Gerten, D., Lucht, W., Hostert, P., Migliavacca, M. and Nemani, R.: A high-resolution approach to estimating ecosystem respiration at continental scales using operational satellite data, Glob. Chang. Biol., doi:10.1111/gcb.12443, 2014.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., De Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., Mu, Q., Mueller, B., Oleson, K., Papale, D., Richardson, A. D., Roupsard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S. and Zhang, K.: Recent decline in the global land evapotranspiration trend due to limited moisture supply, Nature, doi:10.1038/nature09396, 2010.

Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., Arneth, A., Camps-Valls, G., Ciais, P., Friedlingstein, P., Gans, F., Ichii, K., Jain, A. K., Kato, E., Papale, D., Poulter, B., Raduly, B., Rödenbeck, C., Tramontana, G., Viovy, N., Wang, Y. P., Weber, U., Zaehle, S. and Zeng, N.: Compensatory water effects link yearly global land CO 2 sink changes to temperature, Nature, 541(7638), 516–520, doi:10.1038/nature20780, 2017.

Kampf, S., Markus, J., Heath, J. and Moore, C.: Snowmelt runoff and soil moisture dynamics on steep subalpine hillslopes, Hydrol. Process., 29(5), 712–723, doi:10.1002/hyp.10179, 2015.

Keenan, T. F., Migliavacca, M., Papale, D., Baldocchi, D., Reichstein, M., Torn, M. and Wutzler, T.: Widespread inhibition of daytime ecosystem respiration, Nat. Ecol. Evol., 3(3), 407–415, doi:10.1038/s41559-019-0809-2, 2019.

Knowles, J. F., Blanken, P. D. and Williams, M. W.: Wet meadow ecosystems contribute the majority of overwinter soil respiration from snow-scoured alpine tundra, J. Geophys. Res. G Biogeosciences, 121(4), 1118–1130, doi:10.1002/2015JG003081, 2016.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K. and Herrnegger, M.: Rainfall–runoff modelling using Long Short-

Term Memory (LSTM) networks, Hydrol. Earth Syst. Sci., 22(11), 6005–6022, doi:10.5194/hess-22-6005-2018, 2018.

Lasslop, G., Reichstein, M., Papale, D., Richardson, A., Arneth, A., Barr, A., Stoy, P. and Wohlfahrt, G.: Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: Critical issues and global evaluation, Glob. Chang. Biol., 16(1), 187–208, doi:10.1111/j.1365-2486.2009.02041.x, 2010.

Livingston, G. P. and Hutchinson, G. L.: Enclosure-based measurement of trace gas exchange: applications and sources of error., 1995.

Ma, Y., Liu, S., Song, L., Xu, Z., Liu, Y., Xu, T. and Zhu, Z.: Estimation of daily evapotranspiration and irrigation water efficiency at a Landsat-like scale for an arid irrigation area using multi-source remote sensing data, Remote Sens. Environ., 216(August), 715–734, doi:10.1016/j.rse.2018.07.019, 2018.

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U. and Gascon, F.: Sen2Cor for Sentinel-2., 2017.

McCabe, M. F., Aragon, B., Houborg, R. and Mascaro, J.: CubeSats in Hydrology: Ultrahigh-Resolution Insights Into Vegetation Dynamics and Terrestrial Evaporation, Water Resour. Res., 53(12), 10017–10024, doi:10.1002/2017WR022240, 2017.

Metzger, S., Junkermann, W., Mauder, M., Butterbach-Bahl, K., Trancón Y Widemann, B., Neidl, F., Schäfer, K., Wieneke, S., Zheng, X. H., Schmid, H. P. and Foken, T.: Spatially explicit regionalization of airborne flux measurements using environmental response functions, Biogeosciences, 10(4), 2193–2217, doi:10.5194/bg-10-2193-2013, 2013.

Migliavacca, M., Reichstein, M., Richardson, A. D., Mahecha, M. D., Cremonese, E., Delpierre, N., Galvagno, M., Law, B. E., Wohlfahrt, G., Andrew Black, T., Carvalhais, N., Ceccherini, G., Chen, J., Gobron, N., Koffi, E., William Munger, J., Perez-Priego, O., Robustelli, M., Tomelleri, E. and Cescatti, A.: Influence of physiological phenology on the seasonal pattern of ecosystem respiration in deciduous forests, Glob. Chang. Biol., 21(1), 363–376, doi:10.1111/gcb.12671, 2015.

Mohanty, B. P., Cosh, M. H., Lakshmi, V. and Montzka, C.: Soil moisture remote sensing: State-of-the-science, Vadose Zo. J., doi:10.2136/vzj2016.10.0105, 2017.

Mu, Q., Zhao, M. and Running, S. W.: MODIS Global Terrestrial Evapotranspiration (ET) Product (MOD16A2/A3), Algorithm Theor. Basis Doc., 2013.

NASA: Moderate Resolution Imaging Spectroradiometer (MODIS) Overview, Nasa, 2008.

Ng, G., Bedford, D. and Miller, D.: A mechanistic modeling and data assimilation framework for Mojave Desert ecohydrology, Water Resour. Res., 4662–4685, doi:10.1002/2014WR015281.Received, 2014.

Noormets, A., Desai, A. R., Cook, B. D., Euskirchen, E. S., Ricciuto, D. M., Davis, K. J., Bolstad, P. V., Schmid, H. P., Vogel, C. V., Carey, E. V., Su, H. B. and Chen, J.: Moisture sensitivity of ecosystem respiration: Comparison of 14 forest ecosystems in the Upper Great Lakes Region, USA, Agric. For. Meteorol., doi:10.1016/j.agrformet.2007.08.002, 2008.

Olah, C.: Understanding LSTM Networks, https//colah.github.io/posts/2015-08-Understanding-LSTMs/, 2015, 2015.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, J., Subin, Z. M., Swenson, S. C., Thornton, P. E., Bozbiyik, A., Fisher, R. A., Heald, C. L., Kluzek, E., Lamarque, J.-F., Lawrence, P. J., Leung, L. R., Lipscomb, W., Muszala, S., Ricciuto, D. M., Sacks, W. J., Sun, Y., Tang, J. and Yang, Z.-L.: Technical Description of version 4.5 of the Community Land Model (CLM)., 2013.

Omernik, J. M.: Perspectives on the nature and definition of ecological regions., Environ. Manage., doi:10.1007/s00267-003-5197-2, 2004.

Omernik, J. M. and Griffith, G. E.: Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework, Environ. Manage., doi:10.1007/s00267-014-0364-1, 2014.

Oyler, J. W., Dobrowski, S. Z., Ballantyne, A. P., Klene, A. E. and Running, S. W.: Artificial amplification of warming trends across the mountains of the western United States, Geophys. Res. Lett., doi:10.1002/2014GL062803, 2015.

Paca, V. H. da M., Espinoza-Dávalos, G. E., Hessels, T. M., Moreira, D. M., Comair, G. F. and Bastiaanssen, W. G. M.: The spatial variability of actual evapotranspiration across the Amazon River Basin based on remote sensing products validated with flux towers, Ecol. Process., 8(1), doi:10.1186/s13717-019-0158-8, 2019.

Pedersen, S. H., Liston, G. E., Tamstorf, M. P., Abermann, J., Lund, M. and Schmidt, N. M.: Quantifying snow controls on vegetation greenness, Ecosphere, doi:10.1002/ecs2.2309, 2018.

PRIESTLEY, C. H. B. and TAYLOR, R. J.: On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters, Mon. Weather Rev., doi:10.1175/1520-0493(1972)100<0081:otaosh>2.3.co;2, 1972.

Pumpanen, J., Kolari, P., Ilvesniemi, H., Minkkinen, K., Vesala, T., Niinistö, S., Lohila, A., Larmola, T., Morero, M., Pihlatie, M., Janssens, I., Yuste, J. C., Grünzweig, J. M., Reth, S., Subke, J. A., Savage, K., Kutsch, W., Østreng, G., Ziegler, W., Anthoni, P., Lindroth, A. and Hari, P.: Comparison of different chamber techniques for measuring soil CO_2 efflux, Agric. For. Meteorol., doi:10.1016/j.agrformet.2003.12.001, 2004.

Qiao, Z., Xu, X., Zhao, M., Wang, F. and Liu, L.: The application of a binary division procedure to the classification of forest subcategories using MODIS time-series data during 2000–2010 in China, Int. J. Remote Sens., doi:10.1080/01431161.2016.1176269, 2016.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J. M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D. and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm, Glob. Chang. Biol., doi:10.1111/j.1365-2486.2005.001002.x, 2005.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N. and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566(7743), 195–204, doi:10.1038/s41586-019-0912-1, 2019.

Ren, H., Cromwell, E., Kravitz, B. and Chen, X.: Using Deep Learning to Fill Spatio-Temporal Data Gaps in Hydrological Monitoring Networks, Hydrol. Earth Syst. Sci. Discuss., (May), 1–20, doi:10.5194/hess-2019-196, 2019.

Rungee, J., Bales, R. and Goulden, M.: Evapotranspiration response to multiyear dry periods in the semiarid western United States, Hydrol. Process., doi:10.1002/hyp.13322, 2019.

Ryu, Y., Baldocchi, D. D., Kobayashi, H., Van Ingen, C., Li, J., Black, T. A., Beringer, J., Van Gorsel, E., Knohl, A., Law, B. E. and Roupsard, O.: Integration of MODIS land and atmosphere products with a coupled-process model to estimate gross primary productivity and evapotranspiration from 1 km to global scales, Global Biogeochem. Cycles, 25(4), 1–24, doi:10.1029/2011GB004053, 2011.

Seneviratne, S. I., Lüthi, D., Litschi, M. and Schär, C.: Land-atmosphere coupling and climate change in Europe, Nature, doi:10.1038/nature05095, 2006.

Sloat, L. L., Henderson, A. N., Lamanna, C. and Enquist, B. J.: The Effect of the Foresummer Drought on Carbon Exchange in Subalpine Meadows, Ecosystems, 18(3), 533–545, doi:10.1007/s10021-015-9845-1, 2015.

Strachan, S., Kelsey, E. P., Brown, R. F., Dascalu, S., Harris, F., Kent, G., Lyles, B., McCurdy, G., Slater, D. and Smith, K.: Filling the Data Gaps in Mountain Climate Observatories Through Advanced Technology, Refined Instrument Siting, and a Focus on Gradients, Mt. Res. Dev., 36(4), 518–527, doi:10.1659/mrd-journal-d-16-00028.1, 2016.

Suleau, M., Moureaux, C., Dufranne, D., Buysse, P., Bodson, B., Destain, J. P., Heinesch, B., Debacq, A. and Aubinet, M.: Respiration of three Belgian crops: Partitioning of total ecosystem respiration in its heterotrophic, above- and below-ground autotrophic components, Agric. For. Meteorol., doi:10.1016/j.agrformet.2011.01.012, 2011.

Teuling, A. J., Van Loon, A. F., Seneviratne, S. I., Lehner, I., Aubinet, M., Heinesch, B., Bernhofer, C., Grünwald, T., Prasse, H. and Spank, U.: Evapotranspiration amplifies European summer drought, Geophys. Res. Lett., doi:10.1002/grl.50495, 2013.

Thornton, P. E., Thornton, M. M., Mayer, B. W., Wei, Y., Devarakonda, R., Vose, R. S. and Cook, R. B.: Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3, ORNL DAAC, Oak Ridge, Tennessee, USA, 2017.

Tran, A. P., Rungee, J., Faybishenko, B., Dafflon, B. and Hubbard, S. S.: Assessment of spatiotemporal variability of evapotranspiration and its governing factors in a mountainous watershed, Water (Switzerland), 11(2), doi:10.3390/w11020243, 2019.

U.S. Environmental Protection Agency: Level III Ecoregions of the Continental United States, Environ. Prot., 2003.

Viviroli, D. and Weingartner, R.: “Water towers”—A global view of the hydrological importance of mountains, in Advances in Global Change Research., 2008.

Viviroli, D., Dürr, H. H., Messerli, B., Meybeck, M. and Weingartner, R.: Mountains of the world, water towers for humanity: Typology, mapping, and global significance, Water Resour. Res., 43(7), 1–13, doi:10.1029/2006WR005653, 2007.

Vogelmann, J. E., Howard, S. M., Yang, L., Larson, C. R., Wylie, B. K. and Van Driel, N.: Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat thematic mapper data and ancillary data sources, Photogramm. Eng. Remote Sensing, 2001.

Wainwright, H. M., Steefel, C., Trutner, S. D., Henderson, A. N., Nikolopoulos, E. I., Wilmer, C. F., Chadwick, K. D., Falco, N., Schaettle, K. B., Brown, J. B., Steltzer, H., Williams, K. H., Hubbard, S. and Enquist, B. J.: Satellite-derived foresummer drought sensitivity of plant productivity in Rocky Mountain headwater catchments: spatial heterogeneity and geological-geomorphological control, Environ. Res. Lett., doi:10.1088/1748-9326/ab8fd0, 2020.

Wang, B., Zha, T. S., Jia, X., Wu, B., Zhang, Y. Q. and Qin, S. G.: Soil moisture modifies the response of soil respiration to temperature in a desert shrub ecosystem, Biogeosciences, 11(2), 259–268, doi:10.5194/bg-11-259-2014, 2014.

Webb, R. W., Fassnacht, S. R. and Gooseff, M. N.: Hydrologic flow path development varies by aspect during spring snowmelt in complex subalpine terrain, Cryosphere, 12(1), 287–300, doi:10.5194/tc-12-287-2018, 2018.

Wieder, W. R., Knowles, J. F., Blanken, P. D., Swenson, S. C. and Suding, K. N.: Ecosystem function in complex mountain terrain: Combining models and long-term observations to advance process-based understanding, J. Geophys. Res. Biogeosciences, doi:10.1002/2016JG003704, 2017.

Williams, C. A. and Albertson, J. D.: Soil moisture controls on canopy-scale water and carbon fluxes in an African savanna, Water Resour. Res., doi:10.1029/2004WR003208, 2004.

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. and Wang, Y.-P.: Improving land surface models with FLUXNET data, Biogeosciences Discuss., doi:10.5194/bgd-6-2785-2009, 2009.

Wilson, K. B., Hanson, P. J., Mulholland, P. J., Baldocchi, D. D. and Wullschleger, S. D.: A comparison of methods for determining forest evapotranspiration and its components: Sap-flow, soil water budget, eddy covariance and catchment water balance, Agric. For. Meteorol., 106(2), 153–168, doi:10.1016/S0168-1923(00)00199-4, 2001.

Xiao, J., Ollinger, S. V., Frolking, S., Hurtt, G. C., Hollinger, D. Y., Davis, K. J., Pan, Y., Zhang, X., Deng, F., Chen, J., Baldocchi, D. D., Law, B. E., Arain, M. A., Desai, A. R., Richardson, A. D., Sun, G., Amiro, B., Margolis, H., Gu, L., Scott, R. L., Blanken, P. D. and Suyker, A. E.: Data-driven diagnostics of terrestrial carbon dynamics over North America, Agric. For. Meteorol., doi:10.1016/j.agrformet.2014.06.013, 2014.

Xu, L., Baldocchi, D. D. and Tang, J.: How soil moisture, rain pulses, and growth alter the response of ecosystem respiration to temperature, Global Biogeochem. Cycles, 18(4), 1–10, doi:10.1029/2004GB002281, 2004.

Xu, T., Guo, Z., Liu, S., He, X., Meng, Y., Xu, Z., Xia, Y., Xiao, J., Zhang, Y., Ma, Y. and Song, L.: Evaluating Different Machine Learning Methods for Upscaling Evapotranspiration from Flux Towers to the Regional Scale, J. Geophys. Res. Atmos., 123(16), 8674–8690, doi:10.1029/2018JD028447, 2018.

Zhang, L., Potter, N., Hickel, K., Zhang, Y. and Shao, Q.: Water balance modeling over variable time scales based on the Budyko framework - Model development and testing, J. Hydrol., 360(1–4), 117–131, doi:10.1016/j.jhydrol.2008.07.021, 2008.

Zhang, Y., Kong, D., Gan, R., Chiew, F. H. S., McVicar, T. R., Zhang, Q. and Yang, Y.: Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017, Remote Sens. Environ., 222(May 2018), 165–182, doi:10.1016/j.rse.2018.12.031, 2019.

Appendix

1. ET and R_{ECO} Estimation over Time at other Fluxnet sites

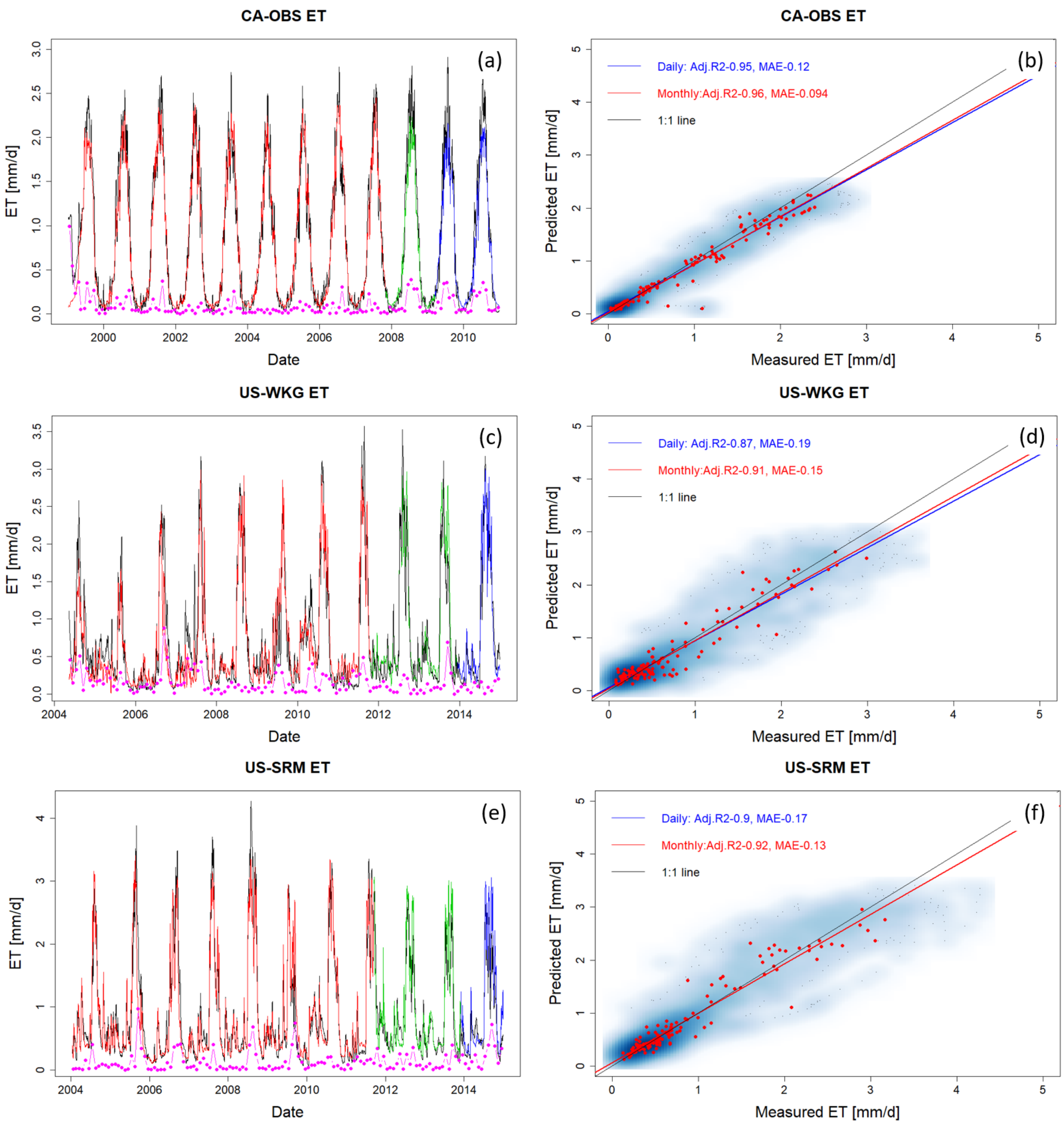

Figure A1: ET estimation with data from selected FLUXNET sites at CA-OBS, US-Wkg, and US-SRM. Panels (a), (c), and (e) present daily estimations of ET with red, green, and blue lines representing data used for training, validation, and prediction, respectively, and the black line representing the eddy covariance measurement. Pink points describe monthly mean difference between HPM estimation and measured data. Panels (b), (d), and (f) show the scatter plots of daily (blue) and monthly (red) ET. Darker blue clouds represent greater density of data points.

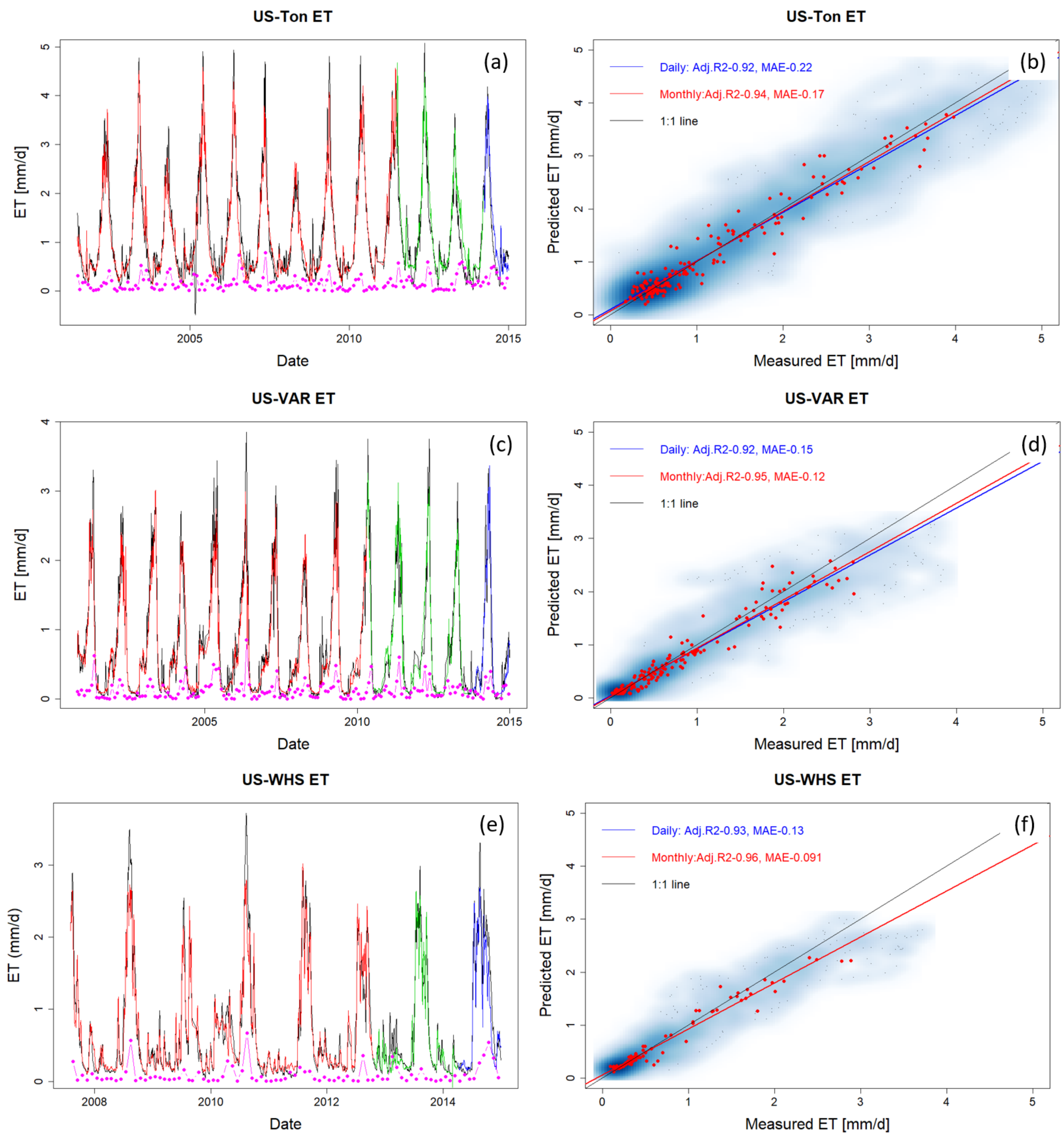

Figure A2: ET estimation with data from selected FLUXNET sites at US-Ton, US-Var, and US-Whs. Panels (a), (c), and (e) present daily estimations of ET with red, green, and blue lines representing data used for training, validation, and prediction, respectively, and the black line representing the eddy covariance measurement. Pink points describe monthly mean difference between HPM estimation and measured data. Panels (b), (d), and (f) show the scatter plots of daily (blue) and monthly (red) ET. Darker blue clouds represent greater density of data points.

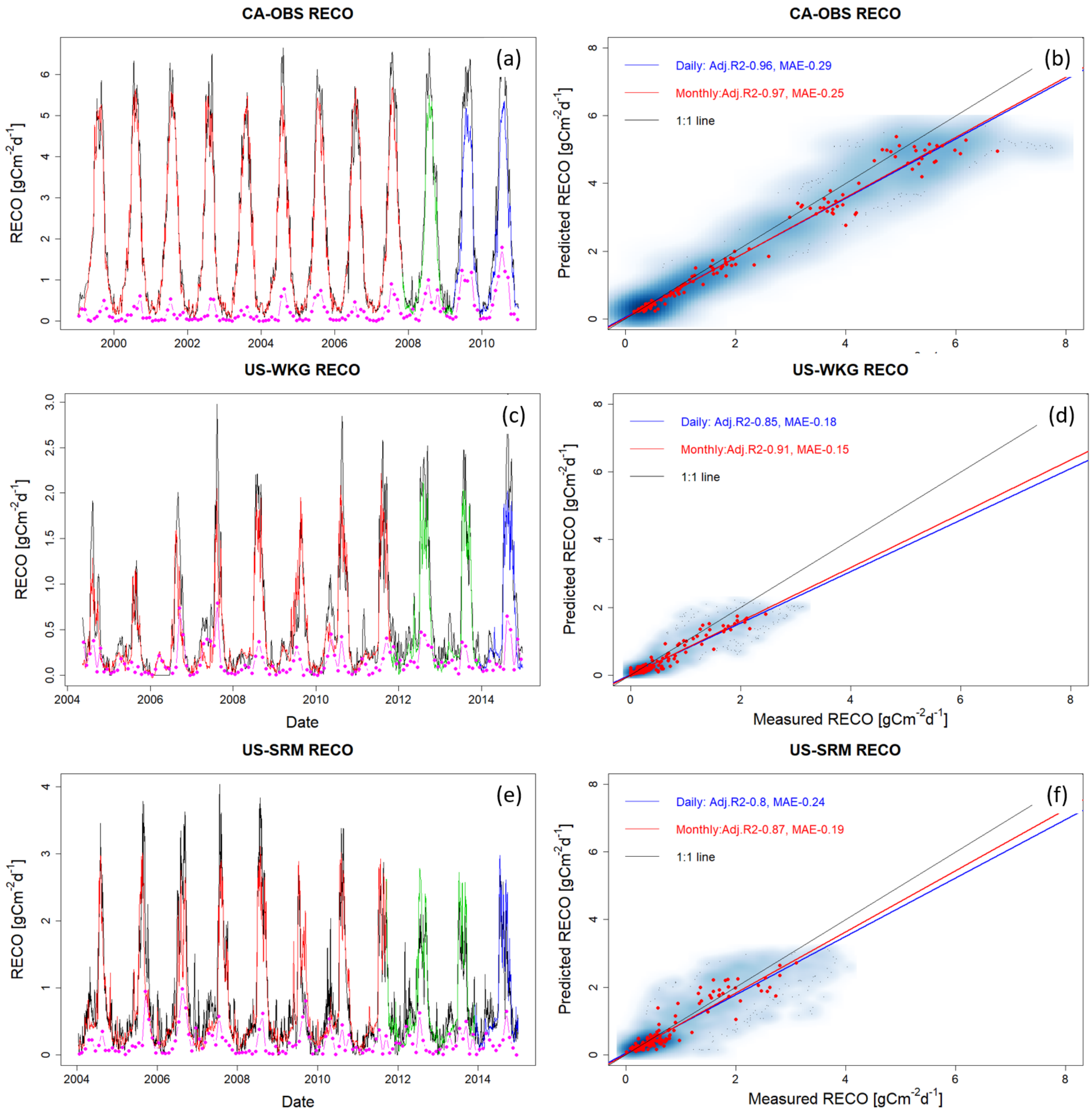

Figure A3: R_{ECO} estimation with data from selected FLUXNET sites at CA-OBS, US-Wkg, and US-SRM. Panels (a), (c), and (e) present daily estimations of R_{ECO} with red, green, and blue lines representing data used for training, validation, and prediction, respectively, and the black line is eddy covariance measurement. Pink points describe the monthly mean difference between HPM estimation and measured data. Panels (b), (d), and (f) show the scatter plots of daily (blue) and monthly (red) R_{ECO}. Darker blue clouds represent greater density of data points.

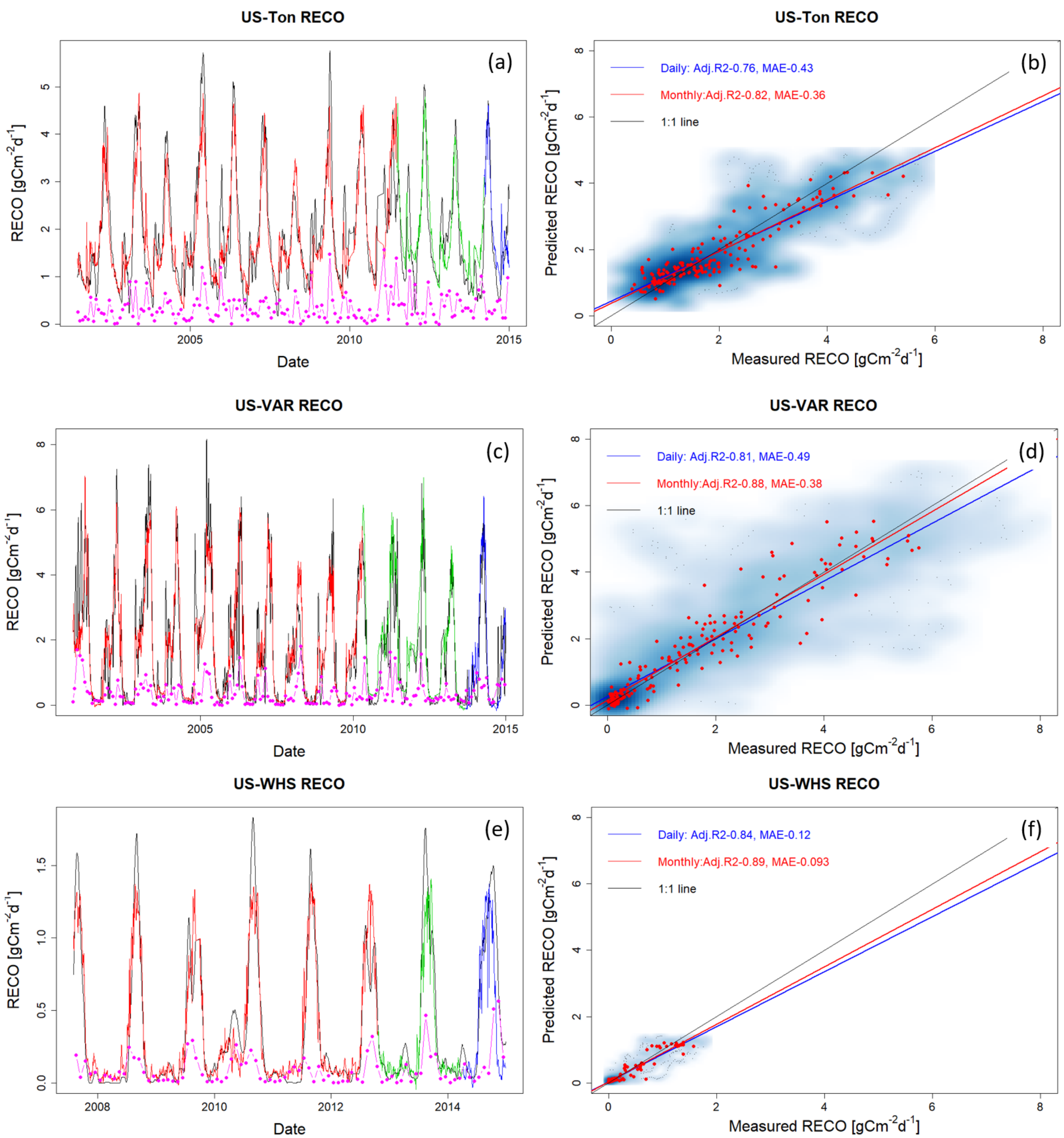

Figure A4: R_{ECO} estimation with data from selected FLUXNET sites at US-Ton, US-Var, and US-Whs. Panels (a), (c), and (e) present daily estimations of R_{ECO} with red, green, and blue lines representing data used for training, validation, and prediction, respectively, and the black line representing the eddy covariance measurement. Pink points describe monthly mean difference between HPM estimation and measured data. Panels (b), (d), and (f) show the scatter plots of daily (blue) and monthly (red) R_{ECO}. Darker blue clouds represent greater density of data points.

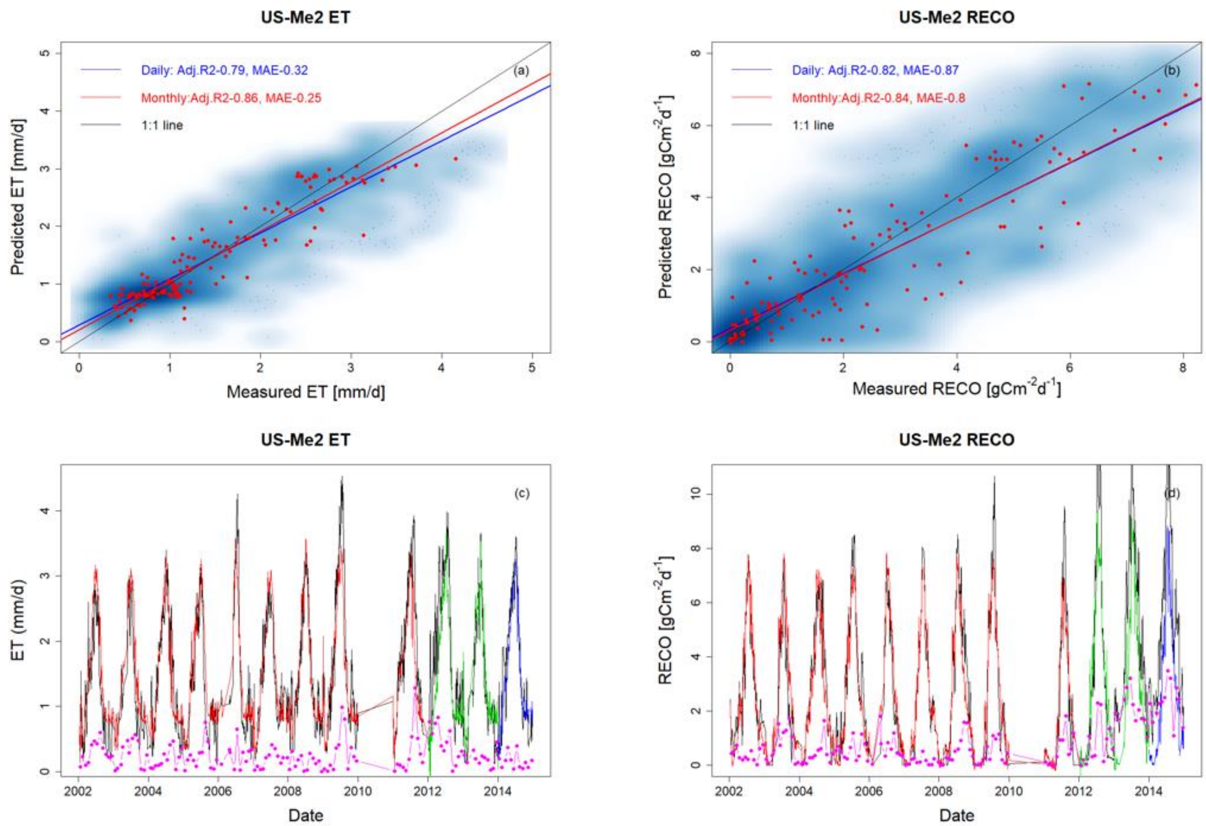

Figure A5: ET and R_{ECO} estimation at US-Me2. Panels (a) and (b) show the scatter plots of daily (blue) and monthly (red) ET and R_{ECO}. Darker blue clouds represent greater density of data points. Panels (c), and (d) present daily estimations of R_{ECO} with red, green, and blue lines representing data used for training, validation, and prediction, respectively, and the black line representing the eddy covariance measurement. Pink points describe monthly mean difference between HPM estimation and measured data.

2. Tested NEE Estimation over Time at CA-OAS and US-NR1

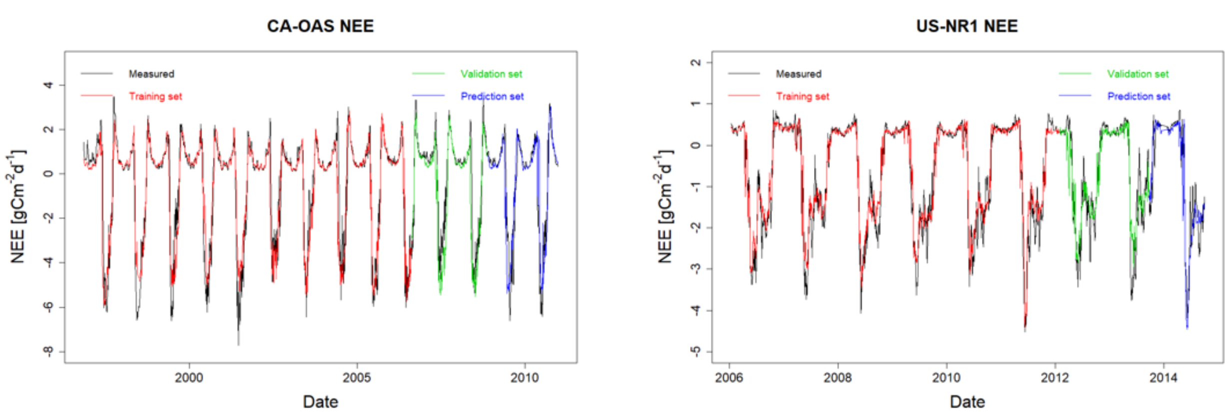

Figure A6. HPM estimate of NEE at CA-OAS and US-NR1. R^2 between estimation and measurements are 0.87, 0.83 and 0.81 at CA-OAS; 0.94, 0.88 and 0.90 at US-NR1 for the training set, validation set and prediction set, respectively. Model inputs include air temperature, soil temperature, sn, precipitation and radiation.