A Deep-Learning Hybrid-Predictive-Modeling (HPM) 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: Climate change is reshaping vulnerable ecosystems, leading to uncertain effects on ecosystem dynamics, including evapotranspiration (ET) and ecosystem respiration (R_{eco}). However, accurate estimation of ET and R_{eco} still remains challenging at sparsely monitored watersheds where data and field instrumentation are limited. 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 ET and R_{eco} in high space-time resolution. HPM relies on a deep learning algorithm, long short-term memory (LSTM), and requires only air temperature, precipitation, radiation, normalized differences vegetation index (NDVI) and soil temperature (when available) as input variables. We tested and validated HPM estimation results in different climate regions and developed four use cases to demonstrate the applicability and variability of HPM at various FLUXNET sites and Rocky Mountain SNOTEL sites in Western North America. To test the limitations and performance of HPMs in mountainous watersheds, an expanded use case focused on the East River Watershed, Colorado, USA. The results indicate 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. The study documents that HPM increases our capability to estimate ET and R_{eco} and enhances process understanding at sparsely monitored watersheds.

1. Introduction:

Climate change has a profound influence on global and regional energy, water and carbon cycling, including evapotranspiration (ET) and ecosystem respiration (R_{eco}). 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 as well as watershed responses to abrupt disturbances and gradual climate changes, 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 total respiration 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). While increases in R_{eco} may contribute to accelerating global warming through positive feedbacks to the atmosphere (Cox et al., 2000; Gao et al., 2017; IPCC, 2019; Suleau et al., 2011), estimating and monitoring R_{eco} over relevant spatiotemporal scales is challenging. As described below, there are many different strategies for measuring and estimating ET and R_{eco}, each of which has advantages and limitations. This study is motivated by 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 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). 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). 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 eddy covariance fluxes is on the order of hundreds of meters, and the temporal resolution of the measurements ranges from hours to decades (Wilson et al., 2001). Tower-based in-situ measurements of fluxes have been integrated into the global AmeriFlux (http://ameriflux.lbl.gov/) and FLUXNET (https://FLUXNET.fluxdata.org/) networks. Eddy covariance towers are usually installed at valley bottoms of mountainous watersheds (Strachan et al., 2016). Data from flux towers should also be used carefully as flux footprints may vary significantly across sites and through time depending on site-specific information, turbulent states of the atmosphere and underlying surface characteristics (Chu et al., 2021). Given the cost and efforts required to install and maintain a flux tower, eddy covariance towers are typically sparse and may not capture complex fluxes at sites with complex terrains, such as montane environments. Though measurements from a single flux tower may not capture heterogeneity in ET and R_{eco} due to complex terrains, they can support the development of statistical or physical-based models integrated with other types of data to provide ET and R_{eco} estimation as we describe herein.

Physically-based numerical models, which 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), such as the Community Land Model (CLM, Oleson et al., 2013). Performance of these models depends on the accuracy of inputs and parameters, such as soil type and leaf area index, which can be difficult to obtain at a 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) suggested that process-based models may not represent transpiration accurately due to challenges in simulating the uneven hydraulic distribution caused by complex terrain. Semi-analytical formulations are also commonly used to estimate ET, including 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). However, these conceptual uncertainties, in addition to data sparseness and data uncertainty, still limit the applicability of these approaches.

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} (Abatzoglou et al., 2014; Daggers et al., 2018; Mohanty et al., 2017; Paca et al., 2019). 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 indices from the MODIS dataset—and used the rate-temperature curve and strong correlations between terrestrial carbon exchange and air 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 data quality is subject to cloud 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 influence water and energy dynamics of an ecosystem.

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. 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, air temperature, and precipitation. 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 occur 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 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 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 can then be applied to extract the spatiotemporal patterns for ET and R_{eco} prediction. The integration of multi-model and multi-data can increase our modeling capability to estimate ET and R_{eco} and enhance our process understanding of ecosystem water and carbon cycling under climate change.

In this study, we developed a hybrid predictive modeling approach (HPM) to 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 flexibly integrate direct measurements from flux towers and/or physically-based model results (e.g., CLM) and utilize deep learning long-short term memory recurrent neural network (LSTM) to establish statistical relationships among fluxes, meteorological and remote sensing inputs. Once developed, the corresponding HPM can be used as a modeling tool to estimate ET and R_{eco} over space and time. We developed four use cases to demonstrate the applicability of HPM based on site-specific data and model availability. 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. We place a particular focus on mountainous sites, given their regional and global importance yet challenges associated with ET and R_{eco} in these regions, as described above.

2.1 FLUXNET Stations and Ecoregions

Nine FLUXNET stations, which cover a wide range of climate and elevations, were selected for this study (Table 1 and Figure 1). 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 were maintained to capture watershed dynamics at 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 four 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, USA), 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 one-dimensional (vertical) 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	Latitude, Longitude	Elevation (m)	Mean Annual air temperature (℃)	Mean Annual Precipitation (m)	Climate Koeppen	Vegetation IGBP	Ecoregion (Level II)	Period of Record
US-NR1	(40.0329, -105.5464)	3050	1.5	800	Dfc	ENF	Western Cordillera	2000-2014
CA-Oas	(53.6289, -106.1978)	530	0.34	428.53	Dfc	DBF	Boreal Plain	1997-2010
CA-Obs	(53.9872, -105.1178)	628.94	0.79	405.6	Dfc	ENF	Boreal Plain	1999-2010
US-SRM	(31.8214, -110.8661)	1120	17.92	380	Bsk	WSA	Western Sierra Madre Piedmont	2005-2015
US-Ton	(38.4316, -120.9660)	177	15.8	559	Csa	WSA	Mediterranean California	2002-2015
US-Var	(38.4133, -120.9507)	129	15.8	559	Csa	GRA	Mediterranean California	2002-2015
US-Whs	(31.7438, -110.0522)	1370	17.6	320	Bsk	OSH	Western Sierra Madre Piedmont	2008-2015
US-Wkg	(31.7365, -109.9419)	1531	15.64	407	Bsk	GRA	Western Sierra Madre Piedmont	2005-2015
US-Me2	(44.4523, -121.5574)	1253	6.28	523	Csb	ENF	Western Cordillera	2012-2015
ER-BT*	(38.894, -106.945)	3096	2.38	821	Dfc	N/A	Western Cordillera	1995-2017
ER-SP*	(39.02, -107.05)	3261	2.46	1064	Dfc	N/A	Western Cordillera	1995-2017
ER-PK*	(38.49, -106.34)	3280	1.97	574	Dfc	N/A	Western Cordillera	1995-2017

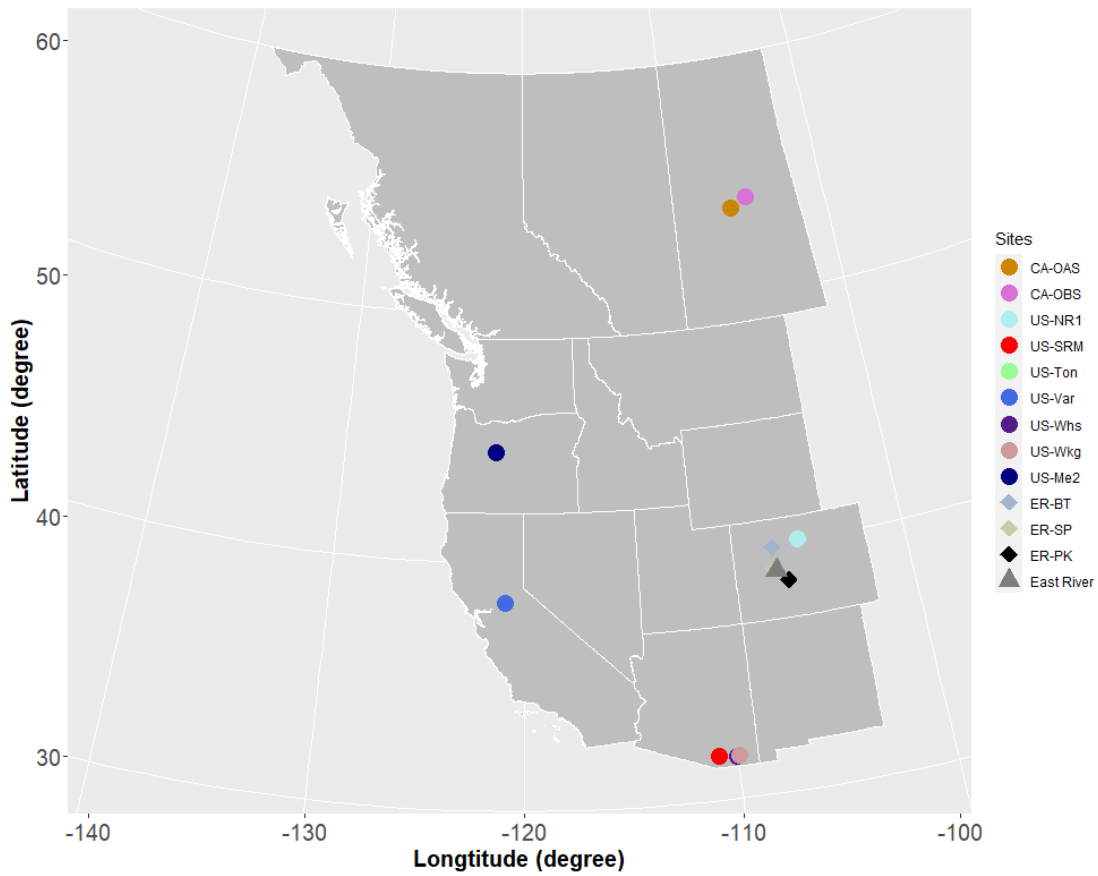

Figure 1. Location of sites considered in this study. Note: US-Ton and US-Var; US-Whs and US-Wkg are closed to each other. East River Watershed is located next to ER-BT. The white lines delineate Western US states and Canadian provinces. Circles represent FLUXNET sites, diamonds represent SNOTEL sites and triangle represents the East River Watershed.

2.3 East River Watershed Characteristics 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 mean annual air temperature in the East River is ~2.4°C, with average daily air temperatures of −7.6°C and 13.4°C in December and July respectively (Kakalia et al., 2020) and 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

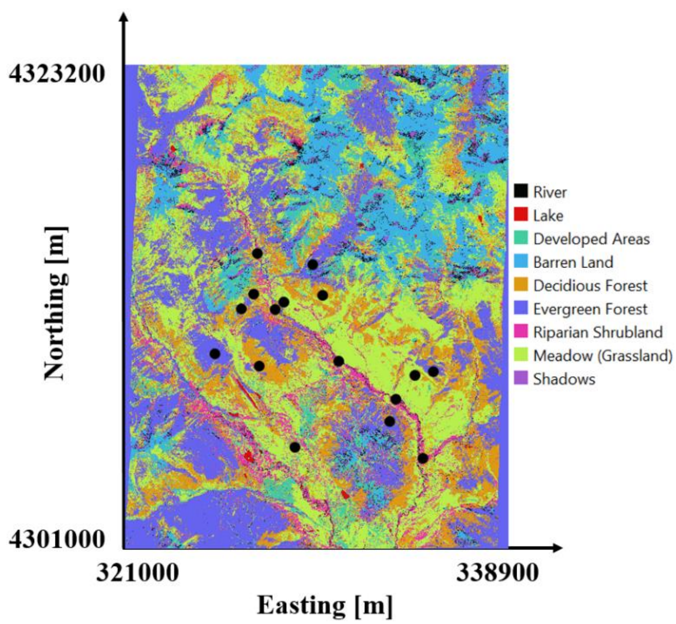

Figure 2: Vegetation classification of the East River, CO Watershed from Falco et al. (2019). East River sites selected in this study are denoted by black circles.

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, air temperature, radiation and NDVI, as inputs to the HTM model. Soil temperature was used when available. 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). 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). At the three selected SNOTEL stations, air temperature data at these three SNOTEL stations were processed following Oyler et al. (2015) and radiation data was obtained from DAYMET. CLM models were generated following Tran et al. (2019) for the SNOTEL stations and US-NR1. At the East River Watershed sites, data were obtained from DAYMET. NDVI time series were calculated from the red band and near-infrared band from Landsat 5, 7 and 8 images at all sites. 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) using deep-learning-based module (fully connected deep neural networks and long short-term memory recurrent neural networks). Long short-term memory (LSTM, Hochreiter & Schmidhuber, 1997) is a type of recurrent neural network (RNN) capable of learning temporal dependence without suffering from optimization difficulties (e.g., vanishing errors). An LSTM layer consists of memory blocks and unique cell states that are controlled by three multiplicative units, including the input, output and forget gates. These gates regulate the flow of information and decide which data in a sequence is important to keep or throw away. Through the LSTM structure, even information from the earlier time steps can make its way to later time steps, reducing the effects of short-term memory and thus capturing long-term dependence. LSTM has been previously used to capture such 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). More information about the LSTM-RNN method is provided by Olah (2015).

HPM modules include input attributes, model development, validation, and prediction. Based on data availability, input features are obtained from flux towers, gridded meteorological data, and remote sensing data; all data are preprocessed for gap filling, smoothing, and updating. In the HPM model development module, individual HPM models can be trained in two different ways based on data availability: with data obtained from flux towers ("data-driven HPM") or with outputs from physically-based models ("mechanistic HPM"). Seventy percent of these data are used for training LSTM to learn the interactions among input features, ET, and R_{eco}, until a pre-defined "stopping criteria" (e.g., root mean squared error, RMSE) is met, indicating subsequent training would lead to minimal improvement. 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.

In the validation module, we implemented a validation procedure that uses the remaining 30 % of the data to assess model performance. Estimation outputs from the trained HPM models are also compared with other ET and R_{eco} data obtained from other independent sites or mechanistic models within the same ecoregion. Statistical measures such as 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.

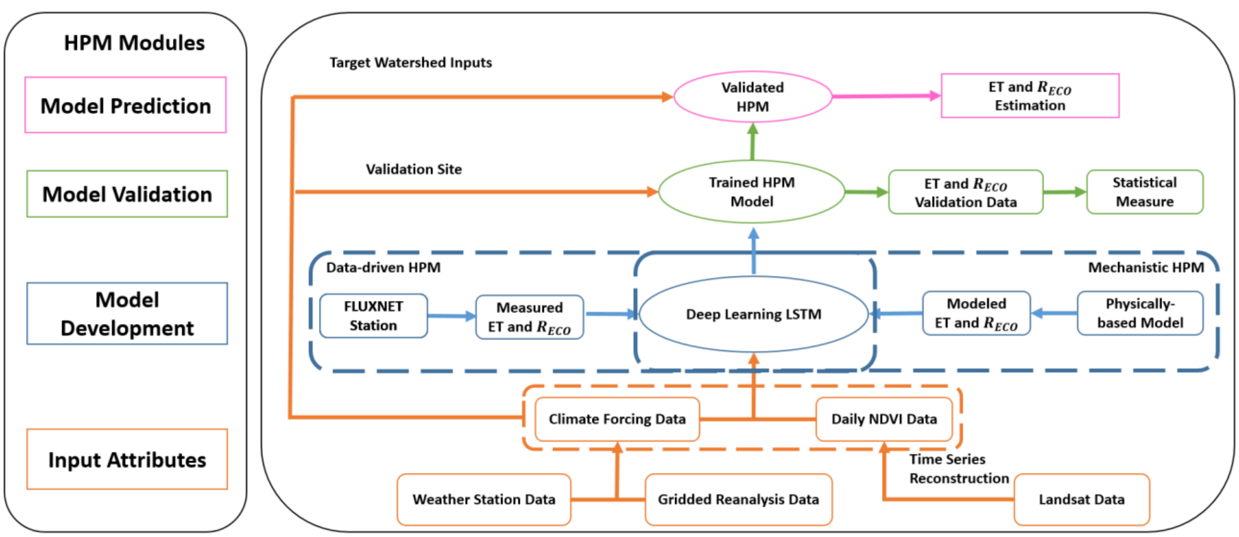

Figure 3: Hybrid Predictive Model (HPM) Framework. The HPM model mainly consists of four modules: Input Attributes, Model Development, Model Validation and Model Prediction, represented by rectangles with colors. Arrows represent the

linkages among different modules. Choices of data-driven HPM or mechanistic HPM depend on the ecoregion of target watershed and data availability.

3.2 Feature Selection

At sparsely monitored watersheds, only weather reanalysis data and remote sensing data are commonly available. Thus, we mainly considered air temperature, radiation, precipitation, vegetation indices (e.g., NDVI) and variables inferred from these data as inputs for HPM. Soil temperature when available is used at FLUXNET sites. 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 snow-influenced mountainous watersheds, we separated precipitation data into snow precipitation (when air temperature < 0) and rainfall precipitation (when air temperature > 0), which is in line with what has been used in hydrological models such as CLM (Oleson et al., 2013). Knowles et al. (2016) discovered a significant correlation between day of peak snow accumulation, snowmelt and air temperature. To capture snow related dynamics (e.g., snowmelt), we constructed a categorical variable (*sn*) based on air and soil temperature thresholds. Note: this may not be needed if snow data becomes available and at sites where snow is rarely present.

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

3.2.2 Vegetation information

We reconstructed daily NDVI values based on meteorological forcing data (e.g., air temperature, precipitation, radiation) using LSTM to increase the temporal coverage of NDVI. 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). Though not ideal, as satellites continue to advance and more training data becomes available, the accuracy of NDVI temporal reconstruction is expected to increase.

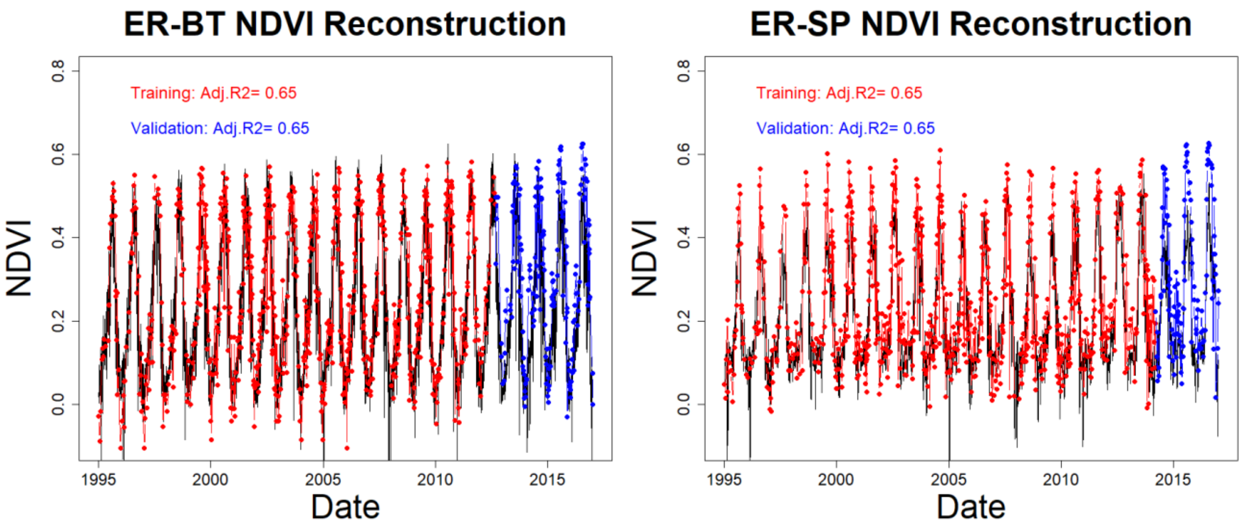

Figure 4: Temporal reconstruction of NDVI at ER-BT (left) and ER-SP (right). Black lines represent reconstructed daily NDVI. Red points are used for training and blue points are used for validation

3.3 Use Cases

We developed four different use cases to demonstrate the applicability of HPMs based on site-specific data and model availability. Use case 1 focuses on ET and R_{eco} in the time domain, where a HPM is trained on direct measurements from flux tower. A 70%-20%-10% training-validation-prediction split of the data was used. These HPMs are useful for time series gap filling and future prediction. Use case 2 and use case 3 have emphasis on providing ET and R_{eco} over space, where use case 2 uses data-driven HPM and use case 3 utilizes mechanistic HPM. Data-driven HPM is trained with data from flux tower and mechanistic HPM is trained upon outputs from a mechanistic model (e.g., CLM). These HPMs are usually trained at well monitored watersheds where either flux data is available or data support the development of a mechanistic model. After training, these HPMs integrate meteorological and remote sensing inputs to provide ET and R_{eco} at target sparsely monitored watersheds within the same ecoregion. For both use case 2 and 3, we validated the HPM estimations against data from other sites within the same ecoregion. Use case 4 focuses on the East River Watershed, where we demonstrate how HPM can increase our understanding of ecosystem fluxes and explore the limitations of HPM in mountainous watersheds. Use case 4 estimations were validated against data extracted from other studies.

4. Results

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 Fig. 5, A1-A4 and Table 3, reveal that the HPM approach was effective for estimating ET and R_{eco}. The long-term trends in ET and R_{eco} are well captured by HPM. However, short-term fluctuations in ET and R_{eco} during the summer periods are also not well captured by HPM. For example, at US-Ton and US-Var, we observed an increasing discrepancy in summer month

ET and R_{eco}. This is mainly caused by insufficient training for summer extremes. At US-Me2, we observed significant increasing errors in the validation set, especially for R_{eco} that are caused by significant differences in raw data between 2002-2010 (data used for training) and post-2011 (data used for validation).

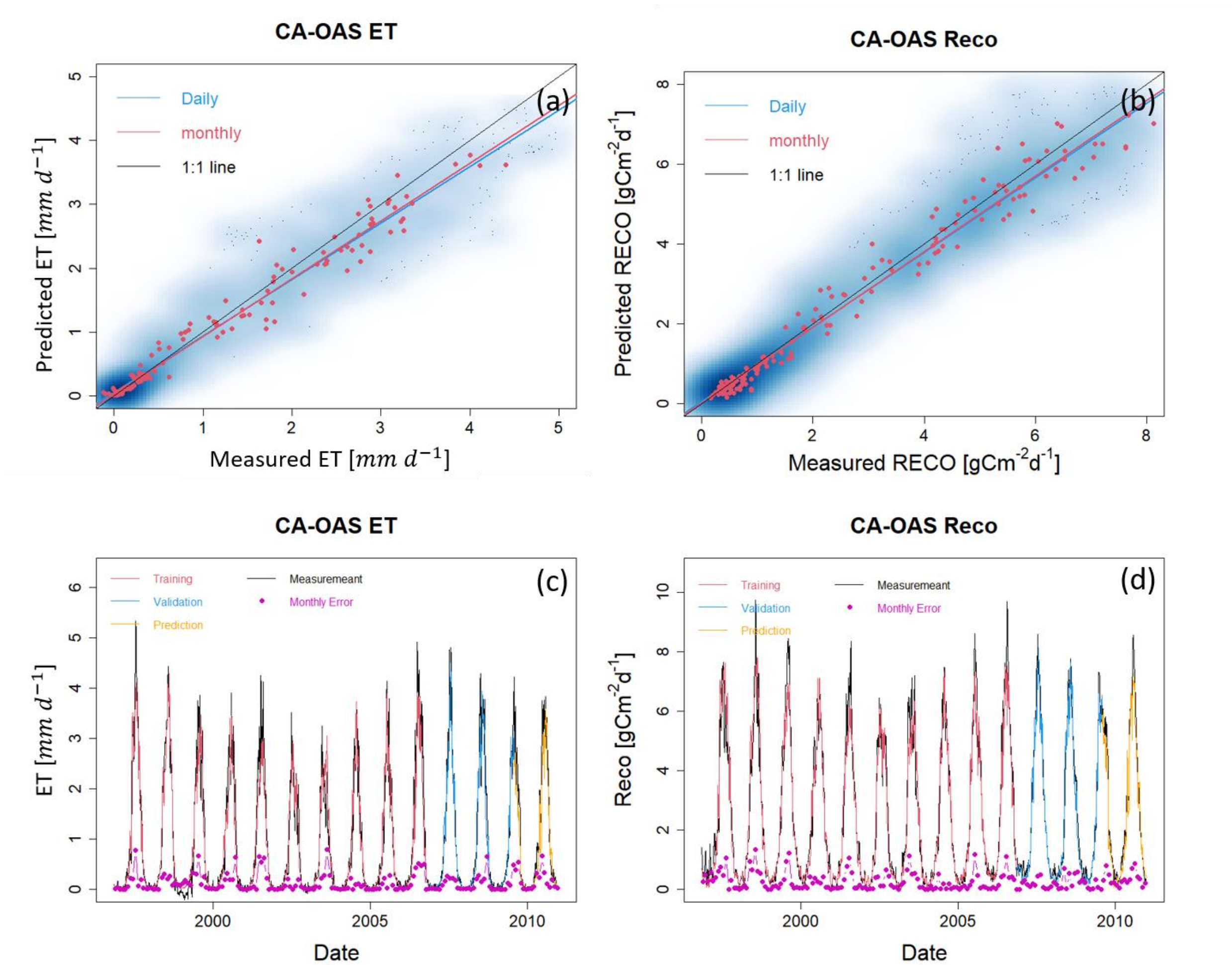

Figure 5: ET and R_{eco} estimation with data from FLUXNET sites at CA-OAS. Panels (a) and (b) show the scatter plots of daily (blue) and monthly (red) ET and R_{eco} between HPM estimation and FLUXNET data. Darker blue clouds represent greater density of data points. Panels (c) and (d) present the daily HPM estimation of ET and R_{eco} separated by training, validation and prediction sets. Pink points depict monthly error between HPM estimation and FLUXNET data. Results for other sites are included in supplementary materials below (Fig. A1, A2, A3 and A4).

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

Site ID	Train MAE -ET $[mm\ d^{-1}]$	Test MAE - ET $[mm\ d^{-1}]$	Train Adj. R^2 - ET	Test Adj. R^2 - ET	Train MAE $-\mathbf{R_{eco}}$ $[gCm^{-2}d^{-1}]$	Test MAE $-\mathbf{R_{eco}}$ $[gCm^{-2}d^{-1}]$	Train Adj. R^2 $-\mathbf{R_{eco}}$	Test Adj. R^2 $-\mathbf{R_{eco}}$
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
US-Me2	0.36	0.43	0.81	0.75	0.75	0.83	0.88	0.85

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

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. 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 6 and A5 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. 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.

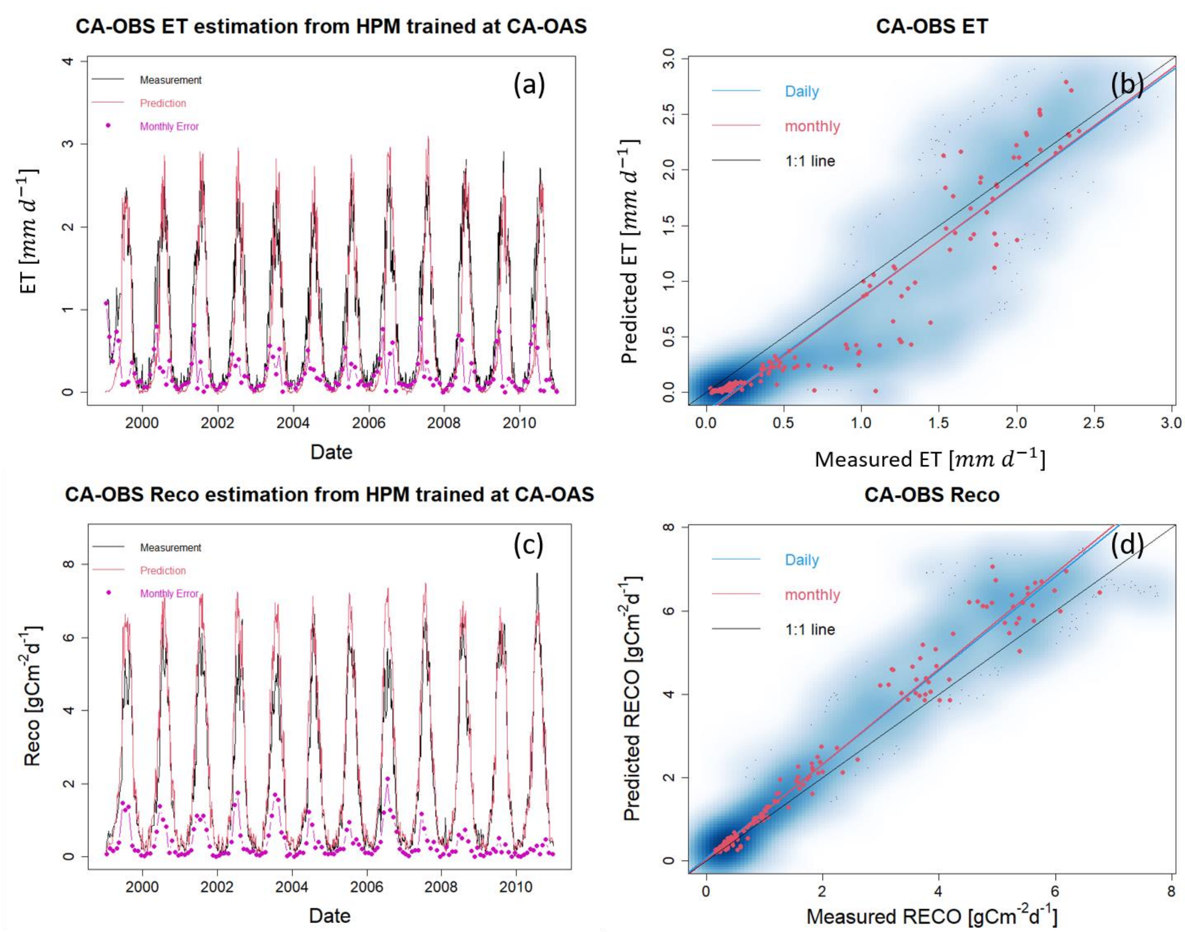

Figure 6. ET and R_{eco} estimation at CA-Obs using HPM trained at CA-Oas. Other sites are presented in Fig. A5.

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

Mechanistic HPM, which is trained upon physically-based model simulations, provides an avenue for estimating fluxes in ecoregions where flux towers are not available. Consistent results between measured ET and CLM-estimated ET at US-NR1 (adjusted $R^2 = 0.88$; $k = 0.95$, Fig. S1) indicate independent CLM simulations can be effectively used to develop the mechanistic HPM. We applied mechanistic HPM trained with 1-Dimensional (vertical) CLM developed at ER-BT (Tran et al., 2019) to estimate ET at sites classified as part of the western Cordillera 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. Figure 7 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)**$[mm\ d^{-1}]$	**ET Adj.** R^2	$\mathbf{R_{eco}}$ **MSE(monthly)**$[gCm^{-2}d^{-1}]$	$\mathbf{R_{eco}}$ **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	-	-

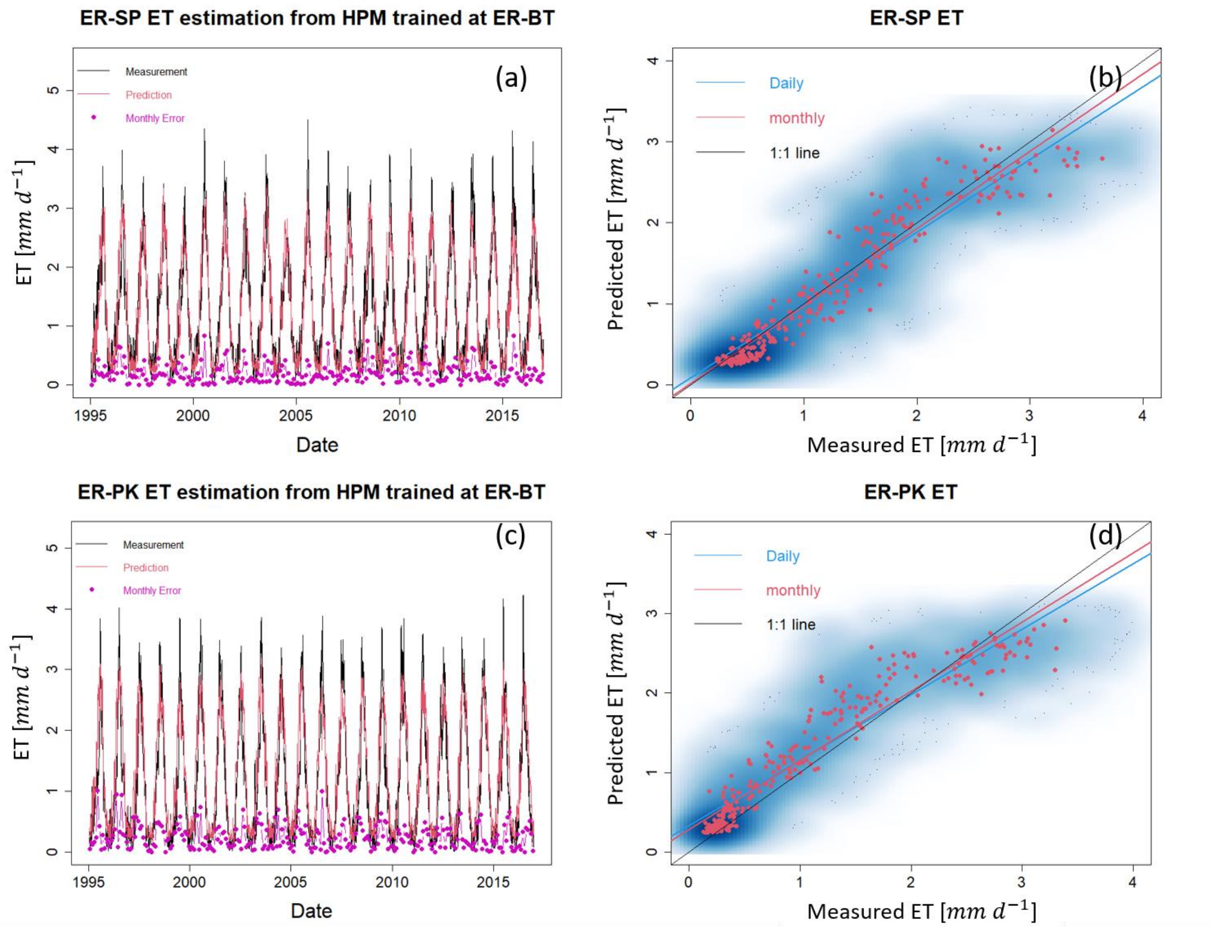

Figure 7. HPMs trained with CLM simulation at ER-BT are used to estimate ET at ER-SP and ER-PK. Panels (a) and (c) display the time series of HPM estimation of ET (red lines), and independent CLM estimation at ER-SP and ER-PK. Panels (b) and (d) show the scatter plots of daily (blue) and monthly (red) ET at these sites. Darker blue clouds represent greater density of data points.

4.4 Use Case 4: HPM approach improved our prediction capability and process understanding at the East River Watershed

With the proposed HPM approach (e.g., mechanistic HPM), we were able to estimate ET and R_{eco} at selected locations at the East River Watershed, CO, USA with only meteorological forcings and remote sensing data. Our estimations are comparable to other independent studies, such as Mu et al. (2013) (Fig. S2) and Berryman et al. (2018). HPM estimations enhanced our understanding of watershed processes and enabled us to explore the limitations in the developed HPM approach especially at mountainous watersheds.

Physiology differences among vegetation types and dynamic changes in meteorological conditions were well captured by input features and HPM at the East River Watershed. Not surprisingly, the reconstructed NDVI indicated that deciduous forests have the highest peak NDVI followed by grasslands, shrublands and evergreen forests whereas annual variation of NDVI in evergreen forests is smaller than the other vegetation types (Fig. 8). Year 2012 is regarded as a fore-summer drought year with earlier than normal snowmelt, and year 2015 is regarded as a normal water year.

The Palmer drought severity index (PDSI) is -5.2 and -1.5 for June and -4.6 and 1.1 for August in 2012 and 2015, respectively. Dynamic changes in meteorological conditions between 2012 and 2015 were also reflected in the reconstructed NDVI time series. We observed an earlier rise of NDVI in 2012: March, April and May mean NDVI values for deciduous forest sites are 0.07, 0.2 and 0.37 compared to 0.06, 0.15 and 0.33 in 2015. Similar trends were observed for other vegetation types during spring months as well. NDVI values remain high during the peak growing season (deciduous forest > grassland > shrubland > evergreen forest) for both 2012 and 2015. However, we observed NDVI declines for grasslands and shrublands since August in 2012 but not until September in 2015. During autumn periods, NDVI declines significantly following the sharp decline in radiation.

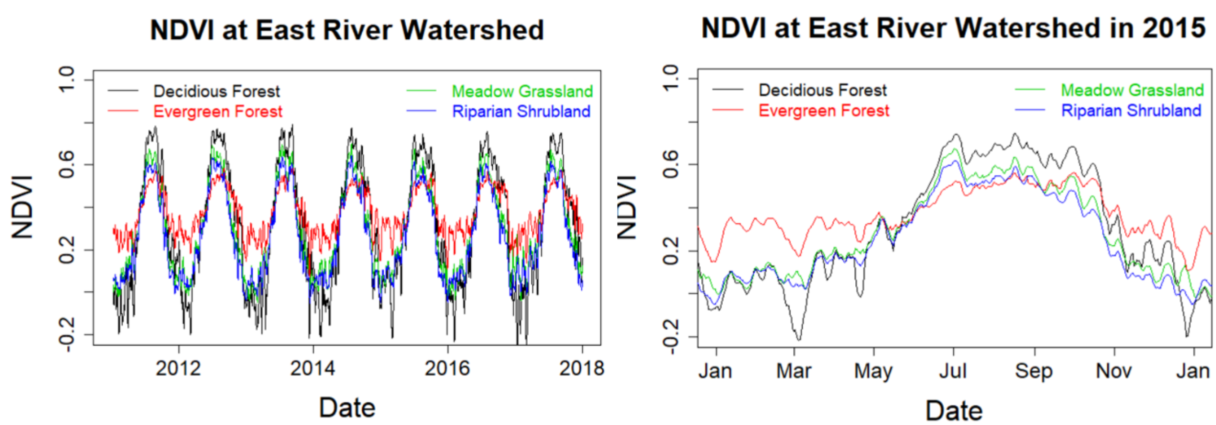

Figure 8: Reconstructed NDVI time series at selected locations in the East River Watershed for 2011 to 2018 (panel a) and for 2015 (panel b, normal water year). Black, red, green, and blue lines represent the time series of NDVI for deciduous forests, meadow grasslands, evergreen forests and riparian shrubland, respectively.

HPM-estimated ET and R_{eco} also show different dynamics with different vegetation types and meteorological conditions. Figure 9a and 9b present the time series of estimated ET and R_{eco} associated with deciduous forests, respectively. Figure 9c and 9d 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., DF: 535 to 573 mm, MS: 534 to 570 mm, RS: 532 to 567 mm and EF: 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}). These results indicate HPM R_{eco} models are sensitive to vegetation types and HPM ET models are mostly constrained by meteorological conditions.

Considering the inter-annual variability in meteorological forcings, we further selected year 2014 (large snow precipitation ~ 587 mm but small rain precipitation ~ 275 mm) in addition to 2012 (drought year) and 2015 (small snow precipitation ~ 383 mm and large rain precipitation ~ 477 mm) to test HPM performance. As HPM does not have the capability to identify snow and monsoon precipitation's contribution to fluxes, we separated annual ET and R_{eco} into pre-June (January-June) and post-July (July-December) to quantify the contribution from snow and monsoon. Earlier snowmelt that occurred in 2012 boosted spring ET and R_{eco} and we observed larger March-mean ET and R_{eco} compared to 2014 and 2015 that are characterized by later snowmelt. Occurrences of fore-summer drought in 2012 led to moisture limiting conditions, resulting in large fluctuations of ET and R_{eco} during May and June. ET fluctuated from 2.9 to 1.9 $mm\ d^{-1}$ during late May, and 3.53 to 2.6 $mm\ d^{-1}$ during early June. However, early occurrence of monsoon in 2012 led to a peak ET in early July. Due to late snowmelt, ET did not significantly fluctuate in 2014 and 2015. However, peak ET shifted towards late July in 2014. Regarding R_{eco} dynamics, fore-summer drought conditions led to variations in R_{eco} from ~ 4 to 6 $gCm^{-2}\ d^{-1}$ in 2012. In 2014, we observed more steady increase of R_{eco} during the early and peak growing seasons. For late-summer and autumn months (August – October), ET decreased steadily in all three years regardless of monsoon precipitation inputs, following the significant decline in radiation. Pre-June ET and R_{eco} (255mm and 217 $gCm^{-2}\ d^{-1}$) were both greater in 2012 compared to 2014 (223 mm and 178 $gCm^{-2}\ d^{-1}$) and 2015 (230 mm and 197 $gCm^{-2}\ d^{-1}$) in deciduous forests. While there were no significant differences in post-July ET among the three years (318, 316 and 306 mm), 2012 was the highest. Within deciduous forests and annually over 2012, 2014 and 2015, ET was 573 mm, 539 mm and 536 mm and R_{eco} was 698 gCm^{-2}, 642 gCm^{-2} and 652 gCm^{-2}, respectively. . Similar trends were observed for other vegetation types.

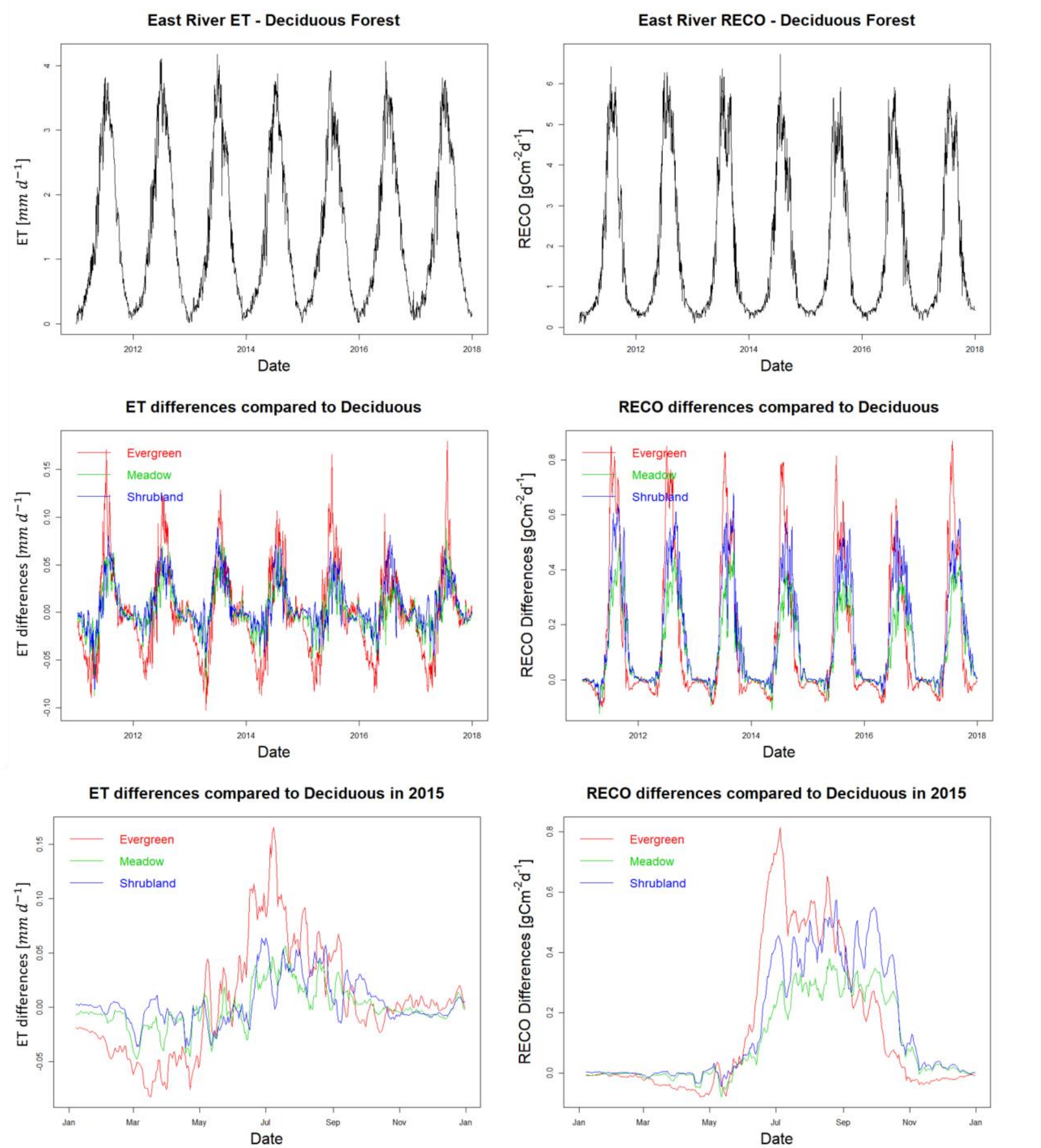

Figure 9: ET (a) and R_{eco} (b) estimation for the deciduous forest site DF1 at the East River Watershed. Panels (c) and (d) show the differences in ET and R_{eco} among various vegetation types and deciduous forest. Red, green, and blue lines represent the differences in evergreen forest, meadow, and riparian shrubland compared to deciduous forest. Panels (e) and (f) zoom into 2015 to better display seasonal variations.

Though HPM estimations allowed us to explore differences in ET and R_{eco} across vegetation types and meteorological forcings heterogeneity, it is necessary to investigate the limitations of HPM approach. Figure 10 shows the absolute value of monthly mean difference in ET (Fig. 10a and Fig. 10b) and R_{eco} (Fig. 10c and Fig. 10d) across

SNOTEL stations (ER-BT, ER-SP and ER-PK) and within selected East River locations. We observed greater differences in air temperature and radiation at the SNOTEL sites and very small differences at the East River sites (Figure S4). June air temperature differences among SNOTEL sites were occasionally over 3 ℃, while the DAYMET data from the East River rarely revealed 0.2 ℃ differences. In addition, a ~80 $W\ m^{-2}$ of radiation differences was 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. Landsat data enabled us to capture NDVI differences at these sites, but we have identified the insufficient resolution of input meteorological forcing data at the East River sites. These results indicate uncertainties in meteorological forcing attributes (e.g., radiation and air temperature) can have a huge influence over HPM ET estimation and HPM R_{eco} model is more sensitive to NDVI datasets. 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.

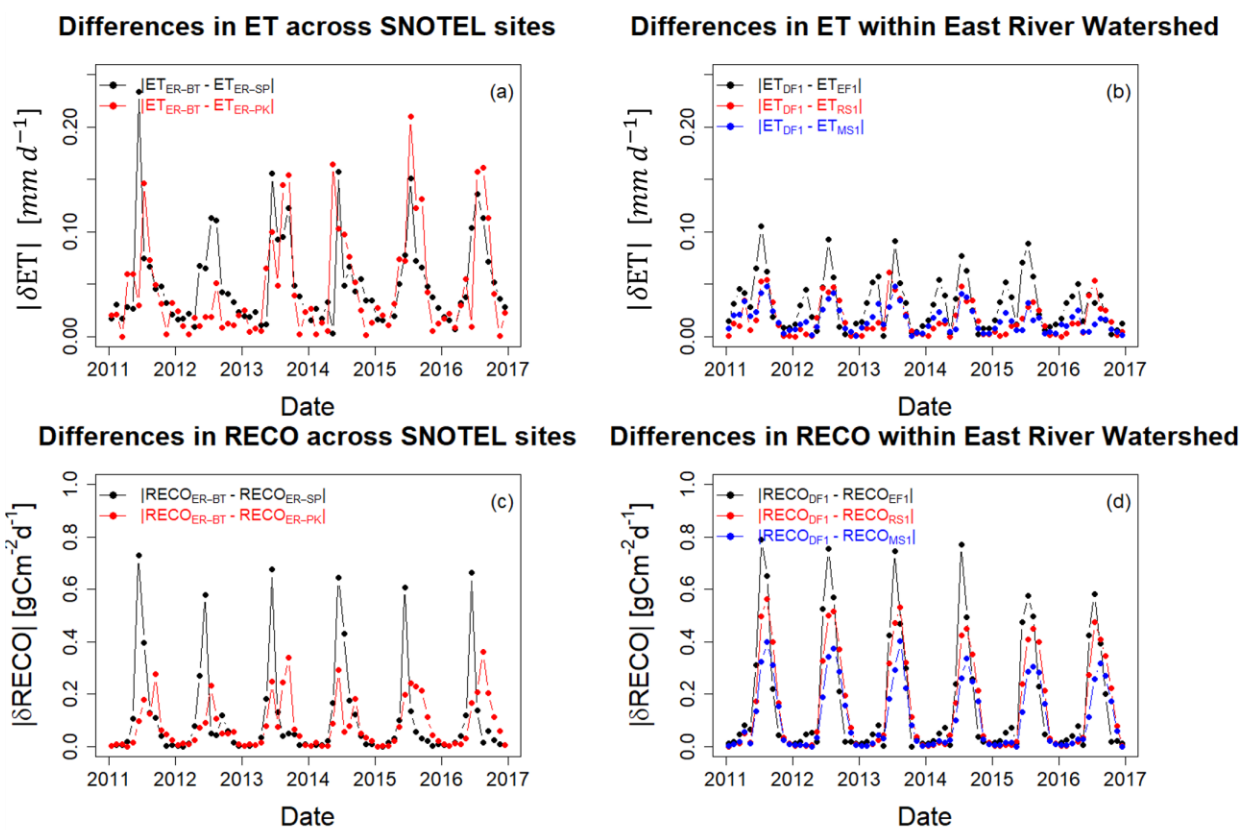

Figure 10. Absolute differences in monthly mean ET and R_{eco} across SNOTEL stations and within East River Watershed. Panels (a) and (c) describe the absolute differences in monthly mean ET and R_{eco} between ER-BT, ER-SP, and ER-PK. Panels (b) and (d) describe the absolute differences in monthly mean ET and R_{eco} within East River Watershed between deciduous forests, evergreen forests, meadow grasslands, and riparian shrublands.

5. Discussion

Our study demonstrates that HPM provides reliable estimations of ET and R_{eco} under various climate and vegetation conditions. The unique gated structures and cell states of LSTM allow HPM to track information from earlier times and decide which information to pass along and which information to forget. This effective configuration allows LSTM to effectively capture the long-term dependencies and ecological memory effects among meteorological forcings, NDVI, ET and R_{eco}. With 70 % of the data used for training (model development), ET and R_{eco} estimation from HPM achieves an average adjusted R^2 of 0.9 compared to flux tower measurements. To demonstrate HPM's applicability for providing ET and R_{eco} estimation at sparsely monitored watersheds, we presented four use cases, including prediction ET and R_{eco} in the time domain, data-driven HPMs and mechanistic HPMs. Results from the four use cases suggest HPM is a powerful approach to estimate ET and R_{eco} at target watersheds requiring only 5 commonly available input data and can advance our understanding of watershed processes.

HPM was capable incorporating information from NDVI time series to delineate the physiological differences among deciduous forests, evergreen forests, shrublands and grasslands. In our study, NDVI data indicated evergreen forests have a longer growing season compared to other vegetation types and 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). 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. 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. Hu et al. (2010) found that extended growing season length resulted in less annual CO_2 uptake at Niwot Ridge, USA. They found increasing growing season length is usually correlated with decreasing snow water storage and decreasing forest carbon uptake. Xu et al. (2020) suggested canopy photosynthetic capacity is the driving force that lead to different resources use efficiencies (RUEs) between deciduous forests and evergreen forests. Novick et al. (2015) focused on the net ecosystem exchange of CO_2 and also suggested seasonality is less important for evergreen forests, where significant amounts of carbon were assimilated outside of active season. These findings are similar to what we found in HPM estimations, where we observed a greater ET and R_{eco} contribution during early and later seasons for evergreen forests compared to deciduous forests that have significantly greater peak ET and R_{eco} during peak growing season. As HPM only requires 5 input features and NDVI is the only variable related with vegetation types, we were not able to perform detailed analysis delinearing the physiological control on ET and R_{eco} dynamics. But we believe HPM models are still useful as they can be provide initial ET and R_{eco} estimation that help with site selection and field campaign designs.

Temporal variability in meteorological conditions also leads to unique ET and R_{eco} responses at the East River Watershed, as shown by HPM estimations. Three years with a diverse combination of snow and rain precipitation were analyzed. In 2012, a year that experienced earlier snowmelt, both ET and R_{eco} increased early in the season. However, earlier growth in vegetation and increasing demand for water resulted in fore-summer drought conditions that led to decreases in ET and R_{eco} in late May and June. In 2014, HPM estimated a steady increase in ET and R_{eco} during spring months following radiation and air temperature trends, with no subsequent significant decline in ET and R_{eco}. This indicates that energy was still the key limiting factor for spring dynamics in 2014, leading to a smaller pre-June ET and R_{eco} compared to 2012. Following an earlier arrival of monsoon in 2012 compared to 2014 and 2015, we observed higher mean ET and R_{eco} in July than in June, which indicates the earlier arrival of monsoon precipitation greatly reduced the moisture limiting condition caused by fore-summer drought and led to subsequent increase in ET and R_{eco}. During late summer and autumn months, radiation declined significantly with ~ 30 % decrease in August and ~ 40 % decrease in September. Though 2012, 2014 and 2015 had diverse monsoon precipitation during these periods, HPM did not estimate significant differences in post-July ET. This result indicates the East River watershed is mainly under energy-limiting rather than moisture-limiting conditions during late-summer and autumn; and timing of monsoon arrival is more important than the absolute amount of monsoon precipitation for ET dynamics. This result is consistent with findings in Carroll et al. (2020). Their study also indicated earlier arrival of summer monsoon was effectively supporting ET and that the monsoon precipitation was quickly consumed by vegetation, whereas later arrival of summer monsoon water mainly contributed to streamflow under energy-limiting conditions.

Uncertainties of HPM models arise from several aspects. First, current choices of only five input features based on data availability may decrease estimation accuracy in certain environments, such as sites with seasonally dry periods. Though the LSTM component within HPMs can capture the memory effects and long-term dependencies of watershed dynamics, rare extreme values are difficult to be captured by LSTM due to insufficient training data for such cases. For example, we observed a decreasing prediction accuracy for ET and R_{eco} estimation at sites that experience drought conditions. Current use of meteorological forcings data and NDVI may not provide sufficient data for LSTM to identify droughts implicitly. Other key variables (e.g., soil moisture) when available can potentially be useful to help LSTM better quantify these rare events and increase model performance. Secondly, parameterization and insufficient spatiotemporal resolution of meteorological data still remain a challenge. 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). However, we did not observe same level of heterogeneities in radiation and air temperature in reanalysis data compared to weather station data (Fig. S4 and S5). Mu et al. (2013) and Zhang et al. (2019) suggested uncertainties in meteorological inputs can result in large errors (i.e., > 20 % MAE) and reduce accuracy by 10 – 30 %. Additionally, HPM is also influenced by remote sensing inputs accuracy, including but not limited to insufficient resolution, cloud conditions, spatial averaging, temporal reconstruction, any other algorithms involved. 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). Finally, errors can stem from the HPM hybrid approaches and conceptual model uncertainties. Any original errors in mechanistic models will be passed onto HPM estimations of ET and R_{eco}. We recommend to train data-driven HPM and mechanistic HPM using long time series (e.g., > 5 years) with high quality data or simulations, which enables HPMs to better memorize long-term dependencies of ecosystem dynamics. Though some of the uncertainties still remain a challenge, efforts have been made to minimize them through the technical advances described herein. Future HPM models can potentially be jointly trained on FLUXNET and process-based simulations to bypass certain limitations and provide more accurate ET and R_{eco} at sparsely monitored watersheds.

6. Conclusion

In this study, we developed and tested a Hybrid Predictive Modeling approach for ET and R_{eco} estimation, with an enhanced focus on a watershed in the Rocky Mountains. We developed individual HPM models at various FLUXNET sites and at sites where data could support 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 and meteorological conditions within the East River Watershed of CO, USA.

Given the promising results of HPM, the approach 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. HPM also has great potential for estimating other parameters important for water and carbon cycles given the right choice of input variables, such as net ecosystem exchange (Figure A6). Thus, we believe the proposed HPM model can improve our prediction capabilities of ET and R_{eco} at sparsely monitored watersheds and advance our understanding of watershed dynamics.

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.

Acknowledgements. 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.

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

Carroll, R. W. H., Gochis, D. and Williams, K. H.: Efficiency of the Summer Monsoon in Generating Streamflow Within a Snow-Dominated Headwater Basin of the Colorado River, Geophys. Res. Lett., 47(23), doi:10.1029/2020GL090856, 2020.

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.

Chu, H., Luo, X., Ouyang, Z., Chan, W. S., Dengel, S., Biraud, S. C., Torn, M. S., Metzger, S., Kumar, J., Arain, M. A., Arkebauer, T. J., Baldocchi, D., Bernacchi, C., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Bracho, R., Brown, S., Brunsell, N. A., Chen, J., Chen, X., Clark, K., Desai, A. R., Duman, T., Durden, D., Fares, S., Forbrich, I., Gamon, J. A., Gough, C. M., Griffis, T., Helbig, M., Hollinger, D., Humphreys, E., Ikawa, H., Iwata, H., Ju, Y., Knowles, J. F., Knox, S. H., Kobayashi, H., Kolb, T., Law, B., Lee, X., Litvak, M., Liu, H., Munger, J. W., Noormets, A., Novick, K., Oberbauer, S. F., Oechel, W., Oikawa, P., Papuga, S. A., Pendall, E., Prajapati, P., Prueger, J., Quinton, W. L., Richardson, A. D., Russell, E. S., Scott, R. L., Starr, G., Staebler, R., Stoy, P. C., Stuart-Haëntjens, E., Sonnentag, O., Sullivan, R. C., Suyker, A., Ueyama, M., Vargas, R., Wood, J. D. and Zona, D.: Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites, Agric. For. Meteorol., 301–302(February), doi:10.1016/j.agrformet.2021.108350, 2021.

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.

Kakalia, Z., Varadharajan, C., Alper, E., Brodie, E., Burrus, M., Carroll, R., Christianson, D., Hendrix, V., Henderson, M., Hubbard, S., Johnson, D., Versteeg, R., Williams, K. and Agarwal, D.: The East River Community Observatory Data Collection: Diverse, multiscale data from a mountainous watershed in the East River, Colorado, , 1–17 [online] Available from: https://doi.org/10.22541/au.160157556.64095872, 2020.

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.

Novick, K. A., Oishi, A. C., Ward, E. J., Siqueira, M. B. S., Juang, J. Y. and Stoy, P. C.: On the difference in the net ecosystem exchange of CO_2 between deciduous and evergreen forests in the southeastern United States, Glob. Chang. Biol., 21(2), doi:10.1111/gcb.12723, 2015.

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.

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.

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.

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.

Visser, A., Thaw, M., Deinhart, A., Bibby, R., Safeeq, M., Conklin, M., Esser, B. and Van der Velde, Y.: Cosmogenic Isotopes Unravel the Hydrochronology and Water Storage Dynamics of the Southern Sierra Critical Zone, Water Resour. Res., doi:10.1029/2018WR023665, 2019.

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.

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, 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, H., Xiao, J., Zhang, Z., Ollinger, S. V., Hollinger, D. Y., Pan, Y. and Wan, J.: Canopy photosynthetic capacity drives contrasting age dynamics of resource use efficiencies between mature temperate evergreen and deciduous forests, Glob. Chang. Biol., 26(11), doi:10.1111/gcb.15312, 2020.

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

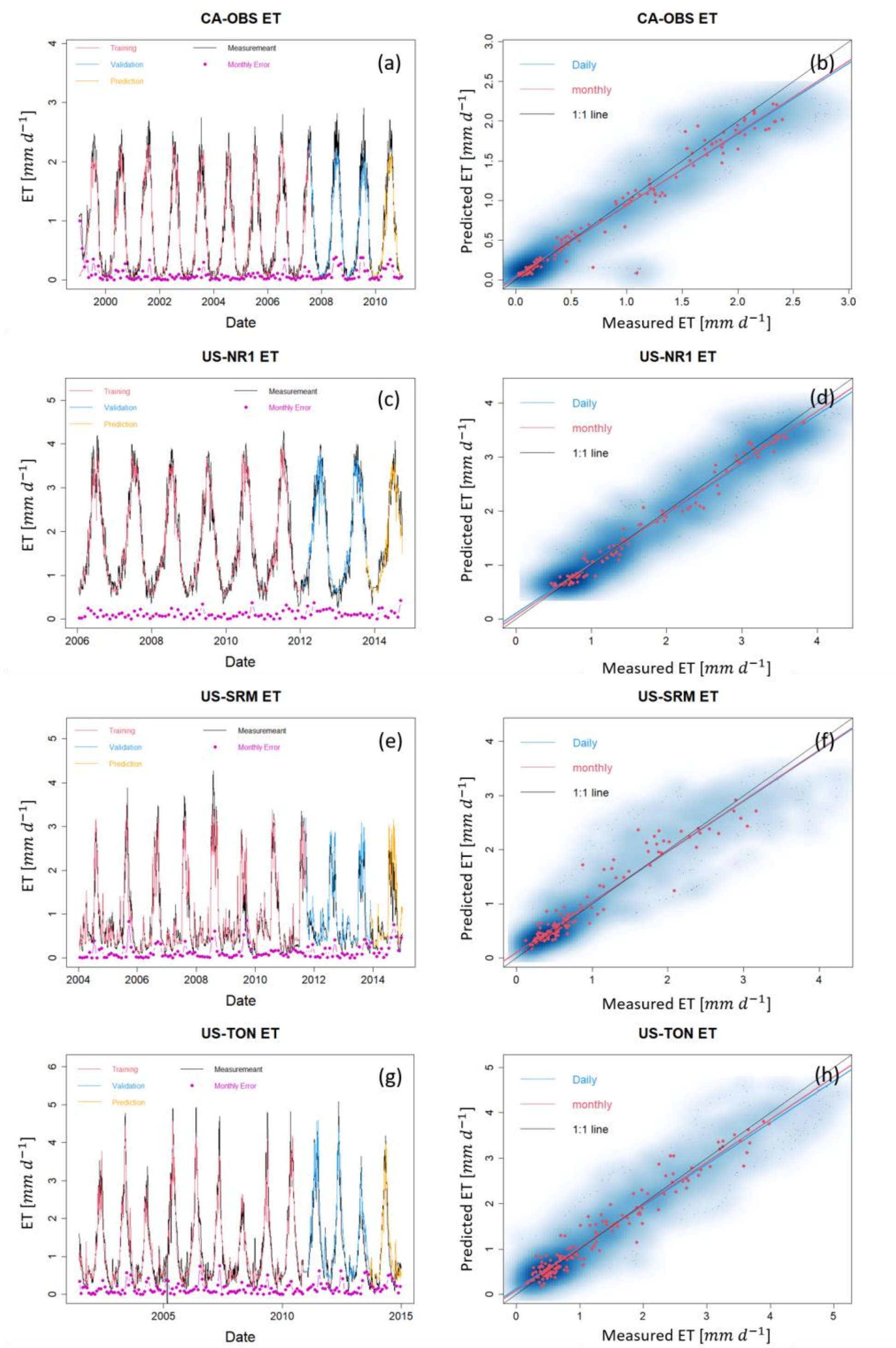

Figure A1: ET estimation with data from selected FLUXNET sites at CA-OBS, US-NR1, US-SRM, and US-Ton. Panels (a), (c), (e) and (g) present daily estimations of ET separated for training, validation, and prediction. Pink points depict monthly error between HPM estimation and FLUXNET data. Panels (b), (d), (f) and (h) show the scatter plots of daily (blue) and monthly (red) ET. Darker blue clouds represent greater density of data points.

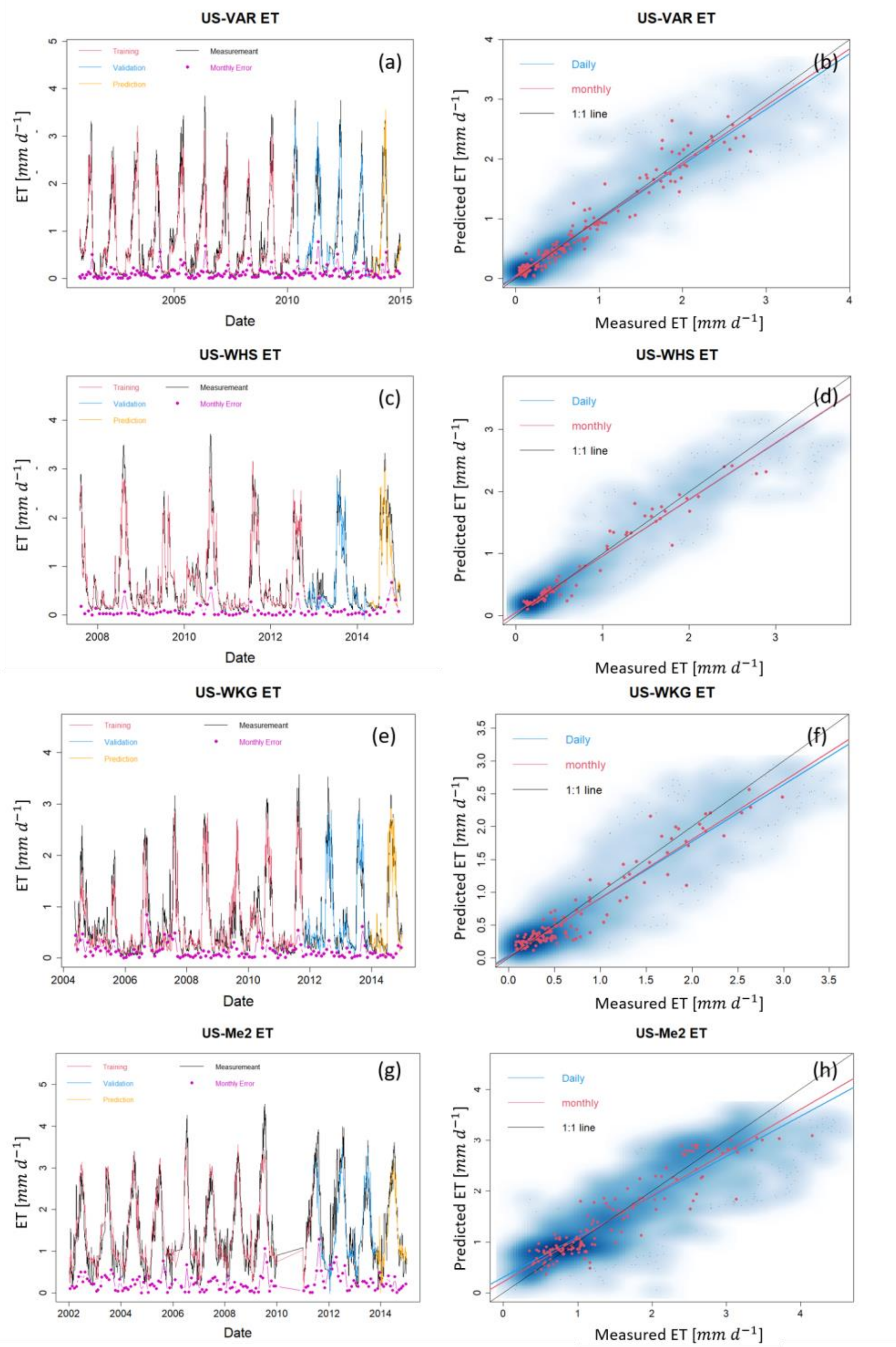

Figure A2: ET estimation with data from selected FLUXNET sites at US-Var, US-Whs, US-Wkg and US-Me2. Panels (a), (c), (e) and (g) present daily estimations of ET separated for training, validation, and prediction. Pink points depict monthly error between HPM estimation and FLUXNET data. Panels (b), (d), (f) and (h) show the scatter plots of daily (blue) and monthly (red) ET. Darker blue clouds represent greater density of data points.

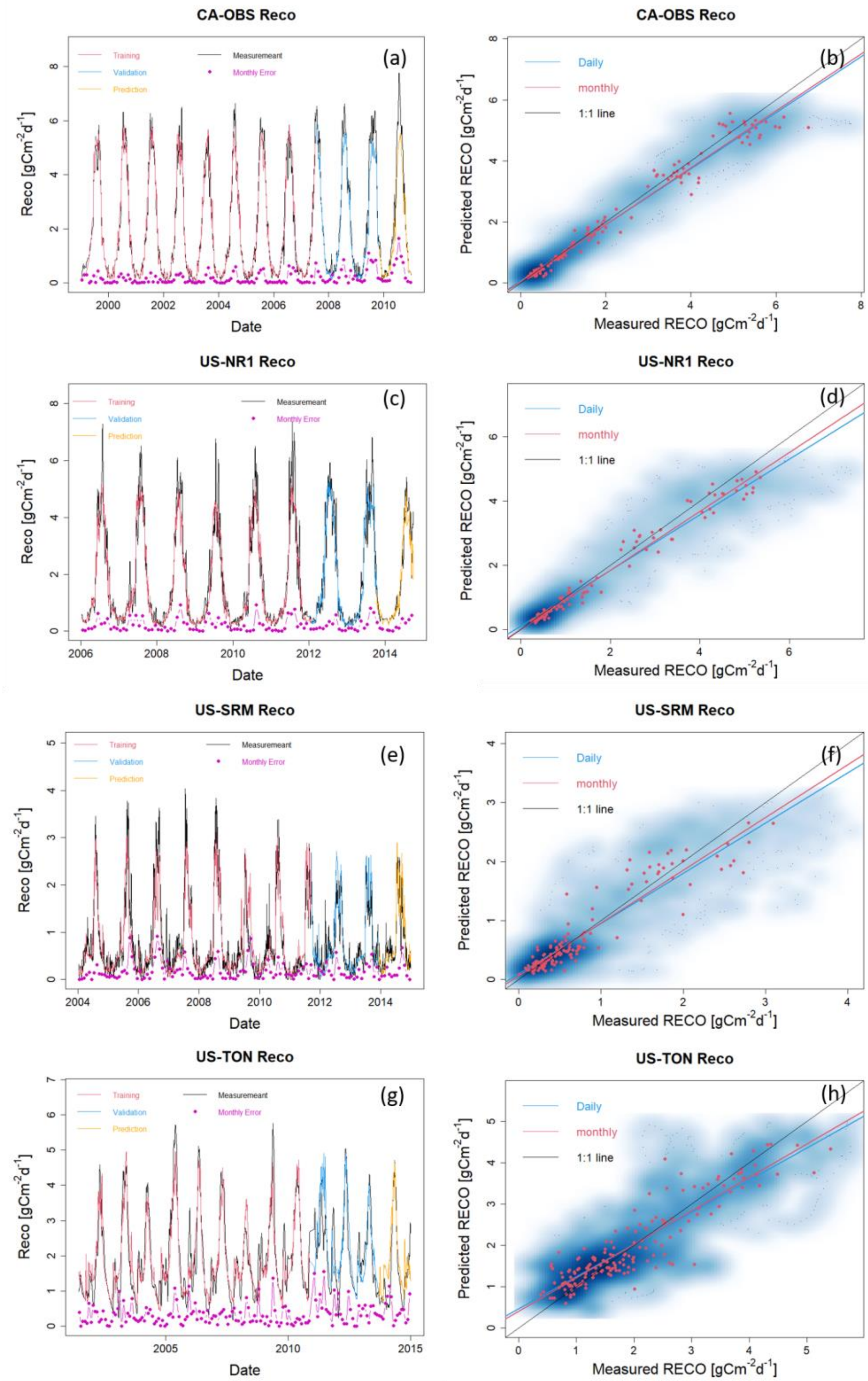

Figure A3: R_{eco} estimation with data from selected FLUXNET sites at CA-OBS, US-NR1, US-SRM, and US-Ton. Panels (a), (c), (e) and (g) present daily estimations of R_{eco} separated for training, validation, and prediction. Pink points depict monthly error between HPM estimation and FLUXNET data. Panels (b), (d), (f) and (h) show the scatter plots of daily (blue) and monthly (red) R_{eco}. Darker blue clouds represent greater density of data points.

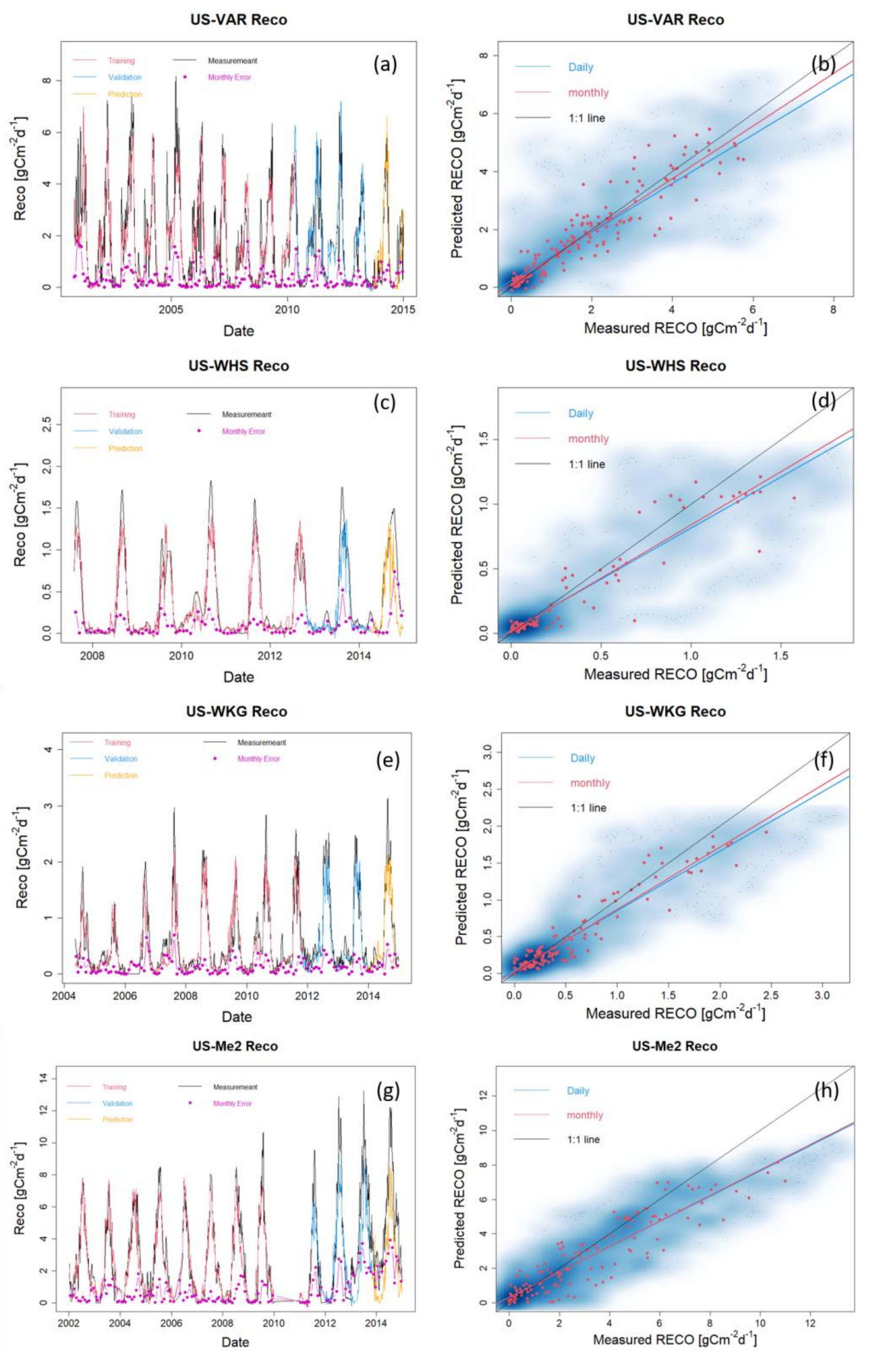

Figure A4: R_{eco} estimation with data from selected FLUXNET sites at US-Var, US-Whs, US-Wkg and US-Me2. Panels (a), (c), (e) and (g) present daily estimations of R_{eco} separated for training, validation, and prediction. Pink points depict monthly error between HPM estimation and FLUXNET data. Panels (b), (d), (f) and (h) show the scatter plots of daily (blue) and monthly (red) R_{eco}. Darker blue clouds represent greater density of data points.

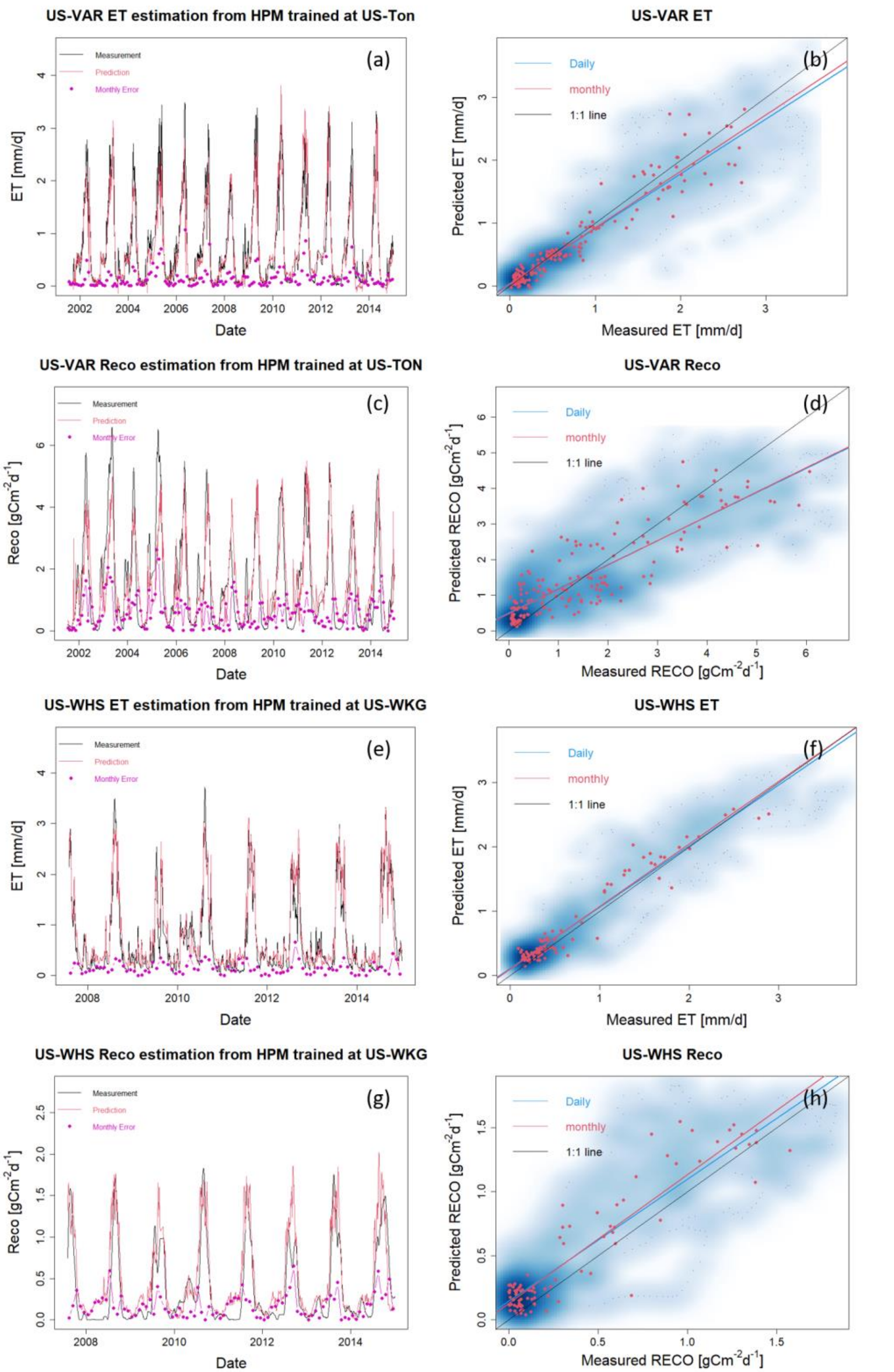

Figure A5: Use case 2. ET and R_{eco} estimation at US-Var and US-Whs from HPM trained at US-Ton and US-Wky, respectively.

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

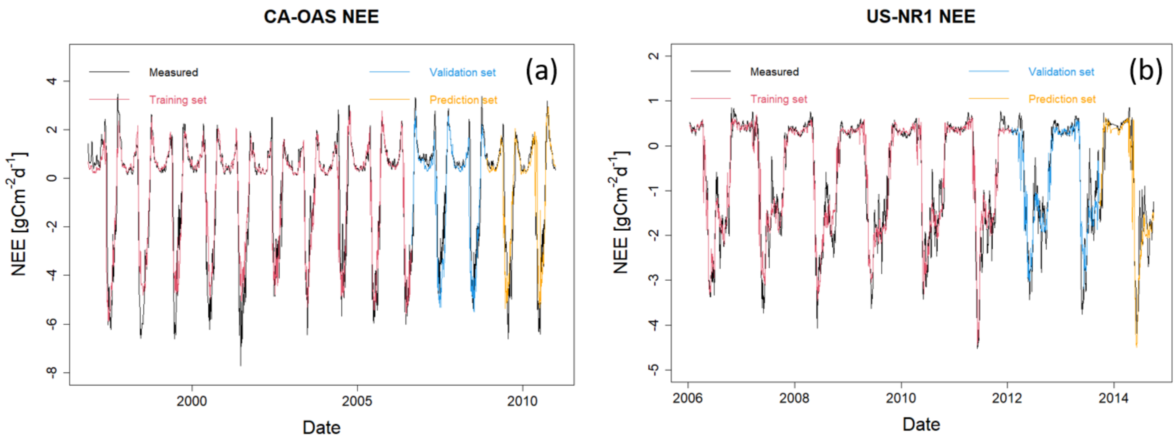

Figure A6. HPM estimate of NEE at CA-OAS and US-NR1. R^2 between estimation and measurements are 0.87, 0.83 and 0.81 at CA-OAS; 0.94, 0.88 and 0.90 at US-NR1 for the training set, validation set and prediction set, respectively. Model inputs include air temperature, soil temperature, sn, precipitation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