Data-driven algorithms to estimate Maize Sap Flow Transpiration based on climatic and soil moisture data

Grazia Tosi
University of Perugia
Marco Legittimo
marco.legittimo@unipg.it
University of Perugia
Francesco Crocetti
University of Perugia
Gabriele Costante
University of Perugia
Jennifer Bertuzzi
University of Perugia
Vergni Lorenzo
University of Perugia
Paolo Valigi
University of Perugia
Francesca Todisco
University of Perugia

Research Article

Keywords: Sap flow prediction, Deep learning, environmental data, soil water content

Posted Date: April 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9236382/v1

Additional Declarations: No competing interests reported.

Data-driven algorithms to estimate Maize Sap Flow Transpiration based on climatic and soil moisture data

Tosi G.[1,0000-0001-8244-8797], **Legittimo M.**[2,3,0000-0002-4532-7678], **Crocetti F.**[2,0000-0001-9134-8368], **Costante G.**[2,0000-0002-8417-9372], **Bertuzzi J.**[1,0009-0004-0767-2675], **Vergni L.**[1,0000-0003-3711-1460], **Valigi P.**[2, 0000-0002-0486-7678], **Todisco F.**[1 0000-0002-0571-6484]

[1] Department of Agricultural, Food and Environmental Sciences, University of Perugia, Italy

[2] Department of Engineering, University of Perugia, Italy

[3] Digitization Centre of Cultural and Environmental Heritage (CeDiPa), Perugia, Italy

Corresponding author: Marco Legittimo, marco.legittimo@unipg.it

Abstract

Purpose Accurate estimation of crop transpiration is essential for optimizing irrigation management and improving water-use efficiency in precision agriculture. However, direct measurement of transpiration is often invasive, costly, and difficult to maintain at large scales. This study proposes a data-driven framework to estimate maize (*Zea mays* L.) sap flow driven by transpiration using widely available climatic and soil moisture data combined with machine learning techniques.

Methods Field experiments were conducted during the 2023 and 2024 growing seasons in central Italy under irrigated silage maize. Meteorological variables, soil water content, and crop growth indicators were used as inputs, while sap flow measurements served as reference outputs. Several machine learning models were evaluated, including Linear Regression, Support Vector Regression (SVR), Decision Tree Regressor, and Multi-Layer Perceptron Regressor (MLPR), using both Point Estimation and Temporal Estimation strategies. Temporal approaches incorporated short-term historical information through feature concatenation and previous-average windows.

Results Results demonstrate that non-linear models, particularly MLPR and SVR, consistently outperform linear and tree-based approaches. The inclusion of short temporal windows (45 minutes to 2 hours) significantly improves predictive accuracy, enhancing reconstruction of the diurnal transpiration pattern. Feature concatenation proved more effective than averaging strategies in capturing soil–plant–atmosphere interactions. Model performance remained robust across two contrasting growing seasons, confirming good generalization capability under interannual variability and data discontinuities.

Conclusion The proposed framework provides a reliable and minimally invasive solution for real-time estimation of maize transpiration, supporting precision irrigation management. These findings highlight the potential of machine learning models as practical decision-support tools for sustainable agricultural water management.

Keywords Sap flow prediction; Deep learning; environmental data; soil water content

Introduction

Optimising water management in agriculture represents a key challenge to ensure the sustainability of agricultural practices, maximise crop productivity and mitigate the effects of climate change. Reducing natural resource consumption and improving crop adaptation to extreme conditions such as drought and high temperatures are crucial objectives.

Maize (Zea mays L.), one of the most widely cultivated cereal crops worldwide, is particularly sensitive to water availability, making the use of advanced methodologies for monitoring transpiration and water demand essential. Its water requirements, which depend on many factors including growth stage, climatic conditions, and soil type, generally range between 500 and 800 mm over the entire crop cycle under optimal conditions.

Transpiration (T), which represents the main component of crop water consumption, can be directly estimated through sap flow measurements, which quantify the movement of water within the plant stem. This estimation provides a direct indication of the crop's actual water use, supporting more accurate and efficient water-resource management in agriculture (Giménez et al., 2013; Jones, 2004; Vadez et al., 2024). Accurate sap flow assessment is particularly important because plant transpiration greatly influences terrestrial water cycle processes, accounting for about 80-90% of total evapotranspiration (Jasechko et al., 2013; Wei et al., 2017). However, transpiration is controlled by complex and non-linear interactions among multiple factors, particularly crop traits and meteorological data (Asbjornsen et al., 2011), making its direct measurement and modelling inherently challenging.

Several methods have been developed to quantify plant water use, including sap-flow techniques, indirect and micrometeorological measurements, as well as empirical, physically based, and machine-learning models.

Direct methods rely basically on thermodynamic, magnetodynamic (nuclear magnetic resonance) and electrodynamic principles. Among these, heat-transfer techniques are the most widely adopted (Granier, 1985; Cermák et al., 2004; Nadezhdina, 1999; Kumar et al., 2022; Steppe et al., 2010) and are commonly applied in a wide range of plant species and numerous research domains to investigate plant hydraulics, improve water management, assess crop water stress, and support irrigation planning. The advantage lies in obtaining continuous transpiration measurements with high temporal resolution. However, in almost all cases, these methods are rather invasive, expensive, time-consuming, and require frequent maintenance (Bush et al., 2010; Ren et al., 2017; Steppe et al., 2015; Wullschleger et al., 2011). To overcome these limitations, several empirical and physical models have been proposed to estimate transpiration. Examples include the SIMDualKc model, the modified Jarvis-Stewart MJS, Multiple linear regression MLR and the Shuttleworth–Wallace S-W (Jiang et al., 2016; Liu et al., 2022; Rosa et al., 2016; Tu et al., 2019; Whitley et al., 2008). However, these models require extensive parameterization and large amounts of site-specific data, which restricts their general applicability.

In recent years, data-driven approaches based on machine learning (ML) and deep learning (DL) have emerged as powerful tools for modelling complex agronomic processes (Benos et al., 2021; Kamarudin et al., 2021; Mana et al., 2024). Their strength lies in their ability to capture the complex relationships between environmental and physiological parameters that regulate these processes. For instance, Convolutional Neural Networks (CNNs) have become a standard for crop disease identification and crop identification, as shown by Agarwal 2020, Banerjee et al. 2024, Coulibaly 2019, Waldchen et al. 2018, and Zhong et al. 2019, who achieved high diagnostic and classification accuracy. In the context of yield prediction, various ML models have been successfully implemented to forecast productivity (Barbosa et al., 2020; Chang et al., 2023; Maya Gopal et al., 2019), while in livestock management, clustering and data processing algorithms like K-Nearest Neighbors (KNN) and Principal Component Analysis (PCA) have been applied to track animal behaviour and health (Borgonovo, 2020; Riaboff, 2020).

In the specific field of water management (Pagano et al., 2023; Sun et al., 2019; Yadav et al., 2024), techniques are often tailored to the nature of the data: Support Vector Machines (SVMs) and Ensemble methods have shown excellent results in estimating transpiration and evapotranspiration (Amir et al., 2021; Yamaç and Todorovic, 2020), while architectures like RNNs, LSTM, and GRU have been introduced to handle the temporal dependencies of plant physiological responses (Fan et al., 2021). The most predicted parameters are crop evapotranspiration ETc, reference evapotranspiration ET0 and transpiration T, especially in forests and crops with mainly high economic value or water-intensive management (Bellido-Jimenez et al., 2023; Li et al., 2023; Yamaç and Todorovic, 2020). However, applications to long-term sap flow prediction in herbaceous crops such as maize or tomato remain still limited. Amir et al. (2021) predicted tomato sap flow based on climate, irrigation and sap flow datasets, applying different ML algorithms; their results showed that random forest (RF) achieved the highest accuracy, with good performance of linear regression (LR) and elastic net regression (ENR). Similarly, Fan et al. (2021) estimated daily maize transpiration using four ML models: SVM, extreme gradient boosting (XGBoost), ANN, and deep neural networks (DNN) models. Their analysis incorporated different combinations of input variables, including weather parameters, soil water content (SWC), leaf area index (LAI), and sap flow. Results indicated that the DNN model provided the most accurate transpiration estimation when all input data were used.

Successful ML model development depends on the availability of high resolution, robust and comprehensive datasets acquired continuously over appropriate spatial and temporal scales. Such data richness is essential to provide algorithms with the variability required for effective learning and to ensure rigorous training and validation. Climatic data such as temperature, humidity, precipitation, solar radiation are commonly employed as input to predict evapotranspiration and transpiration (Ferreira et al., 2020; Xing et al., 2022; Zhang et al., 2025a). Nevertheless, these data are most effective when integrated into eco-physiological frameworks describing soil–plant–atmosphere interactions (Chen et al., 2020; Feng et al., 2017; Tang et al., 2018). This step enables climatic data to be converted into target variables, such as actual evapotranspiration or water fluxes, suitable for machine learning predictions.

Complementary data sources, such as satellite remote sensing and micrometeorological systems, can further enhance ML prediction accuracy by providing valuable information on crop water and physiological status and enriching the dataset's dimensionality and spatial representativeness. Vegetation indices (e.g., NDVI) and thermal parameters offer spatially continuous insights into vegetation condition and water stress (Bounoua et al., 2024; Virnodkar et al., 2020; Gómez et al., 2019; Schwalbert et al., 2020), while eddy covariance and Bowen ratio systems provide direct measurements of energy and water fluxes within the soil-plant-atmosphere continuum (SPAC) supporting model calibration and validation (Dou and Yang, 2018; Fang et al., 2020; Zenone et al., 2022; Zhang et al., 2025b). However, their high cost, operational complexity and maintenance have restricted practical applications. In this context, proximal sensors for direct monitoring of plant water transport, such as sap flow sensors, represent an attractive tool. They provide high-temporal resolution data at relatively low cost suitable for characterizing plant responses to water stress and water uptake dynamics (Amir and Butt, 2025; Li et al., 2022; Zhang et al., 2023). When properly calibrated and integrated with other environmental datasets— including satellite, climate, and micrometeorological data— these measurements can substantially enhance the predictive capability of ML models. The combination of multi-scale data sources enables the development of robust and adaptive models tailored to local agroecological conditions, thereby supporting precision irrigation management and providing actionable decision support to farmers.

The goal of this research activity was to develop data-driven models to derive maize transpiration from easily measurable and widely available data, such as climate and soil moisture, for the smart and sustainable real-time management of irrigation practices. The study focused on irrigated silage maize cultivated at the 'Fondazione per l'Istruzione Agraria' farm (Casalina, Perugia, Italy) within two experimental areas managed using Variable Rate technology. Climate and soil water content data served as reference independent variables to train the ML algorithms, while sap flow as the dependent variable, along with crop development indicators such as growing degree days and days after sowing to account for the crop vegetative phase. A key innovative aspect of this work lies in the combined use of multiple machine learning models—ranging from linear and kernel-based approaches (Elastic Net regression, Support Vector Regression) to tree-based methods (Decision Tree Regressor) and multi-layer neural models (MLPR)—together with temporal estimation frameworks that explicitly encode short-term historical dependencies through feature concatenation or aggregated lag windows. This modelling architecture enables the capture of nonlinear soil–plant–atmosphere interactions and diurnal transpiration dynamics that cannot be reproduced by classical point-based predictors. To assess model generalisation, two independent seasonal datasets (2023 and 2024 growing seasons) were used, allowing both intra- and interannual evaluations under varying environmental conditions.

Materials and methods

Study area

The experimental activity was conducted in 2023 and 2024 in two experimental plots located in central Italy within the 'Fondazione per l'Istruzione Agraria' farm (Casalina, Perugia, Italy). The total area of the parcel analyzed in 2023 was about 30 hectares, 600 m long and 500 m wide (coordinates: 42°56'08.3'' N, 12°23'18.0'' E), while that examined in 2024 was about 18 hectares, with an average length of 570 m and an average width of 310 m (coordinates: 42°55'52.8'' N, 12°22'44.7'' E) (Fig. 1). In both fields, a smaller study area was delineated for sensor installation and focused monitoring (blue dots in Fig. 1). According to previous analyses, the 2023 experimental plot was characterized by a loam soil texture (Santaga et al., 2021), whereas the 2024 plot presented a silty loam texture. Both sites were characterized by a nearly flat topography, with an average slope close to 0%. In both years, corn (*Zea mays* L.) for silage (varieties LG 31,415 in 2023 and MAS576N in 2024) was sown as second crop after pea, in the last decade of June using a precision seed drill, with a planting density of 7.5-8.5 plants per square meter and a row spacing of 75 cm. Harvesting for both years took place in early October. Irrigation was provided by self-propelled hose reel sprinklers, with a 7-day irrigation interval that can be adjusted depending on weather conditions.

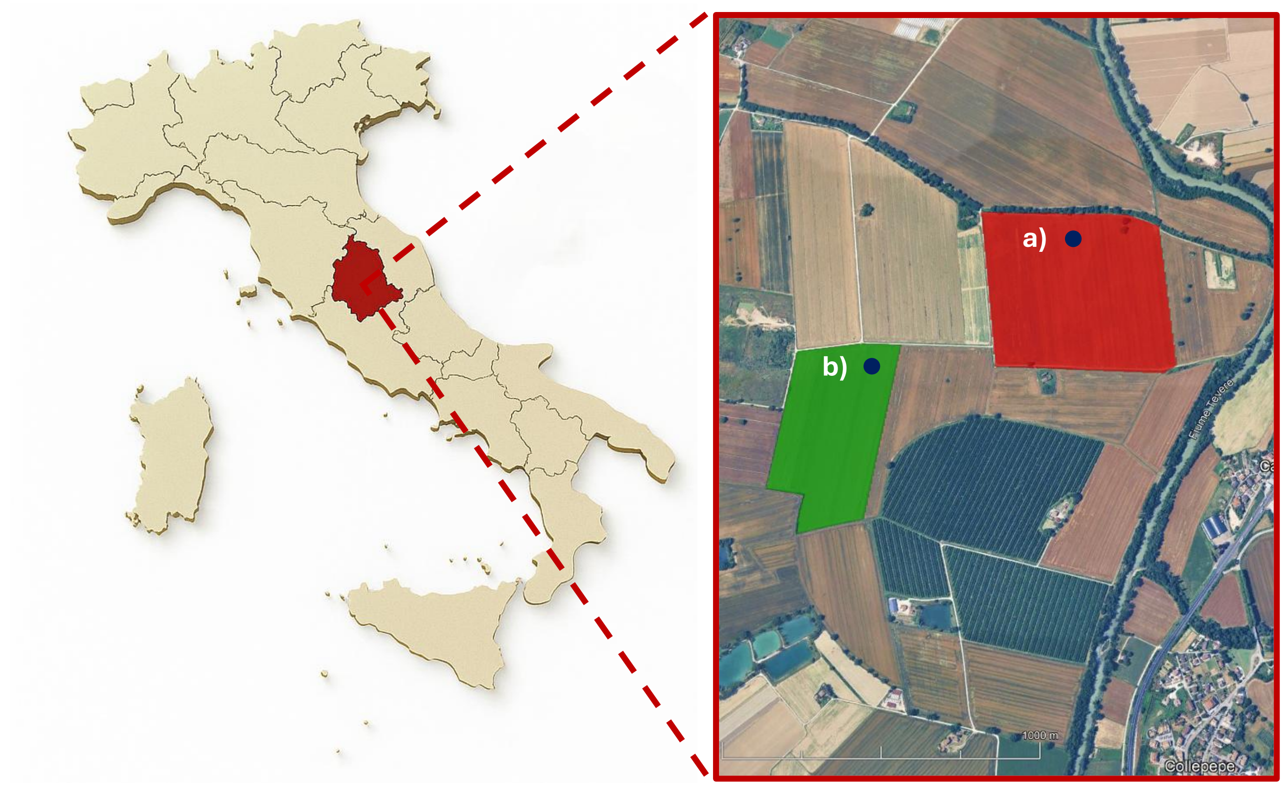


Fig. 1 Framing of the plots (in green) and detail of the study area (in blue) monitored a) in 2023 and b) in 2024

Data collection and dataset consistency

The environmental and physiological variables monitored during the experimental trials are summarized in Table 1 and Table 2, which respectively report external and internal data: descriptions, measurement units, and sampling frequencies.

All the data have been collected during the 2023 and 2024 maize growing seasons, different sensors were installed to monitor external variables, including local climatic conditions and soil water content, along with internal parameters i.e. sap flow (Fig. 2).

Meteorological data were obtained from an automatic weather station near the experimental site. Temperature, relative humidity, wind speed, global and net radiation, and the FAO Penman-Monteith ET0 (Allen et al., 1998) were recorded every 15 minutes.

Precipitation (P) and irrigation (Irr) were collected using a tipping-bucket rain gauge (Spectrum Technologies Inc., Aurora, USA) installed at a height of approximately 2.5 m above the ground, and the datum is the cumulative values in 30 minutes time interval resulting in 2 samples/hour.

Soil water content was monitored by FDR probes (Sentek Drill&Drop Bluetooth, Sentek Technologies, Australia) with sensors at depths of 5-15-25-35-45-55 cm and stored at 60-minute intervals. In July 2023, 4 FDR probes were positioned between two plants along a corn row, spaced about 2-5 meters apart (Fig. 2a). To monitor soil moisture variability between rows, at the end of June 2024 6 probes were installed 10 and 20 cm away from the row, keeping a distance between the probes of 5-10 meters (Fig. 2b). The volumetric water content in the first 60 cm of soil was computed for each probe. Due to the high spatial variability of soil conditions and the intrinsic sensitivity of FDR probes, a max-min normalization was applied to each probe over its entire operating period. This procedure ensures that the probe-specific variability does not introduce bias into the computation of the final averaged soil water content. The normalized data were then averaged to obtain a single value, SWC_avg.

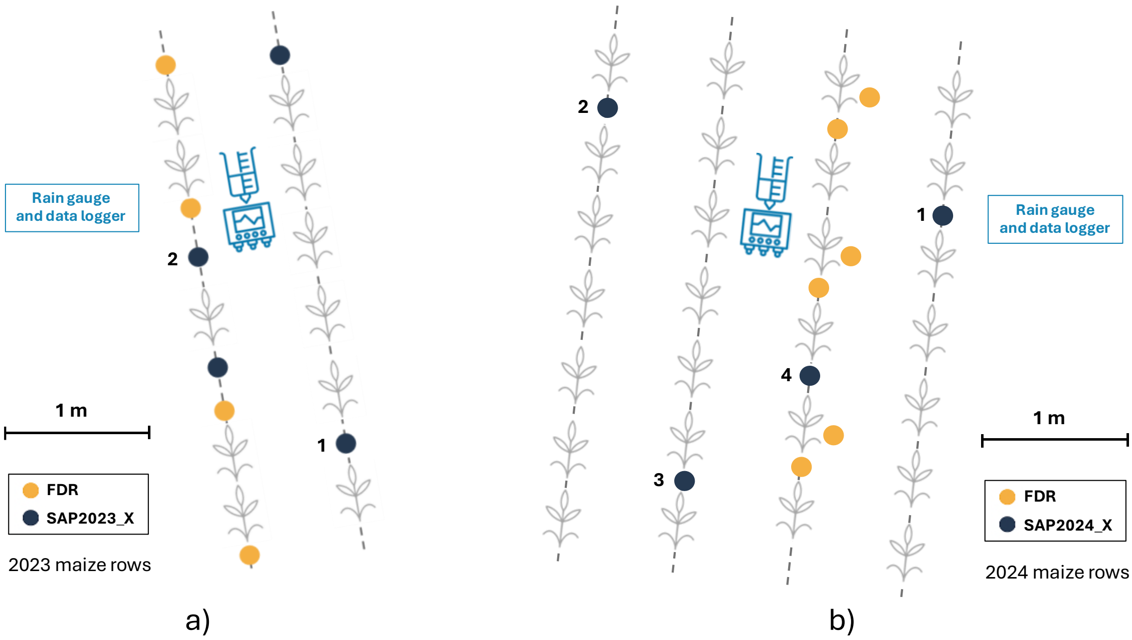


Fig. 2 Experimental plan for the two years monitoring: a) year 2023, 4 Dynagage sap flow sensors and 4 FDR probes positioned between two plants along a row, spaced 2-5 metres apart; b) year 2024, 4 sap flow sensors and 6 FDR probes 10 and 20 cm apart from the row, spaced 5-10 metres

Crop transpiration fluxes were continuously monitored during the irrigation season (August-September), using sap flow heat balance sensors (SAP) produced by Dynamax Inc. (Houston, Texas, USA). This type of sensor is based on heat balance theory and involves applying a constant amount of heat to a small section of the stem by means of a heating strip in the surrounding gauge body (Kišš and Manina, 2019; Lascano et al., 2016). The sensors directly measure the amount of heat carried by the sap through the stem, which is converted into real-time sap flow rate F ($g{\cdot}s^{-1}$) using the equation:

$$F = \frac{(\mathrm{P}_{in} - \mathrm{Q}_v - \mathrm{Q}_r)}{\mathrm{C}_P \cdot \mathrm{dT}} = \frac{\left[\mathrm{P}_{in} - \mathrm{K}_{ST} \cdot \mathrm{A} \cdot \left(\frac{dT_u + dT_d}{dx}\right) - \mathrm{K}_{sh} \cdot \mathrm{CH}\right]}{\mathrm{C}_P \cdot \mathrm{dT}}$$

where P_{in} is the power of heater (W), Q_v is vertical conduction, Q_r is radial heat conduction, C_p is specific heat of water (4.186 $J{\cdot}g^{-1}{\cdot}°C^{-1}$) and dT is temperature increase of sap (°C), K_{ST} is the thermal conductivity of the stem ($W{\cdot}m^{-1}{\cdot}°C^{-1}$), A is the stem cross-sectional area (m^2), dT_u/dx and dT_d/dx ($°C{\cdot}m^{-1}$) are the temperature gradients in the up and down directions, respectively, dx is the spacing between the thermocouple junctions (m), K_{SH} is the sheath conductivity ($W{\cdot}mV^{-1}$), CH is the radial-heat thermopile voltage (mV).
Specifically, four sensors, suitable for plants with a stem diameter of approximately 35 mm, were installed on healthy, representative plants (Fig. 2). In both 2023 and 2024, the probes were deployed in August, during peak vegetation cover and the onset of flowering, when crop water demand is highest. The sensors were positioned 20–30 cm above the ground, between two internodes, following the manufacturer's recommended procedure. The installation procedure required the setting of certain parameters including the type of sensor and the cross-sectional areas of the stem used. The voltage was set at 4.4-5.4 V, with a power between 0.20 and 0.73 W and a temperature range of 0.1 - 13 °C. No stem damage due to heat occurred during the monitoring period. Sap flow measures (g/h) were sampled every 1 minute and averaged over 15-minute time intervals from August through September i.e. frequency 4 samples/hour (Table

1). Due to datalogger malfunctions and technical issues affecting the acquisition system, in the 2023 season, only two sensors (SAP23_1, SAP23_2) were included in the dataset, while for the 2024 season, four sensors (SAP24_1, SAP24_2, SAP24_3, SAP24_4) provided valuable data to be integrated in the dataset. The final sap flow dataset consistency is based on 1371 effective measurements for 2023 and 1472 samples for 2024.
Missing values were handled using various gap-filling and interpolation techniques, as described in Section *"Temporal Alignment and Gap Filling"*.
Two more parameters were added to consider the growth phase of the crop:

- Growing degree-day (GDD): this index expresses the relationship between air temperatures and phenological development of the crop (Liu et al., 2025) and is calculated as the sum of the differences between the average daily temperature (TA_avg) and the vegetative zero (Tz) of the species considered (8°C for maize).
- Days since sowing (DSS): sequential number representing the days elapsed since the sowing date.

As mentioned before, the monitored data and soil-crop dataset structure are presented in Tables 1 and 2, along with the measurement units and time resolution. Table 1 presents the weather data and crop growth variables, while Table 2 lists the soil water content and sap flow data.

Table 1 Meteorological and soil water content variables (external)

Variables	***Description***	***Unit***	***Sampling***
TA_min	Minimum air temperature	°C	4/hour
TA_avg	Average air temperature	°C	4/hour
TA_max	Maximum air temperature	°C	4/hour
RH_avg	Average Relative humidity	%	4/hour
Rg_avg	Average Global radiation	W/m^2	4/hour
Rn_avg	Average Net radiation	W/m^2	4/hour
Wind_avg	Average Wind speed	m/s	4/hour
ET0	Reference evapotranspiration	mm	4/hour
Irr and P	Cumulative irrigation and precipitation from rain gauge	mm	2/hour
DSS	Days elapsed since the sowing date	-	1/day
GDD	Growing degree-day	°C/d	1/day
SWC_avg	Average of normalized soil water content from FDR probes in the 60 cm soil depth	-	1/hour

Table 2 Crop variables (internal)

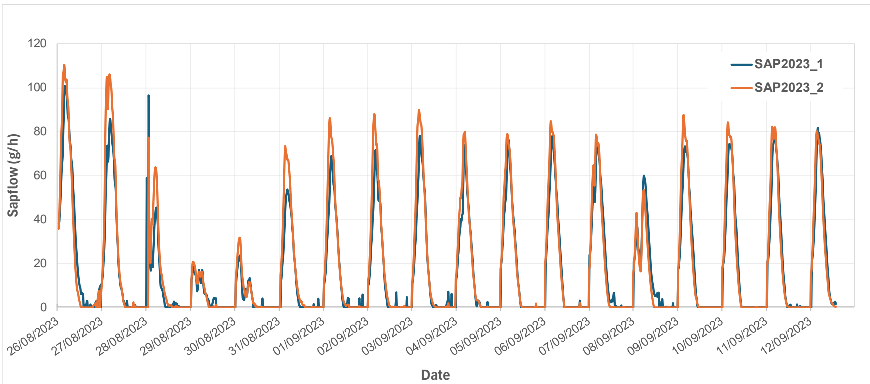
Variables	*Description*	*Unit*	*Sampling*
SAP23_X	Sapflow mass discharge from sensor X in 2023 (X = 1, 2)	g/h	4/hour
SAP24_X	Sapflow mass discharge from sensor X in 2024 (X=1, …, 4)	g/h	4/hour

The collected meteorological, soil and crop data were organised into two separate datasets, one for 2023 and one for 2024, which served as the basis for developing the datasets employed in the implementation of the data-driven model.

Figure 3 shows the behaviour of some of the most significant variables included in the dataset relative to 2023: Sapflow SAP, averaged soil moisture (SWC_avg), ET0, Irr and P. As expected, the dynamics exhibit a close correlation and synchronicity between sap flow and ET0: as evapotranspiration increases, sap flow increases.

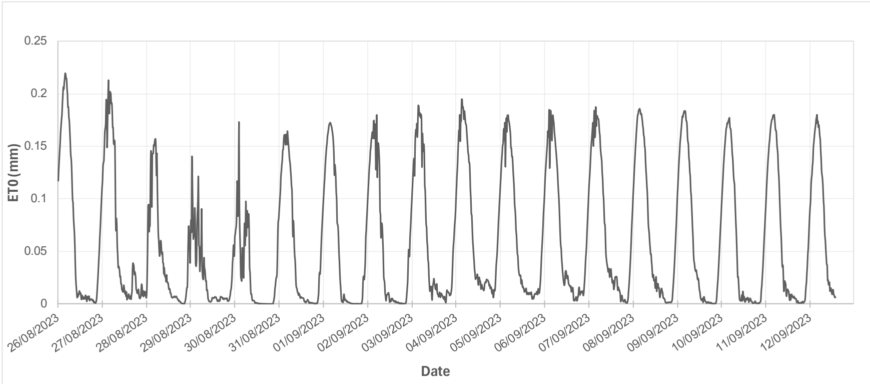
Soil water content, on the other hand, exhibits a dynamic relationship T and ET0. SWC reaches its maximum when T and ET0 are at their daily lowest values, typically at night, and progressively declines throughout the daytime following a characteristic stepwise pattern. The diurnal variation of SWC (ΔSWC) primarily reflects transpiration-driven soil water depletion, highlighting the temporal coupling between soil water availability and crop water uptake over the day–night cycle.

The data-driven models to be developed must be able to recognise the recurring patterns between the climate variables and the plant's transpiration response (i.e. sap flow) and the relationships between T and irrigation/rainfall and consequently SWC, in order to be able to estimate the plant's water demand from real-time input data.

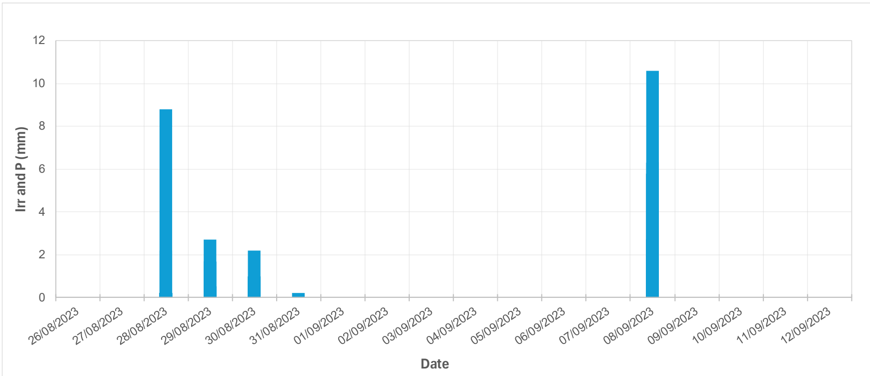


(a)

(b)



(c)



(d)

Fig. 3 Time behaviour of plant water consumption, sapflow, a), volumetric soil water content b), reference evapotranspiration, ET0 c) and Irrigation and precipitation d) during 2023

Data-driven estimation models

The following sections describe the data-driven estimation models adopted in the experiments. Two model types were considered: Point Estimation (PE) models, which predict the output based solely on the current input sample, and Temporal Estimation (TE) models, which additionally incorporate information from the previous N time steps to enhance prediction accuracy.

For each monitored year (2023 and 2024), the objective was to estimate the sap-flow variable SAP_avg based on concurrent meteorological (Table 1) and soil moisture (Table 2) data. SAP_avg is defined as the mean value obtained from all functioning sensors at each time step. As detailed in section *"Temporal Alignment and Gap Filling"*, the reference time step used in the analysis is 15 minutes.

Point Estimate (PE)

The Point Estimate (PE) strategy aims to predict the average sap flow value (SAP_avg) using only the current measurements of climatic and soil variables, while explicitly ignoring any information from previous time steps. In this configuration, the model focuses on instantaneous conditions (those observed within the most recent 15-minute sampling window) without accounting for temporal dependencies or lagged effects.

As summarized in Table 3, the input features include only the current external variables, whereas no averaged or concatenated features are introduced.

Different PE models were exploited to select the best predictive model for SAP_avg, comparing various ML algorithms:

- Linear Regression - LR
- Support Vector Regression - SVR
- Decision Tree Regressor - DTR
- Multi-Layer Perceptron Regressor – MLPR

Temporal Estimate (TE)

Strategies are investigated to capture temporal dependencies within the data. In contrast to PE, where models relied only on current sensor readings to estimate target values, the present analysis extends the input by incorporating information from past observations. This approach enables to assess whether previous measurements can be effectively exploited to improve the accuracy of the current SAP_avg prediction.

Two different approaches have been used to account for past observations: the current sample and Previous Averages (TE-PA) and the current sample and Features Concatenation (TE-FC). TE-PA has been built by concatenating each observation with the average of the N previous ones. The parameter N is computed by specifying a temporal interval., such as 45 minutes, or 2, 6 hours, which corresponds to a specific sample number: since the data are recorded at 15-minute intervals, each additional 15 minutes in the temporal window increases N by one unit. Consequently, for temporal windows of 45 minutes, 2 hours, and 6 hours, *N* assumes values of 3, 8, and 24, respectively.
The TE-FC strategy was implemented by concatenating each observation with the previous N ones.

Table 3 provides a comprehensive summary of the input and output features considered in the models, outlining the variables used for model training and prediction.

Table 3 Features for the Point Estimate (PE) and the Temporal Estimate (TE) models

MODELS	Input features (single sample)	Averaged/concatenated features	Output quantity
PE	time, SWC_avg, TA_avg, RH_avg, Rn_avg, Wind_avg, Eto, Irr and P, GDD, DSS	----	SAP_avg
TE	time, SWC_avg, TA_avg, RH_avg, Rn_avg, Wind_avg, Eto, Irr and P, GDD, DSS	SWC_avg, TA_avg, RH_avg, Rn_avg, Wind_avg, Eto, Irr and P	SAP_avg

Model Assessment and Details

Model assessment is performed by using the RMSE and R^2-Score metric defined as:

$$RMSE = \sqrt{\frac{1}{n} \cdot \Sigma_{K=1}^{N} (y_k - \hat{y}_k)^2}$$

$$R^2 = 1 - \frac{\Sigma_{K=1}^{N} (y_k - \hat{y}_k)^2}{\Sigma_{K=1}^{N} (y_k - \bar{y})^2}$$

in which y_k and $\hat{y}_k$ are the observed and the predicted value for sample k, and $\bar{y}$ is the mean of all SAP_avg values.

All experiments were implemented in Pandas for data manipulation and preprocessing, and Python, using the scikit-learn library for model development and hyperparameter tuning.

Dataset management and feature engineering

This section outlines the data preprocessing steps, including time resolution alignment and gap filling.

Temporal Alignment and Gap Filling

As reported in Table 1, measures are sampled at different frequencies, spanning to 1 sample/day to 4 samples/hour. To ensure data consistency, a fixed 15-minutes time resolution has been selected. This section outlines the main steps taken to align the several sampling frequencies.

- *Meteorological Alignment*: since rainfall and irrigation data were recorded as cumulative values every 30 minutes, while other weather variables were sampled every 15 minutes, a splitting procedure was applied. Each 30-minute cumulative value was divided equally: half was assigned to the current 15-minute record and the other half to the preceding 15-minute slot.
- *Soil Moisture Interpolation*: the SWC_awg, originally recorded every 60 minutes, was linearly interpolated to match the 15-minute sampling frequency of the climatic and sap flow data. This step was essential to provide the models with a continuous representation of soil water availability.
- *Target Variable Synchronization*: sap flow measurements, initially collected every minute, were averaged over 15-minute intervals to align with the climatic reference time step.

After merging the weather and soil-crop datasets using date-time keys, missing values due to malfunctioning in SAP sensors were handled. In 2023 dataset a multi-step imputation strategy was applied. Initially, missing values occurring at night—identified by negative net radiation (Rn_avg < 0)—were set to zero. For daytime gaps, values were imputed using the average recorded at the corresponding time across different days. Regular data gaps were observed due to sensor-specific logging cycles. Since these empties remained after the previous steps, they were handled according to their temporal extent:

- For *SAP23_1* (single isolated missing value), the gap was filled using the average of the immediately preceding and subsequent measurements.
- For *SAP23_2* (three consecutive missing samples), linear interpolation was applied.

As illustrated in Fig. 4, the 2024 sap flow measurements exhibit a considerable number of missing (NaN) values, predominantly clustered around sample index 1500, likely due to sensor malfunction or incorrect activity. Unlike the 2023 dataset, for which missing values were filled using data from previous or subsequent days, the extent and distribution of missing entries in 2024 made the same procedure unfeasible. Therefore, linear interpolation was adopted before integrating the sap flow data with the meteorological dataset. After the merge, the remaining missing values corresponding to periods with negative Rn_avg were set to zero, while the residual NaN values, not temporally aligned or too sparse to be reliably filled, were discarded from the analysis.

For visualization purposes, the x-axis of the 2024 plots reports the sample index instead of the corresponding days. This choice was made to preserve readability, since the presence of a large number of missing (NaN) values, subsequently removed during preprocessing, would have resulted in irregular time spacing and a distorted representation of the temporal scale. Consequently, only the sample number is shown on the x-axis, while the y-axis represents the measured or estimated sap flow values (SAP_avg).

After pre-processing and gap filling, the complete dataset consistency is 1696 samples for 2023, while 2731 samples for 2024.

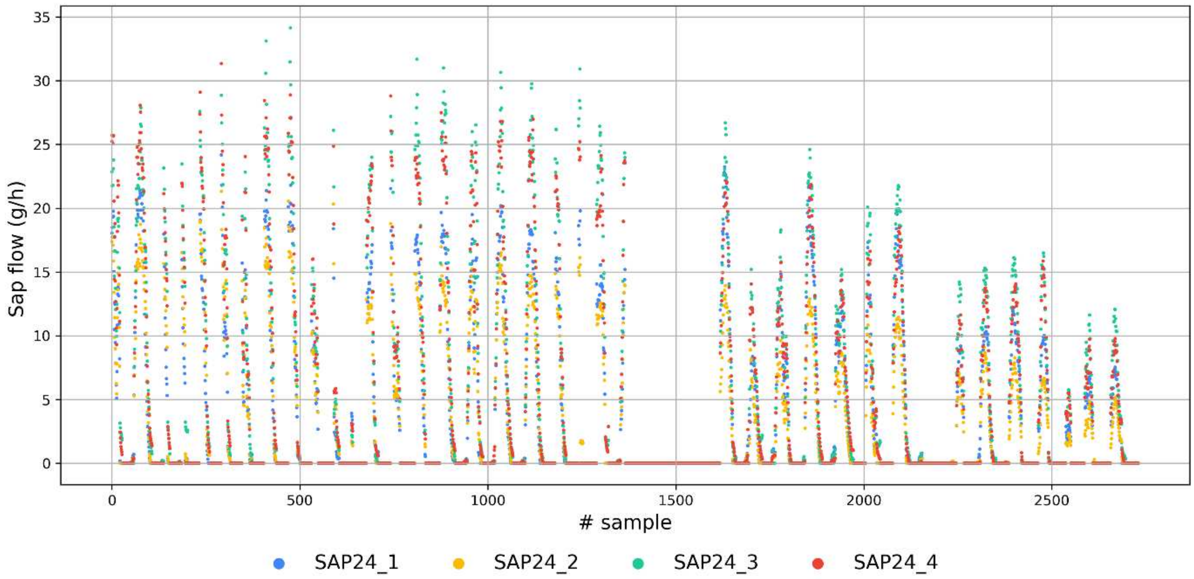


Fig. 4 Sap flow data measured by the four sensors in 2024

Time Encoding and Features Normalization

The collected data is marked with a timestamp: hour and minute of each measurement. To enable the ML models to effectively learn the diurnal periodicity of transpiration, the time of day was transformed using cyclic encoding. Treating time as a linear numerical feature would create a false discontinuity: 23:45 and 00:00 would be treated as numerically distant. Therefore, the timestamp was decomposed into its sine and cosine components. This transformation ensures that the models perceive the proximity between the end of one day and the beginning of the next.

RThe time is split into hours ∈ [0, 23] and minutes ∈ {00, 15, 30, 45}. To aggregate them, minutes are first converted into fractional hours and then combined with the hour component:

$$tempTime = 4 * (hours + (minutes / 60))$$

This produces 96 unique values (24 hours × 4 intervals). To preserves the circular nature of time over the day, a cyclic encoding was then applied:

$$newTime_x = 2\pi sin(tempTime / 95)$$
$$newTime_y = 2\pi cos(tempTime / 95)$$

A min-max normalization was applied to rescale all input features to the range $[-1, +1]$ ensuring consistency with the sine and cosine encodings used for time feature. This transformation preserves the relative distribution of each feature while aligning all inputs to a common scale.

Data partitioning for 2023

The complete 2023 dataset spans just over 18 days. The first 14 days were assigned to the training set, while the remaining 4 days were designated as the test set. Additionally, three days from the training set were reserved for hyperparameter validation; this configuration is referred to as *split1*. As illustrated in Fig. 5, which depicts the SAP_avg values of *split1*, three days within the training set (corresponding to the third, the fourth, and the fifth peaks) exhibit markedly different behaviour compared to other days.

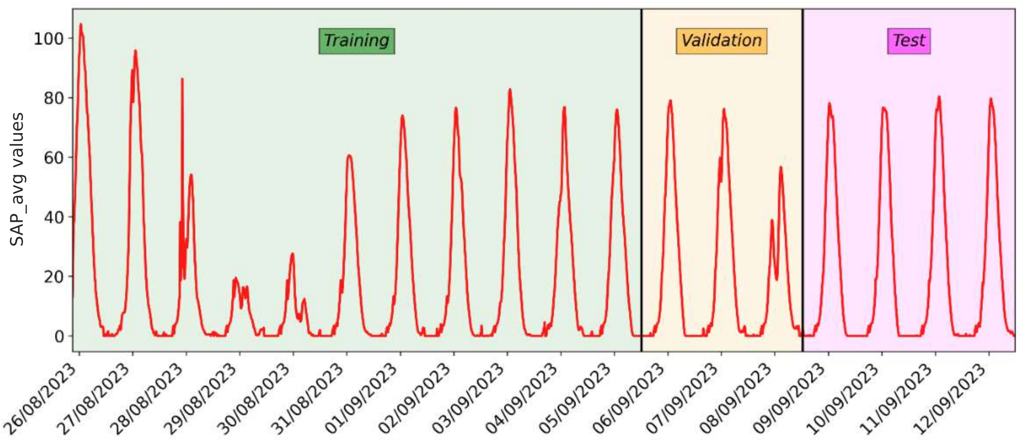
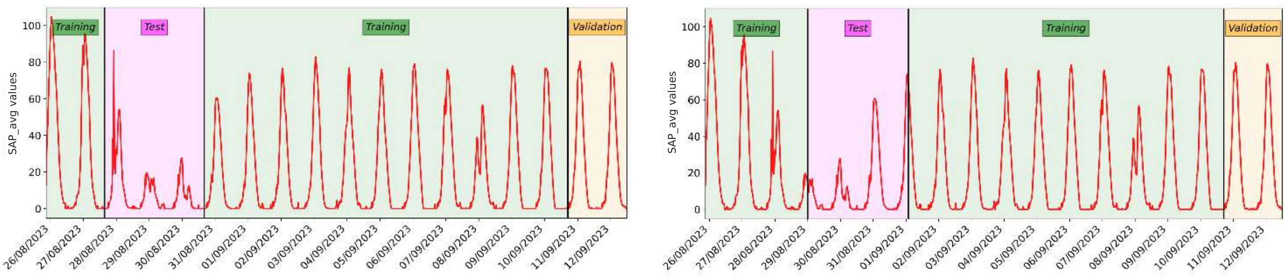


Fig. 5 Train (green), Validation (yellow), and Test (purple) sets (split1) in 2023

To account for these anomalies, two additional data splits were built, *i.e., split2,* and *split3*: in *split2* the test set contains the three days with irregular sap flow values (Fig. 6a), whereas the test set for *split3* includes only half of those irregular measurements, with the remaining part of the set containing regular values (Fig. 6b). These three splits have been designed to evaluate the impact of those irregular patterns on model performance and to assess their generalization capabilities.



(a) (b)

Fig. 6 Train (green), Validation (yellow), and Test (purple) sets for split2 (a) and split 3 (b) in 2023

Data partitioning for 2024

The dataset collected in 2024 was used to assess the generalisation capability of the model across different growing seasons, in additional experiments. These analyses are crucial for evaluating the model's robustness when trained and tested on data acquired under potentially different conditions and spanning an extended plant life cycle.

The 2024 dataset spans a longer temporal period compared to 2023, as evidenced by the mean sap flow values depicted in Fig. 7. As already discussed in Section *"Temporal Alignment and Gap Filling"*, the sap flow (SAP_avg) measurements in 2024 contain a significant number of missing (NaN) values. To address this issue, the dataset was partitioned into two subsets: *Subset 1*, which spans from the beginning to approximately index 1400, and *Subset 2*, which ranges from around index 1600 to the end.

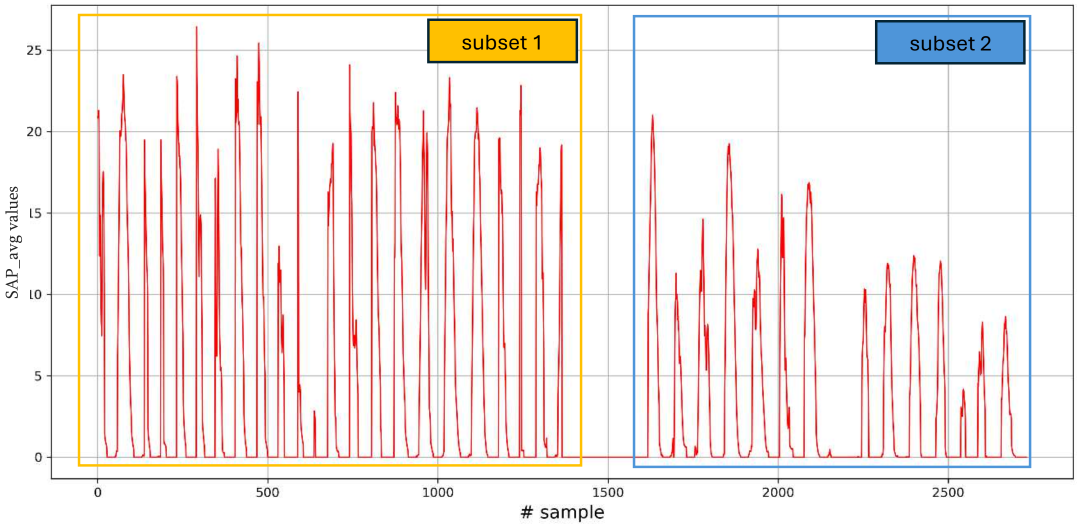


Fig. 7 Mean sap flow values of the four SAP24 sensors

Despite the application of the gap-filling procedure described before and the partitioning strategy, a considerable number of NaN values remained. Due to the incomplete temporal continuity, temporal estimation (TE) experiments, which require incorporating the previous N samples into the prediction process, could not be performed. Consequently, the analysis was restricted to the point estimation (PE) task and two different experiments were implemented:

- *Exp1:* PE models were trained on the 2023 dataset and tested on the 2024 dataset, to directly assess cross-year generalization.
- *Exp2:* PE models were trained on a combined dataset consisting of 2023 and Subset 1 of 2024, and then tested on Subset 2 of 2024, simulating adaptation to partially known 2024 conditions.

Hyperparameters Tuning

After completing the data preprocessing, the hyperparameter tuning procedure was performed prior to model evaluation. This step is essential to identify the optimal configuration of model parameters that influence its learning behaviour. Proper tuning ensures that each algorithm operates under the most suitable conditions, thereby enhancing model performance, stability, and generalization capability. The optimization was carried out through a GridSearchCV (Bishop, 2006) procedure, which systematically explores multiple parameter combinations to select the configuration yielding the best predictive accuracy. Table 4 reports the optimal hyperparameter values obtained through cross-validation on the 2023 dataset for PE case, which are provided here as a representative example. Analogous were performed on PE and TE strategies and on the 2024 dataset. It should be noted that, for the LR model, an Elastic Net (Zou and Hastie, 2005) formulation was adopted to allow L1 and L2 regularization.

Table 4 Best hyperparameters for the 2023 PE case

Hyperparameter	Description	Best Value
LR (Elastic Net)		
α (alpha)	Regularization strength: higher values mean stronger regularization	0.1
L1_ratio	Controls the mix between L1 (Lasso) and L2 (Ridge) penalties	0.9
DTR		
criterion	Function to measure the quality of a split	poisson
max_depth	Maximum depth of the tree	None
max_features	Number of features to consider when looking for the best split	sqrt
min_samples_leaf	Minimum number of samples required to be at a leaf node	1
min_samples_split	Minimum number of samples required to split an internal node	10
SVR		
C	Regularization parameter; controls trade-off between smoothness and accuracy	10
kernel	Specifies the kernel type to be used in the algorithm	rbf
MLPR		
activation	Activation function for the hidden layer	logistic
α (alpha)	L2 penalty (regularization term)	1e-5
hidden_layer_sizes	Neurons per layer	(50, 50)
learning_rate_init	Initial learning rate used during training	0.01
max_iter	Max number of iterations	500
solver	Optimization algorithm	adam

Results

In accordance with the considerations outlined above, the following sections present the results obtained from the application of the two modelling approaches for the 2023 and 2024 datasets, highlighting their predictive performance and the main differences observed between the two monitoring periods.

Results of PE approach for 2023 data

This section summarizes the quantitative results of each PE model and provide a qualitative evaluation regarding the best performing one. Performance metrics reported in Table 5 include the R^2-score on both training and test set, as well as the RMSE (Bishop, 2006).

Table 5 Results of best PE models. The best model for each split is shown in bold.

	Model	RMSE		R^2-score	
		train	test	train	test
split1	LR	10.057	11.246	0.835	0.822
	DTR	2.824	13.379	0.987	0.748
	SVR	5.958	6.568	0.942	0.939
	MLPR	2.910	7.861	0.986	0.913
split2	LR	8.662	9.721	0.894	0.385
	DTR	2.858	8.160	0.989	0.567
	SVR	14.401	9.934	0.708	0.357
	MLPR	3.392	7.318	0.984	0.651
split3	LR	8.985	9.103	0.885	0.590
	DTR	2.307	10.482	0.992	0.456
	SVR	4.946	6.081	0.965	0.817
	MLPR	4.151	5.524	0.976	0.849

Based on the results in Table 5, MLPR and SVR consistently emerge as the best performing models across all dataset splits. In each scenario, at least one of the two achieves the best performance in both RMSE and R^2-score, often outperforming the others by a clear margin.

In split1, both models demonstrate excellent predictive capabilities, while traditional models like LR and DTR show either poor generalization or overfitting. This trend continues in split2, where MLPR proves to be the most robust, especially in handling unseen or irregular data patterns. Even in split3, where the test set contains both regular and irregular sap flow values, MLPR and SVR provide the best performance, confirming their adaptability and stability.

Given their consistently superior results, SVR and MLPR are selected as the best-performing models for sap flow estimation, showing the highest potential for generalizing across diverse conditions.

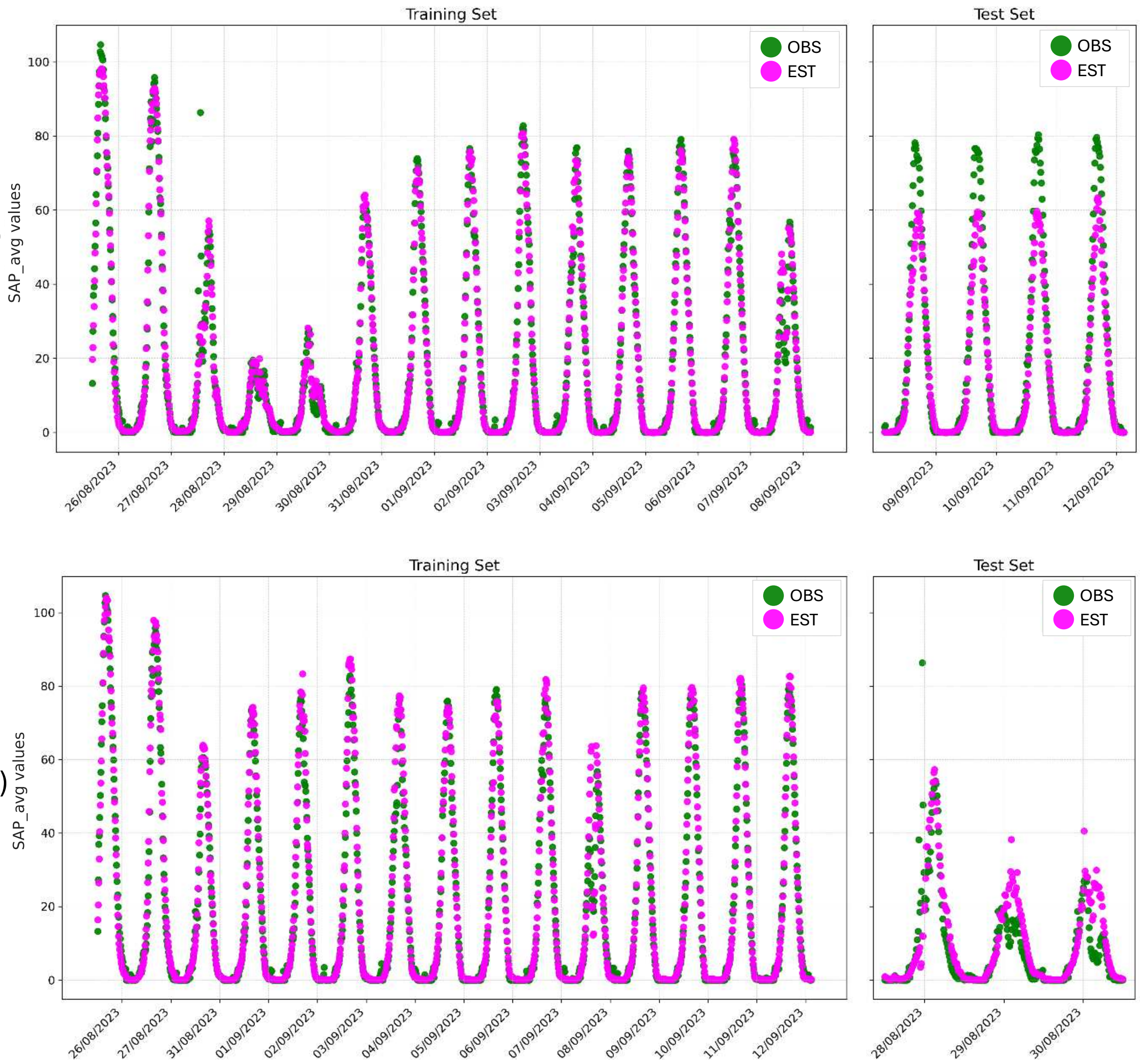
Although SVR and MLPR yield comparable results, MLPR achieves overall a better performance. For this reason, the following graphs shows only the qualitative results of MLPR.

Figure 8 compares the MLPR predictions (EST) with the observed values (OBS) for the three dataset splits. In split1 (Fig. 8a), the model reproduces the temporal pattern of the test set well, except for a few peak values that are slightly underestimated. In split2 and split3 (Figs. 8b–c), the performance degrades due to a general overestimation of several SAP_avg values around midday, which is expected given the greater irregularity of these test sets. It is also worth noting that split2 reports a lower R^2 than split1 despite achieving a slightly lower RMSE. This behaviour, which may appear counter-intuitive, is explained by the markedly reduced variability of SAP_avg in the split2 test set (approximately 0–30 g/h for two days out of three, compared to 0–80 g/h in split1). Because R^2 depends on the variance of the observed data, when the ground truth fluctuates less, the same absolute prediction error represents a larger portion of the total variance, resulting in lower R^2 values even if the RMSE improves.

Split3 represents an intermediate situation: its test set contains both regular and irregular sap-flow patterns, with a variability that is higher than in split2 but still smaller than in split1. Consequently, its R^2 values typically fall between those of split1 and split2, reflecting this intermediate level of signal variability and difficulty.

(a)

(b)



(c)

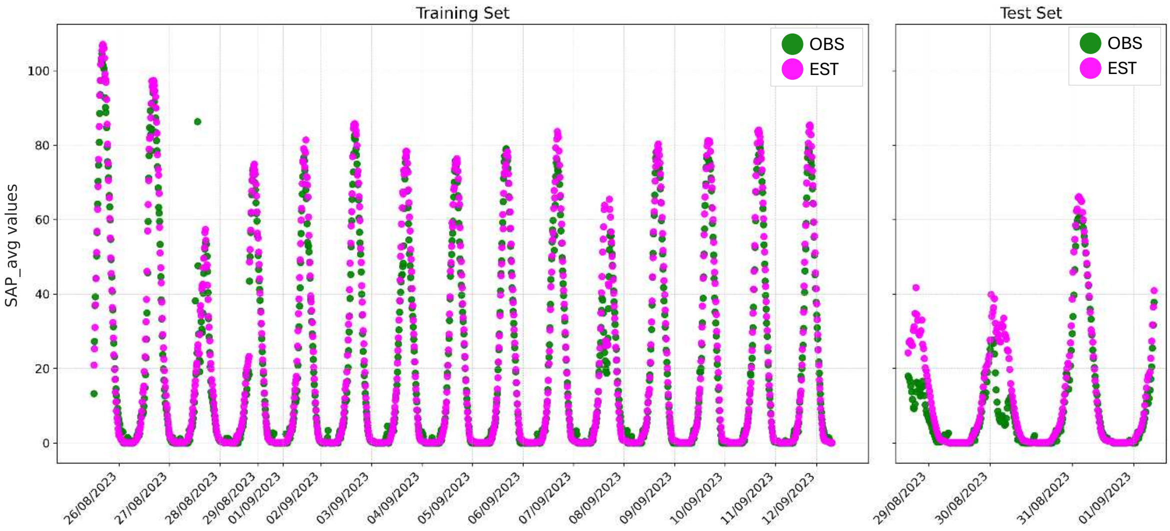


Fig. 8 SAP_avg estimated values, EST, by MLPR model compared against observed values, OBS, for split1 (a), split2 (b) and split3 (c)

Results of TE approach for 2023 data

The following subsections present quantitative and qualitative results for the three time intervals considered in the TE model (45 minutes, 2 hours, and 6 hours). Specifically, table 6 reports the RMSE and R^2-score scores for the SVR and MLPR models, which proved to be the best in the PE approach, while LR and DTR were discarded due to their poor performance (as described in the previous sections). Additionally, three graphs (one for each split) compare the observed values with those estimated using the MLPR model for the three temporal windows. For each split and time interval, results are depicted both for the approach based on concatenating the average of the previous samples (TE-PA) and for the direct concatenation of the features (TE-FC).

Overall, the results obtained across the three temporal windows reveal a clear split-dependent behaviour. Split1 generally provides good predictive performance, but it is not consistently the best case, especially when longer temporal windows are used. Split2 systematically emerges as the most challenging configuration: its low variability and irregular dynamics lead to the highest RMSE values and the lowest R^2 scores across all TE-PA and TE-FC experiments, regardless the window size. In contrast, split3 often achieves the best results, particularly for longer windows and when feature concatenation (TE-FC) is applied. In this setting, the model benefits from the additional temporal context by more effectively reconstructing the overall daily pattern, which results in substantial improvements in R^2 and, in several cases, reductions in RMSE. Increasing the temporal window does not produce a uniform trend: short windows (45 minutes and 2 hours) typically offer the most stable and reliable improvements, whereas longer windows (6 hours) may either enhance or degrade performance depending on the split. This reflects the trade-off between providing meaningful temporal context and introducing overly smoothed or outdated information. Finally, TE-FC generally outperforms TE-PA in terms of R^2, especially with wider windows, indicating a stronger ability to recover the shape of the transpiration signal, although this does not always translate into a corresponding reduction in RMSE.

45 minutes temporal windows for TE-PA and TE-FC strategies

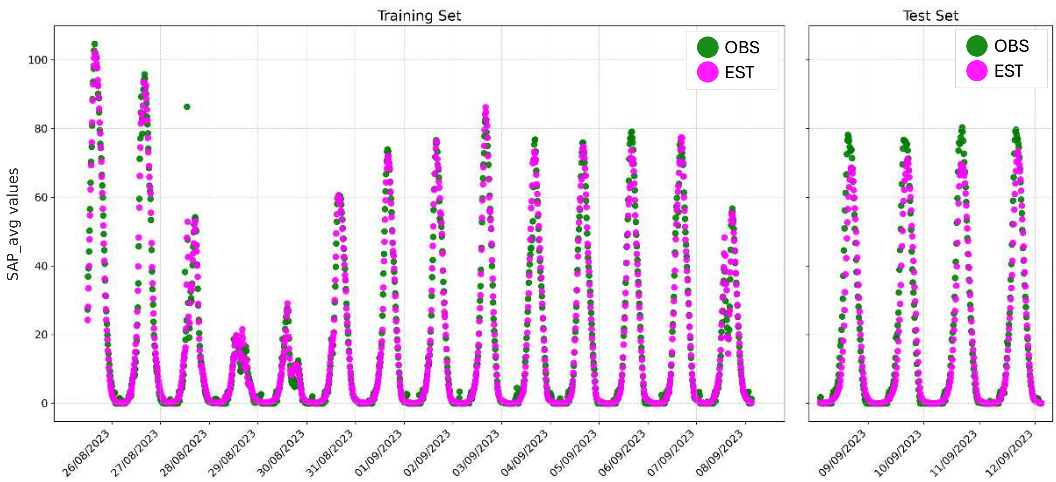
Results obtained by concatenating the average (TE-PA) of the previous 3 samples (corresponding to 45 minutes) along with those derived from direct feature concatenation (TE-FC) are presented below. Quantitative outcomes are summarised in Table 6. The transition from PE to TE-FC improved the R^2 for split3 from 0.849 (see table 5) to 0.959, confirming the importance of temporal context.

Table 6 Results of SVR and MLPR with best hyperparameters found with GridSearchCV.

		Temporal Estimate (TE) models							
		(a) TE-PA				(b) TE-FC			
	Model	RMSE		R2-score		RMSE		R2-score	
		train	test	train	test	train	test	train	test
split1	SVR	5.849	7.481	0.944	0.921	6.079	7.914	0.940	0.912
	MLPR	3.041	5.964	0.985	0.950	3.898	5.377	0.975	0.959
split2	SVR	2.719	10.087	0.990	0.337	2.819	9.523	0.989	0.409
	MLPR	2.826	7.247	0.989	0.658	5.327	6.905	0.960	0.689
split3	SVR	4.967	4.263	0.965	0.912	5.353	3.608	0.959	0.940
	MLPR	2.421	6.334	0.992	0.805	3.022	3.005	0.943	0.959

Figures 9, 10 and 11 depict qualitative results for each split, highlighting the impact of the different test sets.

(a)



(b)

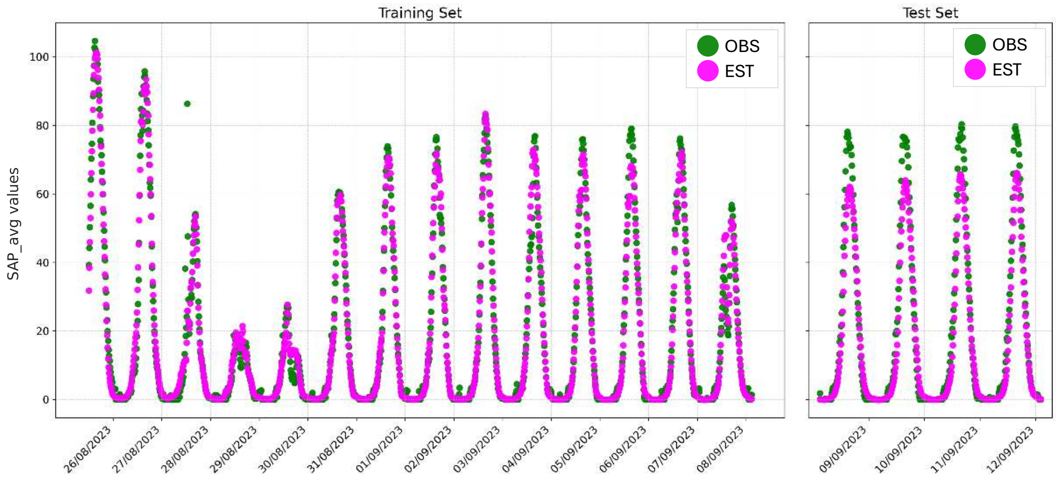


Fig. 9 SAP_avg estimated values, EST, compared against observed ones, OBS, for TE-PA (a) and TE-FC (b) with 45 minutes temporal windows using MLPR model (split1)

(a)

(b)

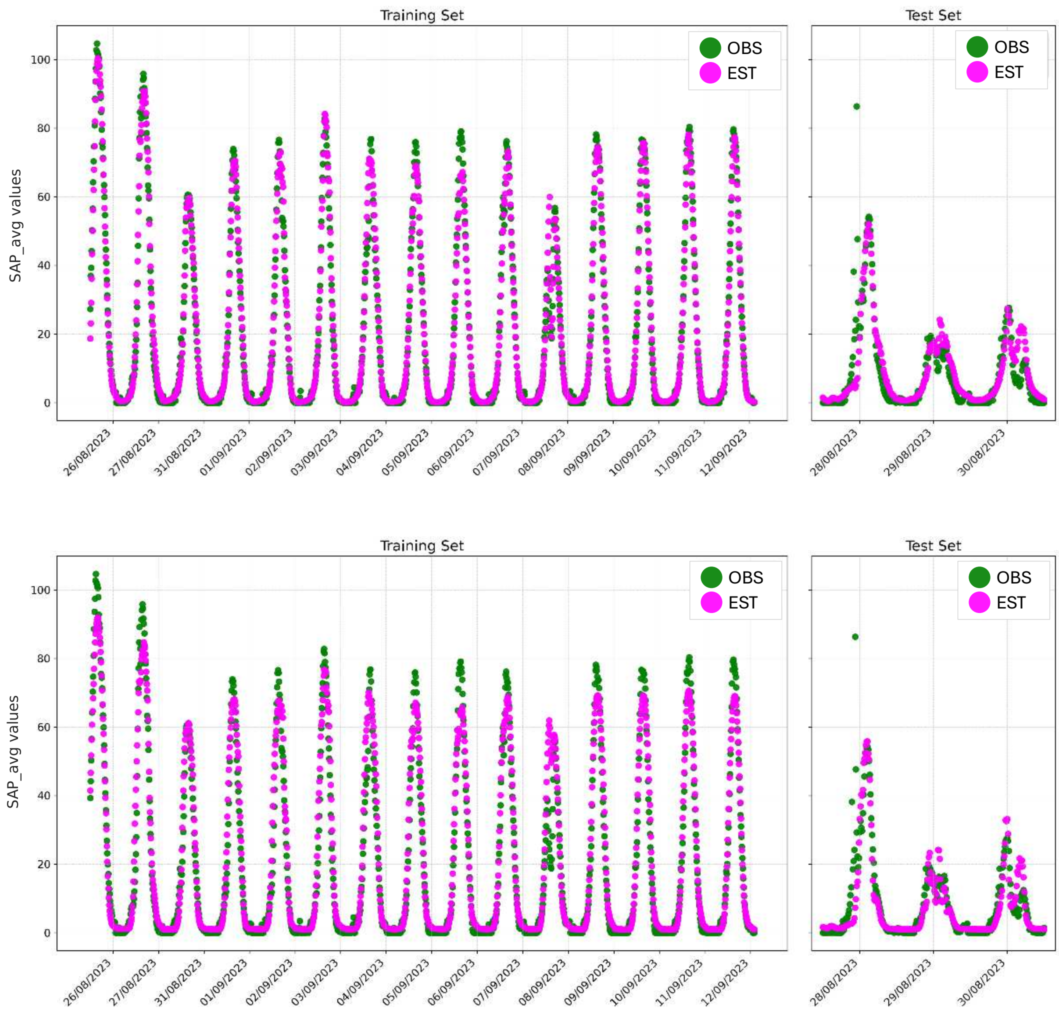


Fig. 10 SAP_avg estimated values, EST, compared against observed ones, OBS, for TE-PA (a) and TE-FC (b) with 45 minutes temporal windows using MLPR model (split2)

(a)

(b)

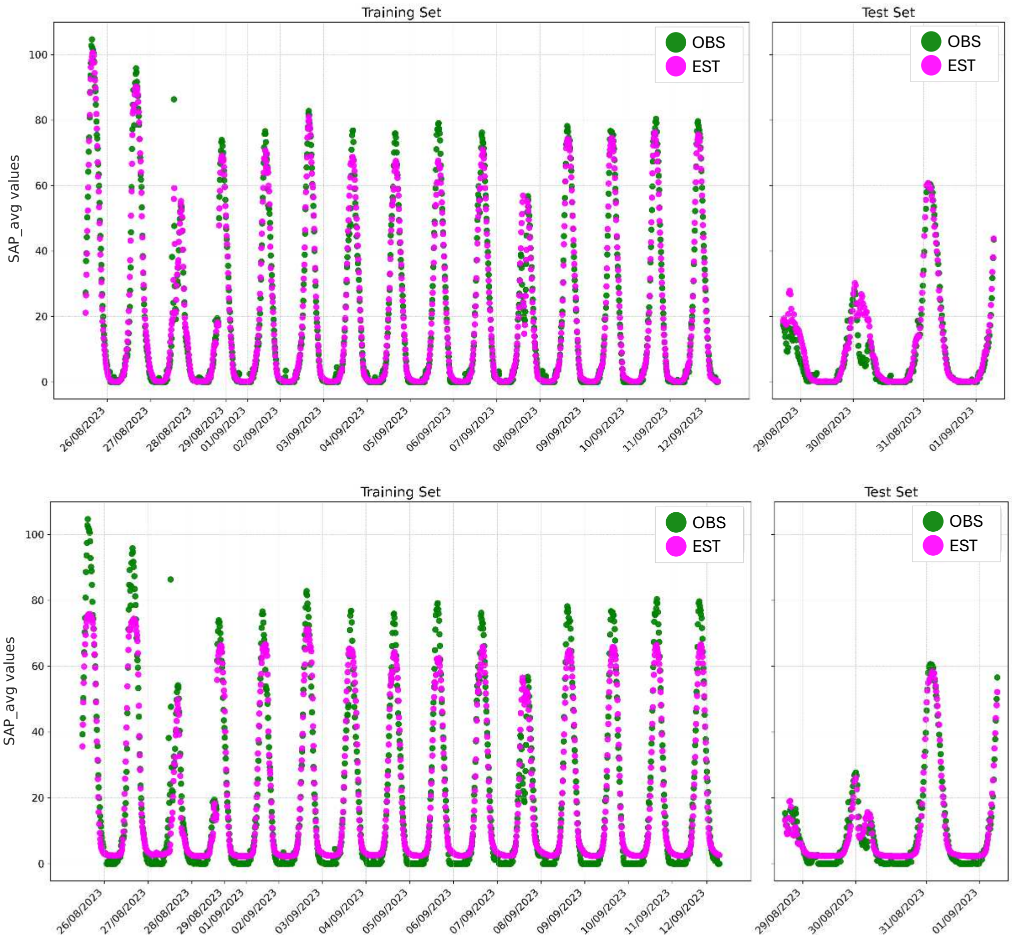


Fig. 11 SAP_avg estimated values EST, compared against observed ones, OBS, for TE-PA (a) and TE-FC (b) with 45 minutes temporal windows using MLPR model (split3)

2 hours temporal windows for TE-PA and TE-FC strategies

The results, obtained by concatenating the average (TE-PA) of the previous 8 samples (2 hours), are shown below along with those obtained by directly concatenating features (TE-FC). Table 7 shows the quantitative results, while Figures 12, 13 and 14 illustrate the qualitative results as the split change.

Table 7 Results of SVR and MLPR with best hyperparameters found with GridSearchCV.

		Temporal Estimate (TE) models							
		(a) TE-PA				(b) TE-FC			
	Model	RMSE		R2-score		RMSE		R2-score	
		train	test	train	test	train	test	train	test
split1	SVR	5.897	7.249	0.943	0.926	6.356	8.310	0.934	0.903
	MLPR	4.067	6.888	0.973	0.933	5.542	9.097	0.950	0.884
split2	SVR	2.485	9.519	0.991	0.409	2.746	8.169	0.989	0.565
	MLPR	2.274	7.347	0.993	0.648	2.833	6.816	0.989	0.697
split3	SVR	5.153	3.553	0.962	0.947	3.391	3.011	0.983	0.967
	MLPR	2.762	3.889	0.989	0.937	4.064	2.816	0.976	0.971

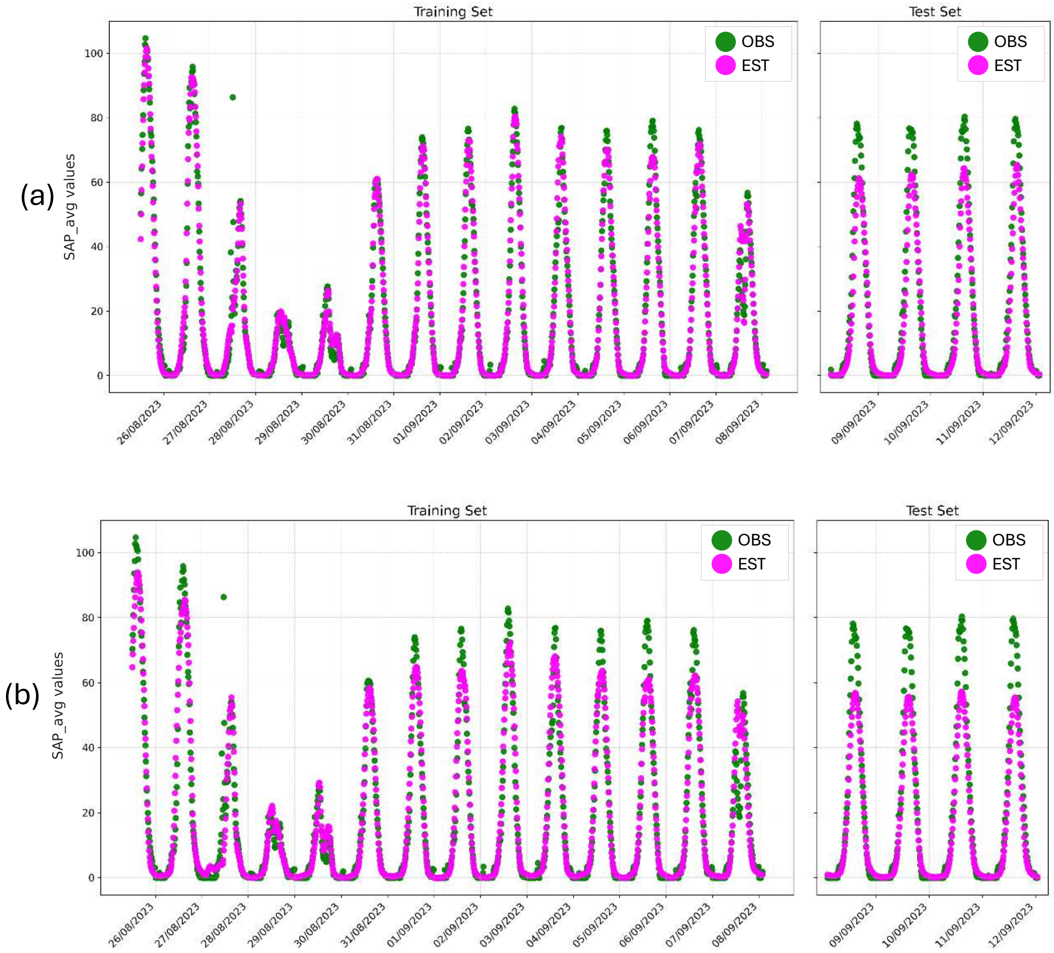
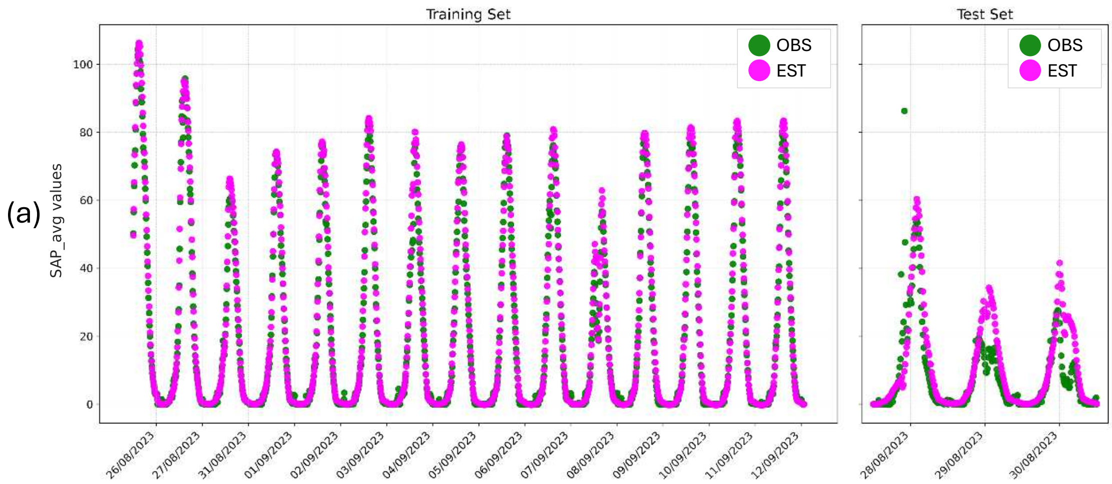


Fig. 12 SAP_avg estimated values, EST, compared against observed ones, OBS, for TE-PA (a) and TE-FC (b) with 2 hours temporal windows using MLPR model (split1)



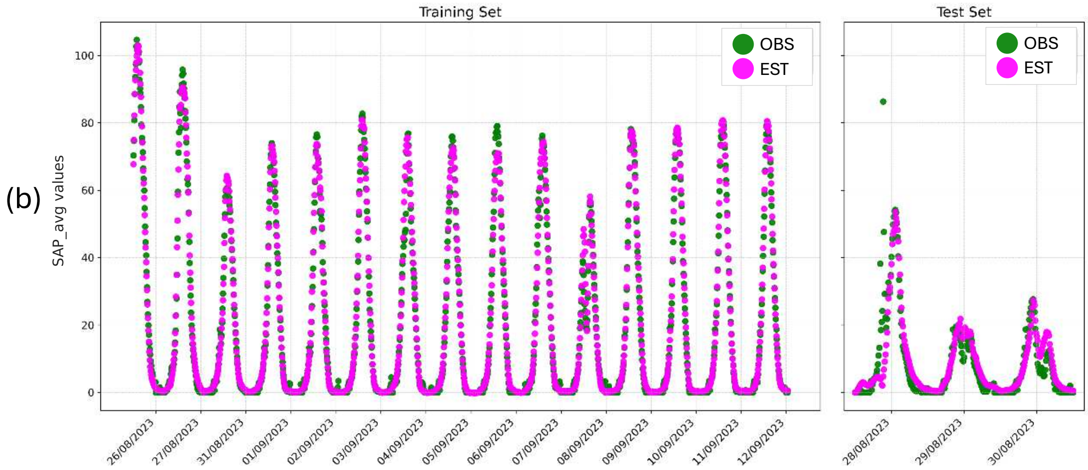


Fig. 13 SAP_avg estimated values, EST, compared against observed ones, OBS, for TE-PA (a) and TE-FC (b) with 2 hours temporal windows using MLPR model (split2)

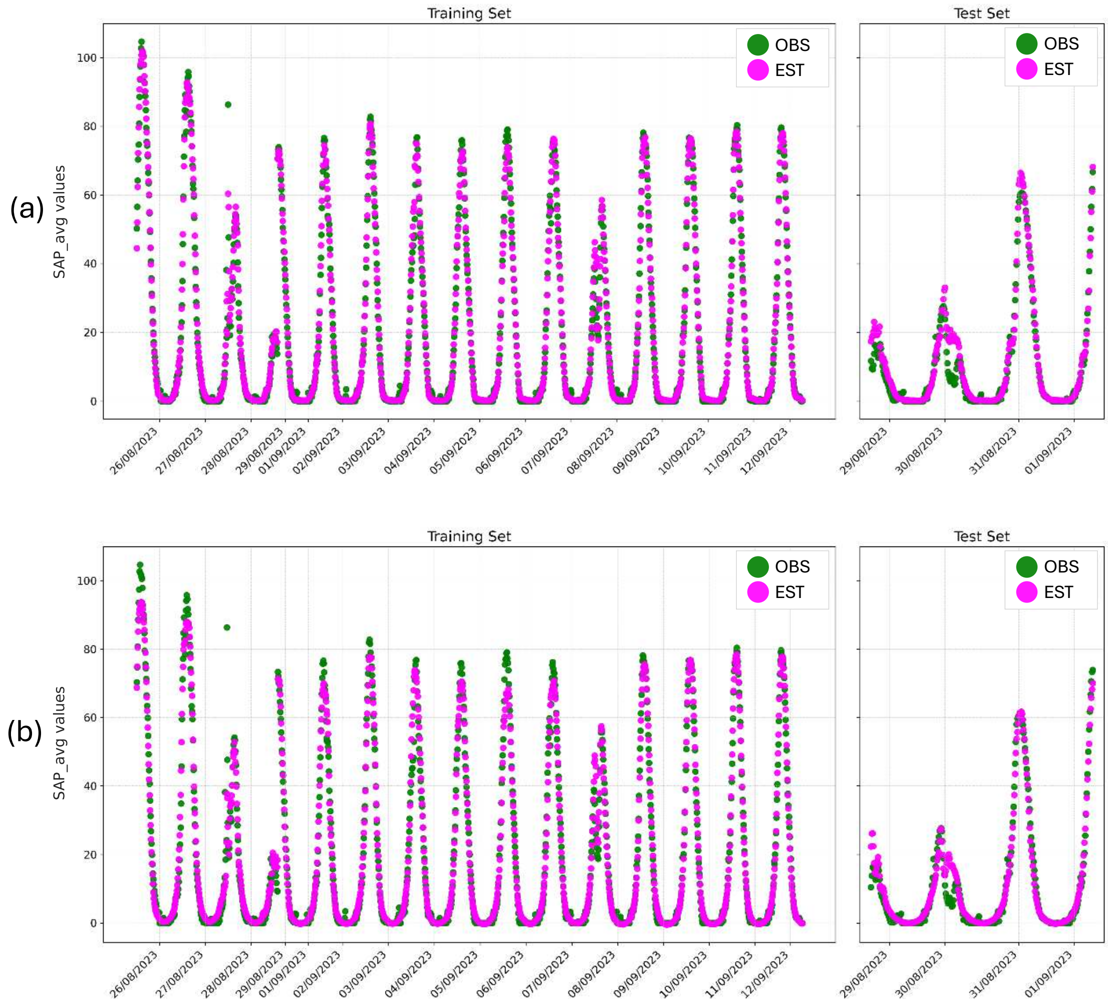


Fig. 14 SAP_avg estimated values, EST, compared against observed ones, OBS, for TE-PA (a) and TE-FC (b) with 2 hours temporal windows using MLPR model (split3)

6 hours temporal windows for TE-PA and TE-FC strategies

The results concatenating the average (TE-PA) of the previous 24 samples (6 hours) are shown below along with those obtained by directly concatenating features (TE-FC). Table 8 reports the quantitative metrics, while Figures 15, 16, 17 depict the corresponding outcomes for each split.

Table 8 Results of SVR and MLPR with best hyperparameters found with GridSearchCV.

		Temporal Estimate (TE) models							
		(a) TE-PA				(b) TE-FC			
	Model	RMSE		R2-score		RMSE		R2-score	
		train	test	train	test	train	test	train	test
split1	SVR	6.322	9.035	0.928	0.886	5.112	8.508	0.952	0.899
	MLPR	3.266	8.772	0.981	0.892	4.069	5.335	0.969	0.960
split2	SVR	2.893	8.932	0.987	0.480	2.450	10.547	0.991	0.275
	MLPR	3.146	7.073	0.985	0.674	2.742	7.281	0.988	0.655
split3	SVR	3.297	4.432	0.983	0.948	3.496	3.773	0.980	0.963
	MLPR	3.065	3.783	0.985	0.962	2.493	2.653	0.99	0.981

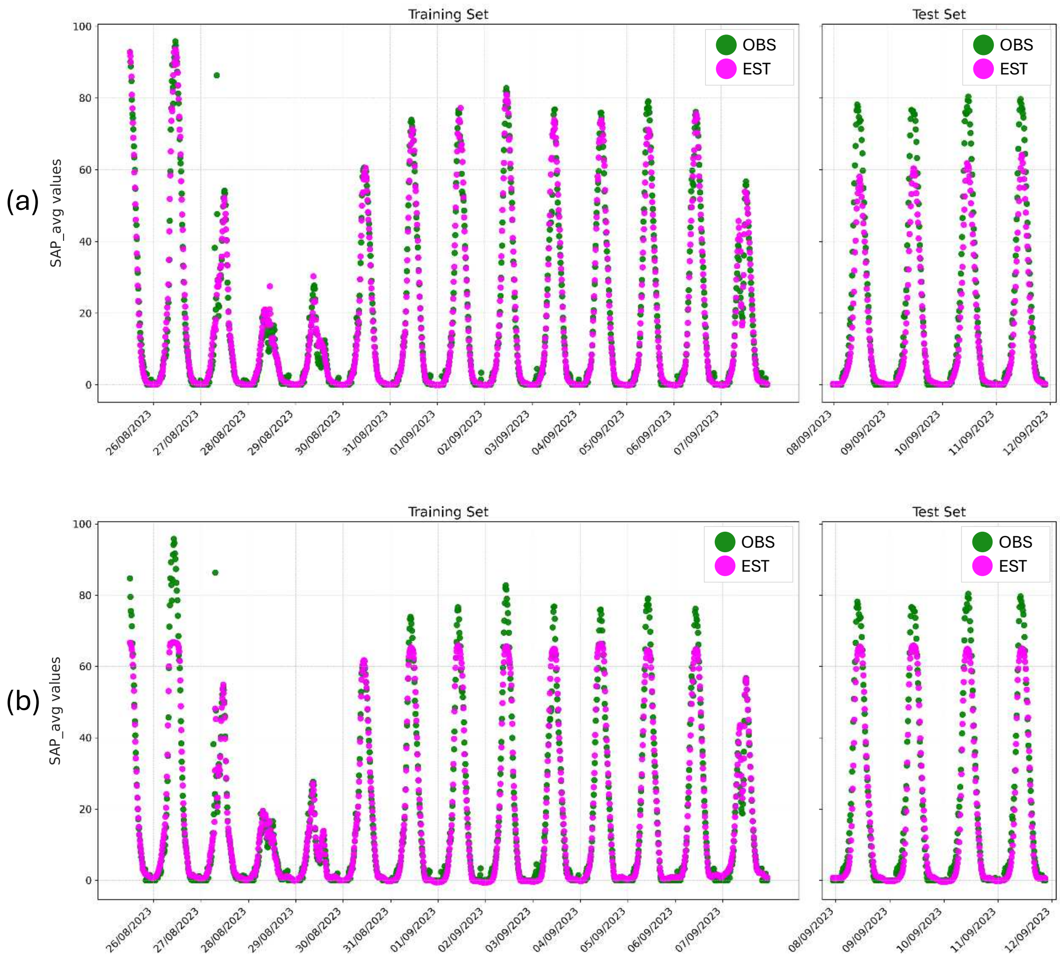


Fig. 15 SAP_avg estimated values, EST, compared against observed ones, OBS, for TE-PA (a) and TE-FC (b) with 6 hours temporal windows using MLPR model (split1)

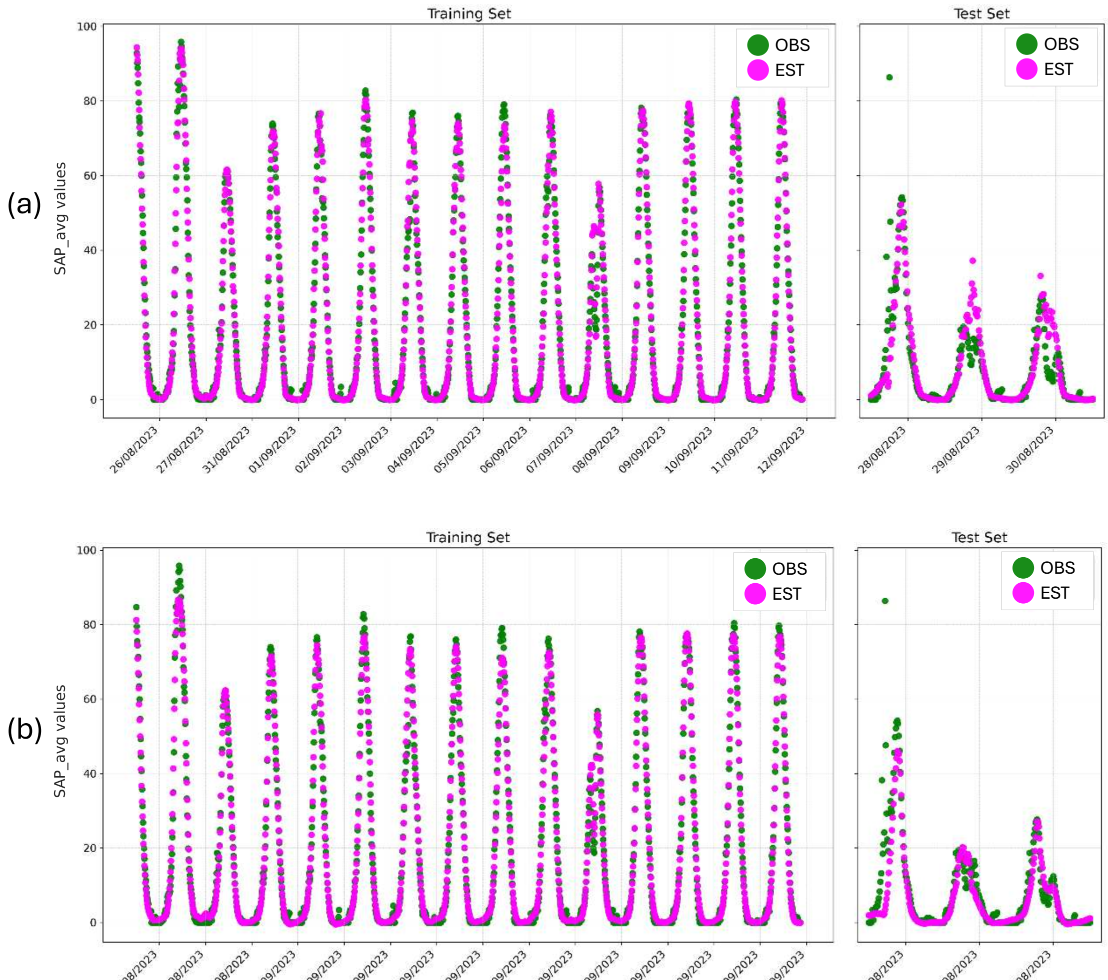
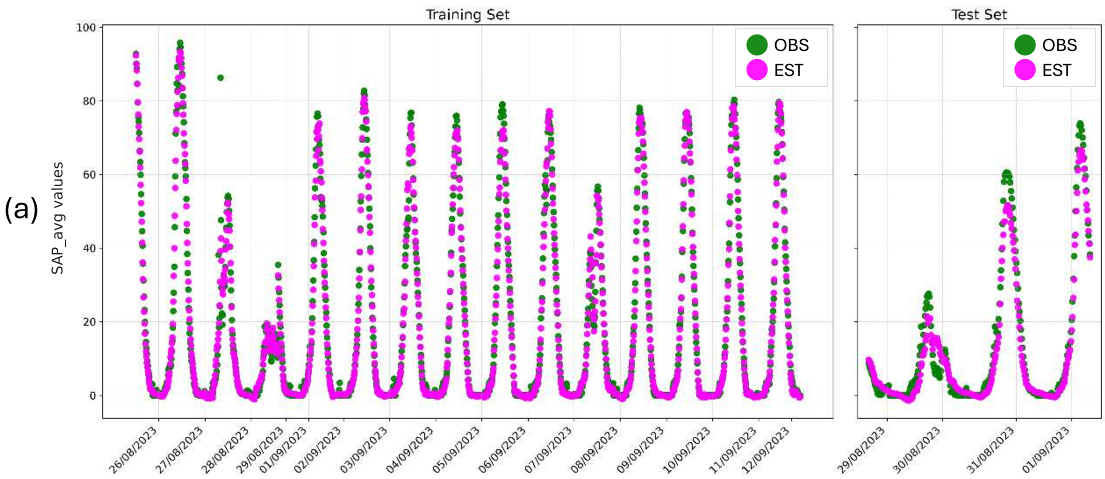


Fig. 16 SAP_avg estimated values, EST, compared against observed ones, OBS, for TE-PA (a) and TE-FC (b) with 6 hours temporal windows using MLPR model (split2)



(b)

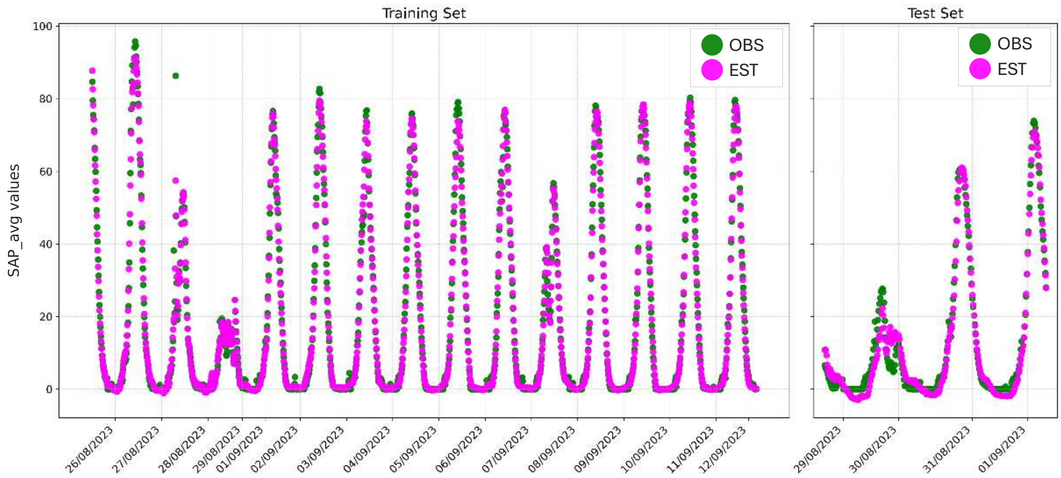


Fig. 17 SAP_avg estimated values, EST, compared against observed ones, OBS, for TE-PA (a) and TE-FC (b) with 6 hours temporal windows using MLPR model (split3)

Experiments with 2024 data

Quantitative results for 2024 are reported in Table 9, and qualitative outcomes for the MLPR model are shown in Figures 18 and 19. The differences observed between RMSE and R^2 across Exp1 and Exp2 (Section *"Data partitioning for 2024"*) are mainly explained by the variability of the target values in the corresponding test sets. In exp1, SAP_avg exhibits a much wider dynamic range, which increases the total variance of the ground truth and makes the R^2 metric less sensitive to large absolute errors. Conversely, the test set of exp2 shows a markedly lower variability, resulting in a smaller total variance; under these conditions, errors of comparable magnitude become proportionally more penalizing in the computation of R^2, leading to higher RMSE but comparatively higher R^2 scores.

Table 9 Results of SVR and MLPR with best hyperparameters found with GridSearchCV

	Model	RMSE		R2-score	
		train	test	train	test
exp1	SVR	4.392	16.535	0.970	0.796
	MLPR	4.261	15.152	0.971	0.828
exp2	**SVR**	5.948	7.431	0.969	0.932
	MLPR	5.291	9.208	0.976	0.896

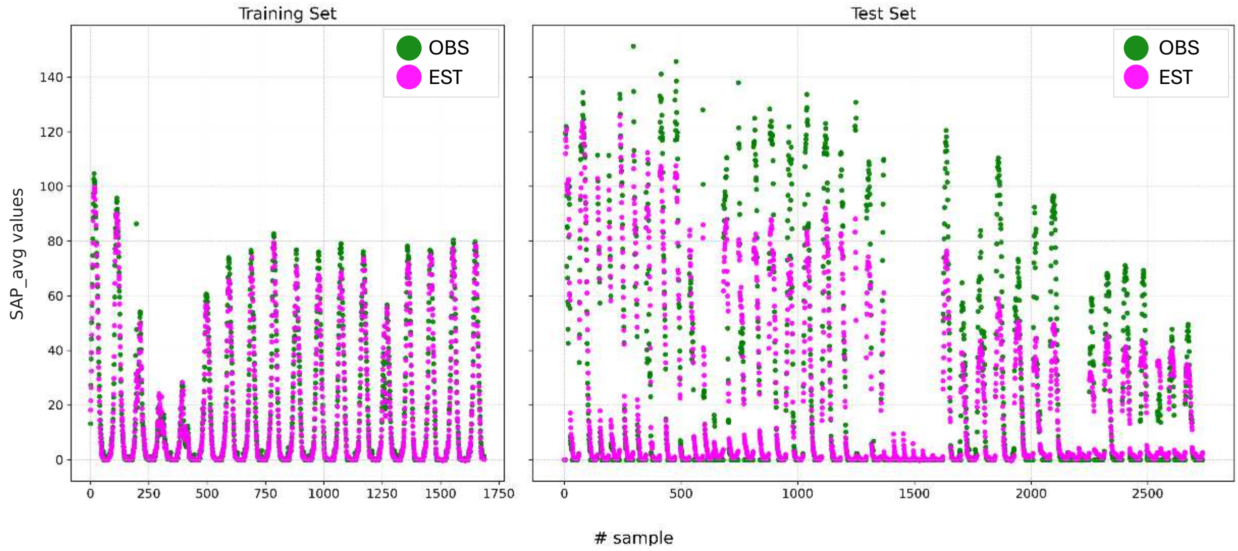


Fig. 18.SAP_avg values compared against observed ones for exp1 with MLPR model

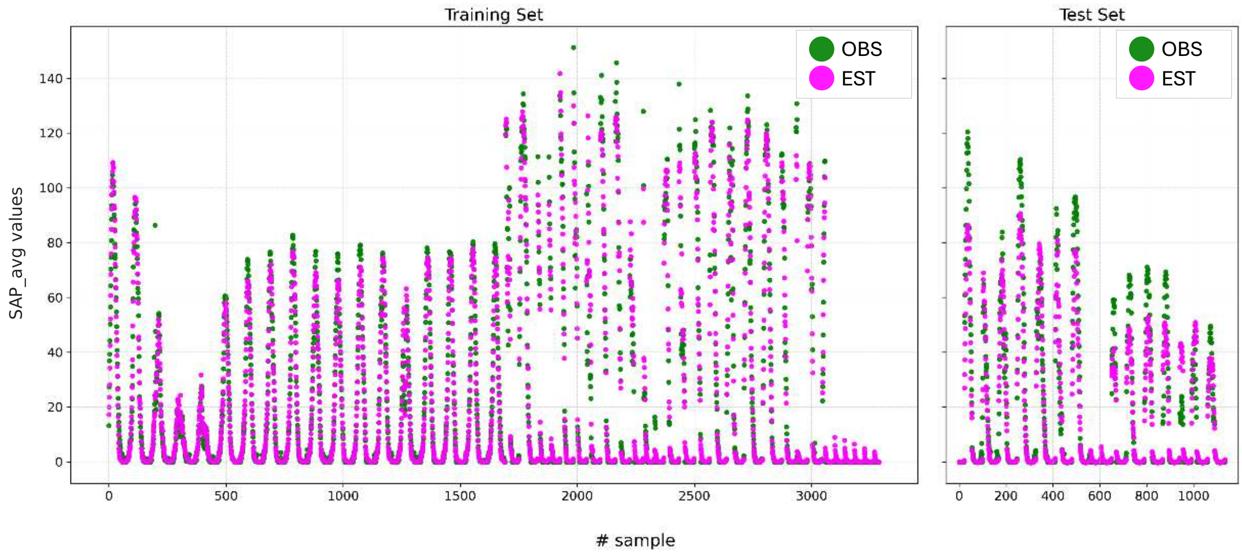


Fig. 19 SAP_avg values compared against observed ones for exp2 with MLPR model

Discussions

2023 data

Considering the results reported in the preceding figures and tables, the influence of historical information on the estimation of SAP_avg becomes clearer when comparing Point Estimate (PE) and Temporal Estimate (TE) strategies. Since LR and DTR showed limited predictive capability in the PE experiments, either because of poor generalization or evident overfitting, only the two best-performing models, MLPR and SVR, were retained for the temporal analyses.

For both MLPR and SVR, past observations affect performance in a way that depends on the temporal window and on the dataset split. Short windows, such as 45 minutes and 2 hours, generally provide stable improvements across most configurations, since they introduce additional temporal context without excessively smoothing the signal. This effect is not uniform across splits. In split1, short windows often yield the best balance between RMSE and R^2, while in split3 longer windows, particularly when using feature concatenation (TE-FC), frequently lead to the highest overall performance. This behaviour is consistent with the smoother temporal structure of the split3 test set, which benefits more from extended temporal information.

The TE-FC approach accentuates these differences. By treating each previous sample as an independent feature, it increases the dimensionality of the input space and allows the model, especially MLPR, to reconstruct the daily pattern of the sap flow signal more effectively. As a result, R^2 can rise substantially when the window is extended to 6 hours, even when RMSE does not improve to the same extent. This indicates that part of the gain comes from a better reproduction of the overall signal shape rather than from a uniform reduction in prediction errors (RMSE). In contrast, TE-PA, which relies on averaging past samples, tends to provide more conservative improvements, with limited benefits from longer windows, especially in split1 and split2.

Clear differences emerge when model performance is evaluated across the different data splits. In split2, both MLPR and SVR consistently achieve the lowest performance, as the reduced variability and irregular dynamics of the test set lead to higher RMSE values and lower R^2 scores across all temporal windows and for both temporal strategies. In split1, the models generally perform well when short temporal windows are used, although this configuration does not systematically yield the best results. In contrast, when evaluated on split3, the models, particularly under the TE-FC strategy with longer temporal windows, often achieve their highest performance. This behaviour can be attributed to the statistical properties of the split3 test set, which combines mixed yet smoother sap-flow dynamics. Under these conditions, the additional temporal context allows the models to better reconstruct the daily transpiration pattern, resulting in improved R^2 values.

Overall, the results show that the benefit of incorporating historical data depends strongly on both the temporal strategy and the characteristics of the test set. Short windows remain the most consistently reliable option, although longer windows combined with TE-FC can provide substantial improvements when the target signal exhibits regular and slowly varying dynamics, as observed in split3.

2024 data

Despite the technical discontinuities in the 2024 dataset, the stable predictive performance of MLPR and SVR confirms the robustness of the proposed framework. The results indicate that the models can generalize effectively across growing seasons when properly trained and that combining data from multiple years enhances adaptability and reliability.

Overall, these results suggest that ML-based models trained on multi-season data can effectively generalize across different growth cycles, provided that missing data are properly handled, and the feature set remains consistent. The differences observed between the two years mainly result from the data completeness rather than the intrinsic behaviour of the models, emphasizing the importance of continuous and high-quality data acquisition for reliable estimation of sap flow and transpiration in field conditions.

Conclusions

This study demonstrated the potential of data-driven frameworks to estimate maize sap flow transpiration using widely available climatic and soil moisture data. The comparative analysis of multiple machine learning models revealed that non-linear architectures, particularly MLPR and SVR, consistently outperform linear and tree-based approaches in both point and temporal estimation tasks.

A key finding of this research is that incorporating short-term historical information, specifically temporal windows of 45 minutes and 2 hours, significantly enhances predictive accuracy by providing relevant temporal context while preserving the intrinsic dynamics of the sap flow signal. In contrast, extending the temporal horizon beyond a few hours (e.g., 3–4 hours) tends to reduce reliability, as the process follows a pronounced 12-hour diurnal cycle with rapid daytime variations and near-zero night-time values.

Specifically, the Feature Concatenation (TE-FC) strategy allowed models to better reconstruct the complex patterns of transpiration, providing a more reliable estimation of crop water demand compared to traditional instantaneous predictors. The results obtained for the 2023 dataset show that concatenation strategies allow the models to better reconstruct the diurnal transpiration pattern, leading to higher R^2 values, although these gains do not always correspond to proportional reductions in RMSE. In contrast, the previous-averages approach provides more conservative improvements, typically yielding more stable error behaviour but limited benefits when longer temporal windows are considered. The effectiveness of temporal information was shown to depend strongly on the statistical properties of the test data, highlighting the importance of dataset characteristics when evaluating model performance.

The analysis conducted on the 2024 dataset confirms the robustness of the selected models under more challenging conditions, including data discontinuities and interannual variability. The 2024 database encompasses measurements collected under markedly different meteorological conditions, soil states, and crop responses, thus providing a broader and more representative characterization of sap flow dynamics that is essential for an effective and unbiased model training process. Despite the restriction to point estimation, MLPR and SVR demonstrated good generalization capability across growing seasons, particularly when multi-year data were combined during training.

Overall, the proposed modelling framework offers an effective and minimally invasive tool for estimating maize sap flow. These findings support the applicability of machine learning-based approaches for real-time irrigation management and decision support in precision agriculture. Future work will focus on extending temporal modelling to longer and more continuous datasets and on integrating additional data sources, such as remote sensing or plant physiological indicators, to further enhance model robustness and scalability.

Statements and Declarations

Author contributions T.G.: Data curation; Formal analysis; Investigation; Methodology; Resources; Supervision; Validation; Visualization; Writing – original draft; Writing – review and editing. L.M.: Data curation; Formal analysis; Investigation; Methodology; Resources; Software; Supervision; Validation; Visualization; Writing – original draft; Writing – review and editing. C.F.: Formal analysis; Investigation; Methodology; Writing – review and editing. C.G.: Conceptualization; Investigation; Methodology; Supervision; Validation; Visualization; Writing – review and editing. B.J.: Data curation; Methodology; Resources; Writing – review and editing. V.L.: Data curation; Methodology; Resources; Writing – review and editing. V.P.: Funding acquisition; Project administration; Supervision. T.F.: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Supervision; Validation; Visualization; Writing – original draft; Writing – review and editing.

Data availability Part of the datasets generated and analyzed during the current study are publicly available at https://github.com/isarlab-department-engineering/Agritech3.1.5FIWARE. The source code used for data processing and analysis will be released upon acceptance of the paper in the GitHub repository https://github.com/isarlab-department-engineering/DD_Maize_Sap_Flow.

Funding This work was carried out within the framework of the project Agritech National Research Center funded by the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR)—MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4—D.D. 1032 17 June 2022, CN00000022).

Declarations

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

References

Agarwal, M., Gupta, S.K., Biswas, K.K. (2020). Development of Efficient CNN model for Tomato crop disease identification. *Sustainable Computing: Informatics and Systems, 28*, 100407. https://doi.org/10.1016/j.suscom.2020.100407.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M. (1998). Crop Evapotranspiration-Guidelines for Computing Crop Water Requirements, Irrigation and Drain. *Paper No. 56. FAO*, Rome, Italy, 300 pp.

Amir, A., Butt, M., Van Kooten, O. (2021). Using machine learning algorithms to forecast the sap flow of cherry tomatoes in a greenhouse. *IEEE Access, 9,* 154183-154193. https://doi.org/10.1109/ACCESS.2021.3127453.

Amir, A., Butt, M. (2025). Improved sap flow prediction: A comparative deep learning study based on LSTM, BiLSTM, LRCN, and GRU with stem diameter data. *Smart Agricultural Technology, 12,* 101105. https://doi.org/10.1016/j.atech.2025.101105.

Asbjornsen, H., Goldsmith, G.R., Alvarado-Barrientos, M.S., Rebel, K., Van Osch, F.P., Rietkerk, M., Chen, J., Gotsch, S., Tobón, C., Geissert, D.R., Gómez-Tagle, A., Vache, K., Dawson, T.E. (2011). Ecohydrological advances and applications in plantwater relations research: A review. *Journal of Plant Ecology, 4(1-2),* 3–22. https://doi.org/10.1093/jpe/rtr005.

Banerjee, S., Reynolds, J., Taggart, M., Daniele, M., Bozkurt, A., Lobaton, E. (2024). Quantifying Visual Differences in Drought-Stressed Maize through Reflectance and Data-Driven Analysis. *AI, 5(2),* 790–802. https://doi.org/10.3390/ai5020040.

Barbosa, A., Trevisan, R., Hovakimyan, N., Martin, N.F., (2020). Modeling yield response to crop management using convolutional neural networks. *Comput. Electron. Agric., 170,* 105197. https://doi.org/10.1016/j.compag.2019.105197.

Bellocchio, E., Ciarfuglia, T. A., Costante, G., and Valigi, P. (2019) Weakly Supervised Fruit Counting for Yield Estimation Using Spatial Consistency. *IEEE Robotics and Automation Letters, 4 (3),* pp. 2348-2355, July 2019. https://doi.org/10.1109/LRA.2019.2903260.

Benos, L., Tagarakis, A.C., Dolias, G., Berruto, R., Kateris, D., Bochtis, D. (2021). Machine Learning in Agriculture: A Comprehensive Updated Review. *Sensors, 21(11)*. https://doi.org/10.3390/s21113758.

Bellido-Jimenez, J.A., Estevez, J., García-Marín, A.P. (2023). Reference evapotranspiration projections in Southern Spain (until 2100) using temperature-based machine learning models *Computers and Electronics in Agriculture, 214,* 108327. https://doi.org/10.1016/j.compag.2023.108327.

Bishop, C.M. (2006). Pattern Recognition and Machine Learning (Information Science and Statistics). Berlin, Heidelberg: Springer-Verlag. ISBN: 0387310738.

Borgonovo, F., Ferrante, V., Grilli, G., Pascuzzo, R., Vantini, S., Guarino, M.A. (2020). Data-Driven Prediction Method for an Early Warning of Coccidiosis in Intensive Livestock Systems: A Preliminary Study. *Animals, 10(4),* 747. https://doi.org/10.3390/ani10040747.

Bounoua, I., Saidi, Y., Yaagoubi, R., Bouziani, M. (2024). Deep Learning Approaches for Water Stress Forecasting in Arboriculture Using Time Series of Remote Sensing Images: Comparative Study between ConvLSTM and CNN-LSTM Models. *Technologies, 12(6),* 77. https://doi.org/10.3390/technologies12060077.

Bush, S.E., Hultine, K.R., Sperry, J.S., Ehleringer, J.R. (2010). Calibration of thermal dissipation sap flow probes for ring-and diffuse-porous trees. *Tree Physiology, 30(12),* 1545–1554. https://doi.org/10.1093/treephys/tpq096.

Čermák, J., Kučera, J., Nadezhdina, N. (2004). Sap flow measurements with some thermodynamic methods, flow integration within trees and scaling up from sample trees to entire forest stands. *Trees, 18,* 529-546. https://doi.org/10.1007/s00468-004-0339-6.

Chang, Y., Latham, J., Licht, M., Wang, L. (2023). A data-driven crop model for maize yield prediction. *Communications Biology, 6,* 439. https://doi.org/10.1038/s42003-023-04833-y.

Chen, Z., Sun, S., Wang, Y., Wang, Q., Zhang, X. (2020). Temporal convolution-network-based models for modeling maize evapotranspiration under mulched drip irrigation. *Computer and Electronics in Agriculture, 169,* 105206. https://doi.org/10.1016/j.compag.2019.105206.

Coulibaly, S., Kamsu-Foguem, B., Kamissoko, D., Traore, D. (2019). Deep neural networks with transfer learning in millet crop images. *Computers in Industry, 108,* 115–120. https://doi.org/10.1016/j.compind.2019.02.003.

Dou, X., Yang, Y. (2018). Evapotranspiration estimation using four different machine learning approaches in different terrestrial ecosystems. *Computers and Electronics in Agriculture, 148*, 95-106. https://doi.org/10.1016/j.compag.2018.03.010.

Fan J., Zheng J., Wu L., Zhang F. (2021) Estimation of daily maize transpiration using support vector machines, extreme gradient boosting, artificial and deep neural networks models. *Agricultural Water Management, 245,* 106547. https://doi.org/10.1016/j.agwat.2020.106547.

Fang, B., Lei, H., Zhang, Y., Quan, Q., Yang, D. (2020). Spatio-temporal patterns of evapotranspiration based on upscaling eddy covariance measurements in the dryland of the North China Plain. *Agricultural and Forest Meteorology, 281,* 107844. https://doi.org/10.1016/j.agrformet.2019.107844.

Feng, Y., Gong, D., Mei, X., Cui, N. (2017). Estimation of maize evapotranspiration using extreme learning machine and generalized regression neural network on the China Loess Plateau. *Hydrology Research, 48(4),* 1156–1168. https://doi.org/10.2166/nh.2016.099.

Ferreira, L.B., da Cunha, F.F. (2020). New approach to estimate daily reference evapotranspiration based on hourly temperature and relative humidity using machine learning and deep learning. *Agricultural Water Management*, *234*, 106113. https://doi.org/10.1016/j.agwat.2020.106113.

Fine, L.,Richard, A., Tanny, J., Pradalier, C.,Rosa, R.,Rozenstein, O. (2022). Introducing State-of-the-Art Deep Learning Technique for Gap-Filling of Eddy Covariance Crop Evapotranspiration Data. *Water, 14(5),* 763. https://doi.org/10.3390/w14050763.

Giménez, C., Gallardo, M., Thompson, R. B. (2013). Plant-water relations. Reference Module *in Earth Systems and Environmental Sciences*. Amsterdam, The Netherlands: Elsevier. https://doi.org/10.1016/B978-0-12-409548-9.05257-X.

Gómez, D., Salvador, P., Sanz, J., Casanova, J.L. (2019). Potato Yield Prediction Using Machine Learning Techniques and Sentinel 2 Data. *Remote Sensing, 11(15),* 1745. https://doi.org/10.3390/rs11151745.

Granier A. (1985). Une nouvelle méthode pour la mesure du flux de sève brute dans le tronc des arbres. *Annales des Sciences forestières EDP Sciences, 42(2),* 193–200. https://doi.org/10.1051/forest:19850204.

Jasechko, S., Sharp, Z.D., Gibson, J.J., Birks, S.J., Y,i Y., Fawcett, P.J. (2013). Terrestrial water fluxes dominated by transpiration. *Nature, 496,* 347–350. https://doi.org/10.1038/nature11983.

Jiang, X.L., Kang, S.Z., Li, F.S., Du, T.S., Tong, L., Comas, L. (2016). Evapotranspiration partitioning and variation of sap flow in female and male parents of maize for hybrid seed production in Arid region. *Agricultural Water Management, 176,* 132-141. https://doi.org/10.1016/j.agwat.2016.05.022.

Jones, H.G. (2004) Irrigation scheduling: advantages and pitfalls of plant-based methods. *Journal of Experimental Botany, 55(407),* 2427–2436. https://doi.org/10.1093/jxb/erh213.

Joshi, A., Pradhan, B., Gite S., Chakraborty, S. (2023). Remote-Sensing Data and Deep-Learning Techniques in Crop Mapping and Yield Prediction: A Systematic Review. *Remote Sensing, 15(8),* 2014. https://doi.org/10.3390/rs15082014.

Kamarudin, M.H., Ismail, Z.H., Saidi, N.B. (2021). Deep Learning Sensor Fusion in Plant Water Stress Assessment: A Comprehensive Review. *Applied Sciences, 11(4),* 1403. https://doi.org/10.3390/app11041403.

Kamilaris, A., Prenafeta-Boldú, F.X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture, 147,* 70–90. https://doi.org/10.1016/j.compag.2018.02.016.

Kišš, V., Manina, M. (2019). Effect of Sap Flow at Optimization Irrigation. *19th International Multidisciplinary Scientific GeoConference SGEM 2019, 19,* 161-168. https://doi.org/10.5593/sgem2019/3.1/S12.021.

Kumar, R., Hosseinzadehtaher, M., Hein, N., Shadmand, M., Jagadish, SVK., Ghanbarian, B. (2022). Challenges and advances in measuring sap flow in agriculture and agroforestry: A review with focus on nuclear magnetic resonance. *Frontiers in Plant Science, 13*. https://doi.org/10.3389/fpls.2022.1036078.

Lascano, R.J., Goebel, T.S., Booker, J., Baker, J.T., Gitz III, D.C. (2016). The Stem Heat Balance Method to Measure Transpiration: Evaluation of a New Sensor. *Agricultural Sciences, 7(9),* 604-620. https://doi.org/10.4236/as.2016.79057.

Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D. (2018). Machine Learning in Agriculture: A Review. *Sensors, 18(8),* 2674. https://doi.org/10.3390/s18082674.

Li, Y., Ye, J., Xu, D., Zhou, G., Feng, H. (2022). Prediction of sap flow with historical environmental factors based on deep learning technology. *Computers and Electronics in Agriculture, 202(C),*107400. https://doi.org/10.1016/j.compag.2022.107400.

Li, Y., Guo, L., Wang, J., Wang, Y., Xu, D., Wen, J. (2023). An Improved Sap Flow Prediction Model Based 652 on CNN-GRU-BiLSTM and Factor Analysis of Historical Environmental Variables. *Forests, 14(7),* 1310. https://doi.org/10.3390/f14071310.

Liu, M., Shi, H., Paredes, P., Ramos, T.B., Dai, L., Feng, Z., Pereira, L.S. (2022). Estimating and partitioning maize evapotranspiration as affected by salinity using weighing lysimeters and the SIMDualKc model. *Agricultural Water Management, 261,* 107362. https://doi.org/10.1016/j.agwat.2021.107362.

Liu, Y.-Y., Su, Y.-C., Sun, P.-W., Dai, H.-Y., Kuo, B.-J. (2025). Stability of Maize Phenology Predictions by Using Calendar Days, Thermal Functions, and Photothermal Functions. *Agriculture, 15(19),* 2020. https://doi.org/10.3390/agriculture15192020.

Mana, A.A., Allouhi, A., Hamrani, A., Rahman, S., El Jamaoui, I., Jayachandran, K. (2024). Sustainable AI-Based Production Agriculture: Exploring AI Applications and Implications in Agricultural Practices. *Smart Agricultural Technology, 7,* 100416. https://doi.org/10.1016/j.atech.2024.100416.

Maya Gopal., P.S., Bhargavi, R. (2019). Performance Evaluation of Best Feature Subsets for Crop Yield Prediction Using Machine Learning Algorithms. *Applied Artificial Intelligence, 33(7),* 621–642. https://doi.org/10.1080/08839514.2019.1592343.

Nadezhdina, N. (1999). Sap flow index as an indicator of plant water status. *Tree Physiology, 19(13)*, 885-891. https://doi.org/10.1093/treephys/19.13.885.

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. *Ecological Informatics 76,* 102133. https://doi.org/10.1016/j.ecoinf.2023.102133.

Ren, R., Liu, G., Wen, M., Horton, R., Li, B., Si, B. (2017). The effects of probe misalignment on sap flux density measurements and in situ probe spacing correction methods. Agricultural and *Forest Meteorology, 232,* 176–185. https://doi.org/10.1016/j.agrformet.2016.08.009.

Riaboff, L., Poggi, S., Madouasse, A., Couvreur, S., Aubin, S., Bédère, N., Goumand, E., Chauvin, A., Plantier, G. (2020). Development of a methodological framework for a robust prediction of the main behaviours of dairy cows using a combination of machine learning algorithms on accelerometer data. *Computers and Electronics in Agriculture, 169*, 105179. https://doi.org/10.1016/j.compag.2019.105179.

Rosa, R.D., Ramos, T.B., Pereira, L.S. (2016). The dual Kc approach to assess maize and sweet sorghum transpiration and soil evaporation under saline conditions: Application of the SIMDualKc model. *Agricultural Water Management, 177,* 77-94. https://doi.org/10.1016/j.agwat.2016.06.028.

Santaga, F. S., Agnelli, A., Leccese, A., Vizzari, M. (2021). Using Sentinel-2 for simplifying soil sampling and mapping: two case studies in Umbria, Italy. *Remote sensing, 13(17),* 3379. https://doi.org/10.3390/rs13173379.

Schwalbert, R.A., Amado, T., Corassa, G., Pott, L.P., Prasad, P.V.V., Ciampitti, I.A. (2020). Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil. *Agricultural and Forest Meteorology, 284*, 107886. https://doi.org/10.1016/j.agrformet.2019.107886.

Steppe, K., De Pauw, D.J., Doody, T.M., Teskey, R.O. (2010). A comparison of sap flux density using thermal dissipation, heat pulse velocity and heat field deformation methods. *Agricultural and Forest Meteorology, 150(7–8),* 1046–1056. https://doi.org/10.1016/j.agrformet.2010.04.004.

Steppe, K., Vandegehuchte, M.W., Tognetti, R., Mencuccini, M. (2015). Sap flow as a key trait in the understanding of plant hydraulic functioning. *Tree Physiology, 35 (4),* 341-345. https://doi.org/10.1093/treephys/tpv033.

Sun, A.Y., Scanlon, B.R. (2019). How can Big Data and machine learning benefit environment and water management: A survey of methods, applications, and future directions. *Environmental Research Letters, 14,* 73001. https://doi.org/10.1088/1748-9326/ab1b7d.

Tang, D, Feng, Y, Gong, D., Weiping H., Ningbo, C. (2018). Evaluation of artificial intelligence models for actual crop evapotranspiration modelling in mulched and non-mulched maize croplands. *Computers and Electronics in Agriculture*, *152*, 375–384. https://doi.org/10.1016/j.compag.2018.07.

Tu, J., Wei, X., Huang, B., Fan, H., Jian, M., Li, W. (2019). Improvement of sap flow estimation by including phenological index and time-lag effect in back-propagation neural network models. *Agricultural and Forest Meteorology*, *276-277*, 107608. https://doi.org/10.1016/j.agrformet.2019.06.007.

Xing, L., Cui, N., Liu, C., Zhao, L., Guo, L., Du, T., Zhan, C., Wu, Z., Wen, S., Jiang, S. (2022). Estimation of daily apple tree transpiration in the Loess Plateau region of China using deep learning models. *Agricultural Water Management, 273,* 107889. https://doi.org/10.1016/j.agwat.2022.107889.

Yadav, A., Muni, L., Upreti, H., Das Singhal, G. (2024). Evaluation of crop water stress index of wheat by using machine learning models. *Environmental Monitoring and Assessment, 196*, 970. https://doi.org/10.1007/s10661-024-13113-z.

Yamaç, S.S., Todorovic, M. (2020). Estimation of daily potato crop evapotranspiration using three different machine learning algorithms and four scenarios of available meteorological data. *Agricultural Water Management, 228,* 105875. https://doi.org/10.1016/j.agwat.2019.105875.

Vadez, V., Grondin, A., Chenu, K., Henry, A., Laplaze, L., Millet, E.J., Carminati, A. (2024). Crop traits and production under drought. *Nature Reviews Earth & Environment, 5,* 211-225. https://doi.org/10.1038/s43017-023-00514-w.

Virnodkar, S.S., Pachghare, V.K., Patil, V.C., Jha, S.K. (2020). Remote sensing and machine learning for crop water stress determination in various crops: a critical review. *Precision Agriculture, 21,* 1121–1155. https://doi.org/10.1007/s11119-020-09711-9.

Wäldchen, J., Rzanny, M., Seeland, M., Mäder, P. (2018). Automated plant species identification—Trends and future directions. *PLoS Computational Biology, 14,* e1005993. https://doi.org/10.1371/journal.pcbi.1005993.

Wei, Z., Yoshimura, K., Wang, L., Miralles, D.G., Jasechko, S., Lee, X. (2017). Revisiting the contribution of transpiration to global terrestrial evapotranspiration. *Geophysical Research Letters, 44(6),* 2792–2801. https://doi.org/10.1002/2016GL072235.

Whitley, R., Zeppel, M., Armstrong, N., Macinnis-Ng, C., Yunusa, I., Eamus, D. (2008). A modified Jarvis-Stewart model for predicting stand-scale transpiration of an Australian native forest. *Plant and Soil, 305*, 35–47. https://doi.org/10.1007/s11104-007-9399-x.

Wullschleger, S. D., Childs, K.W., King, A.W., Hanson, P.J. (2011). A model of heat transfer in sapwood and implications for sap flux density measurements using thermal dissipation probes. *Tree Physiology, 31(6),* 669–679. https://doi.org/10.1093/treephys/tpr051.

Zenone, T., Vitale, L., Famulari, D., Magliulo V. (2022). Application of machine learning techniques to simulate the evaporative fraction and its relationship with environmental variables in corn crops. *Ecological Process, 11,* 54. https://doi.org/10.1186/s13717-022-00400-1.

Zhang, B., Zhang, D., Feng, Z., Zhang, L., Zhang, M., Fu, R., Wang, Z. (2023). Assessment of the Potential of Indirect Measurement for Sap Flow Using Environmental Factors and Artificial Intelligence Approach: A Case Study of Magnolia denudata in Shanghai Urban Green Spaces, *Forests, 14(9),* 1768. https://doi.org/10.3390/f14091768.

Zhang, J., Ding, Y., Zhu, L., Wan, Y., Chai, M., Ding, P. (2025a). Estimating and forecasting daily reference crop evapotranspiration in China with temperature-driven deep learning models. *Agricultural Water Management, 307,* 109268 https://doi.org/10.1016/j.agwat.2024.109268.

Zhang, J., Yang, S., Wang, J., Zeng, R., Zhang, S., Bai, Y., Zhang, J. (2025b). Evapotranspiration Partitioning for Croplands Based on Eddy Covariance Measurements and Machine Learning Models. *Agronomy, 15(3),* 512. https://doi.org/10.3390/agronomy15030512.

Zou, H., Hastie T. (2005). Regularization and Variable Selection Via the Elastic Net. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 67(2),* 301–320, ISSN:1369-7412. https://doi.org/10.1111/j.1467-9868.2005.00503.x.

Zhong, L., Hu, L., Zhou, H. (2019). Deep learning based multi-temporal crop classification. *Remote Sensing of Environment, 221,* 430–443. https://doi.org/10.1016/j.rse.2018.11.032.