Ecological Informatics 85 (2025) 102991

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

A machine learning model using the snapshot ensemble approach for soil respiration prediction in an experimental Oak Forest

S.N. Ferdous [a,*], J.P. Ahire [b,c], R. Bergman [c], L. Xin [d], E. Blanc-Betes [e,f], Z. Zhang [g], J. Wang [h]

[a] Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV, 26505, USA
[b] Biological Systems Engineering Department, University of Wisconsin, Madison, WI, 53706, USA
[c] Forest Products Laboratory, USDA Forest Service, Madison, WI 53726, USA
[d] Department of Computer Science, University at Albany, New York, NY 12222, USA
[e] Institute for Sustainability, Energy, and Environment, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[f] Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[g] Biological Systems Engineering, University of Wisconsin, Madison, WI 53706, USA
[h] Department of Forest Biomaterials, North Carolina State University, Raleigh, NC 27695, USA

ARTICLE INFO

Dataset link: OakForest Dataset (Original data)

Keywords:
Biogeochemical model
CO_2
Forest ecosystem
ML model
Soil carbon

ABSTRACT

Soil respiration plays a key role in regulating ecosystem CO_2 emissions and atmospheric concentrations. However, because it is highly sensitive to the environment and has large temporal and spatial variability, its contribution to the global carbon cycle remains highly uncertain. Recent advances in ecosystem process models have uniquely constrained soil respiration rates, largely based on results from tracing ecosystem CO_2 losses using radioisotope and stable isotope methods. However, the computational requirements of process-based models at large scales are costly, limiting their application and prediction capability. To mitigate this requirement for high computational power for the assimilation of high-throughput data to identify patterns and trends in soil respiration and guide in deploying effective strategies to reduce Greenhouse gas (GHG) emission, the application of machine-learning (ML) models can help reduce the requirement of intensive computational capacity and improve the accuracy of predictions. Here, we developed a novel ML model that integrated a hybrid Prophet-ANN and snapshot ensemble approach. First, we extracted temporal features using the Prophet Forecasting model. We then used the Artificial Neural Network (ANN) regression model to correct the forecasting model errors and perform the final prediction using the snapshot ensemble approach. This model can serve as a surrogate for the ForCent model applied to the Oak Ridge National Laboratory deciduous forest. Our proposed model accurately predicted soil CO_2 flux for the four studied sites: East Low, West Low, East High, and West High. We show the potential of our proposed ML model to significantly improve soil CO_2 flux predictions, which in turn can help us develop more effective strategies for managing soil carbon.

1. Introduction

The forest ecosystem which covers 31 % of the global land area, is one of the most promising solutions to combat global climate change. Forests capture 14.1 PgC yr^{-1} of CO_2 through photosynthesis and release 11.6 PgC yr^{-1} through respiration and fires, resulting in a net storage of 2.5 PgC $yr^{-}1$ (Nunes et al., 2020; Singh et al., 2022). Approximately 40 % of the global soil C stocks reside in forests (Mayer et al., 2020). It functions as a carbon sink, assimilating atmospheric CO_2 into soil and biomass. Through photosynthesis, terrestrial ecosystems uptake large amounts of atmospheric CO_2, which is assimilated into plant biomass and eventually incorporated into soil organic matter pools. However, part of this CO_2 returns to the atmosphere through the ecosystem respiration (R_{eco}), including microbial (heterotrophic) and plant (autotrophic) respiration (R_{het} and R_{aut}, respectively). The net balance of all contributing CO_2 fluxes determines the size and turnover times of the soil organic carbon pool, ultimately defining the carbon sink or source strength of the ecosystem.

Soil respiration (SR) is crucial in regulating net ecosystem CO_2 emissions in forest systems (Heimann and Reichstein, 2008; Lloyd and Taylor, 1994). SR rates are sensitive to changing environments and disturbances (Rodtassana et al., 2021). Due to the magnitude of CO_2

* Corresponding author.
E-mail address: xli48@albany.edu (S.N. Ferdous).

https://doi.org/10.1016/j.ecoinf.2024.102991
Received 12 August 2024; Received in revised form 29 October 2024; Accepted 31 December 2024
Available online 6 January 2025

exchange fluxes of forest ecosystems, small changes in SR can substantially impact atmospheric CO_2 concentrations (Nissan et al., 2023; Schlesinger and Andrews, 2000). Large uncertainty in the determination of SR fluxes results from scale-up of measurement technique and different methods of measurement techniques (Hashimoto et al., 2023; Pumpanen et al., 2004). Reducing this uncertainty would improve understanding of the global carbon cycle and the implications terrestrial ecosystems on climate.

Process-based models are invaluable tools to reduce this uncertainty. Based on our mechanistic understanding of ecosystem processes, process-based models reproduce key processes such as belowground carbon allocation (root growth), soil organic matter (SOM) formation and decomposition, nutrient release, and interactions with environmental parameters such as soil water and nutrient content (Zani et al., 2023). Ecosystem process-based models, such as DayCent, use soil carbon dynamics models with decay constants to calculate decomposition and carbon transfer between pools (Berardi et al., 2020). Hence, a detailed understanding of the complex relationships between forests and CO_2 emissions is essential for developing effective strategies for ecosystem management and mitigating climate change.

DayCent used litter bag techniques, commonly used for studying the decomposition dynamics of SOM (Kurzbesson et al., 2005; Xie, 2020; Xie et al., 2017). However, not devoid of uncertainty, prediction accuracy of DayCent is limited by poor understanding of plant physiology, microbial activity including turnover times, interactions with the environment, and fate of decomposition products (dos Reis Martins et al., 2022). Major limitations of this technique, however, are the underrepresentation of long-term turnover rates and contributions from slow decomposing pools, and the inability to directly assess the fate of nutrients and organic matter compounds derived from decomposition processes. Radiocarbon studies stand out as an instrumental tool for precise measurement of turnover rates for different parts of carbon cycle in soil.

Isotopic approaches with ^{13}C and ^{14}C tracers have been extensively used to identify contribution pathways to main SOM pools, estimate mean residence time of the organic material and track changes in slow decomposing SOM pools (Lynch et al., 2006; Parton et al., 2010), providing a unique opportunity to improve model performance. Parton et al., used data from the Enriched Background Isotope Study (EBIS) ^{14}C to calibrate, develop and test the ForCent model, a forest-centric version of the DayCent model that built seasonal plant growth and root development into the model subroutines. Further, it enhances the determination of carbon flux in different pools by adding a plant-stored carbohydrate pool and a slow surface litter pool (Parton et al., 2010). This improvement provides a new approach for the partitioning between various sources of heterotrophic respiration and the subsequent effect of leaf and root; it offers insights into the complex processes that control the long-term stabilization of SOM. This ForCent model uses EBIS ^{14}C isotopic study to determine turnover rates of different SOM pools with high accuracy. It shows progress in accuracy for predicting critical aspects of soil carbon cycle and organic matter stabilization in different pools.

The present study seeks to enhance the precision and efficacy of the existing approach using ML techniques as a surrogate for the ForCent model. The use of ^{14}C-tracer studies to develop the DayCent model is restricted to a small scale because of the high cost and adverse environmental effects (Andresen et al., 2018). The development of a process-based model is costly and requires high resources, and it is challenging to use for guidance on how climate and management changes affect sustainability and GHG emissions (Lardy et al., 2014; Seppelt and Voinov, 2002). Further, it requires long-term field histories and showed low accuracy when calibration process-based models were used to predict CO_2 emissions (McClelland et al., 2021). The results and the model's ability to predict CO_2 emissions depend on many factors, including field experiments that do not accurately measure emissions, systems that do not accurately parameterize process-based models, and mathematics that do not accurately describe the system's complexity.

The ability of ML can potentially transcend the existing capabilities of the ForCent model; it offers an advanced approach to analyzing the dynamic interplay between litter decomposition, SOM formation, and GHG emissions. Researchers have successfully used different ML models to predict soil GHG emissions (Harsányi et al., 2024; Lu et al., 2021). Recently, ensemble learning (Mendes-Moreira et al., 2012), a machine learning-based approach, was found effective for soil CO_2 flux prediction. Generally, the performance of ensemble models is better than single models as ensemble modeling performs prediction combining multiple model predictions. Classical Random Forest (RF) has been found to be an efficient ML model for agricultural CO_2 flux prediction (Abbasi et al., 2021; Hamrani et al., 2020; Tavares et al., 2018). The RF model has also been successfully applied for other GHG emission prediction such as N_2O emission prediction (Philibert et al., 2013; Saha et al., 2021). Random Forest (RF) and Gradient Boosting (GB) regressor have been identified as the best performing model for soil CO_2 flux prediction (Adjuik and Davis, 2022). (Zhao et al., 2017) predicted biome-specific global soil respiration from 1960 to 2012 using Artificial Neural Network (ANN). ML models combined with unmanned aerial vehicle (UAV) spectral information have been proven to be effective for soil respiration estimation in desertified mining areas (Liu et al., 2023). A stacking-based ensemble model has been proposed for predicting GHG emission from irrigated paddy fields (Jiang et al., 2023). A study on soil CO_2 emission found RF and generalized regression neural network (GRNN) promising for soil respiration prediction of reforested areas in the Brazilian Cerrado site (Alzubaidi et al., 2021). Recently, spatial distribution pattern and driving mechanism of global soil respiration have been analyzed using ML and big data (Liu et al., 2024). (Wang et al., 2024) developed a clustering-based framework, CLUSTER-ARIMA for soil respiration prediction. A new framework, Dish-ECA-Adain-Autoformer (DEAF) has been proposed for soil CO_2 flux prediction (Yang et al., 2024).

In this study, we proposed a novel ensemble-based model for CO_2 flux prediction. Ensemble models, such as RF, have been recognized for their high accuracy in CO_2 flux prediction (Abbasi et al., 2021; Mirzaei et al., 2023). Our model exceeds the performance benchmarks set by RF. The enhanced ability of our model to predict CO_2 flux is attributed to the ensemble approach, which amalgamates predictions from multiple models to capture complex interactions among input variables effectively. Previously, the ML model has been used as a surrogate of DayCent for the agriculture landscape (Nguyen et al., 2019). However, to our knowledge, there are no reports on using ML models to surrogate DayCent for predicting GHG emissions from forest lands. Existing methods for soil respiration prediction ignore the relationship between soil carbon pool and SR. Considering this limitation, we develop a novel ML model that can function as a powerful surrogate in refining the ForCent model, especially in calibrating the turnover rates of different SOM pools using process-based ecosystem models. The objective of this study is to pave the way for more accurate predictions related to the soil carbon cycle in forests and to help develop a toolset to guide decision-making in response to the challenges posed by a changing climate. The model is implemented in Python.

In this study, we propose to estimate SR of forests and CO_2 flux in different soil organic pools using an ML model that works as a surrogate for the ForCent model. The specific objectives of this study are (1) assess the ML model's capability to estimate soil respiration in forests and predict CO_2 flux across different soil carbon pools; (2) propose a novel ML modeling-based approach for CO_2 flux prediction that can be used as surrogate for ForCent; and (3) introduce the usability of new predictor variables (mineral soil respiration, litter respiration by machine learning model) for CO_2 fluxes estimation.

2. Methods

2.1. Description of database

This study utilized data derived from the ForCent model published by Parton et al. (Parton et al., 2010), which focuses on carbon cycling dynamics in forest ecosystems, particularly the partitioning of autotrophic and heterotrophic respiration processes. The primary purpose of the ForCent model was to address issues in the carbon cycle related to the partitioning of autotrophic and heterotrophic respiration processes for forest ecosystems. The data originated from a 65- to 150-year-old Oak Ridge National Laboratory deciduous forest located on ridgetop and upslope positions. Five independent variables were chosen for analysis: air temperature (°C), soil temperature (°C), daily precipitation (cm d^{-1}), daily delta $^{13}C/^{14}C$ values for both surface litter respiration (g C m^{-2} d^{-1}) and mineral soil respiration (g C m^{-2} d^{-1}). CO_2 flux served as the dependent variable. A detailed statistical characterization of the input features is provided in Table 1.

The experiment involved two soil types and two levels of ^{14}C exposure established in 1999. Reciprocal transplants of enriched and near-background litter were implemented at sites (i.e., East, West) with contrasting levels of ^{14}C exposure. Enriched ^{14}C leaf litter was collected from the western site in fall 2000, while background ^{14}C litter was collected from the eastern site during the same period.

Near background and enriched ^{14}C leaf litter were added to the plots in May 2001, with continued additions of elevated and ambient leaf litter (during winter months) for the next two years. Plots in the replicated experimental design included those with (1) ^{14}C enriched.

soil carbon, root litter, and leaf litter; (2) ^{14}C enriched roots, soil carbon, and near background leaf litter; (3) near-background roots, soil carbon, and elevated ^{14}C leaf litter; and (4) near background leaf litter, roots, and soil carbon. The ^{14}C content of surface litter, humus, mineral soil layers, and soil respiration rates were measured from 2001 to 2005. The "East high site" refers to the plot in the eastern location that was subjected to high treatments of ^{14}C enriched leaf litter. "East low site" represents the plot in the eastern location subjected to low or near-background treatments of ^{14}C in leaf litter. Depending on the specific experiment, "high" and "low" refer to several different variables such as the amount of ^{14}C in the litter, soil carbon, or root litter. The same principle is used for the "West high site" and "West low site". These refer to different experimental plots in a western location of the study area that were subjected to high and low treatments, respectively. The diversity in experiment design allows for a range of conditions under which the researchers can study carbon dynamics in the ecosystem. A detailed statistical characterization of the input features is provided in Table 2.

Table 1
Input data specification for CO_2 flux forecasting.

Input		Abbreviation	Description	Unit
Avg. Temp		atemp	Mean of max and min temperature	°C
Soil Temperature		stemp	Average soil temperature near the soil surface	°C
Precipitation		prec	Precipitation for day	Cm d^{-1}
Intermediate Variable	Litter respiration	pdsrfclit	Daily delta $^{13}C/^{14}C$ value for heterotrophic respiration for the surface litter (surface metabolic, structural, som1c, and som2c) prediction using a machine learning model	g Cm^{-2} d^{-1}
	Mineral soil respiration	pdsmnrl	Daily delta $^{13}C/^{14}C$ value for heterotrophic respiration for the mineral soil (soil metabolic, structural, som1c, som2c, and som3c) prediction using a machine learning model	g Cm^{-2} d^{-1}

2.2. Preprocessing

To develop a robust deep-learning model, we needed to pre-process the dataset. The dataset is split into a training set covering the years 1980 to 2000 and a test set spanning 2001 to 2005. The dataset contains a total of 9498 observations. To standardize the dataset, the median was removed, and data were scaled between 25th and 75th quantile range. We also generated a new feature average temperature (atemp) which is calculated by taking the average of minimum temperature (mint) and maximum temperature (maxt).

2.3. Machine learning models

To estimate CO_2 flux, we initially assessed the effectiveness of traditional regression models in predicting CO_2 flux, as elaborated upon in the following section.

2.3.1. Classical regression models

- KNN: K-Nearest Neighbors (KNN) (Maltamo and Kangas, 1998) is a non-parametric supervised ML algorithm popular for both classification and regression. In KNN regression, the algorithm finds the k-nearest neighbors of a new data point and assigns the average (or weighted average) of their target values as the prediction.
- MLP: Multi-layer perceptron (MLP) (Mohamadi et al., 2021; Ramchoun et al., 2016) is a type of ANN used for regression. MLP regression follows a parametric approach that involves training a neural network with multiple layers.
- CART: CART (Gey and Nedelec, 2005) is a decision tree-based regression model that builds a binary tree structure recursively, where each node in the tree represents a decision based on a feature, and the leaves contain the predicted values.
- SVR: Support Vector Regression (SVR) (Noble, 2006) is a regression algorithm based on Support Vector Machines (SVM). SVR focuses on minimizing the error while still staying within a margin around the true output.
- RF: Random Forest (RF) Regression (Segal, 2003) is an ensemble learning method based on the RF algorithm. It extends the idea of decision tree regression by constructing a multitude of decision trees during training and outputting the average prediction of the individual trees for regression tasks.
- LR: Linear Regression (LR) (Seber and Lee, 2003) is a simple and widely used supervised learning algorithm for predicting a continuous outcome variable (also called the dependent variable) based on one or more predictor variables (independent variables). The fundamental idea behind linear regression is to model the relationship between the predictor variables and the outcome variable as a linear eq.
- BR: Bayesian Regression (BR) (Bishop and Tipping, 2003) is a supervised regression model that incorporates Bayesian principles into the linear regression model. Bayesian regression provides a distribution of possible parameter values. This distribution is updated based on prior knowledge (prior distribution) and observed data, resulting in a posterior distribution of parameters.
- GB: Gradient Boosting (GB) Regression (Natekin and Knoll, 2013) is an ensemble learning-based regression model that builds a predictive model in the form of an ensemble of weak learners, usually decision

Table 2
Statistical characteristics of input features for oak forest site.

Features	Min	Max	Mean	Variance	Skewness	Kurtosis
Atemp (°C)	−1.400	30.75	14.4957	53.709	−0.174	9.235
Stemp (°C)	−19.371	32.6522	11.0568	10.4763	0.2363	−1.388
prec (cm d^{-1})	−99.900	12.100	0.31328	2.21845	−38.611	1751.68
pdsrfclit (g Cm^{-2} d^{-1})	98.625	320.779	190.925	50.314	0.5264	−0.602
pdsmnrl (g Cm^{-2} d^{-1})	150.952	286.850	219.307	38.450	4.3195	103.471

trees. The model is built sequentially by adding weak learners to correct the errors of the existing ensemble.

2.3.2. Snapshot ensemble

Ensemble learning (Arbib, 1998; Ferdous et al., 2023; Mohamadi et al., 2022,Sagi and Rokach, 2018) is an ML technique combining multiple models to predict an outcome. Snapshot ensemble (Huang et al., 2017) is an ML technique where several diverse models (snapshot models) are created by converging to multiple local minima. Snapshot models are saved after the completion of each cycle. During the convergence to multiple local minima, the Cyclic Cosine Annealing algorithm (Loshchilov and Hutter, 2017) is used to schedule the learning rate of the model. In this analysis, the learning rate for the schedule used the following formula:

$$\eta = 1 + \frac{\eta_{init}}{2} cos \left(\frac{\pi \, mod \left(t-1, \left\lceil \frac{T}{M} \right\rceil \right)}{\left\lceil \frac{T}{M} \right\rceil} \right) \tag{2}$$

Here, η_{init} is the initial learning rate, η is the updated learning rate, t is the current iteration number, T is the total number of iterations, and M denotes the number of cycles.

2.3.3. Proposed model

To develop an ML model suitable for predicting CO_2 flux, which can also act as a surrogate of ForCent, we followed the design principles of ForCent. Our novel hybrid ensemble model adopts a two-step training approach for CO_2 flux prediction considering both meteorological variables (air temperature, soil temperature, precipitation) and intermediate variables (mineral soil respiration, litter respiration predicted by ML) similar to ForCent. First, we predict surface and mineral soil respiration using supervised ML regressor models. To incorporate temporal information effectively, we perform forecasting. Extracted features from the best forecasting model (in our case, Prophet) are concatenated with input features. Five features (air temperature, soil temperature, precipitation, predicted mineral soil respiration, and predicted litter respiration) along with the extracted Prophet features were selected for the final model simulation on the testing dataset. This selection provided the most robust performance across all algorithmic scenarios examined, evidenced by a higher R^2 value. Nevertheless, the feature set constructed in this two-stage way outperformed training with three set features (air temperature, soil temperature, precipitation) for CO_2 flux prediction. The final prediction is performed by the snapshot ensemble model (Huang et al., 2017). Fig. 1 illustrates the framework of the proposed prediction model.

The proposed deep hybrid model comprises Prophet (Taylor and Letham, 2018) and ANN (Wang, 2003). Prophet is a time series forecasting model based on an additive model used to capture the underlying temporal relationship of data. The Prophet model decomposes a time series into three components: trend, seasonality, and holiday effect, which can be expressed using the following formula:

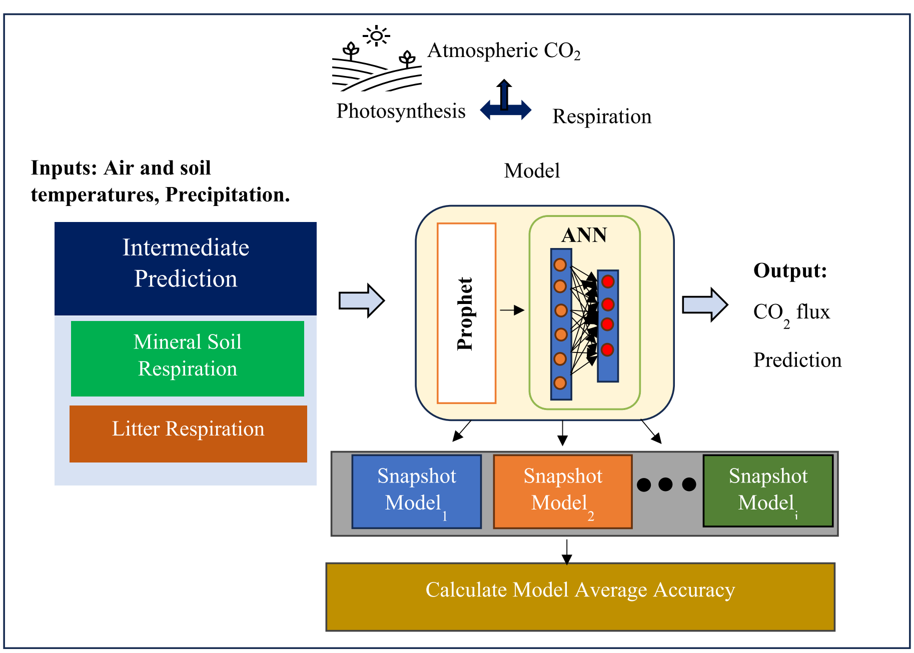

Fig. 1. Methodology framework of a proposed deep hybrid snapshot ensemble model for CO_2 flux forecasting. First, temporal features are extracted using Prophet forecasting model. The candidate feature set is constructed by concatenating original features (air temperature, soil temperature, precipitation), extracted temporal features by Prophet and features predicted by ML (mineral soil respiration and litter respiration). The concatenated input is fed through ANN. Final prediction is performed using a snapshot ensemble approach.

$$Y(t) = g(t) + s(t) + h(t) + \epsilon(t) \quad (1)$$

where, $g(t)$ is the trend function, $s(t)$ represents the periodic seasonal effects, $h(t)$ is the holiday effects, and $\epsilon(t)$ the residuals. The trend represents non-periodic changes in data. Seasonality represents periodic changes (daily, weekly, monthly, yearly) in data. The holiday effect tracks the occurrence of events on irregular intervals, and residuals are the error terms not explained by the model.

The extracted features from the Prophet model (trend, seasonality, additive terms, predicted soil respiration) along with five input features (air temperature, soil temperature, precipitation, predicted mineral soil respiration, and predicted litter respiration) are fed to the ANN. ANN works similarly to the human biological brain system. ANN models the problem as a non-linear approximation. It comprises input, hidden, and output layers interconnected by neurons. The input first passes through the input layer. Then, the hidden layer performs a series of transformations, and activation functions (e.g., sigmoid, tanh, etc.) are applied to compute the hidden representation of features. Finally, the output layer performs specified prediction, in this case, CO_2 flux prediction.

The last step of our proposed model was snapshot ensembles, where diverse snapshot models of hybrid Prophet-ANN were created, and final forecasting was performed by computing the average of the predictions from snapshot models.

3. Model development

This section delineates the development process of our snapshot ensemble model for the "East low site," a methodology that is extendable to the other three sites examined in this study. The model aims to predict one year ahead CO_2 flux, using cross-validation. Initially, the Prophet statistical forecasting model was employed for temporal feature extraction pertinent to CO_2 flux forecasting.

We conducted a preliminary Pearson's correlation analysis to examine the relationship between input and output variables. Fig. 2 illustrates a significant correlation between 'Average_Temperature' and both 'Maximum_Temperature' and 'Minimum_Temperature' prompting the exclusion of the latter two features to focus solely on 'Average_Temperature'. This decision also served to simplify the problem space by reducing its dimensionality. The correlation matrix in Fig. 2 identified air and soil temperatures as pivotal predictors of CO_2 flux, while precipitation was deemed the least influential one. The analysis supports the findings presented in (Joshi et al., 2022). Also, the analysis validates the findings, highlighting strong correlations between CO_2 flux and both air and soil temperatures (Hamrani et al., 2020; Oertel et al., 2016).

3.1. Intermediate variable generation

While designing our soil respiration prediction framework, we considered two intermediate components: litter respiration and mineral soil respiration. This design principle is inspired from ForCent model soil C pool (Parton et al., 2010). In the development of the model's feature set, the original feature set was enhanced with intermediate features (predicted mineral soil respiration and litter respiration) derived from supervised ML regressors. To predict mineral soil respiration and litter respiration, various ML models underwent cross-validation. The box-plot performance, illustrating the comparative efficacy of different ML algorithms for both mineral soil and litter respiration, is depicted in Fig. 3. The RF model emerged as the most effective, accounting for approximately 99 % of the variation in litter respiration prediction, as shown in Fig. 3(a). Conversely, the GB Regression model surpassed other ML models in predicting mineral soil respiration, explaining around 96 % of the variation, as illustrated in Fig. 3(b). Leveraging the highest-performing models, we generated predictions for litter respiration and mineral soil respiration to refine our forecasting model. The annual and 2001–2005 predictions for both litter respiration and mineral soil respiration are presented in Figs. 4 and 5, respectively.

3.2. Temporal feature extraction

To assess the temporal feature extraction capability for CO_2 flux forecasting, a wide range of forecasting models were analyzed:

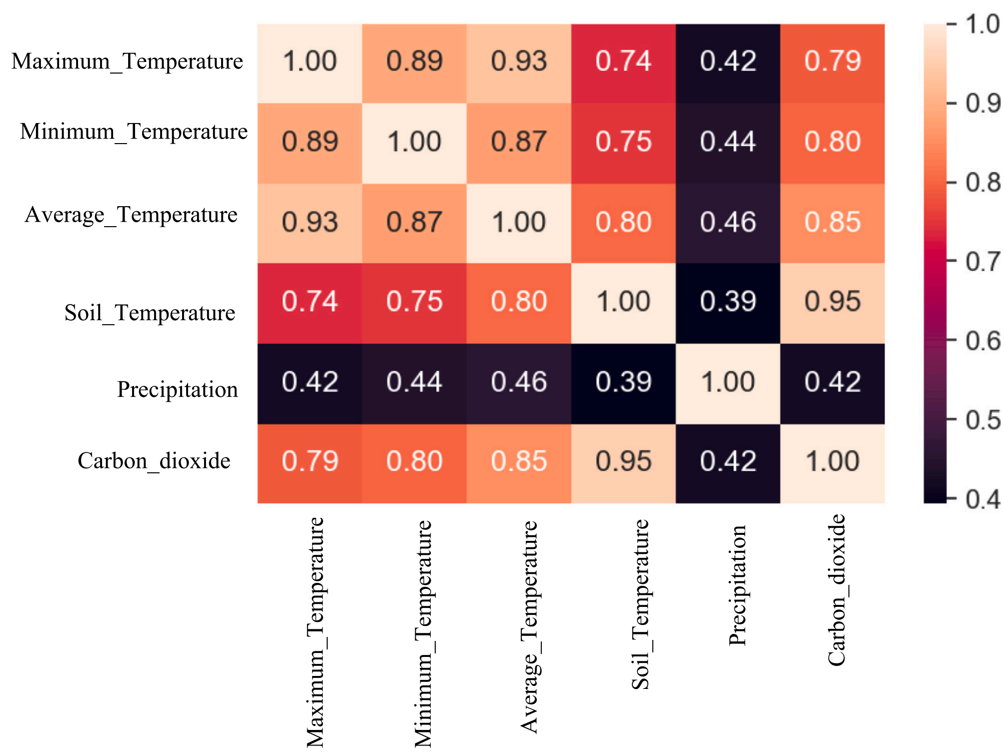

Fig. 2. Cross-correlation matrix showing different correlations between features.

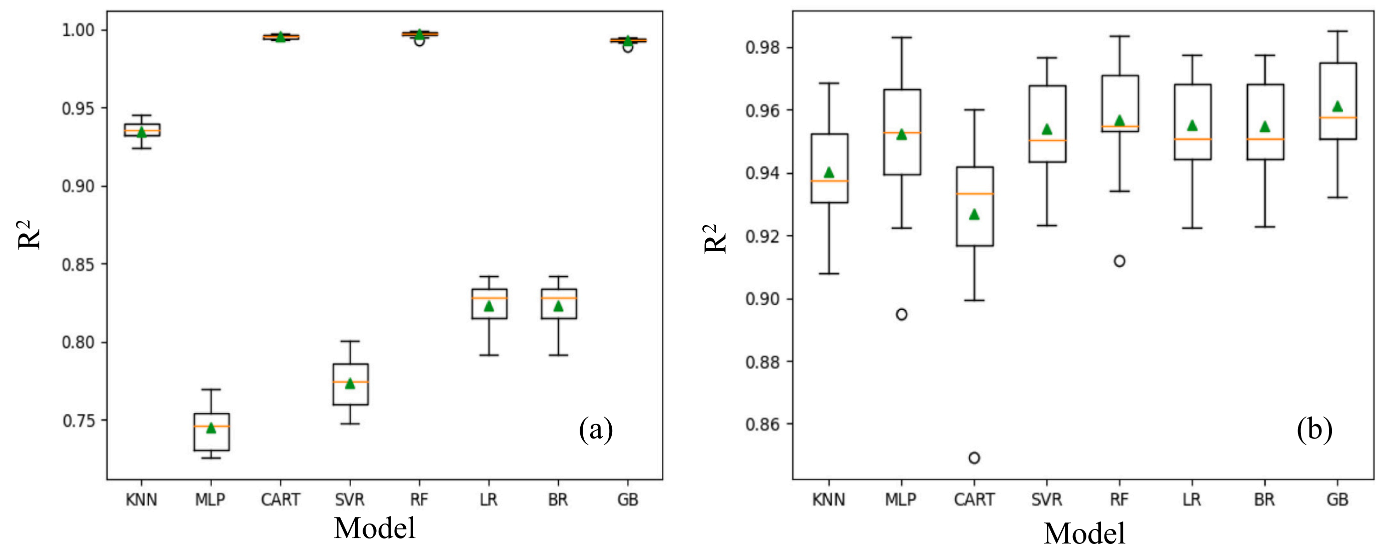

Fig. 3. Box plot performance of predicted values of (a) litter respiration (b) mineral soil respiration for oak forest low east site. A higher R^2 indicates a better-performing model.
K-nearest neighbor: KNN, Multilayer Perceptron (MLP), Decision Tree (CART), Support Vector Regression (SVR), Random Forest (RF), Linear Regression (LR), Bayesian Ridge Regression (BR), Gradient Boosting (GB) regression.

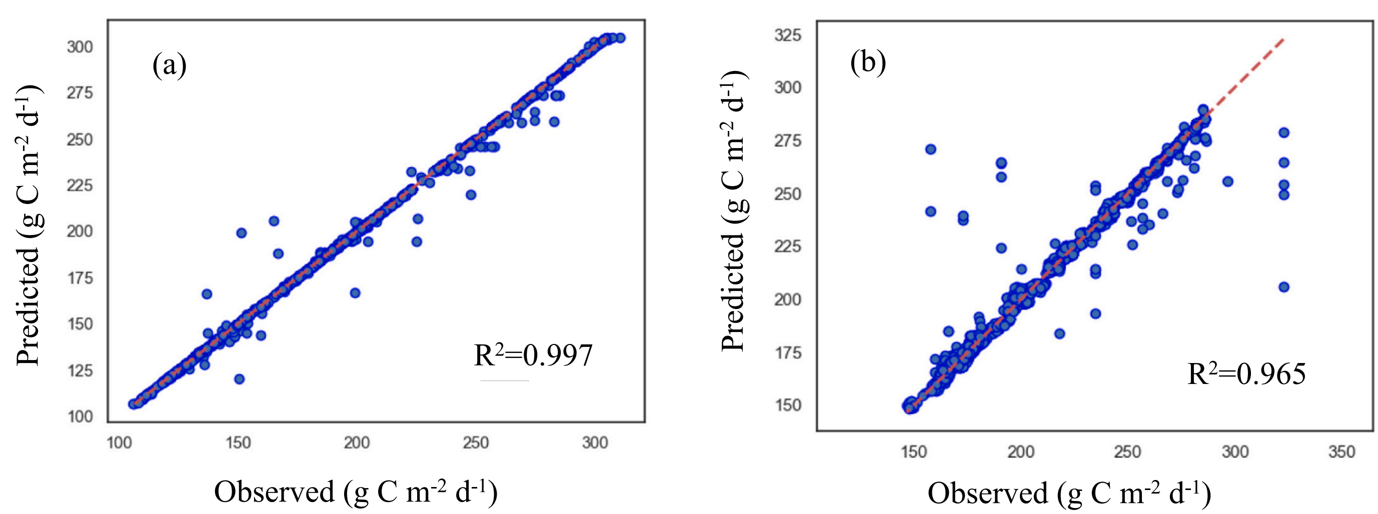

Fig. 4. Comparison of observed and predicted values of (a) litter respiration (b) mineral soil respiration for the oak forest low east site.

- NBeats: Neural Basis Expansion Analysis (NBeats) (Oreshkin et al., 2020) is a simple and efficient deep neural network (DNN) architecture based on multi-layer perceptron (MLP) with backward and forward residual links. It consists of a very deep stack of fully connected layers.
- NHiTS: Neural Hierarchical Interpolation for Time Series (NHiTs) (Challu et al., 2023) is a time series forecasting model that comprises several blocks of MLPs connected in a doubly residual stacking fashion. It performs long-horizon forecasting by applying multi-scale hierarchical interpolation and multi-rate data sampling techniques. The model performs prediction in a memory-efficient way with improved accuracy.
- Prophet: Prophet (Taylor and Letham, 2018) is a time-series forecasting model based on modular regression and Bayesian inference.
- Random Forest: Random Forest (Breiman, 2001) is an ensemble learning-based supervised ML algorithm. It works based on the bagging principle, combining multiple decision trees to make predictions.
- LSTM: Long Short-Term Memory (LSTM) (Gers and Schmidhuber, 2001) is a type of Recurrent Neural Network (RNN) used to learn long-term dependencies addressing vanishing gradient problems in time series data using memory cells and feedback connections. LSTM utilizes three gate structures, namely input, output, and forget gate, to process information.
- TFT: Temporal Fusion Transformer (TFT) (Lim et al., 2020) is a Transformer interpretable time series forecasting model that uses self-attention to model complex temporal time-series patterns. The model comprises novel components such as a static covariate encoder, sequence-to-sequence (LSTM) layer, gating module, and temporal self-attention decoder. The model uses quantile predictions to obtain output intervals across all prediction horizons.
- TCN: Temporal Convolutional Network (TCN) (Bai et al., 2018) TCNs have the advantages of parallelism, low memory requirements for training, and flexible receptive field size. It employs causal convolution and dilation to avoid information leakage from the future to the past.

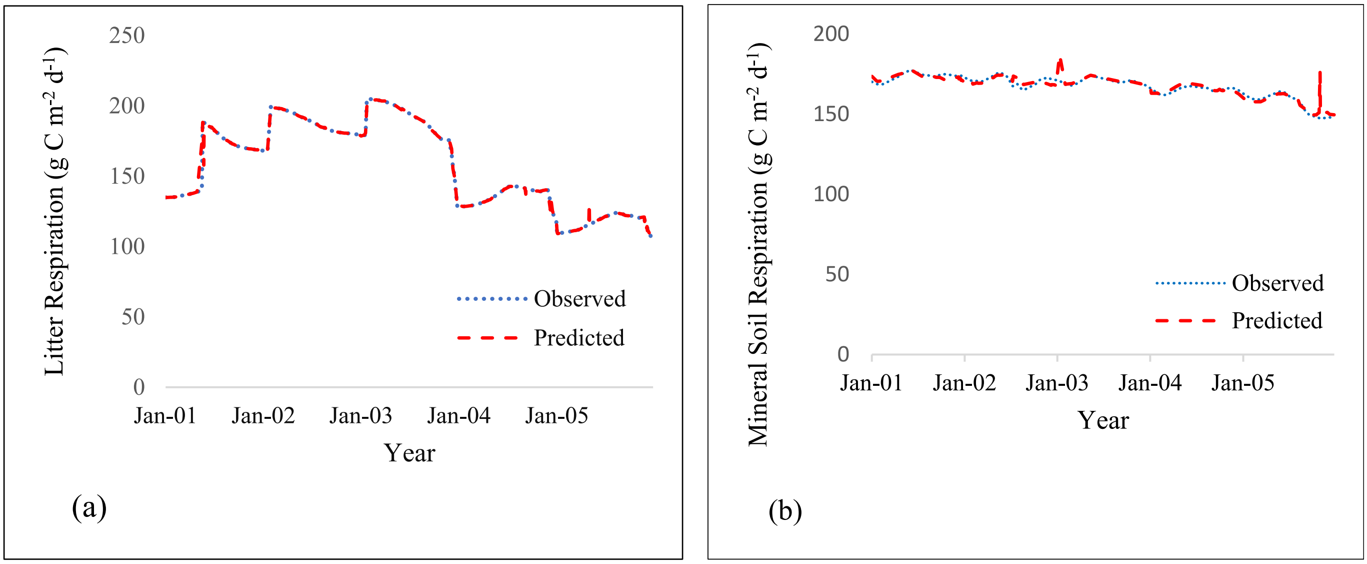

Fig. 5. Comparison of observed and predicted values for (a) litter respiration (b) mineral soil respiration for oak forest low east site from 2001 to 2005.

In this study, the efficacy of various time-series forecasting models was scrutinized to evaluate their capacity for temporal feature extraction, which is crucial for integrating CO_2 flux predictions (Fig. 6). Daily timesteps are considered to perform five-year ahead predictions.

The performance of both traditional and cutting-edge forecasting models, including Prophet, NBeats, NHiTs, TFT, TCN, and LSTM, was assessed with respect to CO_2 flux forecasting for the site under study. The findings unequivocally indicated that the Prophet model outperformed other time-series forecasting models, underscoring its superior capability to encapsulate temporal information effectively. The forecasting outcomes advocated for the Prophet model as the most appropriate model for CO_2 flux forecasting at the studied site, based on considered error metrics (mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE). Furthermore, the Prophet model is characterized by its rapid training process and simplicity, offering a distinct advantage over more complex deep forecasting models.

3.3. Prediction using snapshot ensemble

Finally, the extracted prophet features were concatenated with five features (air temperature, soil temperature, precipitation, predicted mineral soil respiration, and predicted litter respiration) and fed to an ANN to improve predictions. Because the increased number of hidden layers may risk overfitting the model, we trained an ANN with three hidden layers containing 10, 15, and 10 neurons, respectively. For training the hidden layers, 'relu' activation function was used. The model was optimized through an Adaptive Moment Estimation (ADAM) optimizer (Alzubaidi et al., 2021; Zhang, 2018). This hybrid Prophet-ANN model was further refined using a cosine annealing rate schedule algorithm. The model underwent training over 200 epochs, distributed across 20 cycles, with a peak learning rate set at 0.001. Each model is saved upon the completion of a cycle, resulting in a collection of snapshot ensemble models. The final prediction is performed by averaging predictions of snapshot models.

4. Model evaluation metrics

To evaluate the reliability of the proposed model, we estimated the model performance using several metrics or key performance indicators like R^2, RMSE, MAE, and MAPE (Hyndman and Koehler, 2006). R^2 is the goodness-of-measure that indicates fitness to data computed using Eq. 3. RMSE and MAE are scale-dependent error metrics computed using Eq. 4 and Eq. 5, respectively. In RMSE, because errors are squared, relatively

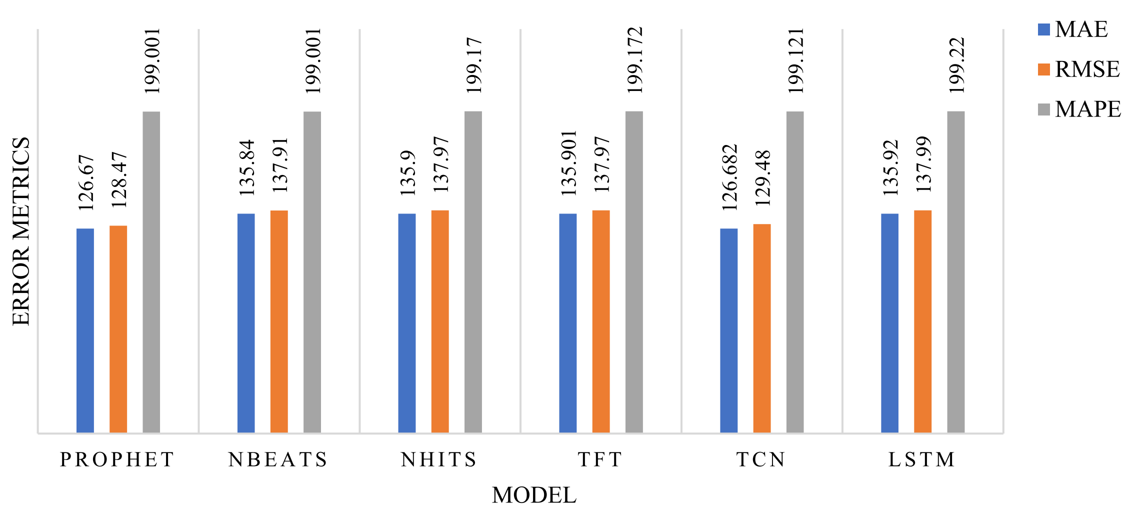

Fig. 6. Comparison of time-series forecasting model performance in terms of error metrics.

large weights are assigned to higher errors compared to MAE. MAPE is the percentage error metric computed using Eq. 6.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (S_a - S_p)^2}{\sum_{i=1}^{n} (S_a - \overline{S_a})^2} \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n} (S_p - S_a)^2} \tag{4}$$

$$MAE = \frac{1}{n}\sum\nolimits_{i=1}^{n} \|S_p - S_a\| \tag{5}$$

$$MAPE = \frac{100}{n}\sum\nolimits_{i=1}^{n} |\frac{S_a - S_p}{S_a}| \tag{6}$$

where n is the number of samples, S_a, S_p, and S_a are the ground truth, predicted values, and actual mean values, respectively.

5. Results and discussion

5.1. Results

To demonstrate the generalization ability of our model, we tested it on four different oak forest sites in the Oak Ridge National Laboratory deciduous forest. The performance of our model is evaluated against eight other ML models to see which one works best. A good model has a high R^2 (which means it can predict outcomes well) and low errors (MAE, RMSE, and MAPE), meaning it makes fewer mistakes. Our method did better than the other ML models we investigated, getting higher R^2 and lower errors for all four places we studied (see Table 3). Of all the models we compared (KNN, MLP, CART, SVR, RF, LR, BR, and GB), the RF model was the top performer. The RF model was very good at predicting what happens at our study sites, predicting correctly about 98.9 % of the time for one site, and similarly high for the others (Table 3). Our new model outperformed the RF model, improving predictions by 0.8 % to 1.3 % for two sites and by about 1 % for the other two sites. This indicates that our model provides more accurate predictions at these sites. Results showing how our model compares to the others are detailed in Table 3. Our study findings yielded estimates of CO_2 flux with high explanatory performance.

The prediction results of the proposed model are illustrated in Fig. 7. A higher R^2 with lower MAE, RMSE and MAPE ensure a better-performing model that can forecast with high certainty and less error. Our proposed model achieved high performance with high R^2 and lower error metrics (MAE, RMSE, and MAPE), suggesting that the proposed model can capture the pattern of data very well and can be used for CO_2 flux prediction. To discern the validity of the proposed model, residuals of modeling the four sites are shown in Fig. 8; the residual values are close to zero and follow a gaussian distribution which ensure the stable prediction of the models.

Employing a hybrid modeling approach and ensemble techniques, the proposed model exhibited high performance. The forecasting inaccuracies of the Prophet model were rectified through integration with the ANN model, and the final prediction was derived by averaging the outcomes of the hybrid Prophet-ANN model, thereby ensuring a robust and diversified forecast. The proposed model counterparts performed better than the others, achieving an R^2 of approximately 0.99 across all four sites under investigation as shown in Fig. 7 (a), (b), (c), and (d), which attests to the model's robustness. Furthermore, the model's performance was notably superior in the East sites compared to the West sites for both high and low regions (Table 3). In comparison with eight other ML models, the proposed model registered the lowest RMSE for all sites (Table 2), with MAPE values of 0.002 and 0.001 for the East low and East high sites, respectively, and 0.007 for both West sites.

When evaluated against other established error metrics (MAE and RMSE), the model demonstrated lower error values; MAE ≈ 0.451 and 0.854 and RMSE ≈ 0.153 and 0.003 for the East low and West low sites, respectively. A similar trend was observed for the East high and West high sites, where the model secured superior predictive performance across R^2, RMSE, MAE, and MAPE metrics. The satisfactory performance of the proposed model across the four sites underscores its robustness and reliability for CO_2 flux prediction.

Table 3
Baseline and proposed model comparison for CO_2 flux forecasting for the experimental oak forest site.

Sites	Model	R^2	MAE	RMSE	MAPE
East low	KNN	0.948	8.001	11.034	0.049
	MLP	0.872	4.918	2.582	0.025
	CART	0.989	3.030	5.004	0.010
	SVR	0.853	5.925	2.662	0.034
	RF	0.989	2.405	3.690	0.010
	LR	0.918	9.832	13.793	0.056
	BR	0.917	9.993	13.982	0.053
	GB	0.983	4.372	6.300	0.020
	Snapshot Prophet-ANN	**0.997**	**0.451**	**0.153**	**0.002**
West low	KNN	0.966	8.023	10.023	0.0425
	MLP	0.864	4.890	2.460	0.037
	CART	0.986	3.041	5.085	0.068
	SVR	0.842	9.922	10.075	0.076
	RF	0.983	2.431	3.930	0.014
	LR	0.916	9.874	13.794	0.045
	BR	0.912	9.948	13.985	0.056
	GB	0.971	4.367	6.126	0.034
	Snapshot Prophet-ANN	**0.996**	**0.854**	**0.003**	**0.001**
East high	KNN	0.546	18.679	42.207	0.092
	MLP	0.908	3.416	5.669	0.043
	CART	0.977	4.405	9.713	0.022
	SVR	0.814	5.541	10.757	0.031
	RF	0.983	3.405	8.029	0.018
	LR	0.278	53.397	28.627	0.139
	BR	0.298	28.737	52.572	0.140
	GB	0.892	8.694	20.283	0.045
	Snapshot Prophet-ANN	**0.993**	**2.722**	**6.210**	**0.007**
West high	KNN	0.532	19.341	43.162	0.094
	MLP	0.901	3.120	5.311	0.129
	CART	0.975	3.342	10.634	0.158
	SVR	0.845	7.671	11.326	0.435
	RF	0.981	3.684	10.025	0.142
	LR	0.274	55.461	28.783	0.279
	BR	0.283	28.987	53.438	0.192
	GB	0.862	8.863	25.231	0.078
	Snapshot Prophet-ANN	**0.991**	**3.171**	**8.124**	**0.007**

5.2. Model hyperparameter selection

To gain a more in-depth knowledge of the proposed deep hybrid model, we analyzed the performance of the proposed model, varying different parameters associated with model training. The impact of changing parameters is shown in Fig. 9. We changed the maximum learning rates to 0.1, 0.01, and 0.001 and found that maximum learning of 0.001 helped the model to achieve stable and robust optimization. Also, we tested the model by changing the Prophet interval width. The interval width of 0.95 gave the best-performing model. We also investigated the effect of varying number of cycles on model performance. We found that setting the cycle number to 20 gave the best performance. From the analysis performed, we can conclude that the performance of the model is significantly dependent on appropriate hyperparameter values, which require careful selection of hyperparameter.

5.3. Discussion

Previous studies have used machine learning to predict agricultural

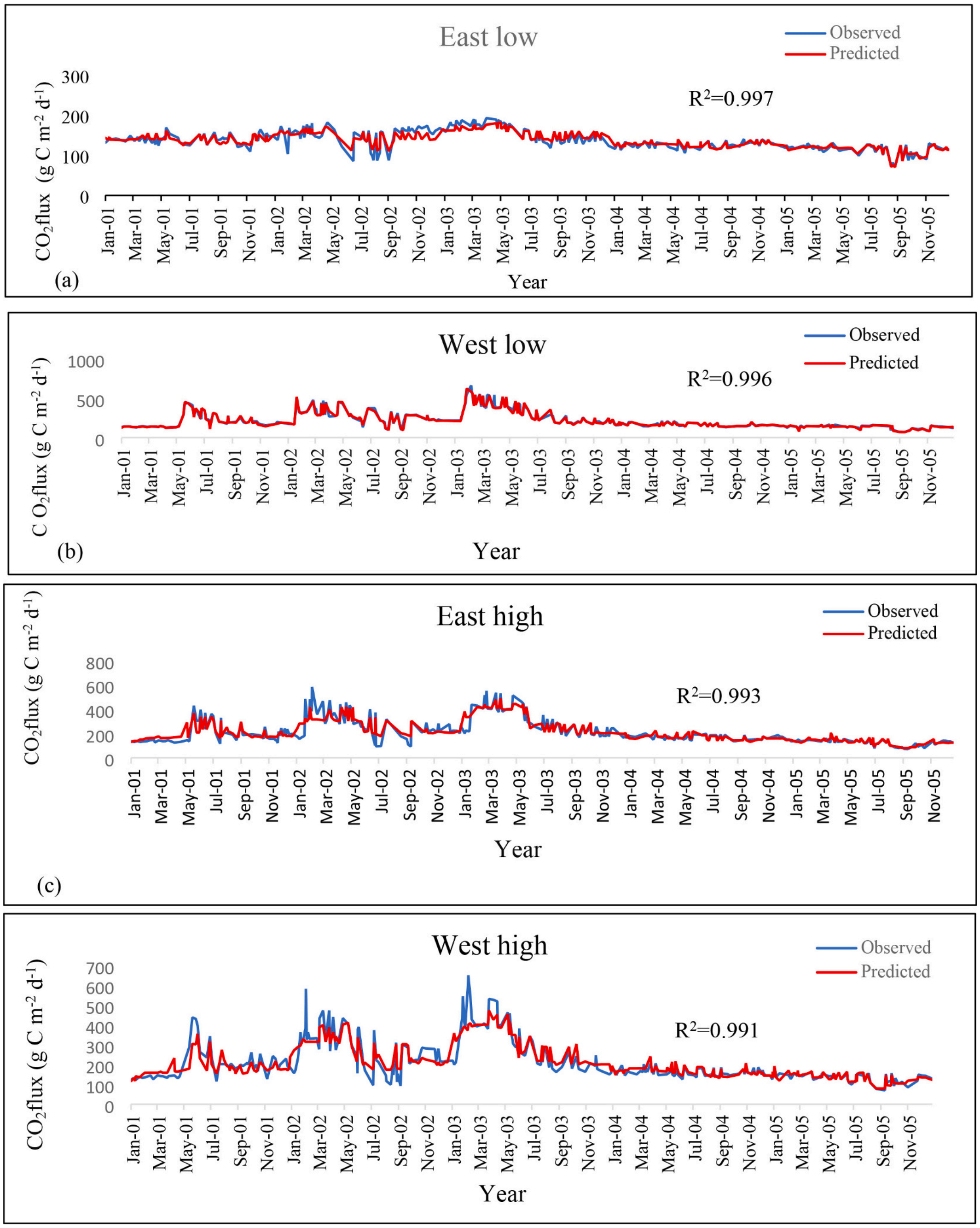

Fig. 7. Comparison of observed and predicted values for soil respiration for four scenarios of oak forest sites (a) East low (b) West low (c) East high (d) West high.

soil CO_2 flux prediction. (Adjuik and Davis, 2022) analyzed the performance of four ML models for daily CO_2–C emissions. This study reported Gradient Boosting as the best-performing model, capturing around 82 % variation of the test set. (Ebrahimi et al., 2019) reported ANN as the best model for two states of soil respiration, basal respiration with $R^2 = 0.66$ and substrate induced respiration with $R^2 = 0.52$ respectively. ANN was also found to be effective for modeling soil CO_2 emissions with $R^2 = 0.80$. (Hamrani et al., 2020) compared the performance of nine machine learning models and elicited LSTM suitable for successfully simulating soil CO_2 flux, capturing around 87 % variation of soil CO_2 flux. (Joshi et al., 2022) demonstrated that the RF model can be utilized to predict daily CO_2–C emissions with 85 % explained variation. However, the current employed models neglect the correlation between soil carbon pool and soil respiration. In this study, we integrated soil carbon pool information by considering the intermediate variables, litter respiration and mineral soil respiration while developing the snapshot ensemble model. Moreover, the performance of our proposed model was comparatively higher than the best-performing model results described in the literature.

This study integrates temporal features extracted by Prophet along with environmental variables and leverages predicted mineral and litter soil respiration from an ML model linking SOM pool with soil respiration. This approach enhances the model's performance significantly. Our proposed model captured around 99 % of the variation for the daily CO_2-C emission forecast, which was higher than the previous works.

The outcomes of our analysis hold significant potential for informed decision-making processes in land management, agriculture (e.g., forestry), and strategies aimed at mitigating climate change. The model's precise predictions of SR can guide farmers and forestland managers in making informed decisions about planting, fertilization, and crop

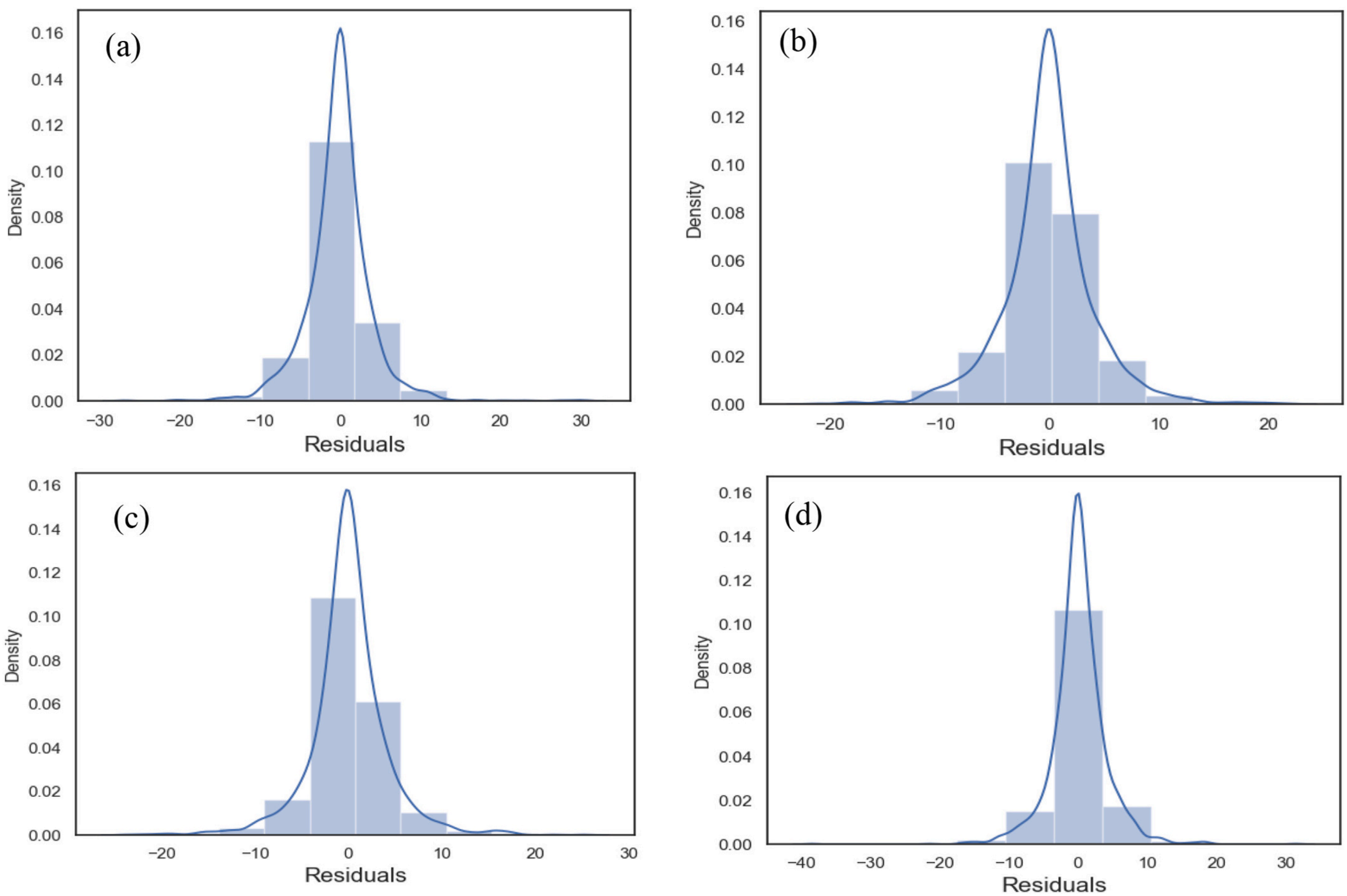

Fig. 8. Model residuals for soil respiration prediction for four scenarios of oak forest sites (a) East low (b) East high (c) West low (d) West high.

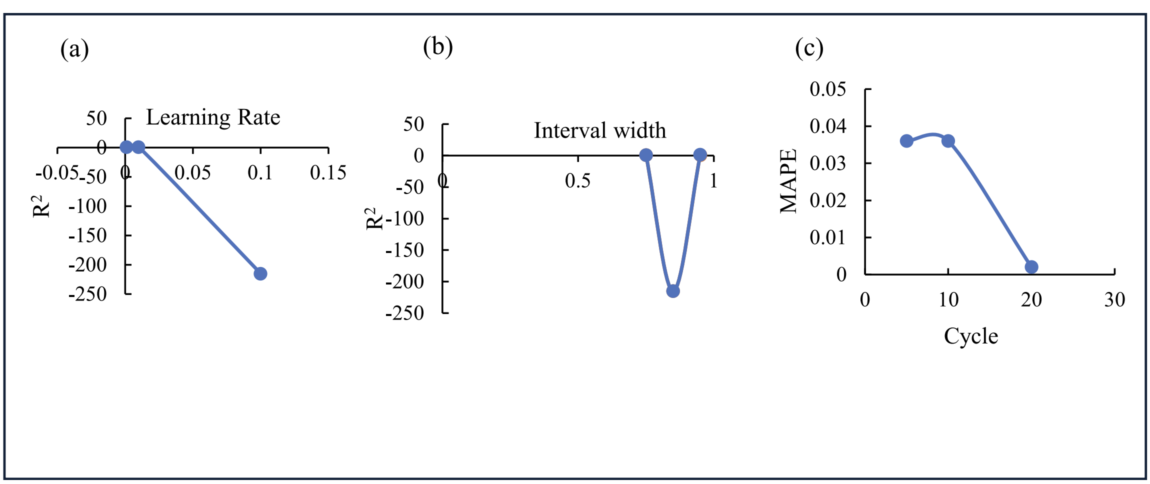

Fig. 9. Impacts of parameter tuning for the proposed snapshot ensemble model: a) learning rate, b) interval width, and c) number of cycles.

rotation practices, thereby enhancing forest sector productivity. The model's accurate estimation of CO_2 can support carbon management policies and promote effective land-use practices, greatly enhancing the practical relevance of this study. Integrating the model insights into policy frameworks allows for more precise carbon budgeting, enabling policymakers to design effective strategies for reducing atmospheric CO_2. By providing real-time insights into soil health and carbon cycling, our model supports sustainable land use planning and contributes to climate change mitigation efforts. The proposed model not only advances scientific understanding but also provides actionable insights for managing carbon footprints and improving environmental resilience.

Machine learning models, due to their reduced computational time compared to biophysical models, offer a viable alternative for predicting soil respiration (Nguyen et al., 2019). In this research, we propose an innovative machine learning model for SR prediction, enabling the examination of temporal information and environmental factors, thus automating the SR system's workflow. Simulating CO_2 flux using the DayCent model on high-performance computing (HPC) infrastructure is computationally intensive, requiring roughly 5 h per simulation. For a region with about 1000 to 5000 grid cells (one grid cell cover minimum

1 km2 region), usage of 64 to 256 cores is generally recommended. The 5 h per run is consider based on a regional-scale study area cover few thousand square kilometers, though the time may vary depending on the length of the simulation period, ecosystem complexity, resolution and extent of the input data, as well as the specific HPC hardware and configuration used. At an HPC usage cost of approximately $0.50 per hour, conducting 1000 simulations for a regional-scale study incurs a total cost of $2500. In contrast, machine learning (ML) models offer a much more efficient alternative. Training an ML model for the same region takes about 5 min on a cloud GPU instance, priced at $2.00 per hour, leading to a minimal cost of $0.16. After training, generating predictions with the ML model takes mere minutes or seconds, making it highly scalable. The prediction phase is computationally inexpensive due to evaluation of different scenarios and optimization. The potential benefits of ML models are important in reducing computational time, energy requirement and costs for large-scale environmental studies.

Analysis of model predictions and the importance of features allows researchers to identify the most impactful factors influencing SR rates. The influence of various input features on CO_2 flux prediction can be discerned through the application of SHapley Additive exPlanation (SHAP) values (Winter, 2002). SHAP values, which serve to interpret the model, highlight the importance of individual input features in evaluating the performance of the forecasting model. The impact of different features on SR, as illustrated in Fig. 10, underscores temporal features (yhat, predicted CO_2 flux forecasting from Prophet time-series data, extracted trend series from Prophet) as crucial for CO_2 flux prediction. SHAP analysis verifies the significant role of temporal features (e.g., yhat_lower, trend_lower) derived from Prophet in CO_2 flux prediction. Additionally, the output from the intermediate variable prediction stage has proven effective, aligning with previous findings on CO_2 flux prediction (Lynch et al., 2006; Parton et al., 2010; Savage et al., 2013). This study highlights a positive correlation between predicted mineral soil respiration and litter respiration, with higher values for both variables enhancing the accuracy of predicted CO_2 flux. Among meteorological variables, soil temperature is identified as a pivotal factor for CO_2 flux prediction, consistent with earlier research in the field (Adjuik and Davis, 2022; Joshi et al., 2022; Mirzaei et al., 2023).

Here, temporal features are predicted CO_2 flux forecasting from Prophet (yhat), uncertainty interval (yhat_lower, yhat_upper), trend, lower and upper range minimum of trend series (trend_lower, trend_upper), combination of seasonality, holiday effect and residuals (additive_terms), lower and upper range minimum of the decomposed additive series (additive_terms_lower, additive_terms_upper) respectively.

We also analyzed our proposed model's sensitivity to different input variables using a partial dependence plot (PDP) (Friedman, 2001). In this work, we used several environmental variables, including air temperature, soil temperature, and precipitation, along with predicted litter and mineral soil respiration, as inputs for CO_2 flux predictions.

The PDP plot in Fig. 11 illustrates the effect of individual predictor variables on the soil CO_2 flux prediction (partial dependence). In agreement with the findings from the meta-analysis performed by Zhang et al., Fig. 11 demonstrates that CO_2 flux prediction increases with the increase in soil temperature and precipitation (Zhang et al., 2023).

To discuss the limitations of this study, we examined model performance specifically for experimental oak forests. It's worth mentioning that the sensitivity of soil respiration varies across ecosystem types (Han et al., 2014). Thus, investigating the model performance in a variety of ecosystems beyond the experimental oak forest could provide a more comprehensive understanding of the model's effectiveness and generalization. Enhancing the model's robustness by accounting for ecosystem diversity can further increase its usefulness, making it more adaptable to a wide range of ecological contexts. Therefore, in future research, we plan to investigate the proposed model performance for diverse ecosystems. Furthermore, variations in soil types and climatic conditions may lead to significant differences in model prediction outcomes, highlighting the need for further research in this area.

The proposed ensemble-based model achieves high performance in CO_2 flux prediction for the studied sites. The trained model delivers accurate predictions within sub-second execution times (less than 0.1 s). This study demonstrates its efficacy as a surrogate for the ForCent model, also the framework's generalizability paves the way for its application as a surrogate for other ecosystem models (e.g., RothC, CEVSA, DNDC) to facilitate efficient CO_2 flux prediction.

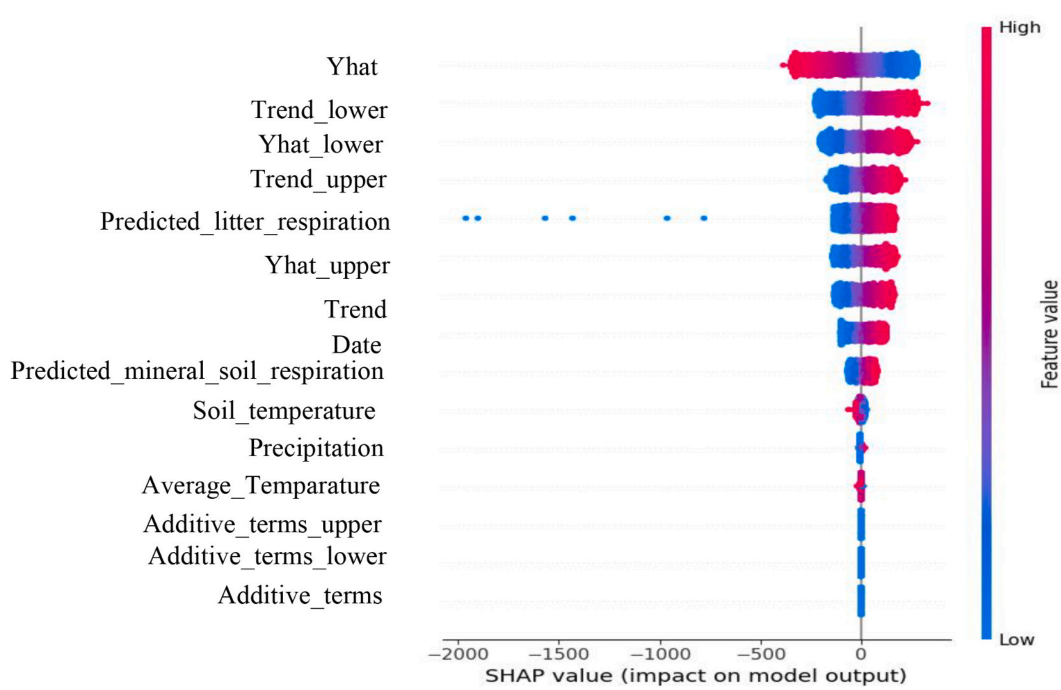

Fig. 10. Feature importance according to the developed model for CO_2 flux prediction.

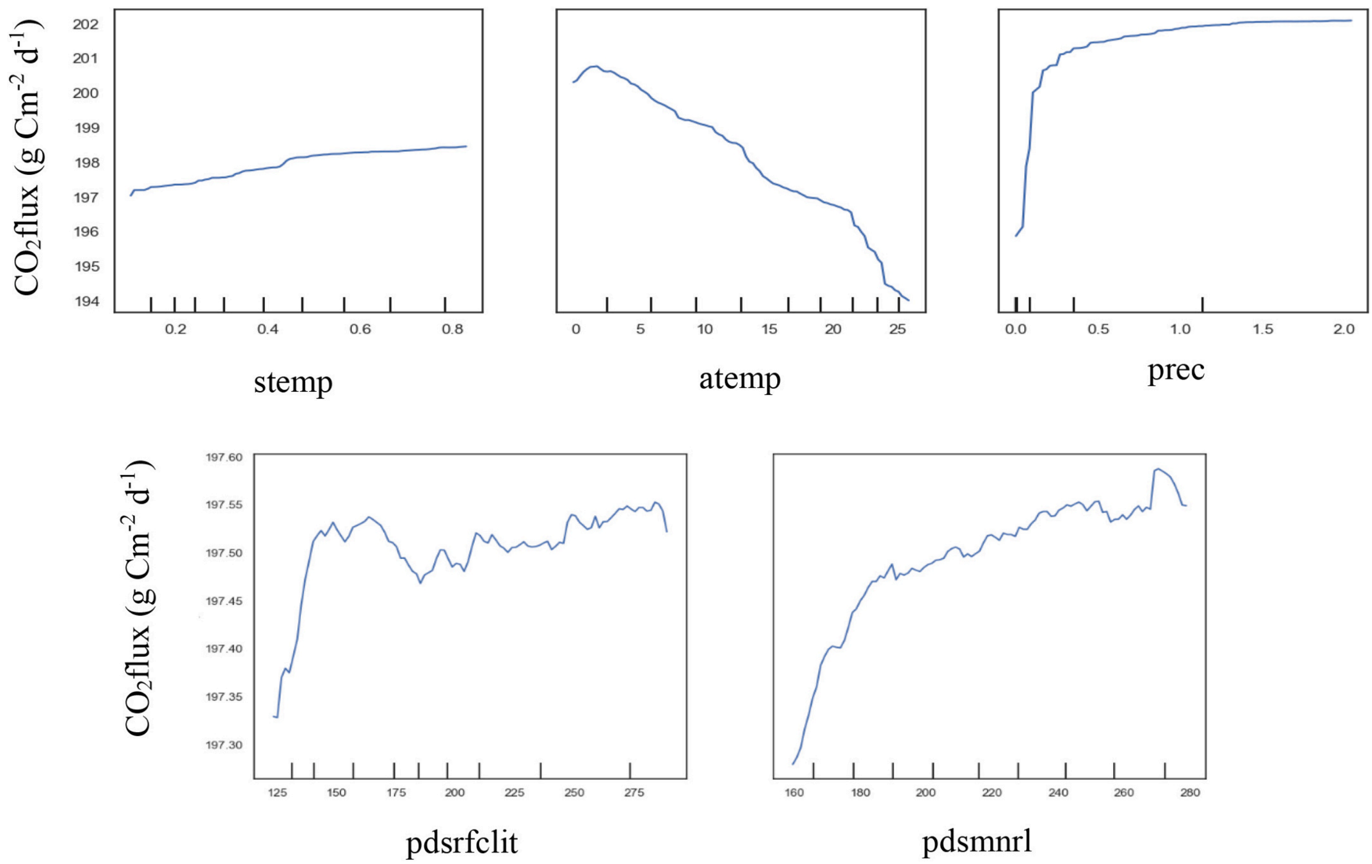

Fig. 11. Partial dependence plot for CO_2 flux prediction in response to the input variables.

6. Conclusion

This study demonstrated the use of a snapshot ensemble modeling technique that can be used as a surrogate for the ForCent model for CO_2 flux prediction. For this study, we used the database from an experimental oak forest for a ^{13}C study to enrich the background isotope. The comparative analysis of other machine-learning models in terms of statistical metrics shows that Random Forest had the best performance in training and predicting soil respiration but less than our proposed snapshot ensemble model. The proposed hybrid snapshot model demonstrates impressive performance, effectively capturing the high variance within the studied data. Results showed the potential use of our ML model to predict trends of surface soil respiration, mineral soil respiration, and total soil respiration. This approach could reduce uncertainties in soil CO_2 flux, aiding in the identification and monitoring of greenhouse gas sources and sinks within ecosystems. In addition, this method may be an effective, low-cost alternative for scale-up of measuring and verifying soil respiration; most of the model attributes are commonly determined in routine soil analysis.

CRediT authorship contribution statement

S.N. Ferdous: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **J.P. Ahire:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **R. Bergman:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition. **L. Xin:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Data curation. **E. Blanc-Betes:** Writing – review & editing, Formal analysis. **Z. Zhang:** Writing – review & editing, Formal analysis. **J. Wang:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Formal analysis.

Declaration of competing interest

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

Data availability

I have shared the link to my data and code in the given github link. OakForest Dataset (Original data)

Acknowledgment

This work is supported by Agriculture and Food Research Initiative Competitive Grant no. 2020-68012-31881 from the USDA National Institute of Food and Agriculture. As some authors of this paper are employees of the USDA FS, the findings and conclusions in this report are those of the author(s) and should not be construed to represent any official USDA or U.S. Government determination or policy. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. government.

Reference

Abbasi, N.A., Hamrani, A., Madramootoo, C.A., Zhang, T., Tan, C.S., Goyal, M.K., 2021. Modelling carbon dioxide emissions under a maize-soy rotation using machine learning. Biosyst. Eng. 212, 1–18.

Adjuik, T.A., Davis, S.C., 2022. Machine learning approach to simulate soil CO2 fluxes under cropping systems. Agronomy 12, 197. https://doi.org/10.3390/agronomy12010197.

Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L., 2021. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J. Big Data 8, 1–74.

Andresen, L.C., Domínguez, M.T., Reinsch, S., Smith, A.R., Schmidt, I.K., Ambus, P., Beier, C., Boeckx, P., Bol, R., de Dato, G., Emmett, B.A., Estiarte, M., Garnett, M.H., Kröel-Dulay, G., Mason, S.L., Nielsen, C.S., Peñuelas, J., Tietema, A., 2018. Isotopic methods for non-destructive assessment of carbon dynamics in shrublands under long-term climate change manipulation. Methods Ecol. Evol. 9, 866–880. https://doi.org/10.1111/2041-210X.12963.

Arbib, M., 1998. The Handbook of Brain Theory and Neural Networks [WWW Document]. MIT Press. URL https://mitpress.mit.edu/9780262511025/the-handbook-of-brain-theory-and-neural-networks/ (accessed 11.8.23).

Bai, S., Kolter, J.Z., Koltun, V., 2018. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling.

Berardi, D., Brzostek, E., Blanc-Betes, E., Davison, B., DeLucia, E.H., Hartman, M.D., Kent, J., Parton, W.J., Saha, D., Hudiburg, T.W., 2020. 21st-century biogeochemical modeling: Challenges for Century-based models and where do we go from here? GCB Bioenergy 12, 774–788. https://doi.org/10.1111/gcbb.12730.

Bishop, C.M., Tipping, M.E., 2003. Bayesian regression and classification. In: Advances in Learning Theory: Methods, Models and Applications. IOS Press, pp. 267–285.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Challu, C., Olivares, K.G., Oreshkin, B.N., Ramirez, F.G., Canseco, M.M., Dubrawski, A., 2023. NHITS: Neural hierarchical interpolation for time series forecasting. Proc. AAAI Conf. Artif. Intell. 37, 6989–6997. https://doi.org/10.1609/aaai.v37i6.25854.

dos Reis Martins, M., Necpalova, M., Ammann, C., Buchmann, N., Calanca, P., Flechard, C.R., Hartman, M.D., Krauss, M., Le Roy, P., Mäder, P., Maier, R., Morvan, T., Nicolardot, B., Skinner, C., Six, J., Keel, S.G., 2022. Modeling N2O emissions of complex cropland management in Western Europe using DayCent: Performance and scope forimprovement. Eur. J. Agron. 141, 126613. https://doi.org/10.1016/j.eja.2022.126613.

Ebrahimi, M., Sarikhani, M.R., Sinegani, A.A.S., Ahmadi, A., Keesstra, S., 2019. Estimating the soil respiration under different land uses using artificial neural network and linear regression models. Catena (Amst) 174, 371–382.

Ferdous, S.N., Li, X., Sahoo, K., Bergman, R., 2023. Toward sustainable crop residue management: a deep ensemble learning approach. Bioresourc. Technol. Rep. 22, 101421. https://doi.org/10.1016/j.biteb.2023.101421.

Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Stat. 1189–1232.

Gers, F.A., Schmidhuber, E., 2001. LSTM recurrent networks learn simple context-free and context-sensitive languages. IEEE Trans. Neural Netw. 12, 1333–1340. https://doi.org/10.1109/72.963769.

Gey, S., Nedelec, E., 2005. Model selection for CART regression trees. IEEE Trans. Inf. Theory 51, 658–670. https://doi.org/10.1109/TIT.2004.840903.

Hamrani, A., Akbarzadeh, A., Madramootoo, C.A., 2020. Machine learning for predicting greenhouse gas emissions from agricultural soils. Sci. Total Environ. 741, 140338. https://doi.org/10.1016/j.scitotenv.2020.140338.

Han, G., Xing, Q., Luo, Y., Rafique, R., Yu, J., Mikle, N., 2014. Vegetation types alter soil respiration and its temperature sensitivity at the field scale in an estuary wetland. PLoS One 9, e91182.

Harsányi, E., Mirzaei, M., Arshad, S., Alsilibe, F., Vad, A., Nagy, A., Ratonyi, T., Gorji, M., Al Dalahme, M., Mohammed, S., 2024. Assessment of advanced machine and deep learning approaches for predicting CO2 emissions from agricultural lands: insights across diverse agroclimatic zones. Earth Syst. Environ. 1–17.

Hashimoto, S., Ito, A., Nishina, K., 2023. Divergent data-driven estimates of global soil respiration. Commun. Earth Environ. 4, 1–8. https://doi.org/10.1038/s43247-023-01136-2.

Heimann, M., Reichstein, M., 2008. Terrestrial ecosystem carbon dynamics and climate feedbacks. Nature 451, 289–292. https://doi.org/10.1038/nature06591.

Huang, G., Li, Y., Pleiss, G., Liu, Z., Hopcroft, J.E., Weinberger, K.Q., 2017. Snapshot Ensembles: Train 1, get M for free.

Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. Int. J. Forecast. 22, 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001.

Jiang, Z., Yang, S., Smith, P., Pang, Q., 2023. Ensemble machine learning for modeling greenhouse gas emissions at different time scales from irrigated paddy fields. Field Crop Res. 292, 108821. https://doi.org/10.1016/j.fcr.2023.108821.

Joshi, D.R., Clay, D.E., Clay, S.A., Moriles-Miller, J., Daigh, A.L.M., Reicks, G., Westhoff, S., 2022. Quantification and machine learning based N2O–N and CO2–C emissions predictions from a decomposing rye cover crop. Agron. J. https://doi.org/10.1002/agj2.21185.

Kurzbesson, C., Couteaux, M., Thiery, J., Berg, B., Remacle, J., 2005. A comparison of litterbag and direct observation methods of Scots pine needle decomposition measurement. Soil Biol. Biochem. 37, 2315–2318. https://doi.org/10.1016/j.soilbio.2005.03.022.

Lardy, R., Bachelet, B., Bellocchi, G., Hill, D.R.C., 2014. Towards vulnerability minimization of grassland soil organic matter using metamodels. Environ. Model Softw. 5 (2), 38–50. https://doi.org/10.1016/j.envsoft.2013.10.015.

Lim, B., Arik, S.O., Loeff, N., Pfister, T., 2020. Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting. https://doi.org/10.48550/arXiv.1912.09363.

Liu, Y., Lin, J., Yue, H., 2023. Soil respiration estimation in desertified mining areas based on UAV remote sensing and machine learning. Earth Sci. Inf. 16, 3433–3448.

Liu, J., Hu, J., Liu, H., Han, K., 2024. Global soil respiration estimation based on ecological big data and machine learning model. Sci. Rep. 14, 13231.

Lloyd, J., Taylor, J.A., 1994. On the temperature dependence of soil respiration. Funct. Ecol. 8, 315–323. https://doi.org/10.2307/2389824.

Loshchilov, I., Hutter, F., 2017. SGDR: Stochastic Gradient Descent with Warm Restarts. https://doi.org/10.48550/arXiv.1608.0398.

Lu, H., Li, S., Ma, M., Bastrikov, V., Chen, X., Ciais, P., Dai, Y., Ito, A., Ju, W., Lienert, S., 2021. Comparing machine learning-derived global estimates of soil respiration and its components with those from terrestrial ecosystem models. Environ. Res. Lett. 16, 054048.

Lynch, D.H., Voroney, R.P., Warman, P.R., 2006. Use of 13C and 15N natural abundance techniques to characterize carbon and nitrogen dynamics in composting and in compost-amended soils. Soil Biol. Biochem. 38, 103–114. https://doi.org/10.1016/j.soilbio.2005.04.022.

Maltamo, M., Kangas, A., 1998. Methods based on k-nearest neighbor regression in the prediction of basal area diameter distribution. Can. J. For. Res. 28, 1107–1115. https://doi.org/10.1139/x98-085.

Mayer, M., Prescott, C.E., Abaker, W.E.A., Augusto, L., Cécillon, L., Ferreira, G.W.D., James, J., Jandl, R., Katzensteiner, K., Laclau, J.-P., Laganière, J., Nouvellon, Y., Paré, D., Stanturf, J.A., Vanguelova, E.I., Vesterdal, L., 2020. Tamm review: influence of forest management activities on soil organic carbon stocks: a knowledge synthesis. For. Ecol. Manag. 466, 118127. https://doi.org/10.1016/j.foreco.2020.118127.

McClelland, S.C., Paustian, K., Williams, S., Schipanski, M.E., 2021. Modeling cover crop biomass production and related emissions to improve farm-scale decision-support tools. Agric. Syst. 191, 103151. https://doi.org/10.1016/j.agsy.2021.103151.

Mendes-Moreira, J., Soares, C., Jorge, A.M., de Sousa, J.F., 2012. Ensemble approaches for regression: a survey. ACM Comput. Surv. https://doi.org/10.1145/2379776.2379786.

Mirzaei, M., Anari, M.G., Diaz-Pines, E., Saronjic, N., Mohammed, S., Szabo, S., Mousavi, S.M.N., Caballero-Calvo, A., 2023. Assessment of soil CO2 and NO fluxes in a semi-arid region using machine learning approaches. J. Arid Environ. 211, 104947.

Mohamadi, S., Nasrabadi, N.M., Doretto, G., Adjeroh, D.A., 2021. Human age estimation from gene expression data using artificial neural networks. In: 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 3492–3497. https://doi.org/10.1109/BIBM52615.2021.9669893.

Mohamadi, S., Doretto, G., Adjeroh, D., 2022. Deep active ensemble sampling for image classification. In: Proceedings of the Asian Conference on Computer Vision, p. 45314547.

Natekin, A., Knoll, A., 2013. Gradient boosting machines, a tutorial. Front. Neurorobot. 7.

Nguyen, T.H., Nong, D., Paustian, K., 2019. Surrogate-based multi-objective optimization of management options for agricultural landscapes using artificial neural networks. Ecol. Model. 400, 1–13. https://doi.org/10.1016/j.ecolmodel.2019.02.018.

Nissan, A., Alcolombri, U., Peleg, N., Galili, N., Jimenez-Martinez, J., Molnar, P., Holzner, M., 2023. Global warming accelerates soil heterotrophic respiration. Nat. Commun. 14, 3452. https://doi.org/10.1038/s41467-023-38981-w.

Noble, W.S., 2006. What is a support vector machine? Nat. Biotechnol. 24, 1565–1567. https://doi.org/10.1038/nbt1206-1565.

Nunes, L.J.R., Meireles, C.I.R., Pinto Gomes, C.J., Almeida Ribeiro, N.M.C., 2020. Forest contribution to climate change mitigation: management oriented to carbon capture and storage. Climate 8, 21. https://doi.org/10.3390/cli8020021.

Oertel, C., Matschullat, J., Zurba, K., Zimmermann, F., Erasmi, S., 2016. Greenhouse gas emissions from soils—A review. Geochemistry 76, 327–352.

Oreshkin, B.N., Carpov, D., Chapados, N., Bengio, Y., 2020. N-BEATS: Neural Basis Expansion Analysis for Interpretable Time Series Forecasting. https://doi.org/10.48550/arXiv.1905.10437.

Parton, W.J., Hanson, P.J., Swanston, C., Torn, M., Trumbore, S.E., Riley, W., Kelly, R., 2010. ForCent model development and testing using the Enriched Background Isotope Study experiment. J. Geophys. Res. Biogeosci. 115. https://doi.org/10.1029/2009JG001193.

Philibert, A., Loyce, C., Makowski, D., 2013. Prediction of N2O emission from local information with Random Forest. Environ. Pollut. 177, 156–163.

Pumpanen, J., Kolari, P., Ilvesniemi, H., Minkkinen, K., Vesala, T., Niinistö, S., Lohila, A., Larmola, T., Morero, M., Pihlatie, M., Janssens, I., Yuste, J.C., Grünzweig, J.M., Reth, S., Subke, J.-A., Savage, K., Kutsch, W., Østreng, G., Ziegler, W., Anthoni, P., Lindroth, A., Hari, P., 2004. Comparison of different chamber techniques for measuring soil CO2 efflux. Agric. For. Meteorol. 123, 159–176. https://doi.org/10.1016/j.agrformet.2003.12.001.

Ramchoun, H., Ghanou, Y., Ettaouil, M., Idrissi, M.A.J., 2016. Multilayer perceptron: architecture optimization and training. Int. J. Interact. Multimedia Artif. Intell. 4, 26–30.

Rodtassana, C., Unawong, W., Yaemphum, S., Chanthorn, W., Chawchai, S., Nathalang, A., Brockelman, W.Y., Tor-ngern, P., 2021. Different responses of soil respiration to environmental factors across forest stages in a Southeast Asian forest. Ecol. Evol. 11, 15430–15443. https://doi.org/10.1002/ece3.8248.

Sagi, O., Rokach, L., 2018. Ensemble learning: A survey - Sagi - 2018 - WIREs Data Mining and Knowledge Discovery - Wiley Online Library [WWW Document]. URL https://wires.onlinelibrary.wiley.com/doi/full/10.1002/widm.1249 (accessed 11.8.23).

Saha, D., Basso, B., Robertson, G.P., 2021. Machine learning improves predictions of agricultural nitrous oxide (N2O) emissions from intensively managed cropping systems. Environ. Res. Lett. 16, 024004.

Savage, K.E., P.W.J, D.E.A, T.S.E, F.S.D, 2013. Long-term changes in forest carbon under temperature and nitrogen amendments in a temperate northern hardwood forest. Glob. Chang. Biol. 19, 2389–2400.

Schlesinger, W.H., Andrews, J.A., 2000. Soil respiration and the global carbon cycle. Biogeochemistry 48, 7–20. https://doi.org/10.1023/A:1006247623877.
Seber, G.A.F., Lee, A.J., 2003. Linear regression analysis. In: Wiley Series in Probability and Statistics, 1st ed. Wiley. https://doi.org/10.1002/9780471722199.
Segal, M., 2003. Machine Learning Benchmarks and Random Forest Regression.
Seppelt, R., Voinov, A., 2002. Optimization methodology for land use patterns using spatially explicit landscape models. Ecol. Model. 151, 125–142. https://doi.org/10.1016/S0304-3800(01)00455-0.
Singh, T., Arpanaei, A., Elustondo, D., Wang, Y., Stocchero, A., West, T.A.P., Fu, Q., 2022. Emerging technologies for the development of wood products towards extended carbon storage and CO2 capture. Carbon Capture Sci. Technol. 4, 100057.
Tavares, R.L.M., de Oliveira, S.R.M., de Barros, F.M.M., Farhate, C.V.V., de Souza, Z.M., La Scala, N., 2018. Prediction of soil CO2 flux in sugarcane management systems using the Random Forest approach. Sci. Agric. 75, 281–287.
Taylor, S.J., Letham, B., 2018. Forecasting at scale. Am. Stat. 72, 37–45. https://doi.org/10.1080/00031305.2017.138008.
Wang, S.-C., 2003. Artificial neural network. In: Wang, S.-C. (Ed.), Interdisciplinary Computing in Java Programming, The Springer International Series in Engineering and Computer Science. Springer US, Boston, MA, pp. 81–100. https://doi.org/10.1007/978-1-4615-0377-4_5.
Wang, G., Su, H., Mo, L., Yi, X., Wu, P., 2024. Forecasting of soil respiration time series via clustered ARIMA. Comput. Electron. Agric. 225, 109315. https://doi.org/10.1016/j.compag.2024.109315.
Winter, E., 2002. Chapter 53 The shapley value. In: Handbook of Game Theory with Economic Applications. Elsevier, pp. 2025–2054. https://doi.org/10.1016/S1574-0005(02)03016-3.
Xie, Y., 2020. A meta-analysis of critique of litterbag method used in examining decomposition of leaf litters. J. Soils Sediments 20, 1881–1886. https://doi.org/10.1007/s11368-020-02572-9.
Xie, Yajun, Xie, Yonghong, Xiao, H., Chen, X., Li, F., 2017. Controls on litter decomposition of emergent macrophyte in Dongting Lake Wetlands. Ecosystems 20, 1383–1389. https://doi.org/10.1007/s10021-017-0119-y.
Yang, F., Jia, L., Chen, L., Gao, L., Zang, Y., Zhang, J., Leng, H., 2024. DEAF: An adaptive feature aggregation model for predicting soil CO2 flux. Ecol. Inform. 82, 102759. https://doi.org/10.1016/j.ecoinf.2024.102759.
Zani, C.F., Abdalla, M., Abbott, G.D., Taylor, J.A., Galdos, M.V., Cooper, J.M., Lopez-Capel, E., 2023. Predicting long-term effects of alternative management practices in conventional and organic agricultural systems on soil carbon stocks using the daycent model. Agronomy 13, 1093. https://doi.org/10.3390/agronomy13041093.
Zhang, Z., 2018. Improved adam optimizer for deep neural networks. In: 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). Ieee, pp. 1–2.
Zhang, Z., Li, Y., Williams, R.A., Chen, Y., Peng, R., Liu, X., Qi, Y., Wang, Z., 2023. Responses of soil respiration and its sensitivities to temperature and precipitation: a meta-analysis. Ecol. Inform. 75, 102057. https://doi.org/10.1016/j.ecoinf.2023.102057.
Zhao, Z., Peng, C., Yang, Q., Meng, F., Song, X., Chen, S., Epule, T.E., Li, P., Zhu, Q., 2017. Model prediction of biome-specific global soil respiration from 1960 to 2012. Earths Future 5, 715–729.