



# A novel hybrid model for actual evapotranspiration estimation in data-scarce arid regions: Integrating modified Budyko and machine learning models using deep learning



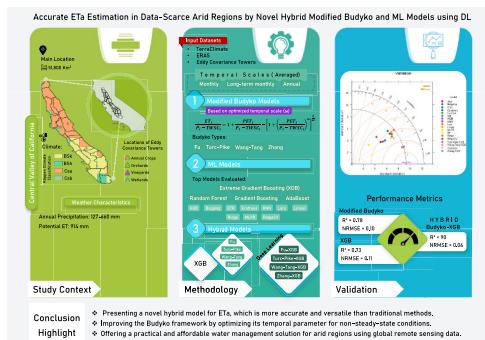
Mahdi Mohammadnezhad <sup>a</sup> , Kamran Davary <sup>a</sup> , Pooya Shirazi <sup>a</sup>, Mohammad Javad Rezvanpour <sup>a</sup>, Seyed Majid Hasheminia <sup>a,\*</sup>

<sup>a</sup> Water Science and Engineering Department, College of Agriculture, Ferdowsi University of Mashhad, P. O. Box: 9177948974, Mashhad, Iran

## HIGHLIGHTS

- Novel hybrid Budyko-ML models using deep learning were developed for accurate ET<sub>a</sub> estimation in data-scarce arid regions.
- The hybrid model significantly outperformed both modified Budyko and pure ML models with superior predictive accuracy.
- Methodological innovation by optimizing the temporal scale of Budyko's parameter for non-steady-state hydrological conditions.
- A scalable and cost-effective solution for water management using globally available remote sensing data.

## GRAPHICAL ABSTRACT



## ARTICLE INFO

**Keywords:**  
 Actual evapotranspiration  
 Central valley  
 Climatic data scarcity  
 Eddy covariance towers  
 Hybrid models  
 Machine learning models  
 Modified Budyko model

## ABSTRACT

Accurate and reliable estimation of actual evapotranspiration (ET<sub>a</sub>) is fundamental for hydrological modeling and effective water management. However, this remains a significant challenge due to the complex interplay between climate, soil, and vegetation, especially in data-scarce regions. This study introduces a novel hybrid approach to estimate monthly ET<sub>a</sub> by integrating a modified Budyko framework with an optimized machine learning model (XGBoost among thirteen other models) using deep learning for a watershed under non-steady-state conditions in California's Central Valley. Remote sensing data from ERA5 and TerraClimate datasets were utilized as primary inputs, and Eddy Covariance Towers data as observed data. While the standalone modified Budyko model, particularly the Zhang equation, provided a suitable estimation of ET<sub>a</sub>, the hybrid model consistently and significantly outperformed all standalone models. This demonstrates the superior predictive capability of the hybrid approach, which successfully mitigates the inherent weaknesses of both conceptual/physical and pure data-driven models. The methodological innovation of optimizing the Budyko parameter's temporal scale also improved performance by accounting for non-steady-state conditions of the watershed. A feature importance analysis using SHAP values highlighted the climate, soil, and vegetation indices as primary

\* Corresponding author at: Irrigation & Drainage Engineering, Water Science and Engineering Department, College of Agriculture, Ferdowsi University of Mashhad, Iran.

E-mail addresses: [mohammadnezhad@alumni.um.ac.ir](mailto:mohammadnezhad@alumni.um.ac.ir) (M. Mohammadnezhad), [k.davary@um.ac.ir](mailto:k.davary@um.ac.ir) (K. Davary), [pooya.shirazi.alian@alumni.um.ac.ir](mailto:pooya.shirazi.alian@alumni.um.ac.ir) (P. Shirazi), [rezvanpour@alumni.um.ac.ir](mailto:rezvanpour@alumni.um.ac.ir) (M.J. Rezvanpour), [hasheminia@ferdowsi.um.ac.ir](mailto:hasheminia@ferdowsi.um.ac.ir) (S.M. Hasheminia).

drivers of ET<sub>a</sub>. This research presents a scalable and cost-effective solution that leverages globally available data, making it a pragmatic and universally applicable tool for sustainable water management in developing countries and any data-scarce regions worldwide.

## 1. Introduction

Actual Evapotranspiration (ET<sub>a</sub>) in arid and semi-arid environments is a major component of the hydrological cycle and one of the most important physical processes of the land surface (Awada et al., 2022). In agriculture, quantifying the spatial variability and the temporal dynamics of ET<sub>a</sub> at the field or larger spatial scales is of great importance for irrigation scheduling, identifying water productivity, and water use efficiency (Molden et al., 2010). It is essential to accurately assess evapotranspiration to avoid excessive and insufficient irrigation, ensuring the sustainable utilization of water resources while meeting agricultural demands (Cutting et al., 2024). ET<sub>a</sub> is also a vital component in water resources management. It represents the actual amount of water transferred from land surfaces to the atmosphere through both evaporation and plant transpiration (Wang and Dickinson, 2012). The accurate quantification of ET<sub>a</sub> is essential because it directly influences the water balance of regions that are under water stress conditions, determining how much water is available for agricultural uses, ecosystem sustainability, groundwater recharge, and surface water flow (Hao et al., 2022). Pursuing this goal involves employing models for measuring and predicting evapotranspiration rates or utilizing advanced instruments for direct measurements (Cutting et al., 2024). However, the complex nature and high costs associated with direct measurements have led to the development of estimation models designed for versatile applications across different contexts (Ghiat et al., 2021). In general, ET estimation methods fall into five primary categories: 1. pan evaporation-based, 2. temperature-based, 3. mass transfer-based, 4. radiation-based, and 5. combined methods (Gharbia et al., 2018). Nevertheless, these models differ in assumptions, data requirements, complexity, and reliability (Xiang et al., 2020). The ET<sub>a</sub> estimations based on the energy balance equation have proven to be highly accurate and efficient (Costa et al., 2023) which was used in this study.

The Budyko model is a simple tool widely used within the hydrological community (Moussa and Lhomme, 2016). This model relates actual evapotranspiration (ET<sub>a</sub>) to precipitation (P) through an aridity index (PET/P), where PET represents potential evapotranspiration. The Budyko model uses long-term averages of these variables and simple mathematical equations to make these estimates (Barkhordari et al., 2025). According to Du et al. (2016) and Xiong et al. (2014), the Budyko hypothesis (BH) is useful for examining the impacts of climate change on mean annual runoff and evapotranspiration. Schreiber (1904), Ol'dekop (1911), Pike (1964), and Budyko (1984) have suggested certain empirical equations without parameters. However, these equations do not address recently identified significant catchment properties, such as groundwater system characteristics, vadose zone properties, and surface vegetation, as they consider only climate fluctuations (radiation, precipitation, evapotranspiration, and air temperature) (Du et al., 2016). Efforts have been made to incorporate physical parameters into the empirical equations of the BH, as seen in studies by Fu (1981), Zhang et al. (2001), Yang et al. (2007), and Yang et al. (2008). These physical parameters are challenging to assess because they include various catchment properties (soil, surface vegetation, groundwater, topography, etc.). Nevertheless, the one-parameter equations of the Budyko model offer the flexibility to apply the BH over long periods. Budyko assumed that the mean period of evapotranspiration is influenced by water availability (approximated by precipitation) and atmospheric demand (represented by net radiation).

The Budyko Hypothesis assumes steady-state conditions, which require that the basin is natural and closed, meaning local precipitation is the only source of water for evapotranspiration (Mianabadi et al.,

2020). The BH has been widely used in recent studies to explore interannual variability in precipitation partitioning (Gerrits et al., 2009), trends in runoff (Li et al., 2014; Xiong et al., 2014), changes in evapotranspiration (Sun et al., 2025), and changes in water storage (Istanbulluoglu et al., 2012; Gao et al., 2014). Studies indicate that climate change and human-induced land cover alteration influence the hydrological systems significantly (Barkhordari et al., 2025). Activities such as deforestation, urbanization, groundwater withdrawal, and constructing hydraulic structures disrupt the natural hydrological cycle, creating a new water balance under altered hydroclimatic conditions (Du et al., 2016). For instance, interbasin water transfer projects introduce new water sources for evapotranspiration, and irrigation in dry regions expand croplands, intensifying water availability for evapotranspiration (Bonacci and Andrić, 2010), which is a definite contrast to a closed watershed system assumed by BH. Nowadays, most inhabited basins have been significantly altered by large-scale human activities, making them no longer closed or natural. Consequently, the relationship between the annual ratio of actual evapotranspiration and potential evapotranspiration to precipitation often fails to meet the primary criterion of the Budyko hypothesis, hindering its application in these unclosed basins (Du et al., 2016).

Nevertheless, the accuracy of Budyko-based models can be significantly improved by optimizing the temporal scale of the Budyko parameter and selecting its most stable and steady value, as it encapsulates both catchment properties and climate variability; otherwise, mismatches in temporal resolution may lead to significant biases, particularly in regions with strong seasonal fluctuations.

Machine learning (ML) has emerged as a powerful tool in hydrology, where it is widely used for parameter calibration, data transfer, and parameter forecasting (Granata, 2019). Its effectiveness stems from its ability to learn complex, non-linear relationships between inputs and outputs, even when the underlying hydrological processes are not fully understood (Cheng et al., 2022). Consequently, ML methods have been extensively applied in evapotranspiration research. A substantial body of work, including studies by Torres et al. (2011), Shrestha and Shukla (2015), Patil and Deka (2016), Feng et al. (2017a, 2017b), Nema et al. (2017), Dou and Yang (2018), Mehdizadeh (2018), Xu et al. (2018), and (Granata, 2019), has demonstrated the potential of ML in this domain. Specific examples from the literature highlight the diverse applications of ML models in ET<sub>a</sub> estimation across various agroclimatic conditions. For instance, Tang et al. (2018) utilized support vector machines (SVMs) and artificial neural networks (ANNs) enhanced by genetic algorithms to accurately model ET<sub>a</sub> in a rainfed maize field under different mulching practices. Similarly, a study by Fuentes et al. (2024) developed and validated ML models using micrometeorological and temperature-based data to estimate ET<sub>a</sub> and energy balance components in vineyards accurately. This research demonstrated the potential of ML to enhance precision irrigation and improve water use efficiency under varying agroclimatic conditions. Furthermore, in a large-scale irrigation area, Karahan et al. (2024) developed a new ANN model to estimate daily ET<sub>a</sub> by integrating limited local climatic data with remote sensing parameters. This approach confirmed strong agreement with established surface energy balance model outputs and offered a practical solution for sustainable water management. These studies collectively illustrate the power and versatility of ML models in ET<sub>a</sub> estimation across various settings and scales.

While ML models are effective at extracting intricate information and might outperform process-based hydrological models at large spatial scales, they face a significant challenge due to the high demand for extensive and reliable data (Amani and Shafizadeh-Moghadam, 2023).

This reliance on abundant data severely limits their applicability in data-scarce regions (Sharafi and Mohammadi Ghalei, 2024), where the lack of sufficient training data can compromise a model's accuracy and generalizability. A promising solution to these critical limitations is the integration of globally available remote sensing (RS) datasets. These datasets, such as the TerraClimate and ERA5-Land reanalysis products, provide invaluable spatial and temporal coverage, making them a robust option for hydrological modeling in regions with insufficient or non-existent ground-based observations (Abatzoglou et al., 2018; Muñoz-Sabater et al., 2021).

Considering these challenges, the central innovation of this study is the development of a novel hybrid modeling framework for estimating actual evapotranspiration ( $ET_a$ ) in arid, data-scarce regions and under non-steady-state conditions. This framework holistically combines the physical principles of the Budyko hypothesis with the predictive power of machine learning, thereby addressing critical gaps in the field.

The primary objective of this research is to develop a novel hybrid model that integrates various modified Budyko-type frameworks with an optimal machine learning model using deep learning techniques to achieve superior  $ET_a$  predictions. This approach leverages the strengths of both physical and data-driven methods, creating a more robust and accurate predictive tool. Another key component of this effort is the optimization of the Budyko parameter based on its temporal scale, utilizing terrestrial water storage changes (TWSC) to account for non-steady-state hydrological conditions. This methodology establishes a robust and reliable annually-derived parameter for the accurate estimation of monthly  $ET_a$ , overcoming the limitations of fitting parameters at shorter timescales. Furthermore, this study utilizes globally available remote sensing data, such as the TerraClimate dataset, to ensure the proposed framework is a scalable and applicable tool for sustainable water management strategies worldwide, particularly in regions with limited on-the-ground observational data (Eddy Covariance Towers). A comparative evaluation is also conducted to benchmark the performance of the new hybrid model against contemporary standalone machine learning and classical Budyko-based approaches, providing a comprehensive assessment of its respective strengths and weaknesses in different hydrological contexts. Ultimately, this research aims to provide

a reliable, globally applicable, and innovative tool that advances the determination of  $ET_a$  for effective water resource management.

## 2. Methodology

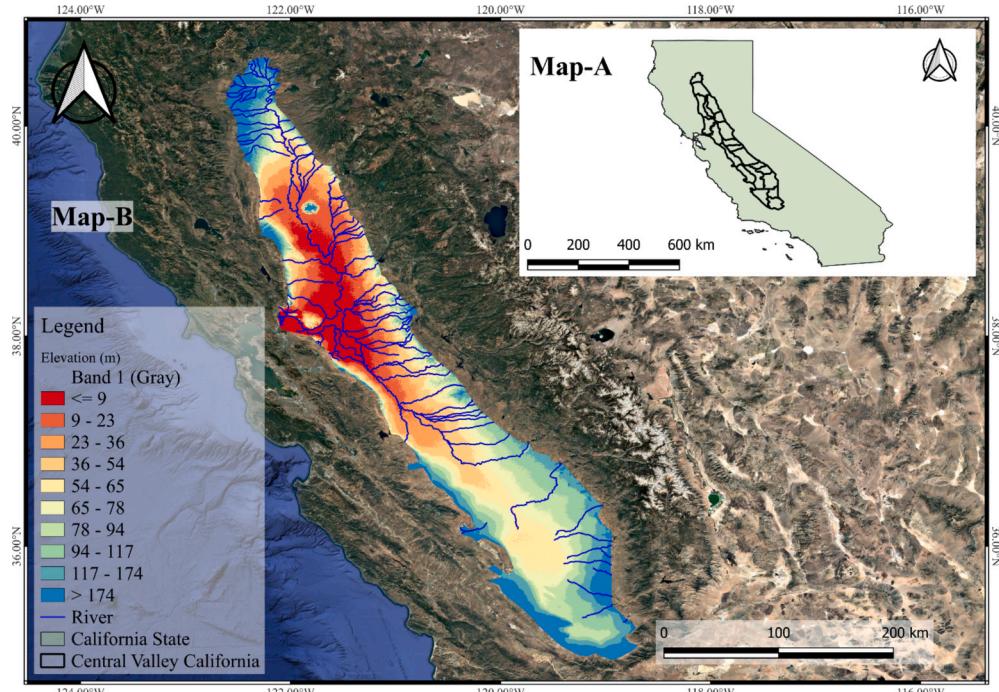
### 2.1. Case study

The Central Valley of California is a vast agricultural region covering over 51,800 km<sup>2</sup> (Fig. 1). It is bounded by the Cascade Range, Sierra Nevada, Tehachapi Mountains, and Coast Ranges/San Francisco Bay, and features minimal surface relief due to millions of years of sedimentation. The valley, drained by the Sacramento and San Joaquin Rivers, can be divided into the Sacramento Valley in the north and the San Joaquin Valley in the south, which includes the San Joaquin and Tulare Basins. The climate ranges from arid to semi-arid, with a hot-summer Mediterranean influence (according to Kottke et al., 2006, Köppen-Geiger climate classification Csa and BSk is the dominant climate classification, as shown in Fig. 2), where the region experiences frequent droughts and floods. Annual precipitation varies significantly, with the Sacramento Valley receiving 330–660 mm while the San Joaquin Valley receives 127–457 mm. The potential evapotranspiration (PET) is relatively high, ranging from 1143 mm in the Sacramento Valley to 1422 mm in the San Joaquin Valley. In general, most of the valley is in a state of perennial water deficiency, where PET exceeds precipitation by as much as 914 mm. Overall, PET exceeds precipitation during the summer, while during the winter seasons, precipitation exceeds PET (Faunt et al., 2009).

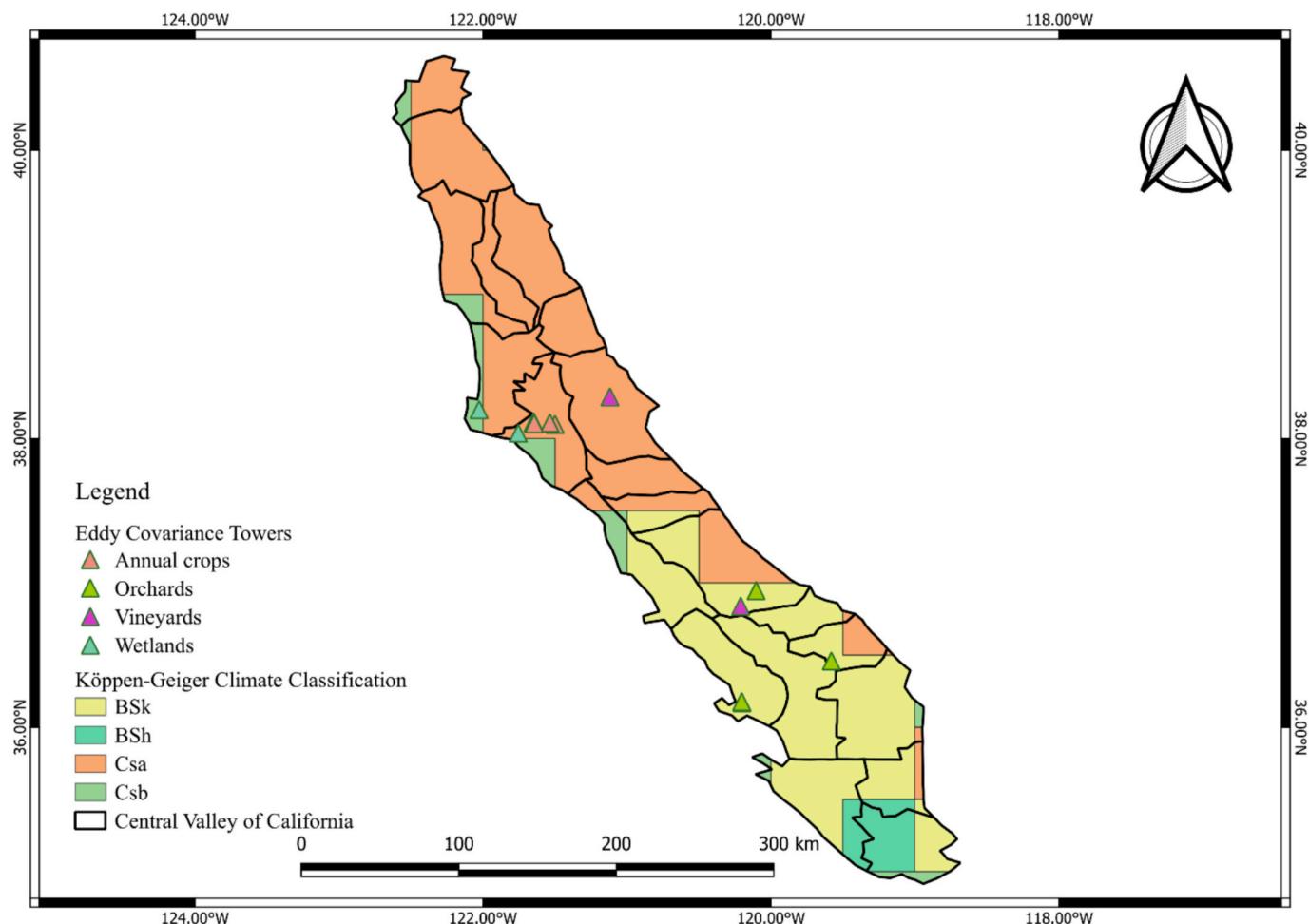
The Central Valley of California was selected for this investigation primarily due to the superior number and optimal spatial and temporal distribution of available Eddy Covariance Towers within the area, compared to other states like Nevada, Colorado, and Arizona. This provided extensive meteorological data essential for this study (Table A.1).

### 2.2. Data

In this study, a multi-source dataset was compiled for the purpose of



**Fig. 1.** The location of the case study in the USA: Map A: State of California, Map B: The Digital Elevation Model and streams of the study area (Central Valley of California).



**Fig. 2.** Map of Köppen-Geiger climate classification for the Central Valley of California (Csa: Hot-summer Mediterranean climate, BSk: Cold semi-arid climate, BSh: Hot semi-arid climate, Csb: Warm-summer Mediterranean climate) and the location of Eddy Covariance stations in the region based on the land cover type.

actual evapotranspiration ( $ET_a$ ) modeling. The primary target variable for model training was observational  $ET_a$  data obtained from Eddy Covariance (EC) tower measurements. Concurrently, a set of predictor variables was derived from these same EC tower measurements, including land use type and local precipitation data. To provide a more comprehensive representation of environmental and climatic conditions, two additional datasets, TerraClimate and ERA5, were integrated as supplementary predictor variables. This integration of local-scale and global-scale data was crucial for training ML models and enhancing the overall understanding of larger-scale hydrological phenomena, thereby contributing to more robust and accurate  $ET_a$  estimations.

Monthly averaged parameters were calculated by determining the mean value of daily data for each specific month and year throughout the study period. For each year, this process yielded 12 distinct monthly values. In contrast, the long-term monthly (LT monthly) averaged parameters were derived by averaging the values of a specific calendar month (e.g., January) over the entire study period from 1993 to 2022. This method produced a total of 12 distinct values, with each value representing the average for a particular month across 30 years.

#### 2.2.1. Eddy covariance data collection

The Eddy Covariance (EC) data utilized in this study were primarily sourced from the AmeriFlux network (<https://ameriflux.lbl.gov/>), a comprehensive collection of environmental monitoring systems across the Americas. Specifically, we accessed post-processed daily and monthly aggregated measurements from 13 EC stations located within California's Central Valley. The core variable extracted for our analysis

was actual evapotranspiration ( $ET_a$ ), which served as the target variable for model training. Additionally, other predictor variables, such as land use type and local precipitation data, were either derived from or directly associated with the respective tower locations. The geographical positions of these 13 stations within the Central Valley are presented in Fig. 2, while a summary of their key characteristics is provided in Table A.1.

#### 2.2.2. TerraClimate data

The TerraClimate dataset (Abatzoglou et al., 2018) was utilized in this study to provide additional climatic inputs. This high-resolution dataset, with an approximate spatial resolution of 4 km, offers monthly climatic water balance data for global terrestrial surfaces spanning from 1958 to 2022. Its generation is based on climatically aided interpolation, which combines high-resolution climatological normals from the WorldClim dataset with coarser monthly data from sources such as CRU Ts4.0 and JRA-55 reanalysis products. For this research, key variables, including precipitation (P), maximum and minimum temperatures, wind speed, vapor pressure, and solar radiation, were extracted from TerraClimate available in the Google Earth Engine platform. Furthermore, P and PET from this dataset were crucially employed to calibrate and fit the parameters in our modified Budyko models. These variables also served as essential predictor variables for the  $ET_a$  estimation using ML models. The reliability of this dataset has been demonstrated through strong validity checks using a large number of station-based observations from networks such as SNOTEL, RAWS, and the Global Historical Climate Network (Elbeltagi

et al., 2024).

### 2.2.3. ERA5 data

The ERA5-Land dataset, a component of the fifth-generation European ReAnalysis (ERA5) produced by the European Centre for Medium-Range Weather Forecasts, was utilized in this study, available in the Google Earth Engine platform. This dataset provides global land component data at an approximate spatial resolution of 9 km, with a temporal coverage from 1950 to the present at a monthly time scale. The ERA5-Land dataset is recognized for its significant improvements over previous versions, including a higher resolution, a more accurate representation of water and energy cycles, and enhanced descriptions of soil moisture and lakes (Muñoz-Sabater et al., 2021). The data, which are continuously updated and available through the Copernicus Climate Change Service, are widely regarded as reliable for various applications in land monitoring, hydrology, and climate modeling. The dataset's reliability for evapotranspiration estimation has been demonstrated in previous research; for instance, a study by Baboli et al. (2024) showed that ERA5 data provided accurate ET estimates, with validation indices demonstrating strong agreement with ground measurements ( $R^2 = 0.94$ , RMSE = 0.98, NSE = 0.86).

All parameters utilized from the ERA5, TerraClimate, and Eddy Covariance (EC) datasets are summarized with their details in Table A.3. To ensure temporal consistency across all datasets, a common study period was established. This was achieved by aligning all data to the temporal range of the EC tower measurements, which represented the shortest time series. Consequently, only data from the ERA5 and TerraClimate datasets that corresponded to the specific months within the EC tower records were used in the analysis.

### 2.3. Budyko model modification

The Budyko framework offers a fundamental conceptual understanding of the long-term water and energy balance at the watershed scale. Initially formulated by Budyko (1958), the hypothesis establishes a simple relationship between long-term mean actual evapotranspiration ( $ET_a$ ), precipitation (P), and the aridity index ( $\phi = PET/P$ ). According to this framework, evapotranspiration is primarily controlled by water availability in arid regions and by energy availability in humid regions (Gunkel and Lange, 2017). The core Budyko equation, which represents the evaporative index ( $ET_a/P$ ), is expressed as Eq. (1).

$$\varepsilon = B(\phi) = \{\phi[1 - \exp(-\phi)]\tanh(\phi^{-1})\}^{0.5} \quad (1)$$

Beyond this classic formulation, various Budyko-type equations have been developed to characterize diverse watershed properties and hydrological partitioning better (Jiang et al., 2015). These modified equations, which are utilized in this study, each include a parameter  $\omega$  that describes watershed characteristics. Their general form can be expressed as  $\varepsilon = B(\Phi|\omega)$ , where the parameter is adjusted to best fit the

**Table 1**

The four single-parameter Budyko-type equations utilized in this study, as outlined by Jiang et al. (2015).

Name of the Budyko-type equation	Expression of the Budyko-type equation
Fu (Fu, 1981)	$B^{TP}(\phi \omega^{TP}) = (1 + \phi^{\omega^{TP}})^{-1/\omega^{TP}}$
Turc-Pike (Turc, 1954; Pike, 1964; Milly and Dunne, 2002; Yang et al., 2008)	$B^F(\phi \omega^F) = 1 + \phi - (1 + \phi^{\omega^F})^{1/\omega^F}$
Zhang (Zhang et al., 2001 and Zhang et al., 2004)	$B^Z(\phi \omega^Z) = (1 + \omega^Z\phi)(1 + \omega^Z\phi + \phi^{-1})^{-1}$
Wang-Tang (Wang and Tang, 2014)	$B^{WT}(\phi \omega^{WT}) = \frac{[1 + \phi - \sqrt{(1 + \phi)^2 - 4\omega^{WT}(2 - \omega^{WT})\phi}]}{[2\omega^{WT}(2 - \omega^{WT})]}$

observed data. The specific equations utilized in this research, along with their general forms and associated parameters, are presented in Table 1.

While the original Budyko framework assumes steady-state conditions and closed basins (Budyko, 1958), its application to most contemporary natural basins is limited. These basins are often significantly altered by human activities and anthropogenic influences, leading to widespread non-steady-state conditions and corrected water balances (Du et al., 2016). Consequently, to account for these complexities, our study utilizes a modified Budyko model that incorporates an optimized parameter ( $\omega$ ) to thoroughly represent the hydrological dynamics of the study area. Under non-steady-state conditions, variations in annual terrestrial water storage changes (TWSC) become a significant factor. In such a context, precipitation (P) is no longer the exclusive source of water for evapotranspiration ( $ET_a$ ), and thus,  $ET_a$  cannot be simply calculated as precipitation minus runoff ( $P - R$ ) (Huang et al., 2022). A key advantage of this approach is its reliance on globally available remote sensing datasets for Budyko parameter fitting. To enhance the robustness and applicability of the model, the Budyko parameter was evaluated across various temporal scales to determine the optimal scale for estimating monthly  $ET_a$ . This adaptation redefines water availability as effective precipitation ( $P_e = P - TWSC$ ). The evaporative and aridity indices are then expressed as  $\frac{ET_i}{P_i - TWSC_i}$  and  $\frac{PET_i}{P_i - TWSC_i}$ , respectively, for a given timescale (i). This approach allows for the estimation of the Budyko parameter ( $\omega$ ) at various timescales under non-steady-state conditions for each equation presented in Table 1. The specific form of the modified Budyko equation used in this study, which incorporates terrestrial water storage changes, is as follows:

$$\frac{ET_i}{P_i - TWSC_i} = 1 + \frac{PET_i}{P_i - TWSC_i} - \left[ 1 + \left( \frac{PET_i}{P_i - TWSC_i} \right)^\omega \right]^{\frac{1}{\omega}} \quad (2)$$

Also, TWSC can be derived from:

$$TWSC_i = P_i - R_i - ET_{ai} \quad (3)$$

In the above equation, the subscript i denotes the index for the specific timescale being considered (e.g., monthly, long-term monthly, and annual averaged). While precipitation ( $P_i$ ) and potential evapotranspiration ( $PET_i$ ) are obtained from direct observations, the actual evapotranspiration ( $ET_{ai}$ ) and the terrestrial water storage changes ( $TWSC_i$ ) are estimated using a hydrological model, as further detailed in Eq. (3).

### 2.4. Machine learning modeling

For this study, fourteen Machine Learning (ML) models were initially considered, representing a diverse set of algorithms including ensemble models (e.g., Random Forest Regressor, AdaBoost Regressor, Bagging Regressor, Extra Trees Regressor, Gradient Boosting Regressor, Hist Gradient Boosting Regressor, and Extreme Gradient Boosting), linear regression models (Ridge, Least Angle Regression (Lars), RidgeCV, and Linear Regression), neural network models (Multi-Layer Perceptron Regressor - MLPR), non-parametric models (K-Nearest Neighbors - KNN), and tree-based models (Decision Tree Regressor - DTR).

#### 2.4.1. Input preprocessing

Prior to model development, a comprehensive preprocessing stage was applied to the input features to ensure data quality, consistency, and temporal alignment. All predictor variables were normalized using the StandardScaler method from the scikit-learn Python library. This standardization process mitigates the influence of differing feature scales by removing the mean and scaling to unit variance, thereby enhancing the convergence and improving the performance of all models sensitive to input feature distributions.

A hybrid imputation strategy was implemented to address missing values within the dataset. Specifically, linear interpolation was applied

to time-dependent variables to preserve temporal continuity. For static or non-temporal features, mean imputation was used to maintain dataset completeness while minimizing information loss.

The observational dataset, consisting of Eddy Covariance (EC) tower measurements, provided the actual evapotranspiration ( $ET_a$ ) values. To ensure temporal consistency with the ERA5 and TerraClimate predictor datasets, a common study period was established by taking the temporal intersection of all datasets. The period from 2009 to 2022, representing the time of overlap among all data sources, was therefore selected for model training and evaluation. This alignment ensured that the input features and reference values were temporally synchronized, which is critical for a robust analysis. Categorical or textual features within the Eddy Covariance (EC) dataset, such as site identification, general classification, and land cover type, were converted into a consistent numeric format. This was achieved using a factorization method, which assigns a unique integer label to each distinct category. This process enabled the proper integration of these categorical variables into the machine learning pipeline.

To enhance model performance and reduce data dimensionality, a model-independent, correlation-based feature selection method was implemented. This approach involves identifying variables that exhibit a strong correlation with the target variable ( $ET_a$ ) while removing those with low predictive power. Redundant features that convey similar information were also eliminated to prevent multicollinearity and decrease model complexity. This process enhances the efficiency and accuracy of the learning algorithms by allowing them to focus on the most relevant predictors.

#### 2.4.2. Hyperparameter tuning

To ensure the best possible prediction performance, hyperparameter optimization was performed for each ML model. This study employed a Grid Search Cross-Validation (GridSearchCV) approach to systematically identify the optimal hyperparameter combinations for each algorithm. A predefined grid of parameter values, established based on a comprehensive literature review, was exhaustively searched. The optimal hyperparameter combination was selected based on maximizing the coefficient of determination ( $R^2$ ) and minimizing another appropriate scoring metric, such as the negative Root Mean Squared Error (RMSE). A K-Fold cross-validation strategy with 10 folds was applied during the GridSearchCV process to ensure robust parameter selection and mitigate overfitting (Table A.2).

#### 2.4.3. Model training, testing, and validation

The compiled dataset, which included actual evapotranspiration ( $ET_a$ ) values and associated predictors from Eddy Covariance (EC) towers, ERA5, and TerraClimate, covered the period from 2009 to 2022. For model development, the data were randomly partitioned into two distinct subsets: a training set (70 %) and a testing set (30 %). This random split ensured that the training set captured a wide range of variability across environmental conditions, while the independent testing set provided an unbiased basis for evaluating the final model's performance.

Each ML model was trained using the designated training dataset. During model development, an internal validation was performed on the training set through cross-validation to assess performance consistency and generalization capability. The final, unbiased evaluation of each trained model was then conducted on the independent testing set to ensure its robustness and to mitigate the risk of overfitting.

#### 2.4.4. Feature selection

To optimize model performance and enhance interpretability, a systematic feature selection process was implemented. This study employed two distinct strategies for feature selection, evaluated across all fourteen ML models.

The first strategy involved a comprehensive scenario-based evaluation using Recursive Feature Elimination (RFE) for some models, while

others, particularly those that do not support RFE directly, were evaluated using a feature selection pipeline. This approach systematically assessed all possible feature subsets, ranging from two features up to the total number of available predictors. The process entailed training and evaluating the model with each feature subset (e.g., resulting in 47 scenarios for the XGBoost Regressor). In other models, a pipeline was utilized to integrate a feature selection step, such as selecting features according to the k's highest scores (SelectKBest method), directly with the regression model, ensuring a consistent and robust workflow. The performance of each scenario was quantified using  $R^2$ , Normalized Root Mean Square Error (NRMSE), and Nash-Sutcliffe Efficiency (NSE). The optimal feature subset for each model was determined by identifying the highest performance across these metrics, with an additional consideration for model parsimony. Models requiring fewer input parameters were preferred when performance differences were negligible, recognizing that increased parameter count can reduce a model's generalizability and practical value.

The second strategy for feature selection involved a comparative analysis based on box plots to visualize the distribution of model performance across different numbers of selected features. Similar to the first approach, this method entailed evaluating all feature subsets for each model. For each specific subset size, the distribution of performance scores (e.g., NRMSE values) from various cross-validation folds was visualized using box plots. This robust comparison of model stability and performance across different feature subset sizes provided a complementary perspective to the single-best-scenario selection. This visual assessment was instrumental in understanding the trade-off between model complexity (number of features) and predictive accuracy across the entire range of scenarios.

#### 2.5. Hybrid Budyko-deep learning model

To enhance the predictive accuracy of the traditional Budyko framework, a novel hybrid modeling approach was developed by integrating it with a deep learning ensemble model. This method leverages the physical principles of the Budyko curve while simultaneously capturing the complex, nonlinear relationships inherent in the Central Valley's water balance that are difficult for conceptual models to fully represent. The core of this concept is a stacked ensemble learning approach. In this framework, the outputs from the four modified Budyko equations and a separate machine learning model are used as input features for a final neural network. This neural network, which functions as a meta-learner, was constructed using a sequential architecture. It consists of an input layer with two nodes, two hidden layers with a Rectified Linear Unit (ReLU) activation function, and a single output layer with a linear activation function. This meta-learner is then trained on these combined predictions to determine the optimal weights and interactions for producing a single, refined output. This hybrid model corrects for the inherent limitations and inaccuracies of each standalone method, thereby yielding more robust and reliable estimations of actual evapotranspiration ( $ET_a$ ), particularly at sub-annual timescales where the Budyko parameter ( $\omega$ ) is known to be unstable.

#### 2.6. Performance metrics

The performance of all models was evaluated by comparing the modeled  $ET_a$  values against the observed  $ET_a$  data from the Eddy Covariance towers. To quantify the agreement between these projected and observed values, three statistical metrics were utilized:  $R^2$ , NSE, and NRMSE. The formulation of these metrics is given in Eqs. (4) to (6), respectively.

$$R^2 = \left[ \frac{\sum_{i=1}^n (P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\left( \sum_{i=1}^n (P_i - \bar{P})^2 \right) \left( \sum_{i=1}^n (O_i - \bar{O})^2 \right)}} \right]^2 \quad (4)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (5)$$

$$NRMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} / \bar{O} \quad (6)$$

## 2.7. Flowchart of the research methodology

The overall workflow and methodological framework of this study are schematically represented in Fig. 3. This flowchart illustrates the sequential steps, from data acquisition and preprocessing to model application and evaluation.

## 3. Results and discussion

### 3.1. Budyko model modification

#### 3.1.1. Annual terrestrial water storage change (TWSC)

Analysis of the annual water balance components revealed relatively small fluctuations in terrestrial water storage changes (TWSC) compared to the substantial variations observed in both actual evapotranspiration ( $ET_a$ ) and runoff (R). To account for the non-steady-state conditions of the watershed, TWSC was estimated from the annual water balance using Eq. (3). The range of TWSC values in the Central Valley was from  $-73.75$  mm to  $79.7$  mm, with a mean absolute value of  $29$  mm. This finding, which shows that annual TWSC is a relatively minor component of the water balance compared to  $ET_a$  and R, is consistent with a global analysis of terrestrial water cycles. Studies have shown that changes in annual precipitation often translate primarily into variations in R and TWSC, with TWSC exhibiting relatively minor fluctuations compared to  $ET_a$  and R on an annual basis (Yin and Roderick, 2020).

However, it is crucial to recognize that the hydrological system of

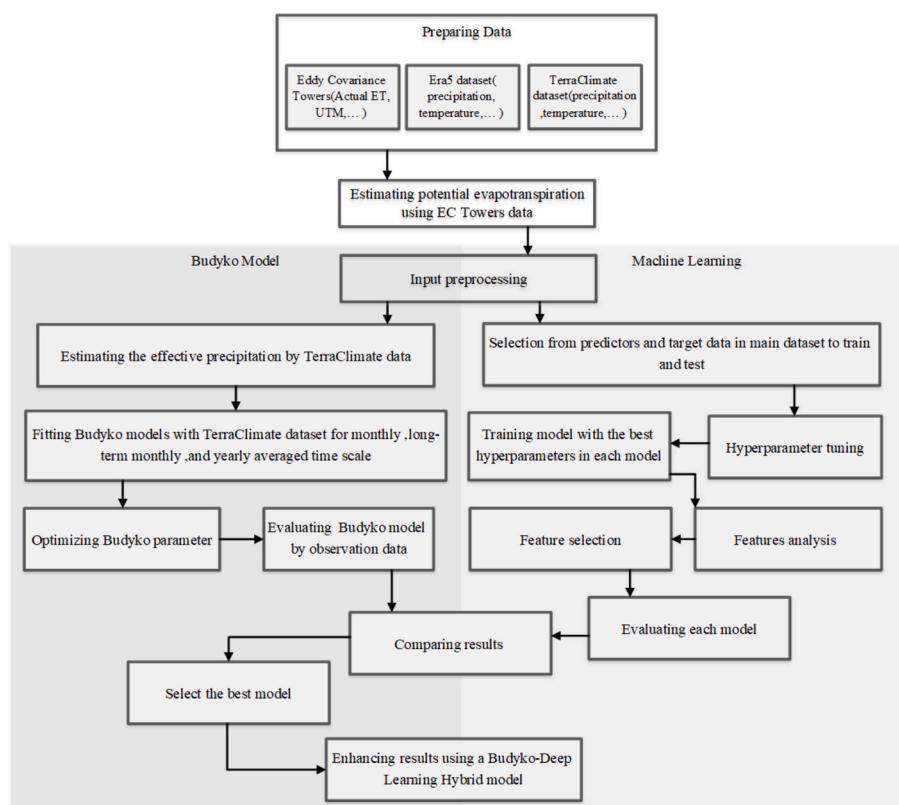
California's Central Valley is currently in a non-steady-state condition. This is particularly evident in semi-arid to semi-humid environments, where complex hydrological patterns are common. The non-steady-state condition in our study area is not solely a climatic phenomenon, but is largely driven by extensive human activities such as inter-basin water transfer, agricultural activities, and climate change. This is further substantiated by the fact that the Central Valley experiences the world's largest volume of land subsidence caused by human activities, a direct indicator of long-term groundwater depletion (Faunt et al., 2024). Given these complexities, the modified Budyko framework, which explicitly accounts for TWSC, provides a more robust and realistic representation of the water balance in this dynamically changing agricultural basin (Fig. 4).

#### 3.1.2. Estimation of $\omega$ using the TerraClimate dataset

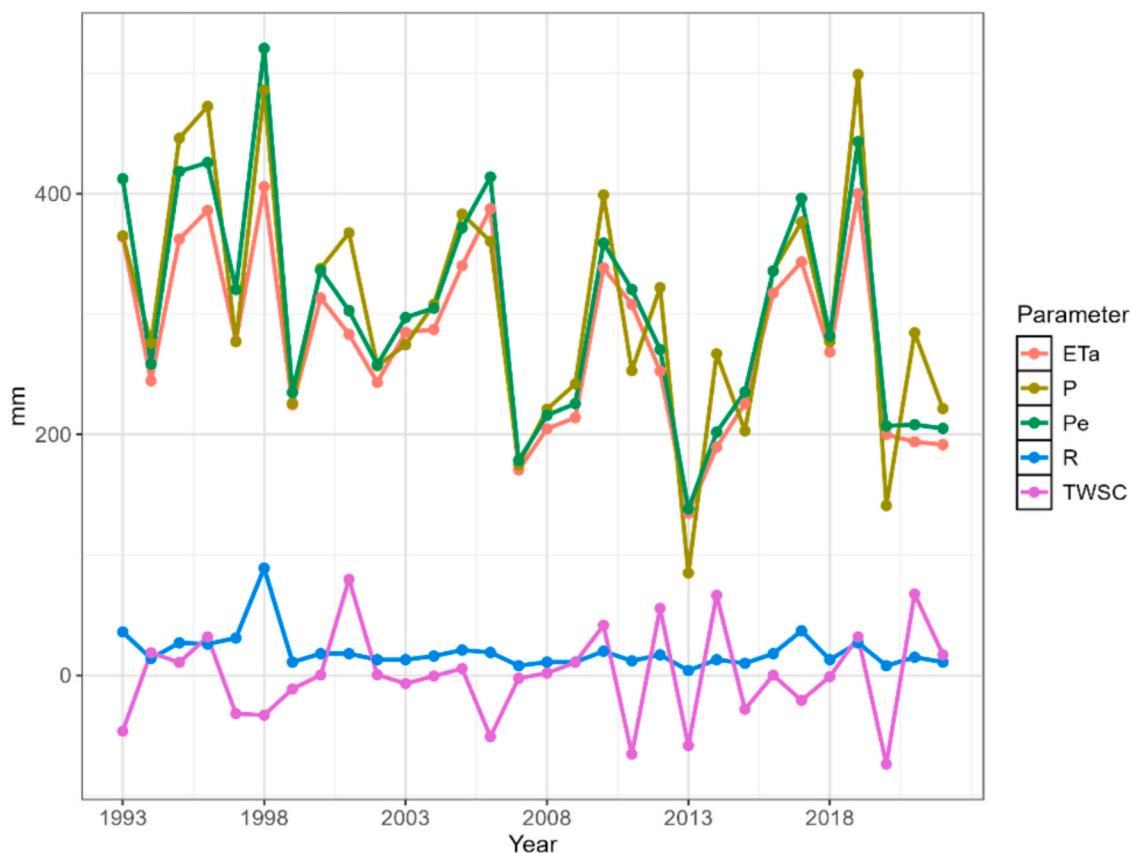
The Budyko parameter ( $\omega$ ) for the Central Valley was derived using long-term mean annual P, R, and  $ET_a$  data collected over 30 years. This was achieved by fitting the modified Budyko equations, as presented in Table 1, to the observed data. The long-term mean evaporative index ( $ET_a/P_e$ ) for the studied watershed is plotted against the aridity index ( $PET/P_e$ ) using these equations (Fig. 5).

As shown in Fig. 5, the relationships between the evaporative and aridity indices exhibit significant variability, particularly at the monthly and long-term monthly (LT monthly) averaged scales. This is in contrast to the more stable relationship observed at the annual scale. The fitted Budyko curves for the monthly and LT monthly averaged datasets, and their corresponding optimized parameters ( $\omega$ ), deviate considerably from the observed data. This indicates that the simulated and observed actual evapotranspiration values are less likely to align closely when using these shorter temporal scales, highlighting the challenge of applying the Budyko framework to sub-annual periods with a single, stable parameter.

A key reason for this instability at sub-annual timescales is the influence of TWSC. At monthly timescales, TWSC cannot be neglected,



**Fig. 3.** Overall workflow of the proposed methodology.



**Fig. 4.** Annual changes of water balance parameters of the TerraClimate Dataset and TWSC in millimeters (ET<sub>a</sub>: Actual Evapotranspiration, P: Precipitation, R: Runoff, TWSC: Terrestrial Water Storage Change, and Pe: Effective Precipitation).

leading to a non-steady state condition for the basin (Barkhordari et al., 2025). The results of this study, which show significant scatter in the monthly Budyko plots (Fig. 5), support findings from other hydrological studies. These studies have demonstrated that neglecting the effects of TWSC can lead to a significant overestimation or underestimation of evapotranspiration variance in both humid and arid climates (Sun et al., 2025). Consequently, the observed scatter in our data is a direct reflection of these non-steady-state conditions, reinforcing the need for a modified approach that accounts for these temporal dynamics.

For instance, the Budyko parameter ( $\omega$ ) for the Fu equation was estimated to be 2.22 at the annual timescale. In contrast, the corresponding  $\omega$  values for the monthly and LT monthly averaged data were 9.3 and 4.1, respectively (Fig. 5). The results of our study show lower accuracy at monthly and seasonal scales compared to annual scales, which aligns with previous research. This confirms that the model's performance in representing the fundamental requirements of Budyko curve fitting is significantly compromised at shorter temporal scales (Wu et al., 2017). Consequently, we chose to use the more stable annual Budyko parameter ( $\omega$ ) to estimate monthly ET<sub>a</sub> in our model, as the monthly averaged values were deemed unacceptable due to their high variability and unphysical values (Fig. 5).

### 3.2. Machine learning models

#### 3.2.1. Hyperparameters tuning

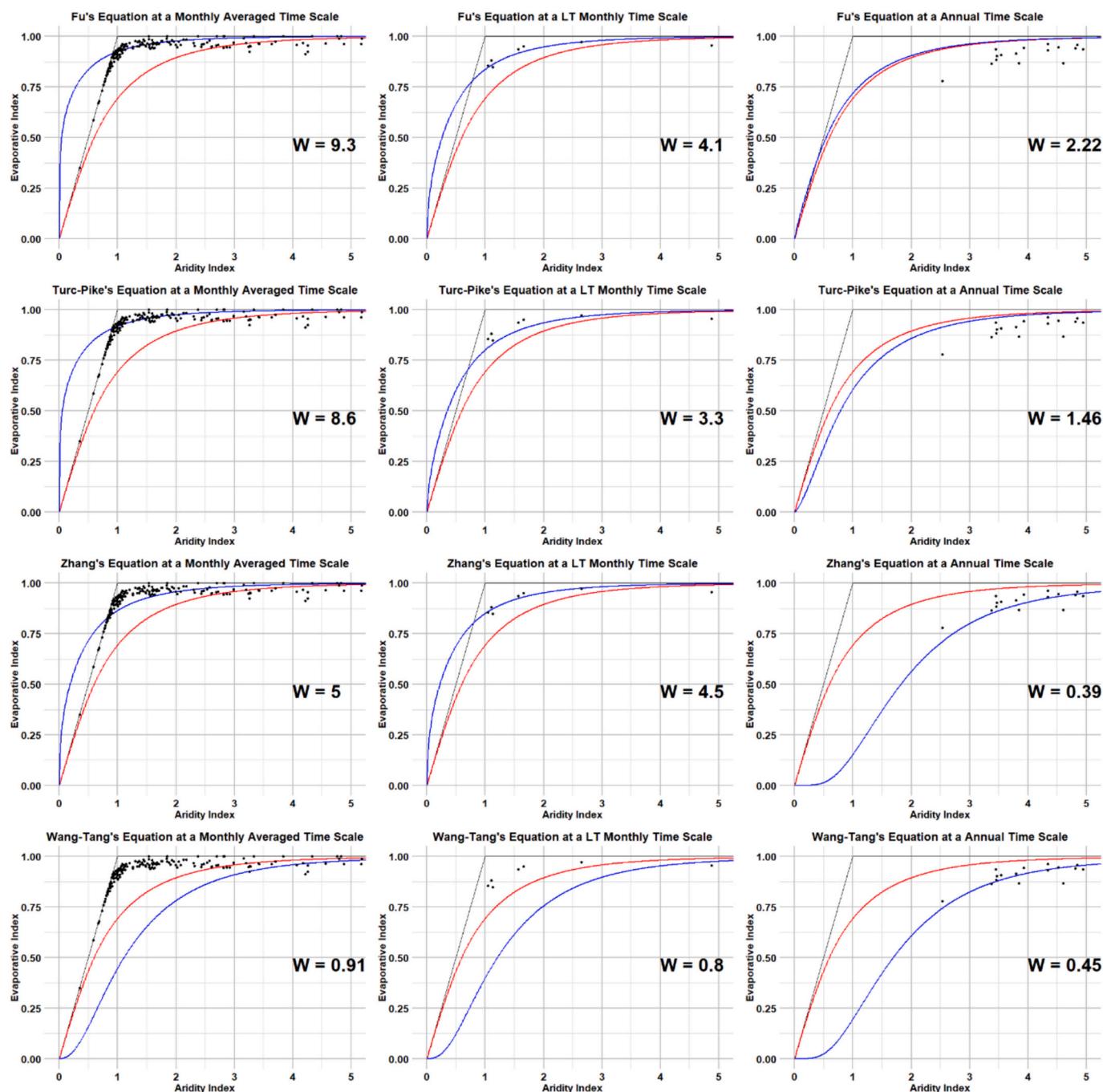
Model performance was optimized by systematically employing the Grid Search CV technique to select the most effective hyperparameters, consistent with established methodologies (Goyal et al., 2017; Granata et al., 2017; Rodriguez-Galiano et al., 2015). This comprehensive hyperparameter tuning is a significant advantage over previous studies that may have used a more limited range of parameter configurations

and fewer models. Our methodology provided a distinct advantage by exploring a broader search space for hyperparameter optimization compared to other studies, such as the work by Pagano et al. (2023), who applied a more constrained grid search for their MLP and RF models. This allowed for the identification of a more robust and finely-tuned parameter set (Table A.2).

The rigorous and exhaustive hyperparameter tuning process employed in this study yielded models with enhanced precision and reliability, enabling the accurate prediction of ET<sub>a</sub> under data-scarce conditions. This methodological approach constitutes a key strength of the work, and the best-performing hyperparameters are presented in Table 2.

#### 3.2.2. Features

**3.2.2.1. Feature analysis.** A comprehensive correlation study was conducted using feature analysis, a crucial prerequisite for identifying and eliminating less-relevant features to streamline model performance. The importance of each feature was quantified by the total reduction in the splitting criterion, a metric known as Gini importance (Ravindran et al., 2021). As shown in Fig. 6, the strongest positive correlations with potential evapotranspiration (PET) were observed for shortwave radiation (Srad) and minimum temperature (Tmin), with Pearson correlation coefficients (R) of 0.97 and 0.93, respectively. This finding is consistent with and supported by a previous study (Williams and Abatzoglou, 2025), which indicates that low precipitation leads to reduced soil moisture. This, in turn, drives a shift in the surface energy balance, resulting in less latent heat flux and more sensible heat flux, thereby increasing PET in alignment with the Budyko framework. In contrast, the correlation between the climate water deficit (Wdef) and observed actual evapotranspiration (ET<sub>a</sub>) was notably weak ( $R = -0.33$ ),



**Fig. 5.** The Budyko framework curves illustrate the relationship between the evaporative index ( $ET_a/P_e$ ) and the aridity index ( $PET/P_e$ ) across different temporal scales: monthly, long-term monthly (LT monthly), and annual. The figure displays the theoretical Budyko curve (red line), the fitted curve (blue line), and the observed data points (black circles). The optimized Budyko parameter,  $\omega$  (represented as  $W$  in the figure), derived for each timescale is explicitly noted. This visual comparison highlights the model's performance and the stability of the parameter across varying averaging periods.

suggesting that data-driven models relying solely on these parameters may be unsuitable for accurate  $ET_a$  estimation in this region. This observation is further supported by the cross panels in Fig. 6, which provides a kernel density estimation for each feature's probability density function. Additionally, Fig. 7 illustrates the Pearson correlation coefficients for all paired features in the ERA5 dataset, where the intensity of color, ranging from dark red ( $R = 1$ ) to dark blue ( $R = -1$ ), offers a rapid visual assessment of correlation strength.

**3.2.2.2. Feature selection.** A rigorous feature analysis was conducted to select the most relevant features and remove those that could

compromise model generalizability. This comprehensive methodology, which evaluated 47 distinct scenarios for each predictor by considering all possible feature combinations, provides a distinct advantage over previous studies. For example, Pagano et al. (2023) used a Violin plot for feature selection, but only explored 12 scenarios across their two machine learning models.

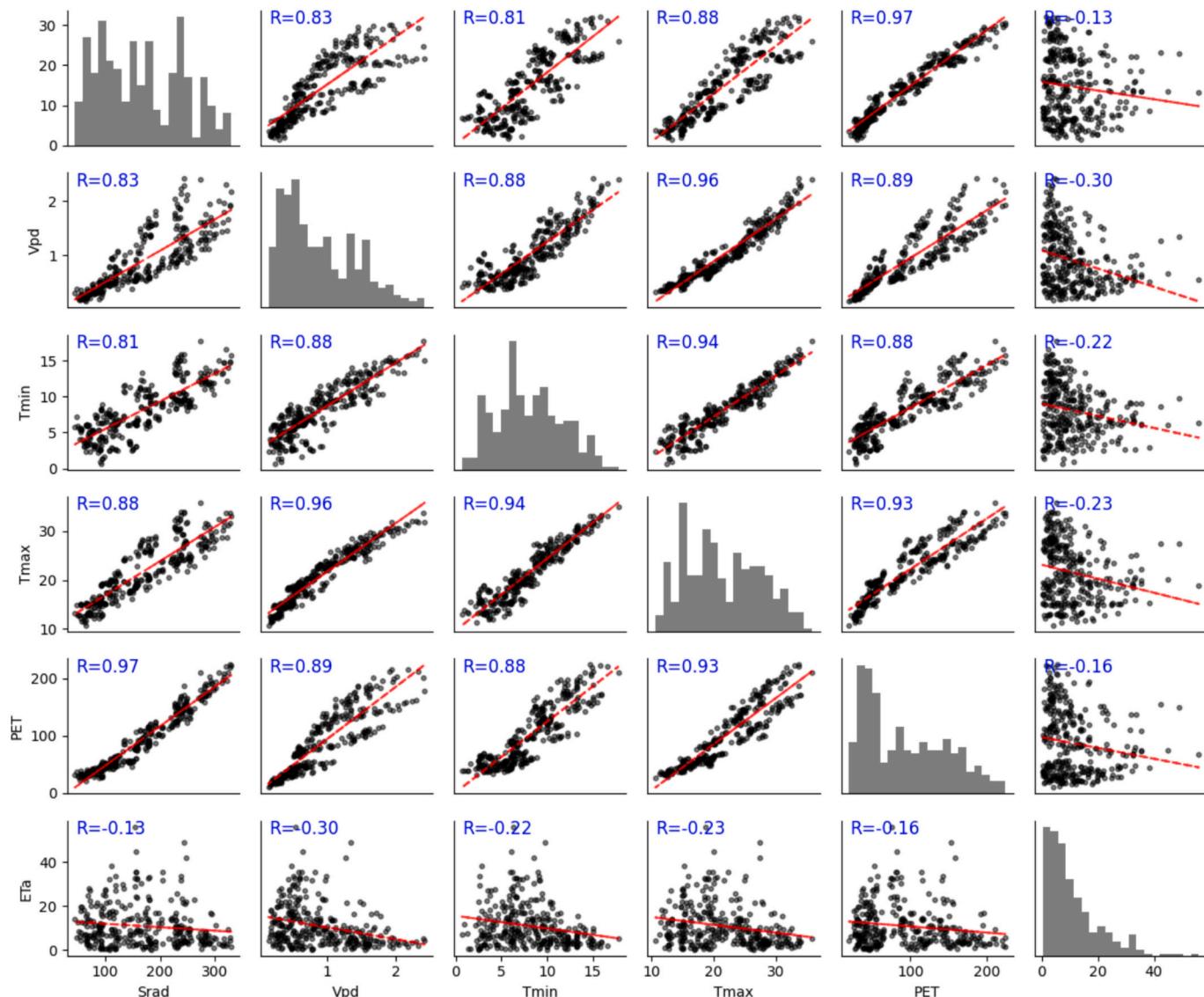
Table 3 summarizes the performance of the various models across different feature combinations, a process crucial for identifying the optimal feature sets for each predictor. The Random Forest (RF) regressor demonstrated that models RF3 and RF7 were the best candidates for  $ET_a$  estimation, achieving the highest  $R^2$  (0.70) and NSE (0.70) with

**Table 2**  
The best hyperparameters selected for each model.

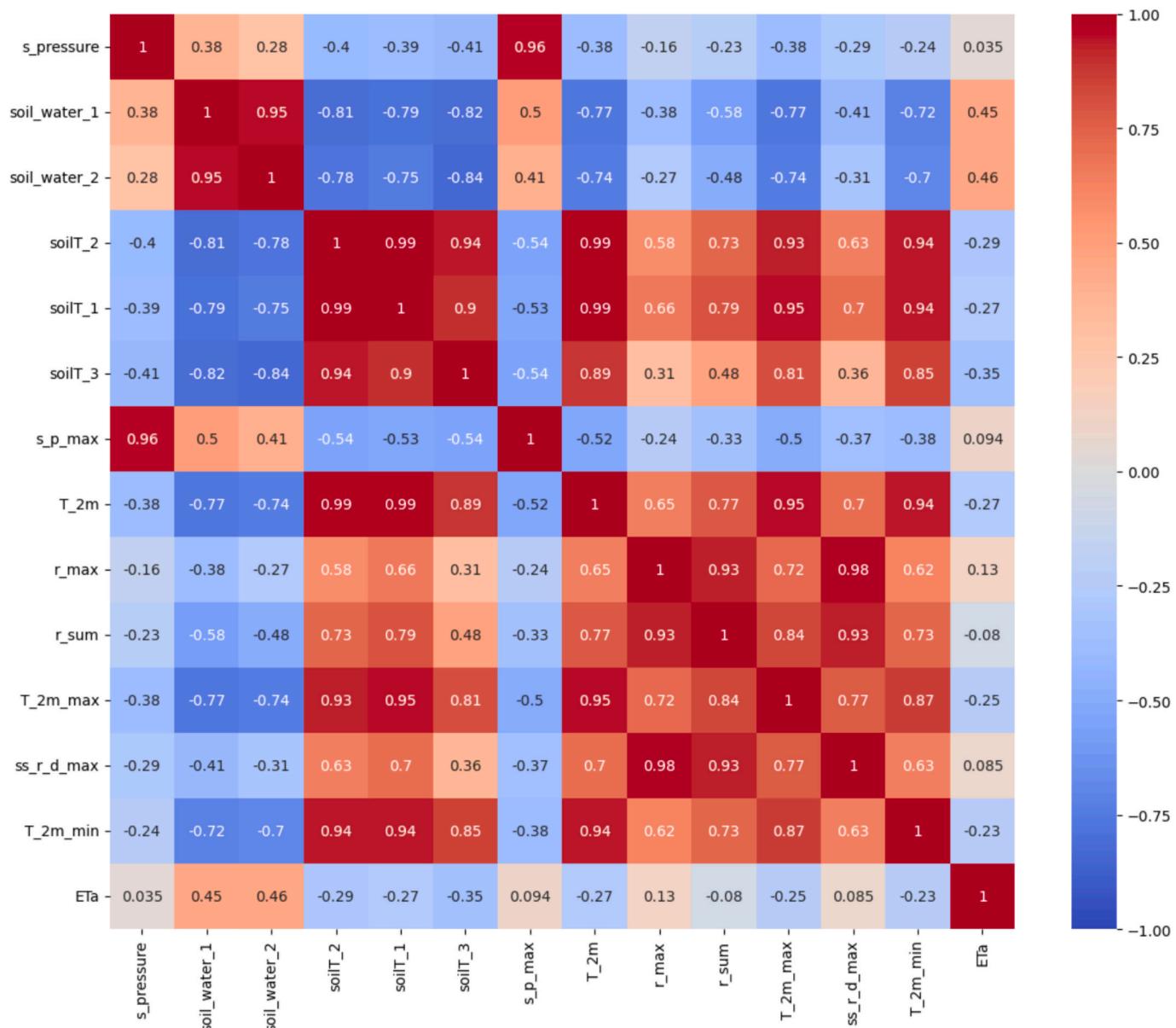
Model	Hyperparameters
Random Forest Regressor (RF)	Max. depth: None Number of trees: 50 Min. samples to split: 5 Best nodes: 20 Bootstrap samples: None
Gradient Boosting Regressor (GB)	Learning rate: 0.1 Sample fraction: 0.75 Boost stages: 200 Max. depth: 6
AdaBoost Regressor (Ada)	Max. estimators: 50 Regressor Weight: 0.25 Loss function: Linear
Extreme Gradient Boosting (XGB)	Number of boosting rounds: 100 Maximum depth of a tree: 3 Step size shrinkage: 0.1 Minimum loss reduction required to make a further partition on a leaf node: 0.2 Subsample ratio of the training instances: 1 Subsample ratio of columns when constructing each tree: 1

a low NRMSE (0.12). Similarly, the Gradient Boosting (GB) predictor's highest scores were associated with models GB5, GB6, and GB9, which yielded an  $R^2$  of 0.64, NRMSE of 0.13, and NSE of 0.63. The AdaBoost (Ada) predictor also showed strong performance with models Ada6, Ada12, and Ada4, recording an  $R^2$  of 0.63, NRMSE of 0.14, and NSE of 0.60. Among all predictors, the Extreme Gradient Boosting (XGB) model, specifically scenario XGB7, achieved the highest overall scores with an  $R^2$ , NRMSE, and NSE of 0.73, 0.11, and 0.73, respectively.

Despite the higher  $R^2$  values (up to 0.84) achieved by the best models in a prior study (Pagano et al., 2023), the comprehensive evaluation of 47 scenarios in this study provides a more robust and reliable foundation for our findings. A key methodological strength of this study is the demonstrable stability of our models, as indicated by consistently low Standard Deviation values (below 8.3 mm per month) for all scenarios, which substantiates the minimal deviation of forecasted  $ET_a$  from actual values. Furthermore, it was found that reducing the number of features often resulted in comparable or even superior model performance, highlighting the value of feature selection in improving efficiency without sacrificing accuracy. For instance, the Random Forest predictor showed similar performance for models RF3 and RF11, despite RF11 using a greater number of features. This principle is further illustrated in



**Fig. 6.** The scatter matrix of the five most correlated features in the TerraClimate dataset. This visualization depicts the pairwise relationships among these features. Refer to Table A.3 for the abbreviations.



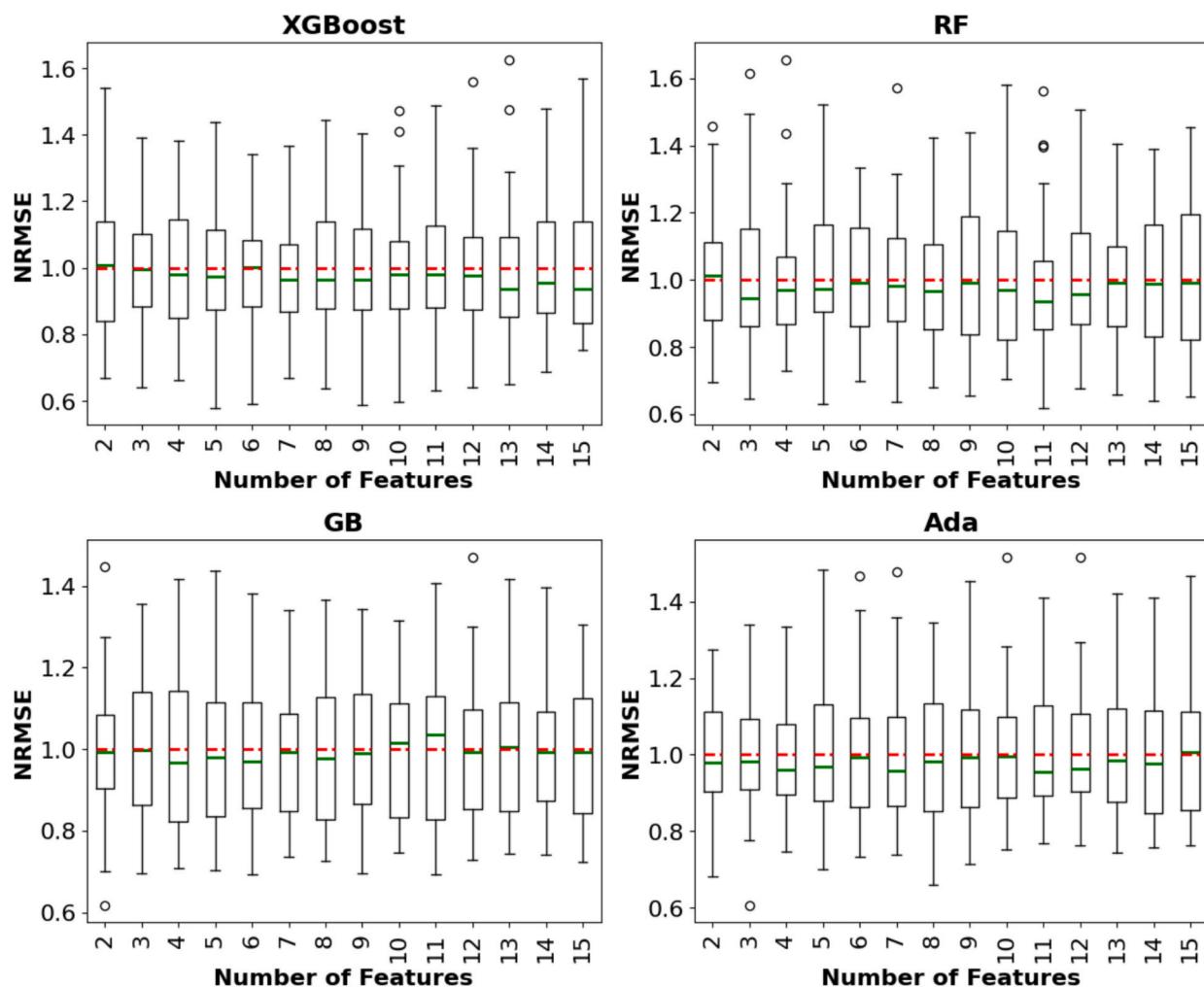
**Fig. 7.** The Correlation matrix of the fifteen most correlated features in the ERA5 dataset. This visualization depicts the pairwise relationships among these features. Refer to Table A.3 for the abbreviations.

**Table 3**  
Comparison of model performance metrics for ET<sub>a</sub> prediction using ML models.

Model	Number of features	R <sup>2</sup>	NRMSE	NSE
XGB	7	0.73	0.11	0.73
RF	3	0.70	0.12	0.70
GB	5	0.68	0.13	0.68
Ada	6	0.64	0.14	0.60
Bagging	14	0.56	0.15	0.54
DTR	14	0.30	0.18	0.29
Extratrees	13	0.63	0.14	0.63
HGB	14	0.51	0.15	0.49
KNN	9	0.47	0.16	0.47
Lars	14	0.59	0.14	0.59
Linear	15	0.46	0.16	0.45
MLPR	4	0.41	0.17	0.40
Ridge	10	0.46	0.16	0.46
RidgeCV	15	0.41	0.17	0.40

**Fig. 8,** which depicts the step-by-step feature selection process. The figure demonstrates that models with a lower number of features and a stabilized performance trend tend to yield better results. For clarity, Fig. 8 presents only the best 15 of the 47 total scenarios that were evaluated.

**3.2.2.3. Sensitivity analysis.** A comprehensive feature importance analysis was conducted on the top 14 machine learning models to elucidate the most influential predictors of ET<sub>a</sub>. The SHapley Additive exPlanations (SHAP) Beeswarm plot, presented in Fig. 9, was utilized to quantify the magnitude and directional impact of each feature on the model's output. The features are ranked by their average absolute SHAP value, providing a clear visual hierarchy of their relative importance. Consistent with prior research, such as the analyses by Liu et al. (2025) and Başaoglu et al. (2021), the results corroborate the paramount importance of precipitation (ppt) and maximum monthly surface net solar radiation (r\_max) as crucial external drivers in the best-performing models (XGB7, RF3, GB5, and Ada6). Furthermore, the significant and



**Fig. 8.** Box plots illustrating the distribution of the Normalized Root Mean Square Error (NRMSE) for various feature counts. The plots compare the performance of the Extreme Gradient Boosting (XGBoost), Random Forest (RF), Gradient Boosting (GB), and AdaBoost (Ada) models. The green and red dashed lines represent the median and mean NRMSE, respectively.

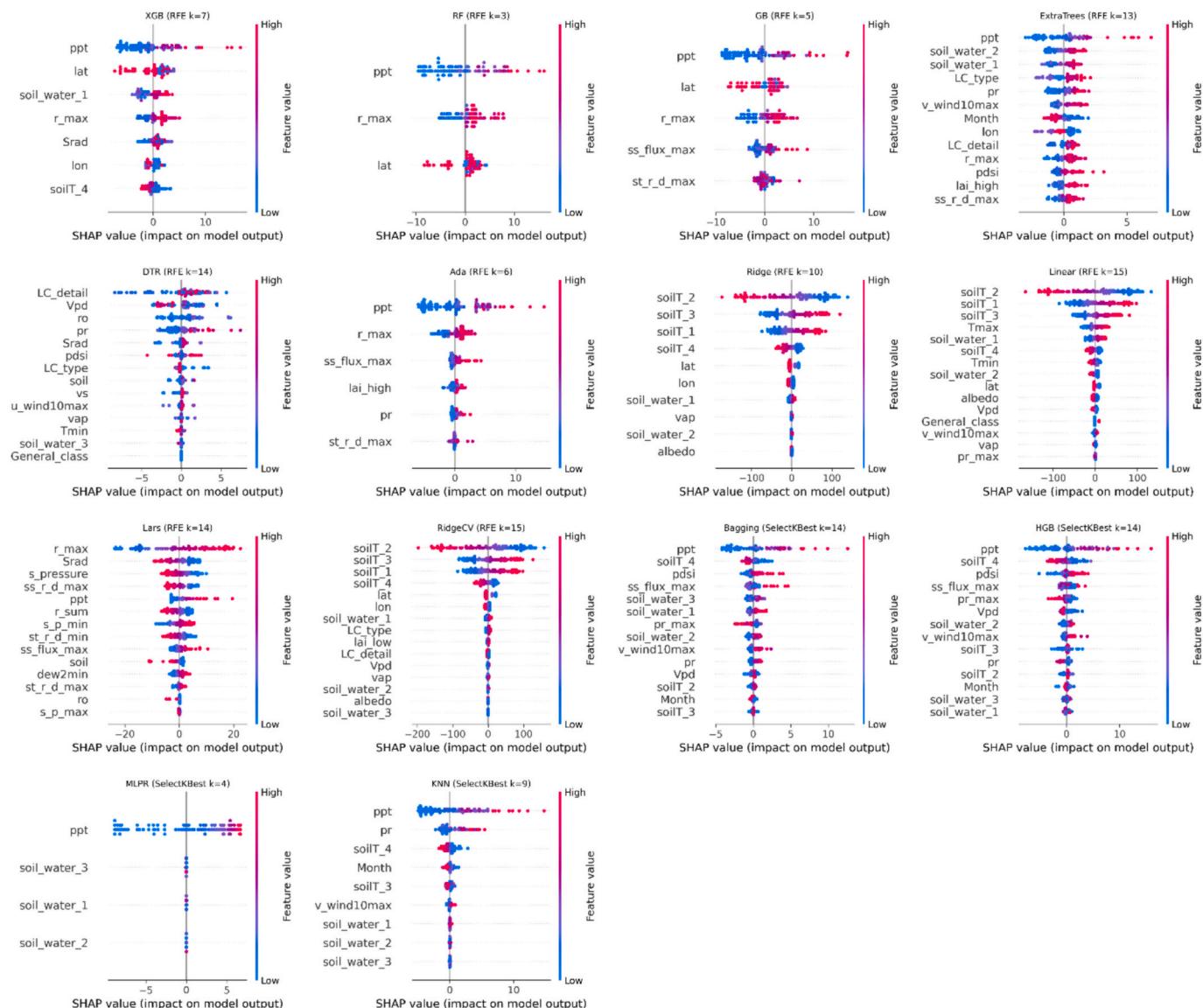
consistent influence of soil water content from various layers (soil\_water\_1 to soil\_water\_4) on model outputs validates the notion that soil moisture dynamics are a critical component of the supply-side of evaporation. This aligns with the findings of Liu et al. (2025), which identified shallow soil moisture as a primary driver of evaporative stress, particularly in water-limited regions. Finally, the substantial impact of vegetation-related features—including land cover type (LC\_type) and leaf area index (lai\_high, lai\_low)—in models such as ExtraTrees13, DTR14, Ada6, Linear15, and RidgeCV15 underscores the conclusion that vegetative factors are critical external drivers that significantly affect ET<sub>a</sub> estimations. This is consistent with a body of research that emphasizes the importance of vegetation in hydrological processes and highlights a strong correlation between ET<sub>a</sub> and vegetation indicators, such as the Normalized Difference Vegetation Index (NDVI) (Yang et al., 2023). Moreover, the Leaf Area Index (LAI) is a well-established indicator of vegetation development used to describe land evaporation, photosynthesis, and carbon sequestration (Alton, 2016; Boussetta et al., 2013), and is a highly accurate predictor of reference evapotranspiration when combined with machine learning techniques (Bachour et al., 2014).

### 3.3. Performance evaluation of predictor models

Monthly actual evapotranspiration (ET<sub>a</sub>) for the study area was estimated using the Budyko framework, with the parameter  $\omega$  derived at

both monthly and annual time scales. The model outputs were subsequently validated against observational data from Eddy Covariance towers, with performance metrics summarized in Table 4 and a visual comparison of simulated versus observed values presented in Fig. 10. A key finding of this analysis is the significant influence of the time scale of the Budyko parameter  $\omega$ ; specifically, using the annual  $\omega$  value for ET<sub>a</sub> estimation markedly enhanced model accuracy. This is evident from the improved performance metrics for each Budyko equation when using the annual scale, as documented in Table 4. The Zhang equation, utilizing the annual  $\omega$  value, emerged as the top-performing method, achieving a robust R<sup>2</sup> of 0.78, an NSE of 0.77, an NRMSE of 0.10, and an RMSE of 3.39 mm. These results confirm that the Budyko models, particularly the Zhang equation, can reliably predict catchment-scale ET<sub>a</sub> values that are highly compatible with EC tower observations. The consistent and stable performance of the Budyko models, achieving an R<sup>2</sup> above 0.77 and an NSE greater than 0.75 across the entire study period, aligns with the literature that values the framework for its strong physical basis and its ability to effectively estimate monthly ET<sub>a</sub> with minimal input data (Sun et al., 2025). This consistency underscores the inherent stability and reliability of the Budyko models when applied to new conditions.

In contrast, the machine learning (ML) models demonstrated exceptional performance during the calibration phase, as detailed in Table 5. For instance, the XGB7 model achieved a robust R<sup>2</sup> of 0.95 and an NSE of 0.94. The scatter plots presented in Fig. 11 provide a visual



**Fig. 9.** SHAP Beeswarm plots for various models and selected features (the abbreviations are presented in Table A.3).

**Table 4**

Performance evaluation of monthly ET<sub>a</sub> estimation using different Budyko equations with monthly and annual  $\omega$  parameters.

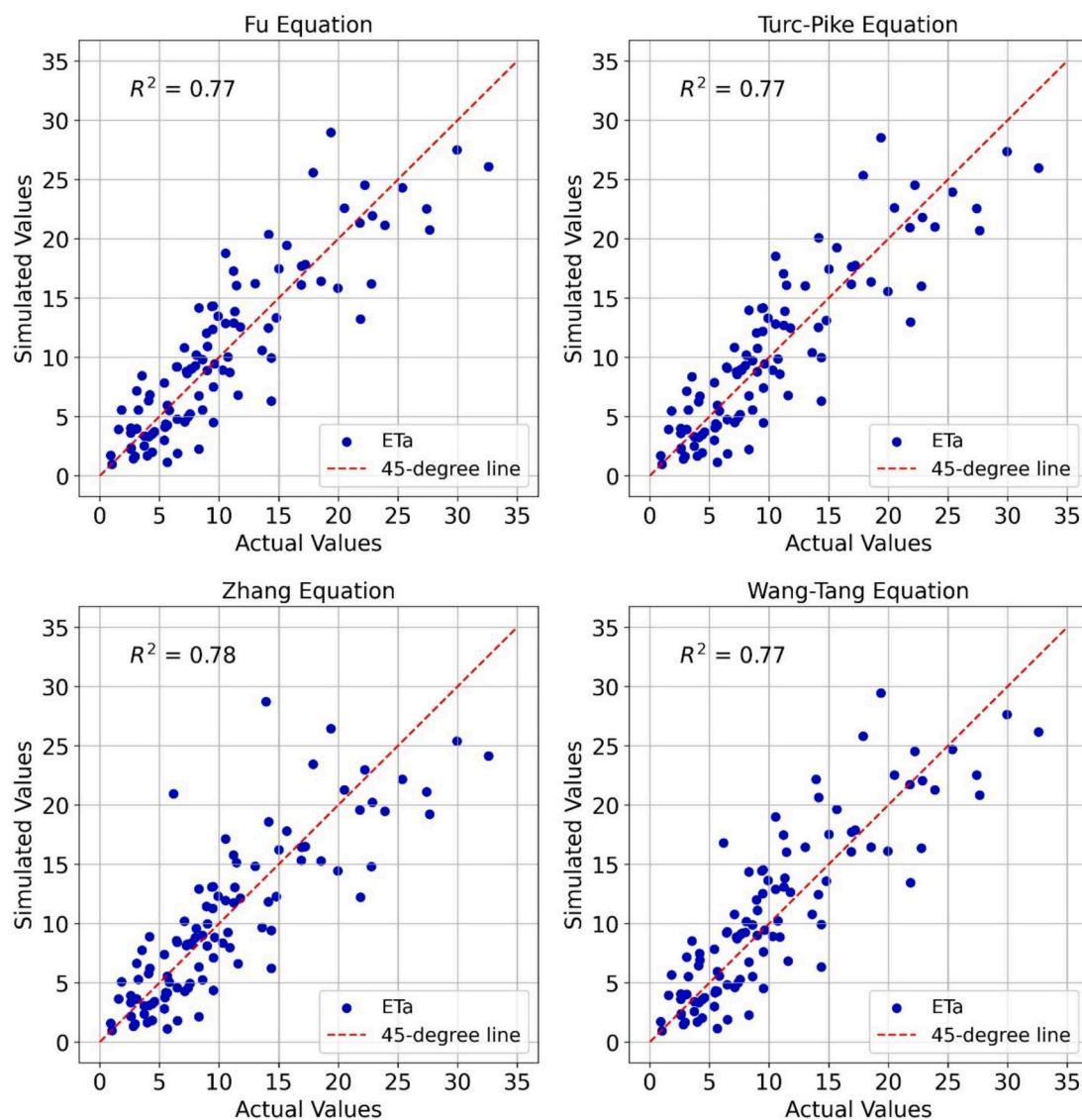
Model	Monthly $\omega$				Annual $\omega$			
	R <sup>2</sup>	NRMSE	NSE	RMSE (mm)	R <sup>2</sup>	NRMSE	NSE	RMSE (mm)
Fu	0.74	0.16	0.50	4.98	0.77	0.11	0.76	3.46
Turc-Pike	0.73	0.16	0.49	4.98	0.77	0.10	0.76	3.44
Zhang	0.75	0.25	-0.33	8.09	0.78	0.10	0.77	3.39
Wang-Tang	0.74	0.15	0.51	4.89	0.77	0.11	0.75	3.49

comparison of the four best-performing ML models (XGB7, RF3, GB5, and Ada6) against observed data, highlighting a distinct shift in performance between the calibration and validation phases. While the ML models performed exceptionally well during calibration, a notable decrease in accuracy was observed during validation. The performance metrics for the XGB7 model, for example, declined to an R<sup>2</sup> of 0.73 and an NSE of 0.73, with other models exhibiting similar patterns. This performance degradation suggests that while pure ML models excel at learning complex, non-linear relationships within a training dataset, they may lack the stability and generalization capabilities inherent to physically-based models like the Budyko framework when confronted

with new data. This is a well-documented limitation in the literature, where pure ML models are often found to have less robustness and generalization under certain conditions, such as sparse data or extreme events, due to their lack of a strong physical foundation (Zhang et al., 2025). This finding validates our observation that, while powerful during the calibration phase, these models may struggle to maintain accuracy when applied to a new set of data.

### 3.4. Hybrid Budyko-ML model performance

The hybrid modeling approach, which integrated the modified



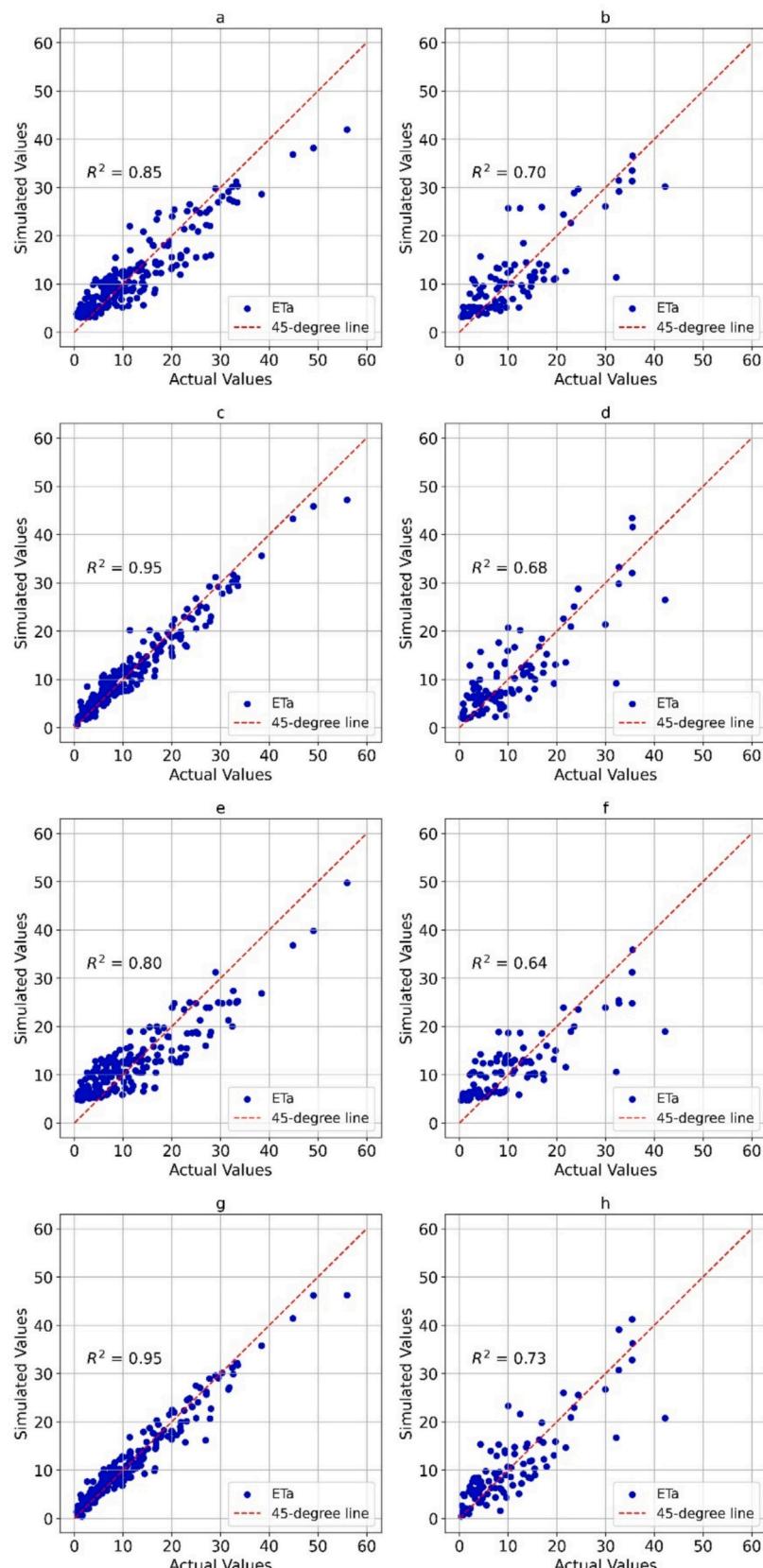
**Fig. 10.** Scatter plot of observed (EC-ET) versus predicted ET<sub>a</sub> values for each equation (Fu, Turc-Pike, Zhang, and Wang-Tang equation, respectively) of the Budyko model (Table 1). The 1:1 line (red) indicates perfect prediction.

**Table 5**  
Performance of the best modified Budyko Model vs. ML models.

Model	R <sup>2</sup>	NRMSE	NSE	RMSE (mm)
Zhang Equation	0.78	0.10	0.77	3.39
XGB7	0.73	0.11	0.73	4.82
RF3	0.70	0.12	0.70	5.28
GB5	0.68	0.13	0.68	5.65
Ada6	0.62	0.14	0.60	5.87

Budyko framework with the optimal machine learning model (XGBoost) using deep learning (DL), significantly improved ET<sub>a</sub> estimation performance, as detailed in Table 6 and shown in Fig. 12. This strategic application of a hybrid approach, combining the Budyko model and machine learning model like XGBoost through a deep learning-based approach, was instrumental in depicting the nonlinear dynamics and complexities of the Central Valley water balance, particularly in complex, data-scarce environments where the assumptions of steady-state conditions are often neglected. The enhanced performance is quantitatively demonstrated by elevated R<sup>2</sup> and NSE, coupled with reduced RMSE and NRMSE values for all hybrid configurations (Table 6). Our

study represents a pioneering effort in coupling a modified Budyko framework with an advanced machine learning model by employing deep learning (DL). Our findings are further corroborated by regional research, such as that by Ahmadi et al. (2025), which demonstrated that applying global learning methodologies to DL models noticeably enhanced forecasting performance and generalization of reference evapotranspiration (ET<sub>0</sub>), even for ungauged basins. Our approach aligns with a growing body of research that supports the use of hybrid models to achieve an optimal balance between physical mechanisms and high estimation accuracy (Zhang et al., 2025). Our findings are also consistent with the literature, which notes that while pure Artificial Intelligence (AI) offers promising alternatives to traditional methods, they can face challenges such as overfitting and a lack of integration with physical processes (Taheri et al., 2025). The approach we used was specifically designed to overcome these limitations. Further research verifies this, highlighting that pure ML methods can be inadequate in data-scarce areas and that future studies should aim to integrate physical processes with data-driven models to improve generalization (Xue et al., 2025). The superior performance of our hybrid model in capturing complex non-linear relationships and improving generalization capabilities is further supported by Tosan et al. (2025). They emphasized the



**Fig. 11.** Scatter plots of predicted versus observed monthly ET<sub>a</sub> values for four top-performing machine learning (ML) models during the calibration and validation phases. The red dashed line represents the 1:1 line of perfect agreement. Subplots illustrate the performance of a: RF model in Calibration, b: RF model in Validation, c: GB model in Calibration, d: GB model in Validation, e: Ada model in Calibration, f: Ada model in Validation, g: XGB model in Calibration, h: XGB model in Validation.

**Table 6**  
Performance of the best hybrid models.

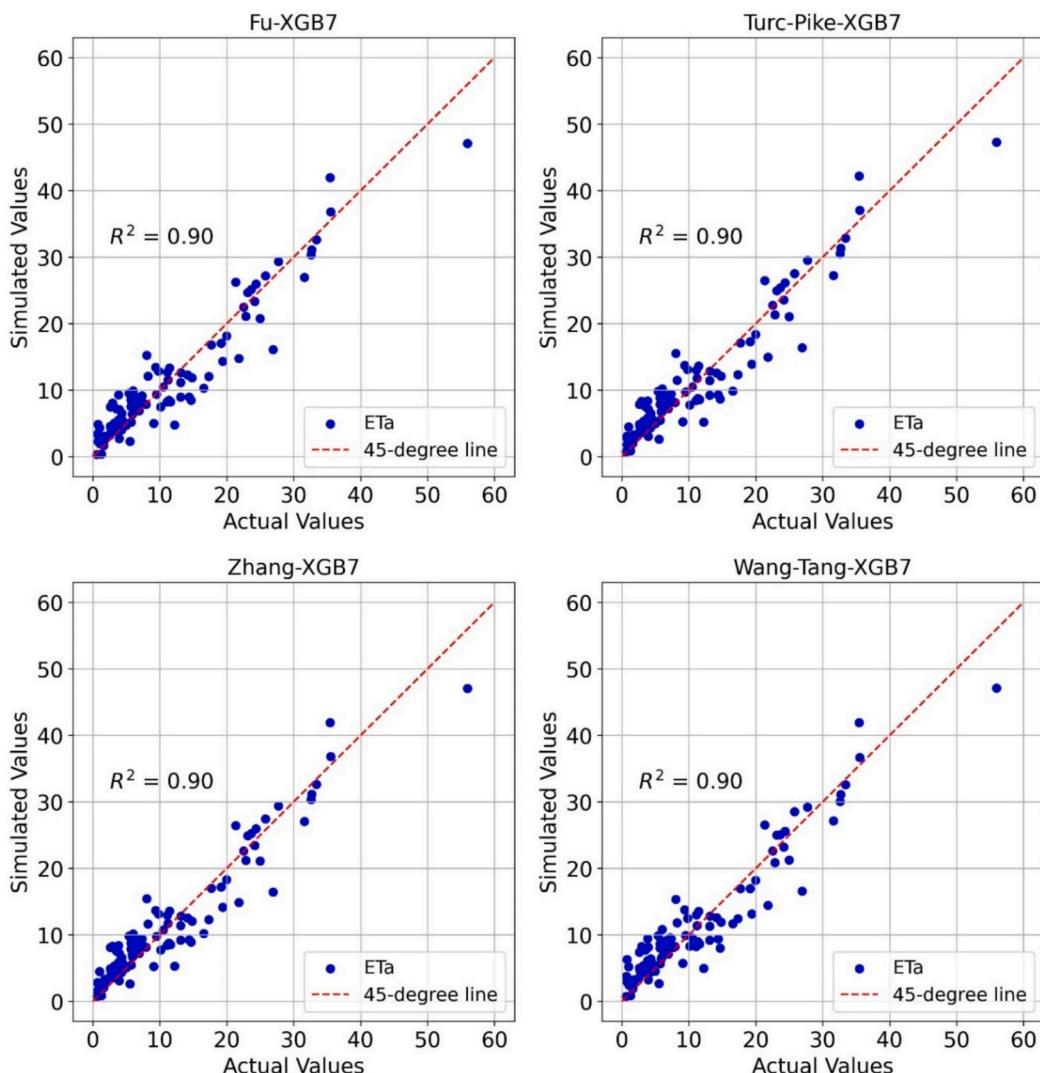
Model	R <sup>2</sup>	NRMSE	NSE	RMSE (mm)
Wang-Tang-XGB7	0.90	0.06	0.90	3.21
Turc-pike-XGB7	0.90	0.06	0.90	3.23
Zhang-XGB7	0.90	0.06	0.90	3.27
Fu-XGB7	0.89	0.06	0.88	3.59

transformative potential of Hybrid Artificial Neural Networks (HANNs) for precise evapotranspiration estimation, particularly in water-scarce and climate-vulnerable regions. By integrating Convolutional Neural Networks (CNNs) for automatic feature extraction and leveraging these hybrid architectures, HANNs offered substantial advantages for optimizing water resource management, especially within the agricultural sector.

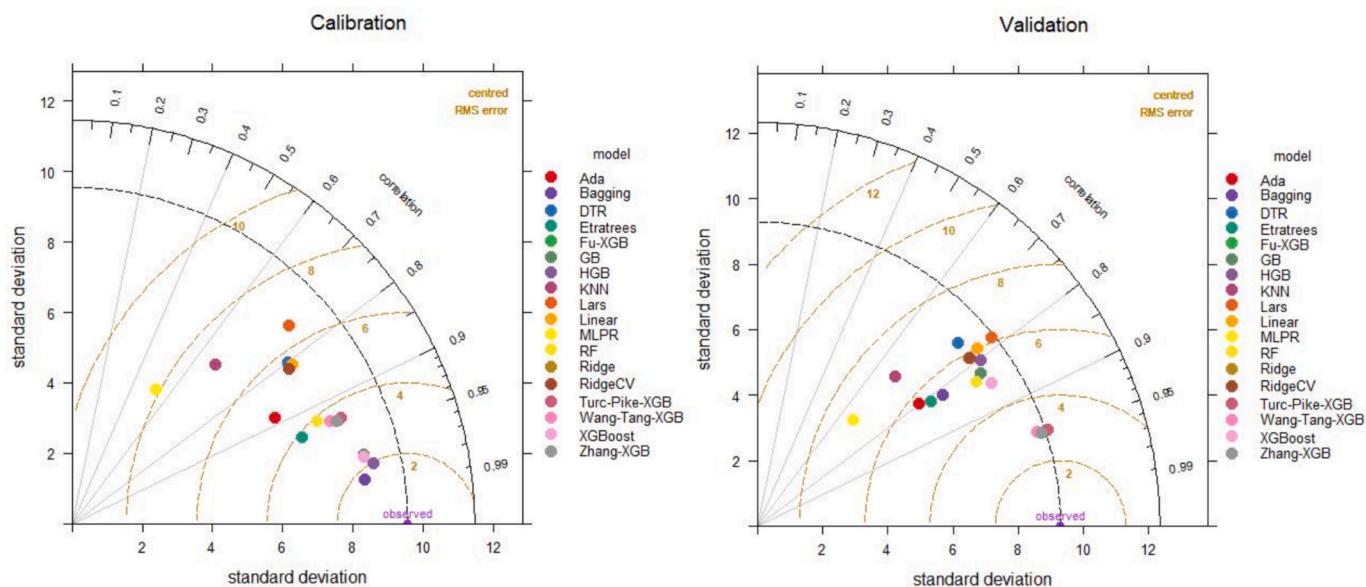
### 3.5. Visual comparison of models' performance

The Taylor diagram is a powerful graphical tool used to visually evaluate the performance of different models by comparing them to a set of observed data. It quantifies the degree of correspondence between modeled and observed behavior across three key statistical metrics: the Pearson correlation coefficient ( $R$ ), RMSE, and the standard deviation

( $\sigma$ ). The proximity of a model's point to the observed data point on the diagram signifies higher accuracy and better model performance. The Taylor diagrams presented in Fig. 13 illustrate the comparative performance of various models during both the calibration and validation phases. During the calibration phase, the pure machine learning (ML) models—specifically XGBoost, GB, HGB, and Bagging—demonstrated the strongest performance. These models are clustered closest to the observed data point, exhibiting high correlation coefficients and a standard deviation closely matching the observations. In the validation phase, some pure ML models continued to show good performance relative to other models, with their correlation coefficients ranging from approximately 0.80 to 0.85 (XGB, RF, GB, and Ada). However, the three hybrid models (Turc-Pike-XGB, Wang-Tang-XGB, and Zhang-XGB) showed a significantly better fit. These models are located closest to the observed point, indicating a substantially higher degree of accuracy and a better match in all three statistical metrics. Their performance, with a correlation coefficient approaching 0.99, demonstrates the superior predictive capability and generalization of the hybrid approach. This finding is consistent with other studies that have shown hybrid models can significantly reduce prediction errors compared to pure machine learning models (Zhang et al., 2025). This empirically demonstrates that the hybrid approach successfully overcomes the inherent limitations of both physically-based and pure data-driven models, providing a more robust and reliable predictive capability when applied



**Fig. 12.** Scatter plot of observed (EC-ET) versus predicted  $ET_a$  values for each hybrid Budyko-XGB7 model. The 1:1 red line indicates perfect prediction.



**Fig. 13.** Combined Taylor diagram illustrating the performance of four Budyko-based models, fourteen pure ML models, and four hybrid models for monthly  $ET_a$  estimation during the calibration and validation phases.

to new data.

#### 4. Conclusions

Accurate and reliable estimation of actual evapotranspiration ( $ET_a$ ) is vital to effective water management strategies in irrigated agriculture and hydrology. However, a significant challenge exists in developing countries and data-scarce regions where insufficient or unreliable climatic data impedes our ability to manage water resources effectively, especially in the face of escalating climate change. This study addressed these challenges by introducing a novel hybrid modeling approach that integrates a modified Budyko framework with the optimal machine learning model (XGBoost) using deep learning. While the standalone modified Budyko model provided a reasonable estimation of  $ET_a$ , which was relatively compatible with Eddy Covariance observations, the hybrid approach consistently and significantly outperformed all stand-alone models. This demonstrates the superior ability of the developed hybrid model to capture complex, non-linear relationships and improve predictive accuracy. This is particularly important because both conceptual/physical and data-driven models, when used alone, possess inherent weaknesses that are effectively mitigated by their combination.

Furthermore, this research introduced key methodological innovations, including the optimization of the Budyko parameter's temporal scale, which significantly improved model performance by accounting for non-steady-state hydrological conditions. A detailed feature importance analysis using SHAP values also underscored the crucial roles of climatic (precipitation, solar radiation), soil (soil moisture dynamics), and vegetation (NDVI, LAI) factors as primary drivers of  $ET_a$ .

In conclusion, this study presents a highly accurate, reliable, and pragmatic method for  $ET_a$  estimation that leverages globally available remote sensing data. This scalable and cost-effective solution serves as a universally applicable tool for sustainable water management, providing a critical resource for arid and data-scarce regions worldwide to address water-related challenges in a climate-constrained era.

#### 5. Further studies

This study lays the foundational groundwork for a novel methodology, and its findings warrant further investigation to establish its

broader applicability and scientific validity. While the current research substantiates the efficacy of the Budyko-XGBoost hybrid model in the specific hydro-climatic regime of the Central Valley, future work is essential to rigorously test its performance and generalizability across diverse climatic conditions, ranging from arid deserts to humid tropical environments. Should this method consistently yield reliable and accurate results across these varied conditions, its scientific robustness and reliability would be established, allowing it to serve as a universal benchmark for  $ET_a$  estimation. The successful application of this model could have transformative implications, particularly in developing countries and data-scarce regions where extensive in-situ data is unavailable. By providing a robust and data-efficient method for  $ET_a$  estimation, this research would significantly contribute to more accurate water resource management, improved hydrological modeling, and a better understanding of the water-energy balance in a changing climate (Hasheminia, 2025; personal communication).

#### CRediT authorship contribution statement

**Mahdi Mohammadnezhad:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation. **Kamran Davary:** Writing – original draft, Validation, Supervision, Project administration, Methodology, Conceptualization. **Pooya Shirazi:** Validation, Supervision, Software, Investigation. **Mohammad Javad Rezvanpour:** Validation, Software, Methodology, Investigation, Formal analysis. **Seyed Majid Hasheminia:** Writing – review & editing, Writing – original draft, Validation, Supervision, Formal analysis, Conceptualization.

#### Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

#### Declaration of competing interest

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

## Acknowledgments

We extend our deepest gratitude to Dr. Kamran Davary, a distinguished faculty member of the Water Science and Engineering

Department and the esteemed Head of the Water and Environment Research Institute at Ferdowsi University of Mashhad. His insightful guidance and profound intellectual contributions were instrumental in the conception and successful completion of this research.

## Appendix A

**Table A.1**

Characteristics of the Eddy Covariance (EC) towers in California's Central Valley. The table provides detailed information for the available observation period, including the general land cover classification of each station as either cropland (CL) or wetland/riparian (WR).

Site ID	Data source	Period	Latitude	Longitude	Elevation	General classification	Land cover details
US-Twt	AmeriFlux	2009–2017	38.10	-121.65	-7	CL	Rice
US-Tw3	AmeriFlux	2013–2018	38.11	-121.64	-9	CL	Alfalfa
US-Bi1	AmeriFlux	2016–2019	38.09	-121.49	-2.7	CL	Alfalfa
US-Bi2	AmeriFlux	2017–2020	38.10	-121.53	-5	CL	Corn
US-Tw2	AmeriFlux	2012–2013	38.10	-121.64	-5	CL	Corn on peat soil
MB_Pch	CSUMB	2012–2015	36.45	-119.58	90	CL	Peach
Almond_High	USDA-ARS	2016–2019	36.16	-120.20	147	CL	Almond
Almond_Low	USDA-ARS	2016–2019	36.94	-120.10	78	CL	Almond
Almond_Med	USDA-ARS	2016–2019	36.17	-120.20	147	CL	Almond
RIP760	USDA-ARS GRAPEX	2017–2018	36.83	-120.21	57	CL	Vineyard
SLM001	USDA-ARS GRAPEX	2017–2018	38.28	-121.11	39	CL	Vineyard
US-Sne	AmeriFlux	2016–2019	38.03	-121.75	-5	WR	Restored wetland
US-Srr	AmeriFlux	2016–2017	38.20	-122.02	8	WR	Brackish tidal marsh

Data were sourced from the AmeriFlux network. (<https://ameriflux.lbl.gov/>).

**Table A.2**

Predefined grids of hyperparameter values for the machine learning models utilized in this study.

Algorithm	Hyperparameter	Values	Reference
Random Forest Regressor	Maximum depth	2, 4, 6, 8, 10, None	Sun et al. (2020)
	Number of trees	10, 50, 100, 200	
	Minimum samples for split	2, 3, 4, 5	
	Best nodes (impurity reduction)	5, 10, 20, None	
	Bootstrap randomness control	22, None	
AdaBoost Regressor	Maximum estimators	10, 50, 100, 200	Chandran and Chithra (2025)
	Regressor weight	0.25, 0.5, 0.75, 1, 10, 50, 100	
	Loss function	linear, square, exponential	
Gradient Boosting Regressor	Learning rate	0.1, 0.25, 0.5, 0.75, 1, 10, 50, 100	Musyimi et al. (2022)
	Sample fraction	0.25, 0.5, 0.75, 1	
	Boosting stages	10, 50, 100, 200	
	Maximum depth	2, 3, 4, 6, 8, 10, None	
Extreme Gradient Boosting	Number of boosting rounds	50, 100, 150	Acharki et al. (2025)
	Maximum depth of a tree	3, 5, 7	
	Step size shrinkage	0.1, 0.01	
	Minimum loss reduction required to make a further partition on a leaf node	0, 0.1, 0.2	
	Subsample ratio of the training instances	0.8, 1	
	Subsample ratio of columns when constructing each tree	0.8, 1	

**Table A.3**

All parameters used in this study.

Number	Parameter	Description	Units	Dataset
1	ID	Stations ID		EC Towers
2	Year	Year		same
3	Month	Month		same
4	ET <sub>a</sub>	Observed ActualET in EC Towers	mm	ECTowers
5	ppt	Observed Precipitationin EC Towers	mm	ECTowers
6	eto_hargreaves	Calculated reference ET with EC Towers Data	mm	Calculated
7	lat	Latitude	degree	same
8	lon	Longitude	degree	same
9	dew2	Temperature to which the air, at 2 m above the surface of the Earth, would have to be cooled for saturation to occur.	K	ERA5
10	T_2m	Temperature of air at 2 m above the surface of land, sea or in-land waters.	K	ERA5
11	soilT_1	Temperature of the soil in layer 1 (0–7 cm) of the ECMWF Integrated Forecasting System.	K	ERA5
12	soilT_2	Temperature of the soil in layer 2 (7–28 cm) of the ECMWF Integrated Forecasting System.	K	ERA5

(continued on next page)

Table A.3 (continued)

Number	Parameter	Description	Units	Dataset
13	soilT_3	Temperature of the soil in layer 3 (28–100 cm) of the ECMWF Integrated Forecasting System.	K	ERA5
14	soilT_4	Temperature of the soil in layer 4 (100–289 cm) of the ECMWF Integrated Forecasting System.	K	ERA5
15	soil_water_1	Volume of water in soil layer 1 (0–7 cm) of the ECMWF Integrated Forecasting System.	Volume fraction	ERA5
16	soil_water_2	Volume of water in soil layer 2 (7–28 cm) of the ECMWF Integrated Forecasting System.	Volume fraction	ERA5
17	soil_water_3	Volume of water in soil layer 3 (28–100 cm) of the ECMWF Integrated Forecasting System.	Volume fraction	ERA5
18	soil_water_4	Volume of water in soil layer 4 (100–289 cm) of the ECMWF Integrated Forecasting System.	Volume fraction	ERA5
19	albedo	Is a measure of the reflectivity of the Earth's surface.		
20	r_sum	Amount of solar radiation (also known as shortwave radiation)	J/m <sup>2</sup>	ERA5
21	s_pressure	Pressure (force per unit area) of the atmosphere on the surface of land, sea and in-land water.	Pa	ERA5
22	lai_high	One-half of the total green leaf area per unit horizontal ground surface area for high vegetation type.	Area fraction	ERA5
23	lai_low	One-half of the total green leaf area per unit horizontal ground surface area for low vegetation type.	Area fraction	ERA5
24	dew2min	minimum dewpoint_temperature_2m value each month	K	ERA5
25	dew2max	maximum dewpoint_temperature_2m value each month	K	ERA5
26	T_2m_min	minimum temperature_2m value each month	K	ERA5
27	T_2m_max	maximum temperature_2m value each month	K	ERA5
28	r_max	maximum surface_net_solar_radiation value each month	J/m <sup>2</sup>	ERA5
29	ss_flux_max	maximum surface_sensible_heat_flux value each month	J/m <sup>2</sup>	ERA5
30	ss_r_d_max	maximum surface_solar_radiation_downwards value each month	J/m <sup>2</sup>	ERA5
31	st_r_d_min	minimum surface_thermal_radiation_downwards value each month	J/m <sup>2</sup>	ERA5
32	st_r_d_max	maximum surface_thermal_radiation_downwards value each month	J/m <sup>2</sup>	ERA5
33	u_wind10max	maximum u_component_of_wind_10m value each month	m/s	ERA5
34	v_wind10max	maximum v_component_of_wind_10m value each month	m/s	ERA5
35	s_p_min	minimum surface_pressure value each month	Pa	ERA5
36	s_p_max	maximum surface_pressure value each month	Pa	ERA5
37	pr_max	maximum total_precipitation value each month	m	ERA5
38	date	Date		same
39	ET <sub>a</sub>	Actual evapotranspiration, derived using a one-dimensional soil water balance model	mm	TerraClimate
40	Wdef	Climate water deficit, derived using a one-dimensional soil water balance model	mm	TerraClimate
41	pdsi	Palmer Drought Severity Index		TerraClimate
42	PET	Reference evapotranspiration (ASCE Penman-Montieth)	mm	TerraClimate
43	pr	Precipitation accumulation	mm	TerraClimate
44	ro	Runoff, derived using a one-dimensional soil water balance model	mm	TerraClimate
45	soil	Soil moisture, derived using a one-dimensional soil water balance model	mm	TerraClimate
46	Srad	Downward surface shortwave radiation	W/m <sup>2</sup>	TerraClimate
47	Tmax	Minimum temperature	°C	TerraClimate
48	Tmin	Maximum temperature	°C	TerraClimate
49	vap	Vapor pressure	kPa	TerraClimate
50	Vpd	Vapor pressure deficit	kPa	TerraClimate
51	vs	Wind-speed at 10 m	m/s	TerraClimate
52	delta_s	Terrestrial Water Storage Change	mm	Calculated
53	time	Date		same
54	Site_ID	Factorized Site_ID		Calculated
55	General_class	Factorized General_classification		Calculated
56	LC_detail	Factorized Land_cover_details		Calculated
57	LC_type	Factorized Land_cover_type		Calculated

## Data availability

The pre-processed input data and scripts employed in the analysis, modeling, and visualization are available at a GitHub repository ([https://github.com/mahdimohammadnezhad/Estimation\\_ETa](https://github.com/mahdimohammadnezhad/Estimation_ETa)). All data used in this study are available from their respective sources. For further guidance on data preparation, please get in touch with the corresponding author.

## References

- Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Sci. Data 5 (1), 170191. <https://doi.org/10.1038/sdata.2017.191>.
- Acharki, S., Raza, A., Vishwakarma, D.K., Amharref, M., Bernoussi, A.S., Singh, S.K., Al-Ansari, N., Dewidar, A.Z., Al-Othman, A.A., Mattar, M.A., 2025. Comparative assessment of empirical and hybrid machine learning models for estimating daily reference evapotranspiration in sub-humid and semi-arid climates. Sci. Rep. 15 (1), 2542. <https://doi.org/10.1038/s41598-024-83859-6>.
- Ahmadi, A., Daccache, A., He, M., Namadi, P., Bafti, A.G., Sandhu, P., Bai, Z., Snyder, R.L., Kadir, T., 2025. Enhancing the accuracy and generalizability of reference evapotranspiration forecasting in California using deep global learning. J. Hydrol.: Reg. Stud. 59, 102339.
- Alton, P.B., 2016. The sensitivity of models of gross primary productivity to meteorological and leaf area forcing: a comparison between a Penman–Monteith ecophysiological approach and the MODIS light-use efficiency algorithm. Agric. For. Meteorol. 218, 11–24.
- Amani, S., Shafizadeh-Moghadam, H., 2023. A review of machine learning models and influential factors for estimating evapotranspiration using remote sensing and ground-based data. Agric. Water Manag. 284, 108324. <https://doi.org/10.1016/j.agwat.2023.108324>.
- Awada, H., Di Prima, S., Sirca, C., Giadrossich, F., Marras, S., Spano, D., Pirastru, M., 2022. A remote sensing and modeling integrated approach for constructing continuous time series of daily actual evapotranspiration. Agric. Water Manag. 260, 107320.
- Baboli, N., Ghamarnia, H., Hafezparast Mavaddat, M., 2024. Estimating wheat evapotranspiration through remote sensing utilizing GeeSEBAL and comparing with lysimetric data. Appl Water Sci 14 (9), 193. <https://doi.org/10.1007/s13201-024-02248-6>.
- Bachour, R., Walker, W.R., Ticlavilca, A.M., McKee, M., Maslova, I., 2014. Estimation of spatially distributed evapotranspiration using remote sensing and a relevance vector machine. J. Irrig. Drain. Eng. 140 (8), 4014029.
- Barkhordari, H., Dastjerdi, P.A., Nasseri, M., 2025. Development of a framework for estimating regional gridded streamflow and actual evapotranspiration datasets: fusing Budyko and water balance closure methods using remotely sensed ancillary data. J. Hydrol. 133456 <https://doi.org/10.1016/j.jhydrol.2025.133456>.

- Başagaoğlu, H., Chakraborty, D., Winterle, J., 2021. Reliable evapotranspiration predictions with a probabilistic machine learning framework. *Water* 13 (4), 557.
- Bonacci, O., Andrić, I., 2010. Impact of an inter-basin water transfer and reservoir operation on a karst open streamflow hydrological regime: an example from the Dinaric karst (Croatia). *Hydrol. Process.* 24 (26), 3852–3863.
- Boussetta, S., Balsamo, G., Beljaars, A., Kral, T., Jarlan, L., 2013. Impact of a satellite-derived leaf area index monthly climatology in a global numerical weather prediction model. *Int. J. Remote Sens.* 34 (9–10), 3520–3542.
- Budyko, M.I., 1958. The Heat Balance of the Earth's Surface. US Department of Commerce.
- Budyko, M.I., 1984. *Evaporation Under Natural Conditions*. Gidrometeorizdat, St. Petersburg, Russia, 8335.
- Chandran, D., Chithra, N.R., 2025. Predictive performance of ensemble learning boosting techniques in daily streamflow simulation. *Water Resour. Manag.* 39 (3), 1235–1259. <https://doi.org/10.1007/s11269-024-04029-x>.
- Cheng, S., Cheng, L., Qin, S., Zhang, L., Liu, P., Liu, L., Xu, Z., Wang, Q., 2022. Improved understanding of how catchment properties control hydrological partitioning through machine learning. *Water Resour. Res.* 58 (4), e2021WR031412. <https://doi.org/10.1029/2021WR031412>.
- Costa, T.S., Filgueiras, R., Dos Santos, R.A., Cunha, F.F. da, 2023. Actual evapotranspiration by machine learning and remote sensing without the thermal spectrum. *PLoS One* 18 (5), e0285535.
- Cutting, N.G., Kaur, S., Singh, M.C., Sharma, N., Mishra, A., 2024. Estimating Crop Evapotranspiration in data-scarce regions: a comparative analysis of eddy covariance, empirical and remote-sensing approaches. *Water Conserv. Sci. Eng.* 9 (2), 65.
- Dou, X., Yang, Y., 2018. Evapotranspiration estimation using four different machine learning approaches in different terrestrial ecosystems. *Comput. Electron. Agric.* 148, 95–106. <https://doi.org/10.1016/j.compag.2018.03.010>.
- Du, C., Sun, F., Yu, J., Liu, X., Chen, Y., 2016. New interpretation of the role of water balance in an extended Budyko hypothesis in arid regions. *Hydrol. Earth Syst. Sci.* 20 (1), 393–409. <https://doi.org/10.5194/hess-20-393-2016>.
- Elbeltagi, A., Heddam, S., Katipoğlu, O.M., Alsumaie, A.A., Al-Mukhtar, M., 2024. Advanced long-term actual evapotranspiration estimation in humid climates for 1958–2021 based on machine learning models enhanced by the RReliefF algorithm. *J. Hydrol.: Reg. Stud.* 56, 102043. <https://doi.org/10.1016/j.ejrh.2024.102043>.
- Faunt, C.C., Hanson, R., Belitz, K., Schmid, W., Predmore, S.P., Rewis, D.L., McPherson, K., 2009. Groundwater availability of the central valley aquifer, California, professional paper 1766. In: US Geological Survey, 10.
- Faunt, C.C., Traum, J.A., Boyce, S.E., Seymour, W.A., Jachens, E.R., Brandt, J.T., Snead, M., Bond, S., Marcelli, M.F., 2024. Groundwater sustainability and land subsidence in California's Central Valley. *Water* 16 (8). <https://doi.org/10.3390/w16081189>.
- Feng, Y., Jia, Y., Cui, N., Zhao, L., Li, C., Gong, D., 2017a. Calibration of Hargreaves model for reference evapotranspiration estimation in Sichuan basin of Southwest China. *Agric. Water Manag.* 181, 1–9. <https://doi.org/10.1016/j.agwat.2016.11.010>.
- Feng, Y., Peng, Y., Cui, N., Gong, D., Zhang, K., 2017b. Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. *Comput. Electron. Agric.* 136, 71–78. <https://doi.org/10.1016/j.compag.2017.01.027>.
- Fu, B.P., 1981. On the calculation of the evaporation from land surface. *Sci. Atmos. Sin.* 5 (1), 23.
- Fuentes, S., Ortega-Farfás, S., Carrasco-Benavides, M., Tongson, E., Viejo, C.G., 2024. Actual evapotranspiration and energy balance estimation from vineyards using micro-meteorological data and machine learning modeling. *Agric. Water Manag.* 297, 108834.
- Gao, H., Hrachowitz, M., Schymanski, S.J., Fenicia, F., Sriwongsitanon, N., Savenije, H.G., 2014. Climate controls how ecosystems size the root zone storage capacity at catchment scale. *Geophys. Res. Lett.* 41 (22), 7916–7923.
- Gerrits, A.M.J., Savenije, H.H.G., Veling, E.J.M., Pfister, L., 2009. Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model. *Water Resour. Res.* 45 (4). <https://doi.org/10.1029/2008WR007308>.
- Gharbia, S.S., Smullen, T., Gill, L., Johnston, P., Pilla, F., 2018. Spatially distributed potential evapotranspiration modeling and climate projections. *Sci. Total Environ.* 633, 571–592.
- Ghiat, I., Mackey, H.R., Al-Ansari, T., 2021. A review of evapotranspiration measurement models, techniques and methods for open and closed agricultural field applications. *Water* 13 (18), 2523.
- Goyal, M.K., Sharma, A., Katsifarakis, K.L., 2017. Prediction of flow rate of karstic springs using support vector machines. *Hydrol. Sci. J.* 62 (13), 2175–2186.
- Granata, F., 2019. Evapotranspiration evaluation models based on machine learning algorithms—a comparative study. *Agric. Water Manag.* 217, 303–315. <https://doi.org/10.1016/j.agwat.2019.03.015>.
- Granata, F., Papirio, S., Esposito, G., Gargano, R., De Marinis, G., 2017. Machine learning algorithms for the forecasting of wastewater quality indicators. *Water* 9 (2). <https://doi.org/10.3390/w9020105>.
- Gunkel, A., Lange, J., 2017. Water scarcity, data scarcity and the Budyko curve—an application in the Lower Jordan River Basin. *J. Hydrol.: Reg. Stud.* 12, 136–149. <https://doi.org/10.1016/j.ejrh.2017.04.004>.
- Hao, Y., Baik, J., Tran, H., Choi, M., 2022. Quantification of the effect of hydrological drivers on actual evapotranspiration using the Bayesian model averaging approach for various landscapes over Northeast Asia. *J. Hydrol.* 607, 127543. <https://doi.org/10.1016/j.jhydrol.2022.127543>.
- Huang, P., Song, J., Wu, Q., Sun, H., Mao, R., Cheng, D., Zhang, J., Shi, Y., 2022. The ecohydrological effects of climate and landscape interactions within the Budyko framework under non-steady state conditions. *Catena* 217, 106481.
- Istanbulluoglu, E., Wang, T., Wright, O.M., Lenters, J.D., 2012. Interpretation of hydrologic trends from a water balance perspective: the role of groundwater storage in the Budyko hypothesis. *Water Resour. Res.* 48 (3).
- Jiang, C., Xiong, L., Wang, D., Liu, P., Guo, S., Xu, C.-Y., 2015. Separating the impacts of climate change and human activities on runoff using the Budyko-type equations with time-varying parameters. *J. Hydrol.* 522, 326–338. <https://doi.org/10.1016/j.jhydrol.2014.12.060>.
- Karahan, H., Cetin, M., Can, M.E., Alsenjar, O., 2024. Developing a new ANN model to estimate daily actual evapotranspiration using limited climatic data and remote sensing techniques for sustainable water management. *Sustainability* 16 (6), 2481.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen-Geiger climate classification updated. *Meteorol. Z.* 15 (3), 259–263. <https://doi.org/10.1127/0941-2948/2006/0130>.
- Li, H., Sivapalan, M., Tian, F., Harman, C., 2014. Functional approach to exploring climatic and landscape controls of runoff generation: 1. Behavioral constraints on runoff volume. *Water Resour. Res.* 50 (12), 9300–9322.
- Liu, Y., Wang, Y., Zhao, Y., Chen, S., Wang, L., Yang, W., Li, X., Li, X., Lei, H., Chang, H., 2025. Evapotranspiration stress intensifies with enhanced sensitivity to soil moisture deficits in a rapidly greening China. *Hydrol. Earth Syst. Sci.* 29 (14), 3379–3404.
- Mehdizadeh, S., 2018. Estimation of daily reference evapotranspiration (ETo) using artificial intelligence methods: offering a new approach for lagged ETo data-based modeling. *J. Hydrol.* 559, 794–812. <https://doi.org/10.1016/j.jhydrol.2018.02.060>.
- Mianabadi, A., Davary, K., Pourreza-Bilondi, M., Coenders-Gerrits, A.M.J., 2020. Budyko framework: towards non-steady state conditions. *J. Hydrol.* 588, 125089. <https://doi.org/10.1016/j.jhydrol.2020.125089>.
- Milly, P.C.D., Dunne, K.A., 2002. Macroscale water fluxes 2. Water and energy supply control of their interannual variability. *Water Resour. Res.* 38 (10), 21–24.
- Molden, D., Oweis, T., Steduto, P., Bindraban, P., Hanjra, M.A., Kijne, J., 2010. Improving agricultural water productivity: between optimism and caution. *Agric. Water Manag.* 97 (4), 528–535.
- Moussa, R., Lhomme, J.-P., 2016. The Budyko functions under non-steady-state conditions. *Hydrol. Earth Syst. Sci.* 20 (12), 4867–4879. <https://doi.org/10.5194/hess-20-4867-2016>.
- Muñoz-Sabater, J., Dutra, E., Agusti-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* 13 (9), 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>.
- Musyimi, P.K., Sahbeni, G., Timár, G., Weidinger, T., Székely, B., 2022. Actual evapotranspiration estimation using Sentinel-1 SAR and Sentinel-3 SLSTR data combined with a gradient boosting machine model in Busia county, Western Kenya. *Atmosphere* 13 (11), 1927.
- Nema, M.K., Khare, D., Chandniha, S.K., 2017. Application of artificial intelligence to estimate the reference evapotranspiration in sub-humid Doon valley. *Appl Water Sci* 7 (7), 3903–3910. <https://doi.org/10.1007/s13201-017-0543-3>.
- Ol'dekop, E.M., 1911. On evaporation from the surface of river basins. *Trans. Meteorol. Obs.* 4, 200.
- Pagano, A., Amato, F., Ippolito, M., De Caro, D., Croce, D., Motisi, A., Provenzano, G., Tinnirello, I., 2023. Machine learning models to predict daily actual evapotranspiration of citrus orchards under regulated deficit irrigation. *Eco. Inform.* 76, 102133. <https://doi.org/10.1016/j.ecoinf.2023.102133>.
- Patil, A.P., Deka, P.C., 2016. An extreme learning machine approach for modeling evapotranspiration using extrinsic inputs. *Comput. Electron. Agric.* 121, 385–392.
- Pike, J.G., 1964. The estimation of annual run-off from meteorological data in a tropical climate. *J. Hydrol.* 2 (2), 116–123.
- Ravindran, S.M., Bhaskaran, S.K.M., Ambat, S.K.N., 2021. A deep neural network architecture to model reference evapotranspiration using a single input meteorological parameter. *Environ. Process.* 8 (4), 1567–1599.
- Rodríguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., Chica-Rivas, M., 2015. Machine learning predictive models for mineral prospectivity: an evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geol. Rev.* 71, 804–818.
- Schreiber, P., 1904. Über die Beziehungen zwischen dem Niederschlag und der Wasserführung der Flüsse in Mitteleuropa. *Z. Meteorol.* 21 (10), 441–452.
- Sharafi, S., Mohammadi Ghalei, M., 2024. Revealing accuracy in climate dynamics: enhancing evapotranspiration estimation using advanced quantile regression and machine learning models. *Appl Water Sci* 14 (7), 162. <https://doi.org/10.1007/s13201-024-02211-5>.
- Shrestha, N.K., Shukla, S., 2015. Support vector machine based modeling of evapotranspiration using hydro-climatic variables in a sub-tropical environment. *Agric. For. Meteorol.* 200, 172–184.
- Sun, D., Wen, H., Wang, D., Xu, J., 2020. A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. *Geomorphology* 362, 107201.
- Sun, H., Wang, X., Qu, Y., Lyu, J., Cai, H., Zhang, X., 2025. Interpretation of the Budyko framework for monthly evapotranspiration prediction of individual basins. *Atmos. Res.* 315, 107895.
- Taheri, M., Bigdeli, M., Imanian, H., Mohammadian, A., 2025. An overview of evapotranspiration estimation models utilizing artificial intelligence. *Water* 17 (9), 1384.
- Tang, D., Feng, Y., Gong, D., Hao, W., Cui, N., 2018. Evaluation of artificial intelligence models for actual crop evapotranspiration modeling in mulched and non-mulched maize croplands. *Comput. Electron. Agric.* 152, 375–384. <https://doi.org/10.1016/j.compag.2018.07.029>.

- Torres, A.F., Walker, W.R., McKee, M., 2011. Forecasting daily potential evapotranspiration using machine learning and limited climatic data. *Agric. Water Manag.* 98 (4), 553–562.
- Tosan, M., Gharib, M.R., Attar, N.F., Maroosi, A., 2025. Enhancing Evapotranspiration Estimation: A Bibliometric and Systematic Review of Hybrid Neural Networks in Water Resource Management, 142(2). *Computer Modeling in Engineering & Sciences (CMES)*.
- Turc, L., 1954. Le bilan d'eau des sols: relation entre la precipitations, l'évaporation et l'écoulement. *Ann. Agron. Série A* 5, 491.
- Wang, K., Dickinson, R.E., 2012. A review of global terrestrial evapotranspiration: observation, modeling, climatology, and climatic variability. *Rev. Geophys.* 50 (2).
- Wang, D., Tang, Y., 2014. A one-parameter Budyko model for water balance captures emergent behavior in darwinian hydrologic models. *Geophys. Res. Lett.* 41 (13), 4569–4577. <https://doi.org/10.1002/2014GL060509>.
- Williams, E.L., Abatzoglou, J.T., 2025. Climate change increases evaporative and crop irrigation demand in North America. *Earths Future* 13 (7), e2025EF005931. <https://doi.org/10.1029/2025EF005931>.
- Wu, C., Yeh, P., Xu, K., Huang, G., Wang, P., 2017. Modeling water balance using the Budyko framework over variable timescales under diverse climates. *Hydrol. Earth Syst. Sci. Discuss.* 1–33. <https://doi.org/10.5194/hess-2017-441>.
- Xiang, K., Li, Y., Horton, R., Feng, H., 2020. Similarity and difference of potential evapotranspiration and reference crop evapotranspiration—a review. *Agric. Water Manag.* 232, 106043.
- Xiong, L., Yu, K., Gottschalk, L., 2014. Estimation of the distribution of annual runoff from climatic variables using copulas. *Water Resour. Res.* 50 (9), 7134–7152. <https://doi.org/10.1002/2013WR015159>.
- Xu, T., Guo, Z., Liu, S., He, X., Meng, Y., Xu, Z., Xia, Y., Xiao, J., Zhang, Y., Ma, Y., Song, L., 2018. Evaluating different machine learning methods for upscaling evapotranspiration from flux towers to the regional scale. *J. Geophys. Res. Atmos.* 123 (16), 8674–8690. <https://doi.org/10.1029/2018JD028447>.
- Xue, Y., Zhang, Z., Li, X., Liang, H., Yin, L., 2025. A review of evapotranspiration estimation models: advances and future development. *Water Resour. Manag.* 1–17.
- Yang, D., Sun, F., Liu, Z., Cong, Z., Ni, G., Lei, Z., 2007. Analyzing spatial and temporal variability of annual water-energy balance in nonhumid regions of China using the Budyko hypothesis. *Water Resour. Res.* 43 (4).
- Yang, H., Yang, D., Lei, Z., Sun, F., 2008. New analytical derivation of the mean annual water-energy balance equation. *Water Resour. Res.* 44 (3). <https://doi.org/10.1029/2007WR006135>.
- Yang, L., Shi, L., Li, J., Kong, H., Wu, D., Wei, J., 2023. Effects of climatic conditions and vegetation changes on actual evapotranspiration in Mu Us sandy land. *Water Sci. Technol.* 88 (3), 723–737. <https://doi.org/10.2166/wst.2023.226>.
- Yin, D., Roderick, M.L., 2020. Inter-annual variability of the global terrestrial water cycle. *Hydrol. Earth Syst. Sci.* 24 (1), 381–396. <https://doi.org/10.5194/hess-24-381-2020>.
- Zhang, L., Dawes, W.R., Walker, G.R., 2001. Response of mean annual evapotranspiration to vegetation changes at catchment scale. *Water Resour. Res.* 37 (3), 701–708.
- Zhang, L., Hickel, K., Dawes, W.R., Chiew, F.H.S., Western, A.W., Briggs, P.R., 2004. A rational function approach for estimating mean annual evapotranspiration. *Water Resour. Res.* 40 (2).
- Zhang, C., Zhou, C., Luo, G., Ye, S., Shi, Z., 2025. Physics-constrained machine learning for satellite-derived evapotranspiration in China. *J. Hydrol.* <https://doi.org/10.1016/j.jhydrol.2025.133512>.