Journal Pre-proof

Machine learning for predicting greenhouse gas emissions from agricultural soils

Abderrachid Hamrani, Abdolhamid Akbarzadeh, Chandra A. Madramootoo

PII: S0048-9697(20)33860-2

DOI: https://doi.org/10.1016/j.scitotenv.2020.140338

Reference: STOTEN 140338

To appear in: *Science of the Total Environment*

Received date: 9 March 2020

Revised date: 5 June 2020

Accepted date: 16 June 2020

Please cite this article as: A. Hamrani, A. Akbarzadeh and C.A. Madramootoo, Machine learning for predicting greenhouse gas emissions from agricultural soils, *Science of the Total Environment* (2020), https://doi.org/10.1016/j.scitotenv.2020.140338

Machine Learning for Predicting Greenhouse Gas Emissions from Agricultural Soils

Abderrachid Hamrani [a], *Abdolhamid Akbarzadeh* [a, *], *Chandra A. Madramootoo* [a, †]

[a] Department of Bioresource Engineering, McGill University, Montreal, QC H9X3V9, Canada.

* Corresponding author: email address: hamid.akbarzadeh@mcgill.ca.

† Corresponding author: email address: chandra.madramootoo@mcgill.ca.

ABSTRACT: Machine learning (ML) models are increasingly used to study complex environmental phenomena with high variability in time and space. In this study, the potential of exploiting three categories of ML regression models, including classical regression, shallow learning and deep learning for predicting soil greenhouse gas (GHG) emissions from an agricultural field was explored. Carbon dioxide (CO_2) and nitrous oxide (N_2O) fluxes, as well as various environmental, agronomic and soil data were measured at the site over a five-year period in Quebec, Canada. The rigorous analysis, which included statistical comparison and cross-validation for the prediction of CO_2 and N_2O fluxes, confirmed that the LSTM model performed the best among the considered ML models with the highest R coefficient and the lowest root mean squared error (RMSE) values (R = 0.87 and RMSE = 30.3 $mg.m^{-2}.hr^{-1}$ for CO_2 flux prediction and R = 0.86 and RMSE = 0.19 $mg.m^{-2}.hr^{-1}$ for N_2O flux prediction). The predictive performances of LSTM were more accurate than those simulated in a previous study conducted by a biophysical-based Root Zone Water Quality Model (RZWQM2). The classical regression models (namely RF, SVM and LASSO) satisfactorily simulated cyclical and seasonal variations of CO_2 fluxes (R = 0.75, 0.71 and 0.68, respectively); however, they failed to reasonably predict the peak values of N_2O fluxes (R <0.25). Shallow ML was found to be less effective in

predicting GHG fluxes than other considered ML models ($R < 0.7$ for CO_2 flux and $R < 0.3$ for estimating N_2O fluxes) and was the most sensitive to hyperparameter tuning. Based on this comprehensive comparison study, it was elicited that the LSTM model can be employed successfully in simulating GHG emissions from agricultural soils, providing a new perspective on the application of machine learning modeling for predicting GHG emission to environment.

Keywords: Machine learning; Classical regression; Shallow learning; Deep learning; Greenhouse gas emissions; Agricultural soil.

INTRODUCTION

Anthropogenic greenhouse gas (GHG) emissions have significantly been increased due to the growing use of fossil fuels, deforestation and land use changes. The Intergovernmental Panel on Climate Change (IPCC) has highlighted the impact of GHG emissions, particularly carbon dioxide (CO_2), nitrous oxide (N_2O), troposphere ozone (O_3), methane (CH_4) and chlorofluorocarbons (CFCs) on climate change (Pachauri et al., 2014). The increase in average global GHG concentrations was higher over the last half-century than all previous years (Masson-Delmotte et al., 2018). Anthropogenic activities in the agricultural and forestry sectors, particularly due to intensification and changes in land-use resulted in GHG emissions causing an ecosystem imbalance (Aneja et al., 2009). In Canada, agriculture is considered as one of the major emitters of greenhouse gases (Crippa et al., 2018; Janssens-Maenhout et al., 2017).

Agricultural soils can be both emitters and sinks of GHGs. Climate-smart agronomic strategies, including the proper use of irrigation, tillage, drainage and fertilization, combined with soil characteristics (e.g. organic matter, texture, bulk density, pH and microbial activity) and environmental conditions (e.g. air and soil temperatures and precipitation) all contribute to reducing greenhouse gases. With the challenges of food security, population growth and climate change, agriculture is required to be more innovative in reducing GHGs. Emissions from soils are usually measured by various techniques, like closed chambers and flux towers (Li et al., 2018; Lloyd et al., 2019; Xu et al., 2018; Yu et al., 2013). Various biophysical models have been

successfully developed for simulating GHG emissions, such as the Denitrification-Decomposition model (DNDC) (Gilhespy et al., 2014; Zhang and Niu, 2016), Root Zone Water Quality Model (RZWQM2) (Jiang et al., 2018; Smith et al., 2019), daily version of the CENTURY model (DAYCENT) (Necpálová et al., 2015) and Decision Support System for Agrotechnology Transfer (DSSAT) (Jiang et al., 2019a). These biophysical models, although effective and widely-used, have their own constraints and limitations, mainly because they require: (1) qualified and experienced users having agro-environmental expertise, knowledge and skills; (2) pre-procedures and protocols for model calibration and validation; and (3) availability of various input parameters dependent on site-specific conditions (agricultural fields, forests, savannahs, rangelands and geographical location). In order to overcome these limitations, ML algorithms can be used as an alternative or as a complement to the biophysical approaches described above, to simulate and predict GHG emissions in particular, and environmental phenomena in general. Such models include support vector machine (SVM) (González Costa et al., 2017; Liu et al., 2017; Pham et al., 2019; Yan et al., 2020), random forest (Dou et al., 2019; Hong et al., 2019; Hu et al., 2017; Rodriguez-Galiano et al., 2014; Zhang et al., 2017), artificial neural networks (ANN) (Ganzenmüller et al., 2019; Pradhan et al., 2020; Schmidt et al., 2018; Yang et al., 2020) and deep learning (DL) (Dou et al., 2020; Ng et al., 2020; Wu et al., 2020). In ANN, the algorithm must be extremely flexible to accomplish various tasks. A learning mechanism is used to train each component of the network on how to achieve the desired task. DL has advanced extensively in the past few years (Reichstein et al., 2019). These algorithms are a class of ML that use as processing units, special neural networks with multiple layers for feature extraction and transformation (LeCun et al., 2015). The term “deep” refers to the number of hidden layers comprising the network to draw a distinction between deep and shallow learning (SL) with only one hidden layer. Deep learning (also known as hierarchical learning) is the terminology used for the training process of deep neural networks (DNNs). Compared to shallow neural networks (SNN), DNNs could learn more complicated tasks and therefore accurately simulate complex processes. DNNs have been applied successfully in many fields including

agriculture. The main applications of DNN in agricultural research are: a) image identification and classification; where one can distinguish: plant and weeds (Grinblat et al., 2016; Singh et al., 2016), land cover and crop type classifications (Kussul et al., 2017; Walker, 2016) and fruits or leaf counting (Maldonado and Barbosa, 2016; Rahnemoonfar and Sheppard, 2017), b) predicting future trends; where the following lines of research can be found: estimation of crop yield (Chlingaryan et al., 2018), predicting soil moisture content (Adeyemi et al., 2018; Fang et al., 2017) and weather conditions (Scher, 2018), etc. This review on the application of ML in agriculture research confirms the fact that ML models are modern approaches with high potential and promising prospects (Kamilaris and Prenafeta-Boldú, 2018).

This review of previous studies shows a research gap concerning the use of ML models in the simulation and prediction of GHG emissions from agricultural fields. Therefore, this paper investigates the potential of exploiting ML models to predict soil GHG emissions in agricultural fields. To this end an objective comparison between nine ML models (classical, shallow and deep learning) was conducted to ascertain the most efficient model. The study also identified the predictive accuracy and computational time of all models, and further analyzed their strengths and weaknesses for GHG prediction.

MATERIALS AND METHOD

Study site. Greenhouse gas emissions (N_2O and CO_2) were measured at a drainage and water table management research site in Saint-Emmanuel, in the region of Côteau-du-Lac, Quebec, Canada (Figure 1). Measurements were carried out at three plots (75m ×15m each) during the growing seasons of 2012, 2013, 2014, 2015 and 2017. A full description of the site and agricultural management treatments has been given in recent paper (Jiang et al., 2019b). Soil physical and chemical properties are provided in supplementary material (Table S1).

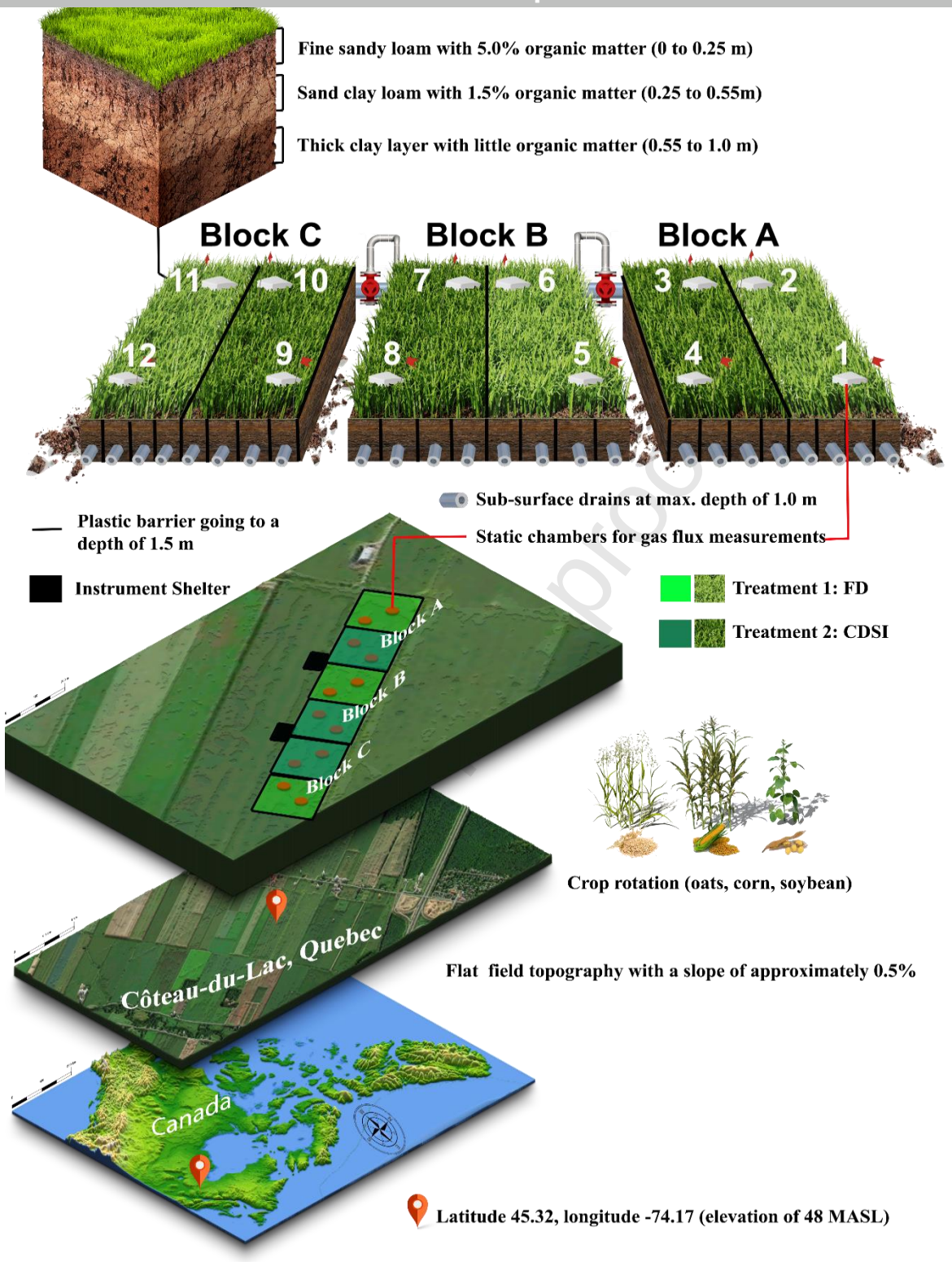

Figure 1. Schematic representation of site description, water treatments, soil and crop types.

Measurements and data collection. Gas samples were collected from static chambers located at the study site and analyzed at McGill University lab using a Bruker 450-GC chromatograph system (Bruker corp., Bremen, Germany). Other measurements were taken at the time of gas sampling i.e. air/soil temperatures and soil volumetric water content. Meteorological data i.e. precipitation, wind speed, humidity and surface pressure were taken from the Environment Canada weather station at Côteau-du-Lac (Station ID – 7011947) and the NASA POWER Project Data Sets (Rosenzweig et al., 2014; Stackhouse et al., 2018). The crop N uptake was calculated from model-based patterns (Heard and Hay, 2006; Ross R. Bender, 2013). A more detailed description of the gas sampling and analysis can be found in (Crézé and Madramootoo, 2019; Edwards et al., 2018). As shown in Figure 2, two groups of experimentally measured data series including four years, from 2012 to 2015, of CO_2 and N_2O fluxes were used to build predictive models through the training of ML and a one-year dataset (2017) was employed to assess the prediction performance of the constructed ML models. The plots were subjected to two water treatments i.e. free drainage (FD) and controlled drainage sub-irrigation (CDSI) (FD in 2012 and 2013 seasons, FD and CDSI in 2014, 2015 and 2017, however only data from FD treatment were considered for 2017). It can be seen in Figure 2 that CO_2 is cyclically emitted throughout the year, while N_2O mostly occurred as sharp peaks in June/July, particularly after fertilizer applications. It can be also noticed that one of the challenges of this study was the limited data availability (with a total amount of 126 data by considering FD and CDSI). The data used in this paper were summarized in supplementary material (Table S2).

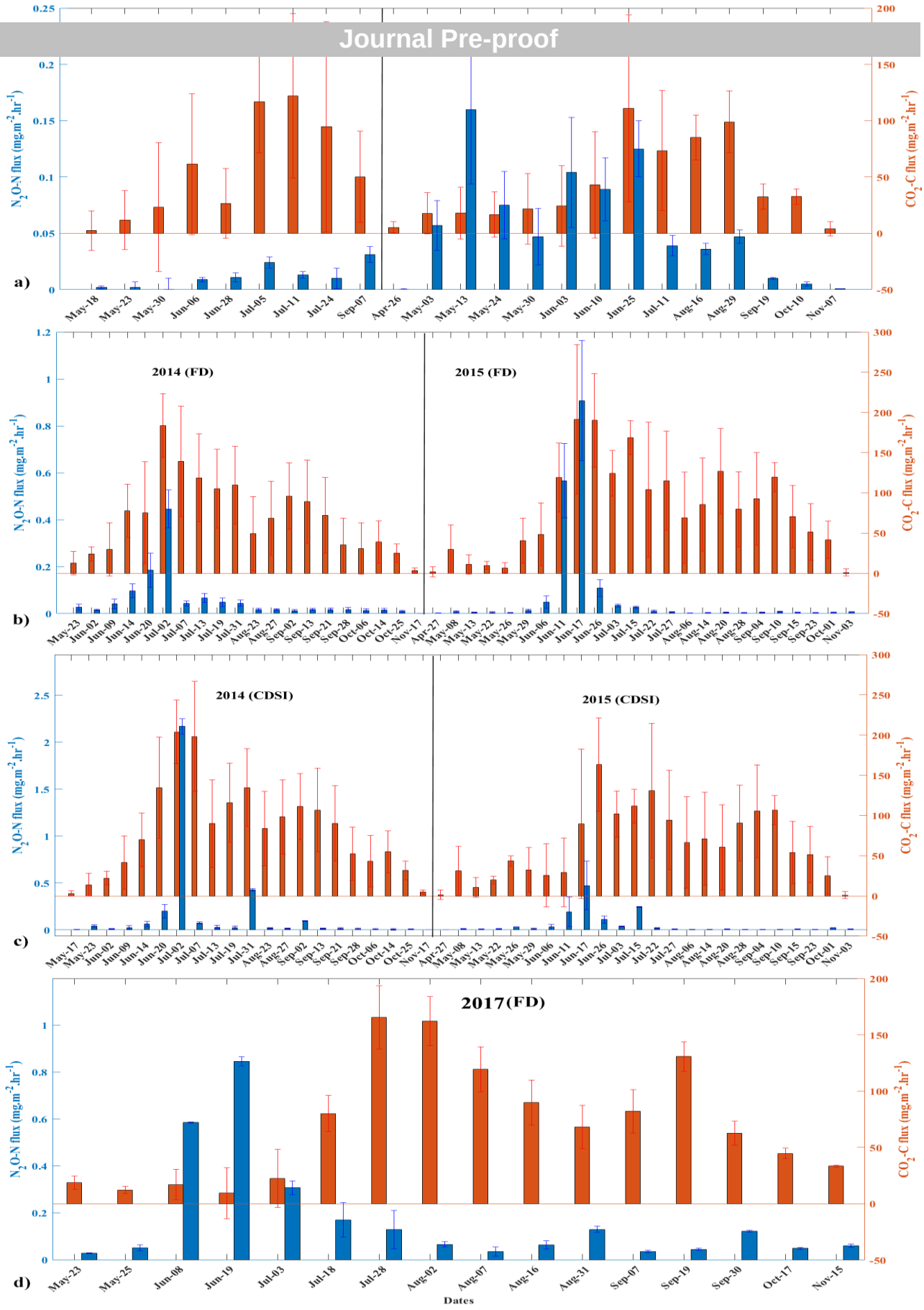

Figure 2. GHG fluxes (CO_2-C and N_2O-N) data during: (a) The growing seasons of 2012-2013 (FD), (b) 2014-2015 (FD), (c) 2014-2015 (CDSI) and (d) 2017 (FD).

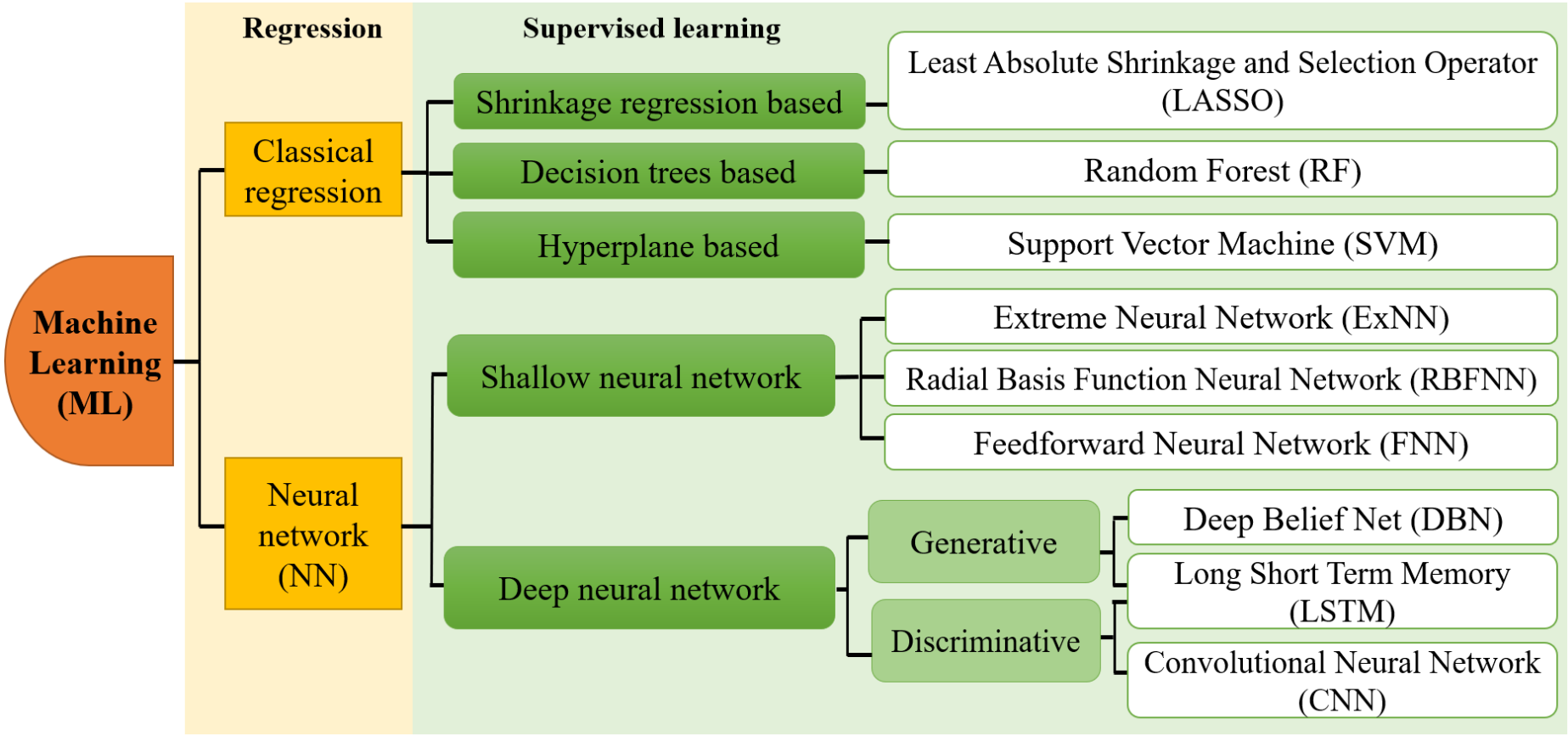

Figure 3. Flowchart presenting ML models used in this study.

regression (i.e. SVM, RF and LASSO), Shallow learning (i.e. FNN, RBFNN and ExNN), Deep learning (i.e. LSTM, DBN and CNN). ML models were implemented in MATLAB® software. The reader may refer to the supplementary material (Table S3) for more detailed coverage of all ML methods.

Support vector machine (SVM). This is a supervised ML classifier model based on the separation of data by hyperplanes (Cortes et al., 1995). SVM has been extensively used for solving regression problems as function approximation and prediction (Chang and Lin, 2011). The SVM constructs linear and non-linear decision boundaries (hyperplanes) in variable space by optimizing a distance-based (margin) formulation on a subset of training data designed as support vectors (Smola and Schölkopf, 2004). In the case of non-linearly separable data, kernel functions (e.g. polynomial and radial basis functions) are used in SVM to transform the data into another form that allows hyperplane construction.

Random Forest (RF). So-called random decision forests, are supervised ML classifiers based on decision trees (Breiman L., 2001). These decision trees are learning ensembles constructed by the bootstrap, aggregation or bagging techniques (Zhang and Ma, 2012). Decision trees grow in parallel by splitting randomly and searching for the optimal combination of all variables. The optimization process does not weigh the model since RF involves only few hyperparameters. These combined features make RF a widely-used ML predictor (Kulkarni and Sinha, 2012).

Least absolute shrinkage and selection operator (LASSO). Was developed initially to improve the ordinary least square framework (Tibshirani, 1996), this extension of least squares optimizes the summation of their absolute values multiplied by a weight parameter that decelerates or accelerates the penalty. LASSO model is almost similar to ridge regression; however, the only difference is that LASSO uses L1 norm and ridge regression uses L2 norm

Feed-forward neural networks (FNN). This model is a neural network that typically includes one hidden layer. The first and the last layers are for inputs and outputs, respectively. The operating principle of FNN is based on the behavior of neurons and synaptic connections, where the input data are processed through three different operations: (1) Multiplication by the synaptic connection weights, (2) Summation with adding the threshold function and (3) Transformation of results by applying the activation function. In the training phase, FNN automatically adjusts their synaptic weights and bias by the back-propagation learning procedure and thereby constructs a unique model corresponding to the input/output relation in the network. However, the standard backpropagation algorithm used for solving this optimization problem has a low rate of convergence, high risk of falling in a local minimum and is computationally expensive (Møller, 1993). Therefore, the Bayesian regularization backpropagation (MacKay, 1992a, 1992b; Schmidt et al., 2018) was used in this study.

Radial basis function neural network (RBFNN). These ML models are FNN with radial basis functions as activation functions. Global approximation, compact scheme and ability to fit continuous and noisy data are features that distinguish RBFNN from other SNN (Yu et al., 2011). The most commonly used RBFNN activation function is the Gaussian function, which possesses only a few hyperparameters. RBFNN's principle of operation does not differ from FNN. However, the training process is tuned by incorporating some clustering techniques such as decision trees (Kubat, 1998), or k-means (Sing et al., 2003).

Extreme neural network (ExNN). Initially proposed by Huang (Huang and Chen, 2007), this algorithm is a modified version of FNN, that uses different training procedure mainly based on two techniques (Huang et al., 2015): (1) Random feature mapping of the hidden layer and (2) Optimization of synaptic weights solved by a least squared error norm. The first technique

the input layer by some nonlinear functions such as sigmoid and radial basis functions (RBFs). The second stage involves the use of a moving least square (MLS) formulation to minimize the approximation error.

Long short-term memory network (LSTM). Initially introduced by Hochreiter and Schmidhuber (Hochreiter and Schmidhuber, 1997), this model belongs to the family of recurrent neural networks (RNN). RNNs are neural networks based on the concept of time layering, which gives them the ability to take temporal series data as inputs and outputs. In more general terms, RNNs are particularly suitable for sequence-to-sequence learning purposes. However, ordinary RNNs suffer from shortcomings, i.e., vanishing and exploding gradients (Sainath et al., 2015). Originally designed to consider long-term dependencies in the learning phase, LSTM can overcome the previous shortcomings of RNNs (Fischer and Krauss, 2018). The standard structure of LSTM networks consists of an input layer, multiple hidden layers and an output layer. In contrast to other RNNs, memory cells, which are the main features of LSTM networks, are incorporated into the hidden layers. In a similar manner to the rules concerning the training phase in SNN, the weights and bias terms are computed using the back-propagation algorithm with gradient descent so that an objective function is minimized across temporal target values.

Deep belief network (DBN). This is a deep learning model, introduced by Hinton et al. (Hinton et al., 2006; Hinton and Salakhutdinov, 2006). The main building blocks of DBN are the Restricted Boltzmann machines (RBMs). A linear regression layer is usually added to DBN for the predicting task. RBM is a variant of Markov random field techniques, where the learning phase is performed by means of probability distribution functions. In one RBM network, neurons of visible layers are connected with neurons of the hidden layers using undirected weighted connections, however, no interconnection (restricted) exists between neurons in the same layer.

classification and model recognition problems, CNN was inspired by the biological organization principles in the animal optical cortex (Daw, 2009; Hubel and Wiesel, 1962). CNN was based on the concept of local connectivity where convolutions are used instead of the standard FNN weighted sums. In general, the structure of CNN consists of piles of alternating convolutional and pooling layers and a logistic regression layer as an output. The convolutional layer uses convolution operations to the input image. The pooling layer is responsible for reducing image dimension by merging neighboring pixels. The notion of local connectivity means that no connection between neurons in the same layer exists and for neurons from different layers, they are interconnected via a weight sharing technique.

Machine learning configurations. To compare the predicting performance of ML models, we conducted several computational simulations, where we fixed a set of hyperparameters (e.g. activation function, kernel function, number of epochs and hidden layers) as indicated in the supplementary material (Table S4). The hyperparameters tuning is an essential step in ML modeling since it controls the behaviors of their training algorithms and hence affects their predictive performance. Several approaches have been developed to wisely select these hyperparameters (Bengio, 2000; Bergstra et al., 2012; Duan et al., 2003; Keerthi, 2002; Probst et al., 2019). However, because each ML model had its particular hyperparameters, it will be very difficult for us to test all possible hyperparameter combinations to get the optimal configurations for the nine ML models. For this reason, we selected in this study various scenarios by fixing a range of hyperparameters covering the most common choices in the literature (Khaledian and Miller, 2020; Luo, 2016). Sensitivity analysis was also conducted for the SNN and DNN models (except ExNN model where the hyperparameters are internally optimized) by varying the number of neurons and epochs in order to examine their effects on the training process (supplementary material Table S5). The details of the statistical metrics (i.e. coefficient of

(MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE)) used to assess the effectiveness of ML models in predicting GHG emissions can be found in the supplementary material (Table S6).

RESULTS AND DISCUSSION

Model input selection. A preliminary analysis was conducted to determine which input parameters have the highest statistical significance on the prediction of GHG emissions during the training period (2012 to 2015 growing seasons). This analysis involved: correlation, stepwise regression, neighborhood component analysis (NCA) and minimum redundancy maximum relevance (MRMR) algorithm. The input parameters (or predictors) included: air (T_{AIR}) and soil (T_{SOIL}) temperatures to emphasize the effect of environmental temperature, soil volumetric water content (VWC), precipitation and humidity to account for atmospheric moisture content, wind speed (V_{WIND}), surface pressure and crop N uptake (N_{Uptake}). The latter accounted for the effect of fertilizer applications and crop types. We found that soil temperature produced the highest Pearson's correlation coefficients (R) and the lowest values of significance coefficients (Sig.) or P-values (R=0.68 and Sig.=$3.9.10^{-16}$ for CO_2 flux and R=0.37 and Sig.=5.10^{-5} for N_2O flux (Figure 4)). However, the values of correlations in the case of N_2O fluxes remain low (with a maximum of R value of 0.37). An obvious correlation was found between soil temperature and air temperature. The same observation is valid concerning the climatic parameters, i.e. precipitation, wind speed, humidity and surface pressure, all strongly correlated with a low significance coefficients. In Table 1, the stepwise regression procedure was used as a selection technique to measure the effect of all predictors on GHG fluxes. The analysis confirmed the finding that air and soil temperatures are the most influential parameters on the emitted GHG fluxes. This is consistent with the findings of other studies (Berglund et al., 2010; Oertel et al.,

were obtained by considering all predictors/inputs (R^2 = 0.96 and RMSE=16.6 $mg.m^{-2}.hr^{-1}$, R^2 = 0.99 and RMSE = 0.04 $mg.m^{-2}.hr^{-1}$ for emitted CO_2 and N_2O fluxes from soils, respectively). The final results from the two feature selection methods i.e. NCA and MRMR were illustrated in Figure 5. The feature weights obtained from the NCA analysis suggested that (classified from the highest to lowest value) air and soil temperatures, humidity, soil volumetric water content, crop N uptake and surface pressure were the most influencing factors on the emission of CO_2 fluxes. According to the NCA analysis, the crop N uptake and to a lesser degree the humidity were the only influential parameters on N_2O fluxes. From this analysis, the crop N uptake was estimated to directly affect N_2O fluxes more than CO_2, which is in line with the findings of several research works (Drury et al., 2014; Linquist et al., 2012; McSwiney and Robertson, 2005; Shcherbak et al., 2014; van Groenigen et al., 2010; van Kessel et al., 2013). The results from MRMR algorithm showed almost similar findings that air and soil temperatures were the most influential factors on CO_2 and N_2O fluxes respectively, except for the precipitation which was found to be important on the CO2 fluxes, a fact that was also reported in other research studies in the literature (Ni et al., 2019; Wu et al., 2013; Yan et al., 2018).

In view of these findings, all input parameters (i.e. air and soil temperatures, soil volumetric water content, precipitation, humidity, wind speed, surface pressure and crop N uptake) were selected for the ML predictive analysis of soil GHG emissions.

Results for CO_2-C flux ($mg.m^{-2}.hr^{-1}$)										
Order of stepwise regression	**R^2**	**RMSE**	**Predictors**							
			T_{AIR}	**T_{SOIL}**	**VWC**	**Precipitation**	**V_{WIND}**	**Humidity**	**Pressure**	**N_{Uptake}**
Linear	0.50	35.4		*						
Quadratic	0.78	26		*	*	*	*	*	*	*
			**	**					**	
Cubic	0.96	16.6	*	*	*	*	*	*	*	*
			**	**	**		**	**	**	**
								***		
Results for N_2O-N flux ($mg.m^{-2}.hr^{-1}$)										
Linear	0.24	0.21	*	*						
Quadratic	0.72	0.14	*	*	*	*	*	*	*	*
				**	**	**	**			**
Cubic	0.99	0.04	*	*	*	*	*	*	*	*
			**	**	**	**	**	**	**	**
								***	***	***
*	Linear terms									
**	Quadratic terms									
***	Cubic terms									

Table 1: Input selection using stepwise regression analysis.

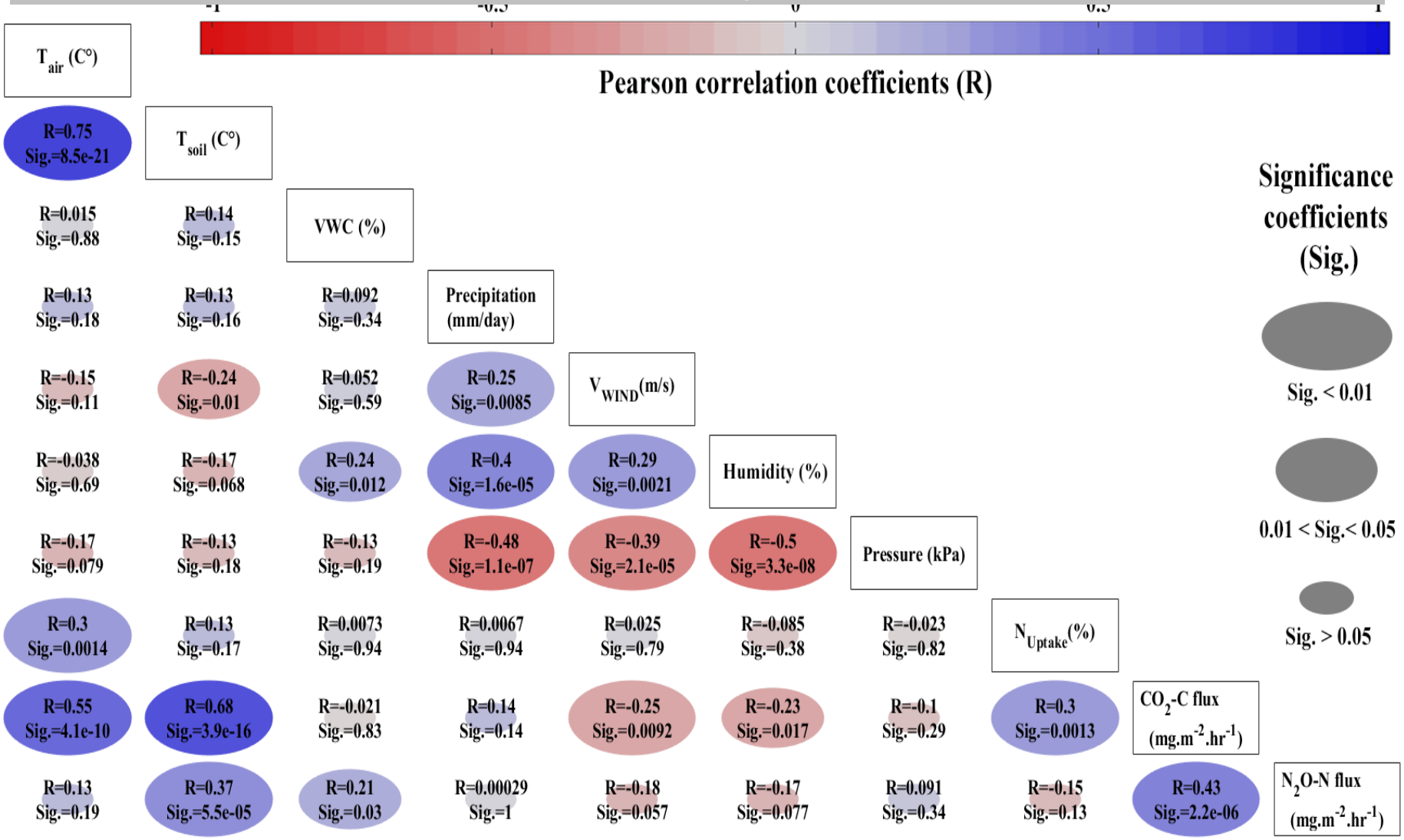

Figure 4. Correlogram of GHG (CO_2 and N_2O) fluxes and all input variables.

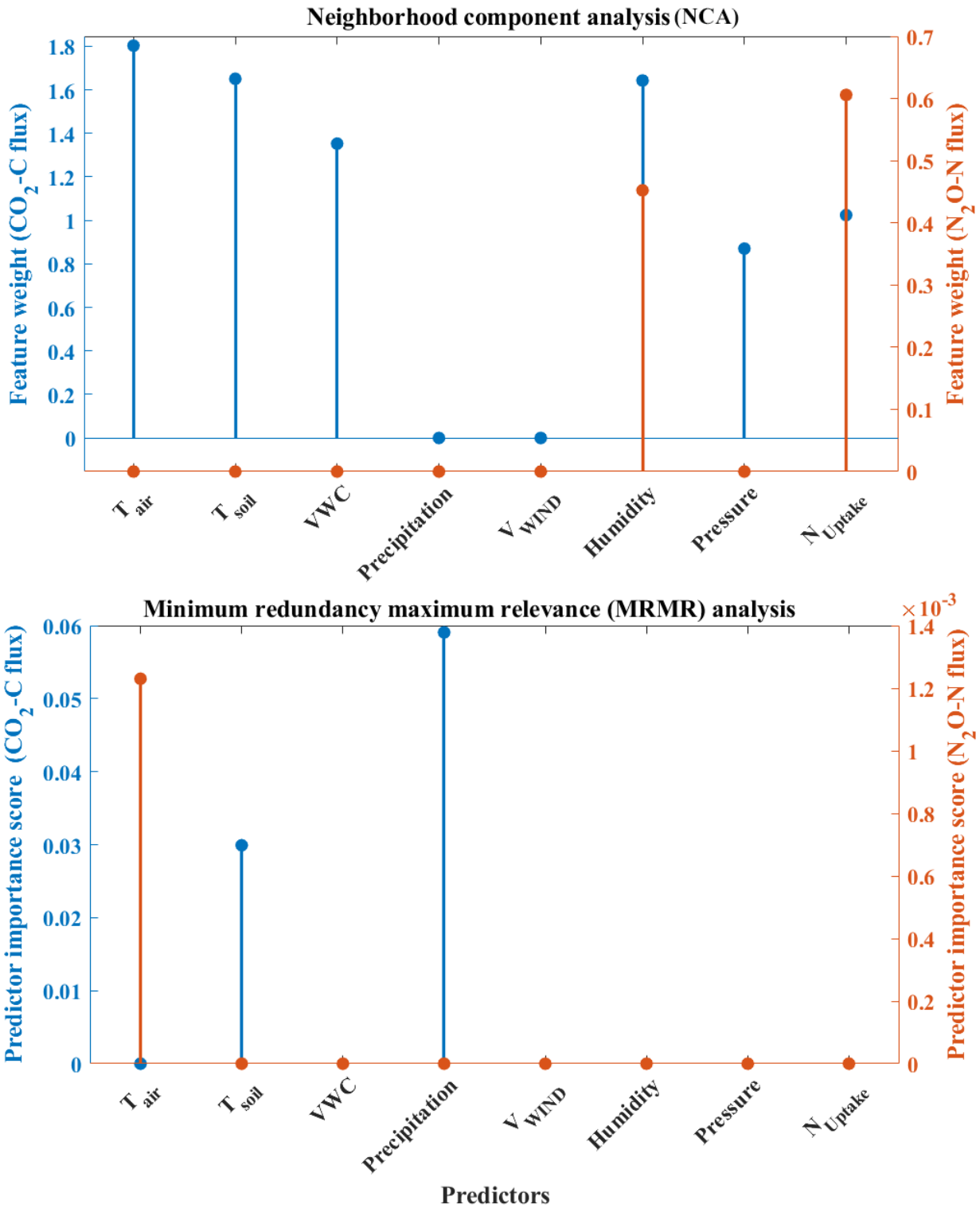

Figure 5. Results of the feature selection analysis by NCA and MRMR algorithms.

Predictive performance: CO_2 fluxes. Figures 6 and 7 show the performance metrics for the CO_2 training and prediction stages obtained from the nine ML models. The box and whisker plots indicate the median value and variability of the performance metrics. The radar plots show the best variant of each ML model, in terms of the best R coefficient. The heatmaps show the ranking of these ML variants.

The results clearly demonstrate that LSTM performs the best among all ML models at both training and prediction stages (R=0.97 and 0.87, respectively). The range of variability of LSTM is narrow (slightly affected by the change in the number of hidden neurons), confirming that LSTM models have the ability to store historical information through their internal state cells, which act as long-term or short-term memory cells. This capacity to store historical information is extremely valuable when we need to predict new data depending on the past data sets. This was found to be true in previous studies (Qi et al., 2019; Zhang et al., 2018). Compared to the simulation results obtained by the biophysical model RZWQM2 in (Jiang et al., 2019b) for the same study site and time intervals for data collection, (the best value of R was 0.84), LSTM was able to give better results with R=0.88. In the same category of deep learning, DBN models performed well in the training phase, but did not perform as well in the prediction phase (R=0.96 and 0.65, respectively) and this is due to their low effectiveness of learning in data series prediction tasks (Kuremoto et al., 2014; Längkvist et al., 2014). CNN performed poorly in both training and prediction stages (R=0.73 and 0.68, respectively), which is mainly due to insufficient amount of measured CO_2 fluxes and input data to be transformed into a 2D image (used in the CNN data-image input layer) (LeCun et al., 2015; Sharma et al., 2019). It is further observed that classical regression models (i.e. RF, SVM and LASSO) performed better compared to the SNN models (i.e. FNN, RBFNN and ExNN), especially in the prediction phase. Among the classical regression models, RF exhibited the best performance metrics, thus ranking

second after the LSTM. The RF model achieved good predictive performances in both training and prediction (R=0.96 and 0.75, respectively) and proved to be a simple, yet an effective decision tree model (Prasad et al., 2018). Next was the LASSO (R=0.80 and 0.71 in the training and prediction phases, respectively) and, to a lesser degree, SVM with acceptable prediction performance (R=0.92 and 0.68 in the training and prediction stages, respectively). SVM exhibited the highest variability compared to the other ML models, which indicates the high sensitivity of SVM to the hyperparameters tuning. SNN models performed poorly overall in their prediction results, with the exception of FNN where acceptable results were obtained (R= 0.94 and 0.68 in the training and prediction stages, respectively). Unlike SNN models, ExNN presents constant results with the lowest variability due to the internal optimization of its hyperparameters. From the heatmap of the prediction performances with regard to all metrics, ML models can be classified from the best to the worst, as follows: (1) LSTM, (2) RF, (3) LASSO, (4) FNN, (5) SVM, (6) CNN, (7) ExNN, (8) RBFNN and (9) DBN.

N_2O fluxes. Unlike the case of CO_2 fluxes, the challenge with N_2O predictions lies within the ability of ML models to capture the peak N_2O fluxes. Figures 8 and 9 show clearly that LSTM was the only ML model that accomplished this challenge and succeeded in capturing the pattern, outperforming all ML models (R=0.99 and 0.86 in the training and predicting stages, respectively) and the RZWQM2 (where the best R was 0.75 in the case of N_2O fluxes) (Jiang et al., 2019b), with less variability compared to the previous CO_2 predictions. The short-term memory cells of LSTM have the ability to detect sharp peak patterns occurring in a short time period. These findings are very much in accordance with those found in (Liu et al., 2019; Yue et al., 2020). DBN and CNN performed slightly better in the prediction phase (R=0.37 and 0.34, respectively) than during the training (R=0.1 and 0.2, respectively). CNN was ranked as the second ML in the testing/prediction stage (when considering all metrics) and this can be

explained by the fact that the peaks in the data appear as darker pixels in the CNN data-image input layer, making it easier for CNN to detect. However, the lack of sufficient training data makes CNN less accurate for predictions. Shallow ML followed the Deep ML, where RBFNN outperformed other shallow ML models in the prediction stage (R=0.34). The RBFNN model is characterized by a compact support and ability to fit a data series, allowing the RBF to capture the short-term pattern of the N_2O peak fluxes. In general, all ML models, except LSTM, failed to produce good prediction results of N_2O fluxes.

From the heatmap of the prediction performances with regard to all metrics, ML models can be ranked from the best to the worst as follows: (1) LSTM, (2) CNN, (3) RBFNN, (4) LASSO, (5) DBN, (6) FNN, (7) ExNN, (8) SVM and (9) RF.

Cross validation analysis: In the previous predictive analysis, the evaluation of ML model performance was based on GHG data split into four years training set (2012 to 2015) and one year predicting or testing set (2017). However, this analysis remained partially reliable since the predictive performance obtained for only one test set can be different to that obtained for another test set. In the cross-validation (CV), the original dataset (training + testing) was divided into partitions in such a way ensuring that every partition had the chance to be appeared in the training and testing set (Rodríguez et al., 2010).

K-Fold cross-validation: Here, the data collected during the five years were split into five subsets (K=5 folds), each subset corresponding to one growing season (Figure 10). In the first iteration, the first fold corresponding to the growing season of 2012, was used to test ML models and the remaining folds (2013, 2014, 2015 and 2017) were used to the training. In the next iteration, the second fold (2013) served for the testing stage while the rest (2012, 2014, 2015 and 2017) was for the training. The process was repeated until reaching the final iteration. It is noticed that the last iteration is the same as in the previous predictive analysis, where GHG data

from the growing season of 2017 were predicted. Thus, the final iteration was not included in the present CV analysis. It should also be noted that the ratios between training/predicting data increases from one iteration to the next (the proportion of predicting data was 8% in the 1st iteration, 13% in the 2nd iteration, 39% in the 3rd iteration, 42% in the 4th iteration). These changes in the ratio were due to the consideration of the two water management treatments (free drainage (FD) and controlled drainage sub-irrigation (CDSI)) in our analysis. As a result, the 5-fold CV used here, can be considered more as a procedure performed to estimate the skill of ML models on predicting missing GHG data.

The results of the prediction stage in terms of the best R^2 and RMSE values for all ML models are presented in Table 2. These results confirm the superiority of the LSTM model (with the highest R^2 and the lowest RMSE) for predicting CO_2 and N_2O fluxes under various cases of training/testing dataset. The mean values in predicting CO_2 and N_2O fluxes obtained by LSTM (R^2=0.77 and 0.72) were better than those obtained from RZWQM2 (Jiang et al., 2019b) (R^2=0.67 and 0.53). These results were very encouraging for the potential utility of ML models in simulating and predicting GHG emissions.

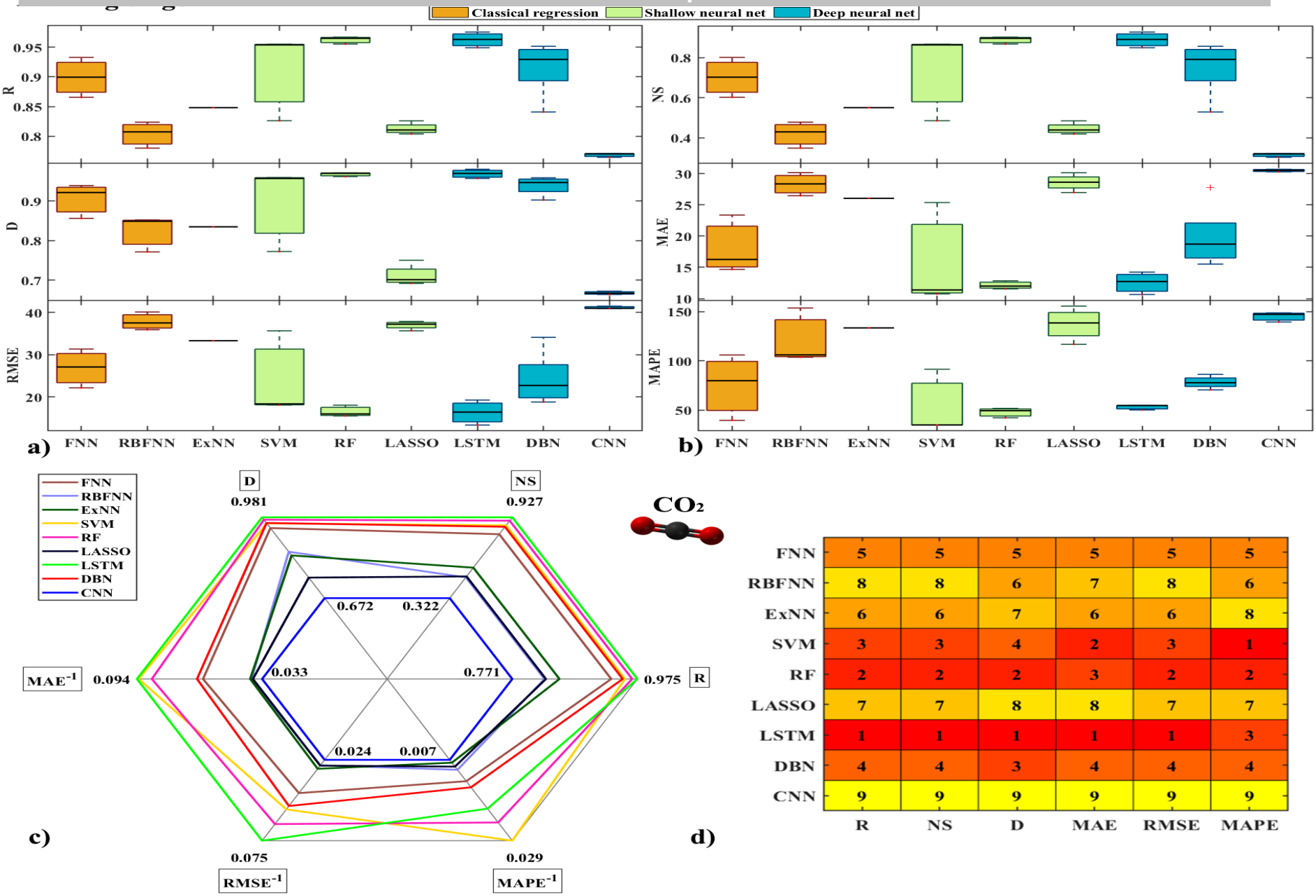

Figure 6. Comparative assessment of the ML models regarding their training performance for CO_2 fluxes according to the statistical metrics: (a-b) Side-by-side boxplots, (c) Radar plot and (d) Heatmap ranking.

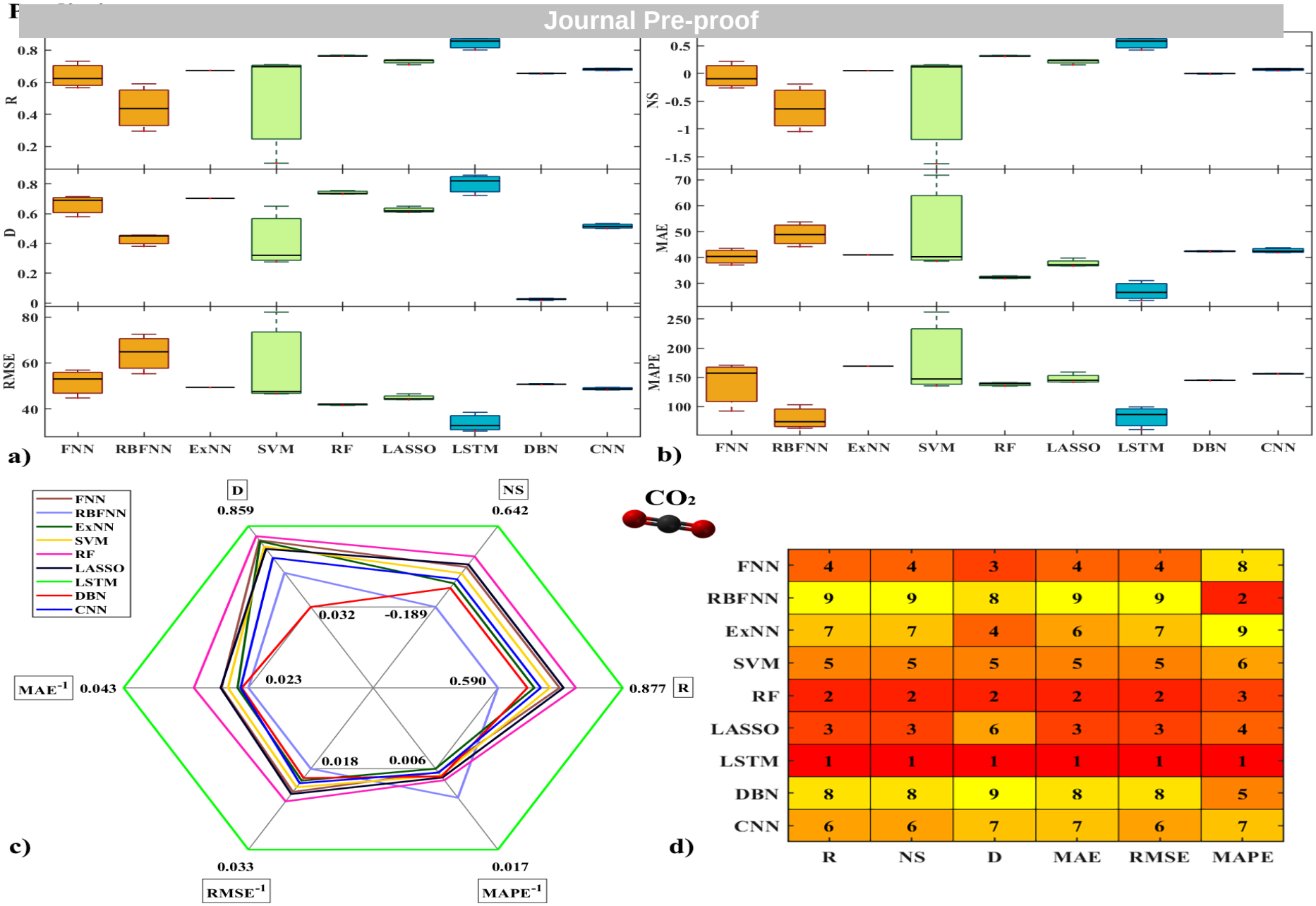

Figure 7. Comparative assessment of the ML models regarding their predicting performance for CO_2 fluxes according to the statistical metrics: (a-b) Side-by-side boxplots, (c) Radar plot and (d) Heatmap ranking.

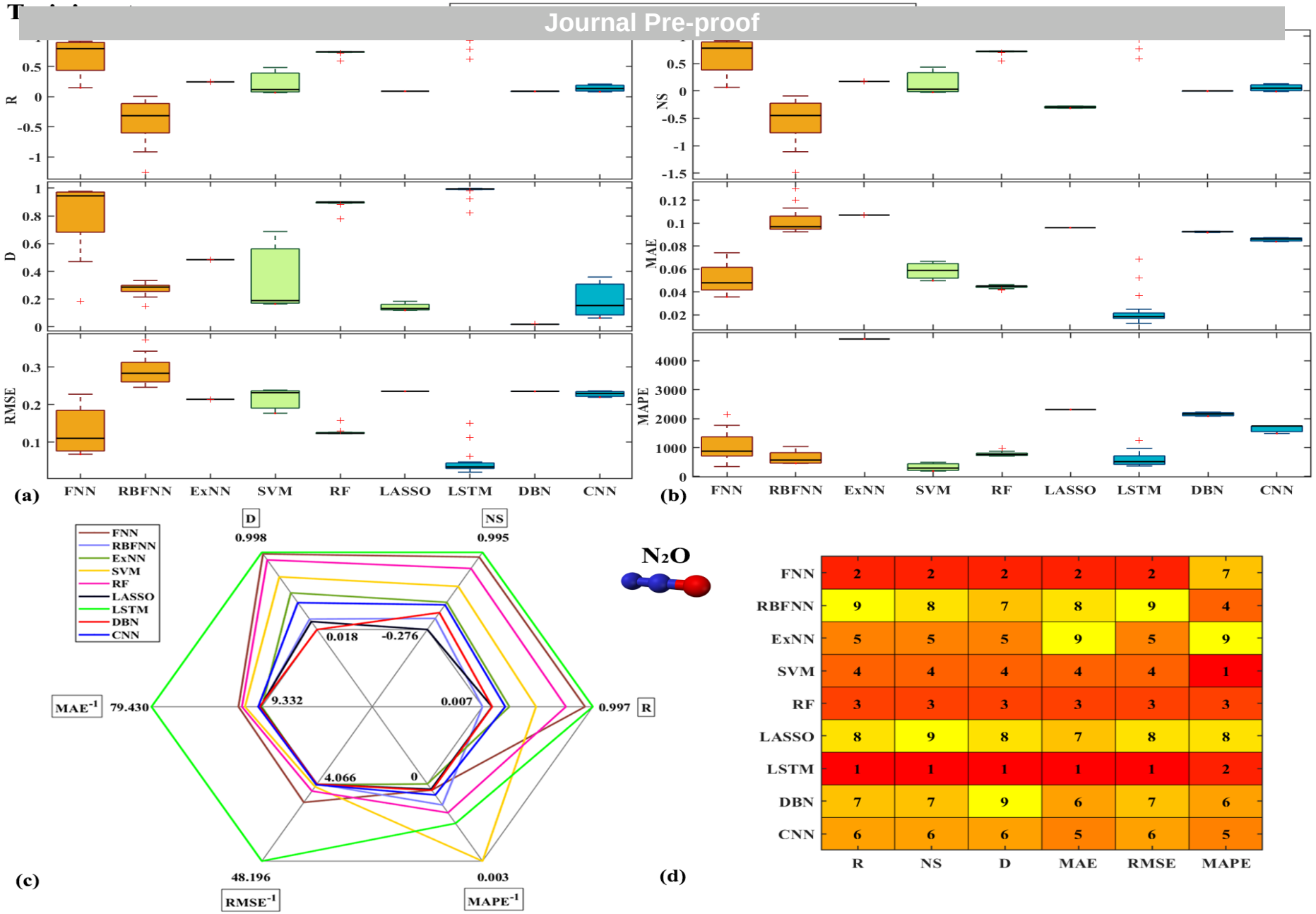

Figure 8. Comparative assessment of the ML models regarding their training performance for N_2O fluxes according to the statistical metrics: (a-b) Side-by-side boxplots, (c) Radar plot and (d) Heatmap ranking.

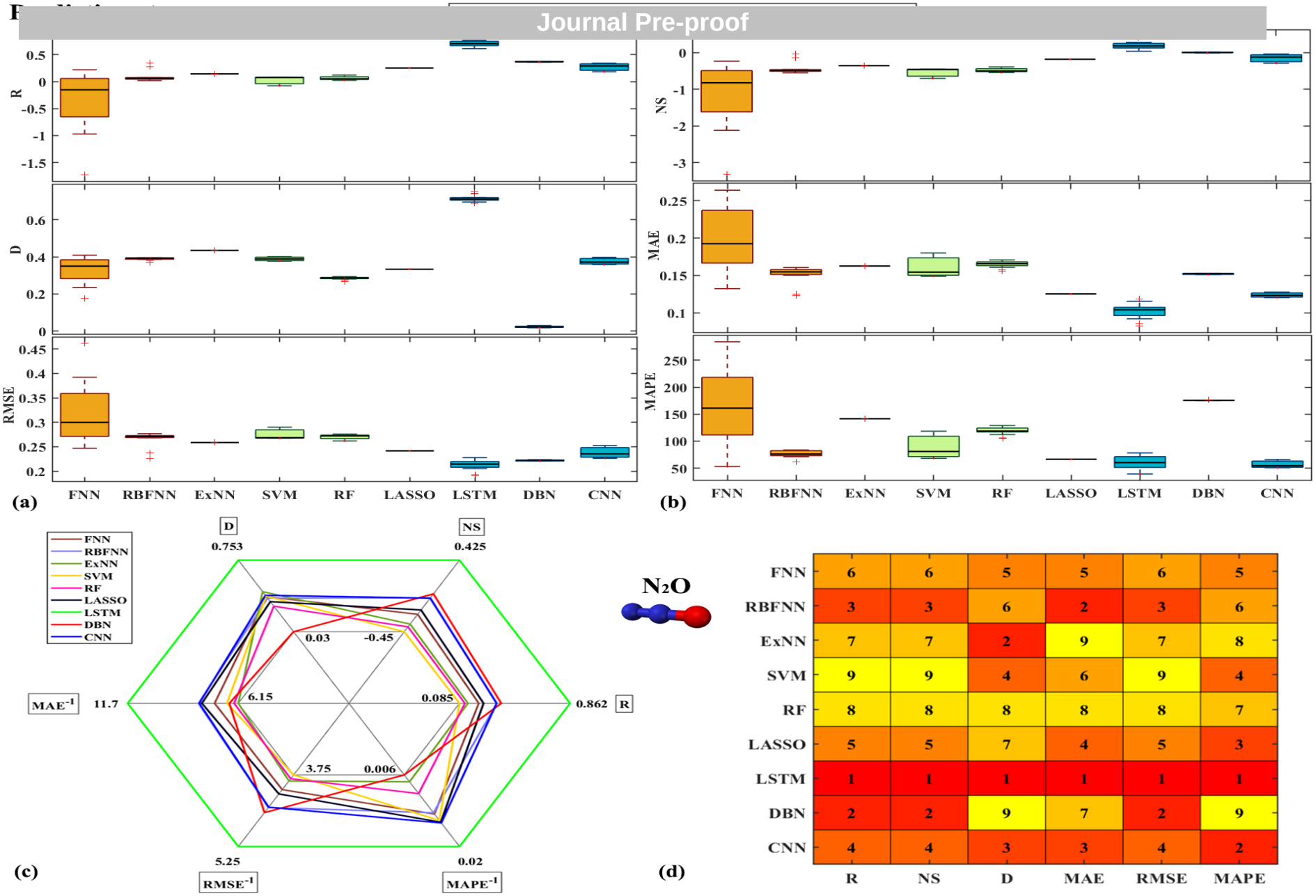

Figure 9. Comparative assessment of the ML models regarding their predicting performance for N_2O fluxes according to the statistical metrics: (a-b) Side-by-side boxplots, (c) Radar plot and (d) Heatmap ranking.

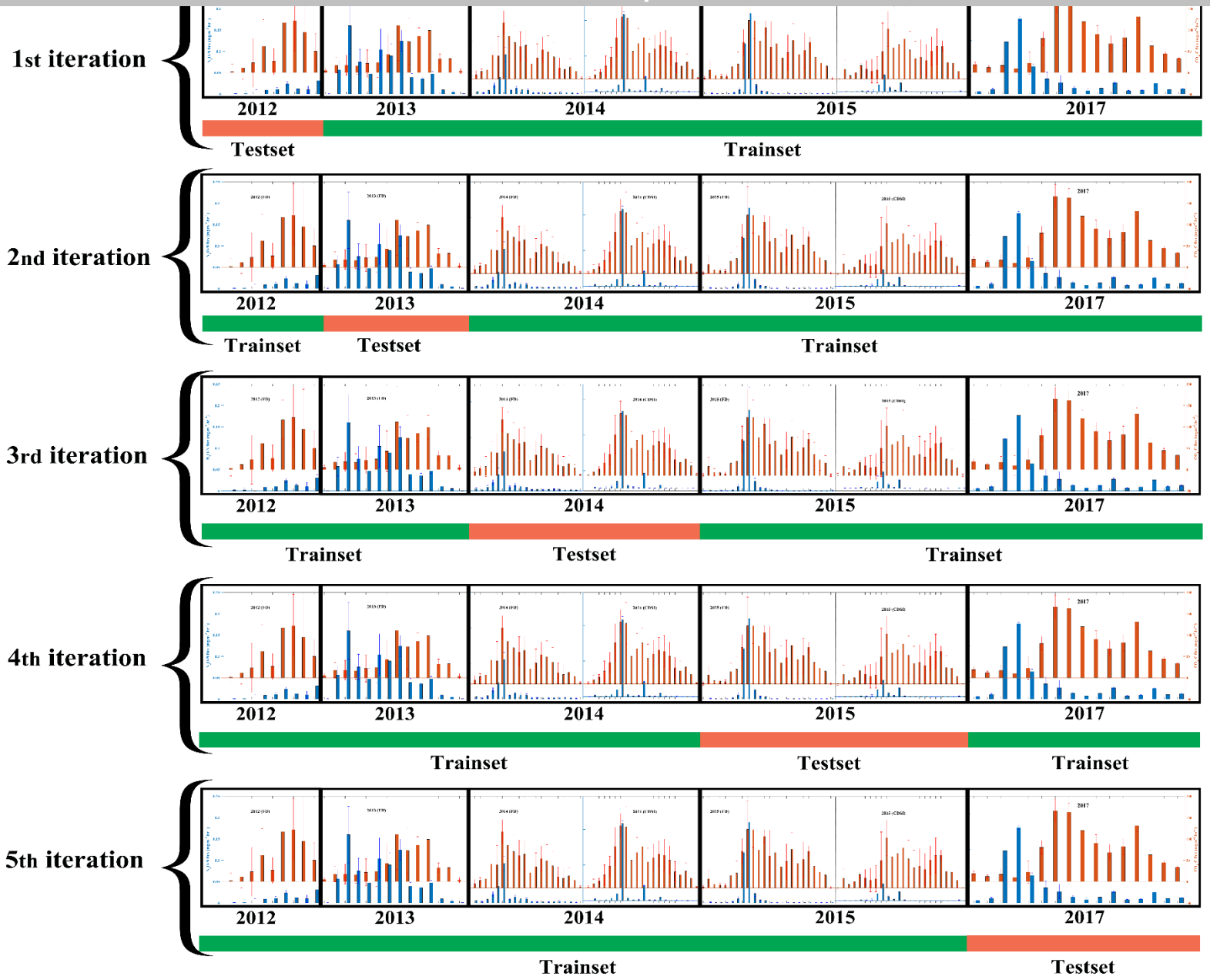

Figure 10. Schematic representation of the 5-fold cross-validation.

Table 2: Results of the K-Folds cross-validation analysis.

CO_2-C fluxes										
Models / **K-Folds (Train-Test ratio)**	**Best R^2(RMSE)**									**Best R^2**
	FNN	**RBFNN**	**ExNN**	**SVM**	**RF**	**LASSO**	**LSTM**	**DBN**	**CNN**	**RZWQM2** (Jiang et al., 2019b)
1st iteration (92-8%)	0.61(36)	0.52(41)	0.61(36)	0.69(31)	0.50(42)	0.67(33)	**0.72(30)**	0.55(40)	0.64(35)	0.71
2nd iteration (87-13%)	0.41(31)	0.34(34)	0.32(34)	0.44(30)	0.45(30)	0.41(31)	**0.67(30)**	0.34(34)	0.36(33)	
3rd iteration (61-39%)	0.61(37)	0.44(47)	0.72(30)	0.79(26)	0.79(25)	0.81(34)	**0.82(25)**	0.45(46)	0.53(41)	0.63
4th iteration (58-42%)	0.79(32)	0.40(60)	0.61(46)	0.71(39)	0.69(40)	0.56(50)	**0.88(23)**	0.48(55)	0.49(54)	
Mean	0.61(34)	0.43(45,5)	0.57(36,5)	0.66(31,5)	0.61(34,2)	0.61(37)	**0.77(27)**	0.46(43,7)	0.51(40,7)	0.67
N_2O-N fluxes										
1st iteration (92-8%)	0.19(0.08)	0.21(0.05)	0.12(0.14)	0.41(0.02)	0.14(0.12)	0.15(0.08)	**0.68(0.01)**	0.35(0.05)	0.22(0.05)	0.50
2nd iteration (87-13%)	0.43(0.09)	0.34(0.04)	0.40(0.12)	0.15(0.05)	0.08(0.06)	0.14(0.05)	**0.67(0.02)**	0.35(0.04)	0.40(0.04)	
3rd iteration (61-39%)	0.05(0.11)	0.05(0.09)	0.04(0.12)	0.10(0.09)	0.02(0.15)	0.01(0.10)	**0.75(0.01)**	0.06(0.09)	0.05(0.09)	0.56
4th iteration (58-42%)	0.18(0.16)	0.01(0.21)	0.07(0.19)	0.20(0.16)	0.07(0.19)	0.01(0.21)	**0.76(0.09)**	0.01(0.20)	0.01(0.21)	
Mean	0.21(0.11)	0.15(0.10)	0.16(0.14)	0.22(0.08)	0.08(0.13)	0.08(0.11)	**0.72(0.03)**	0.19(0.10)	0.17(0.10)	0.53

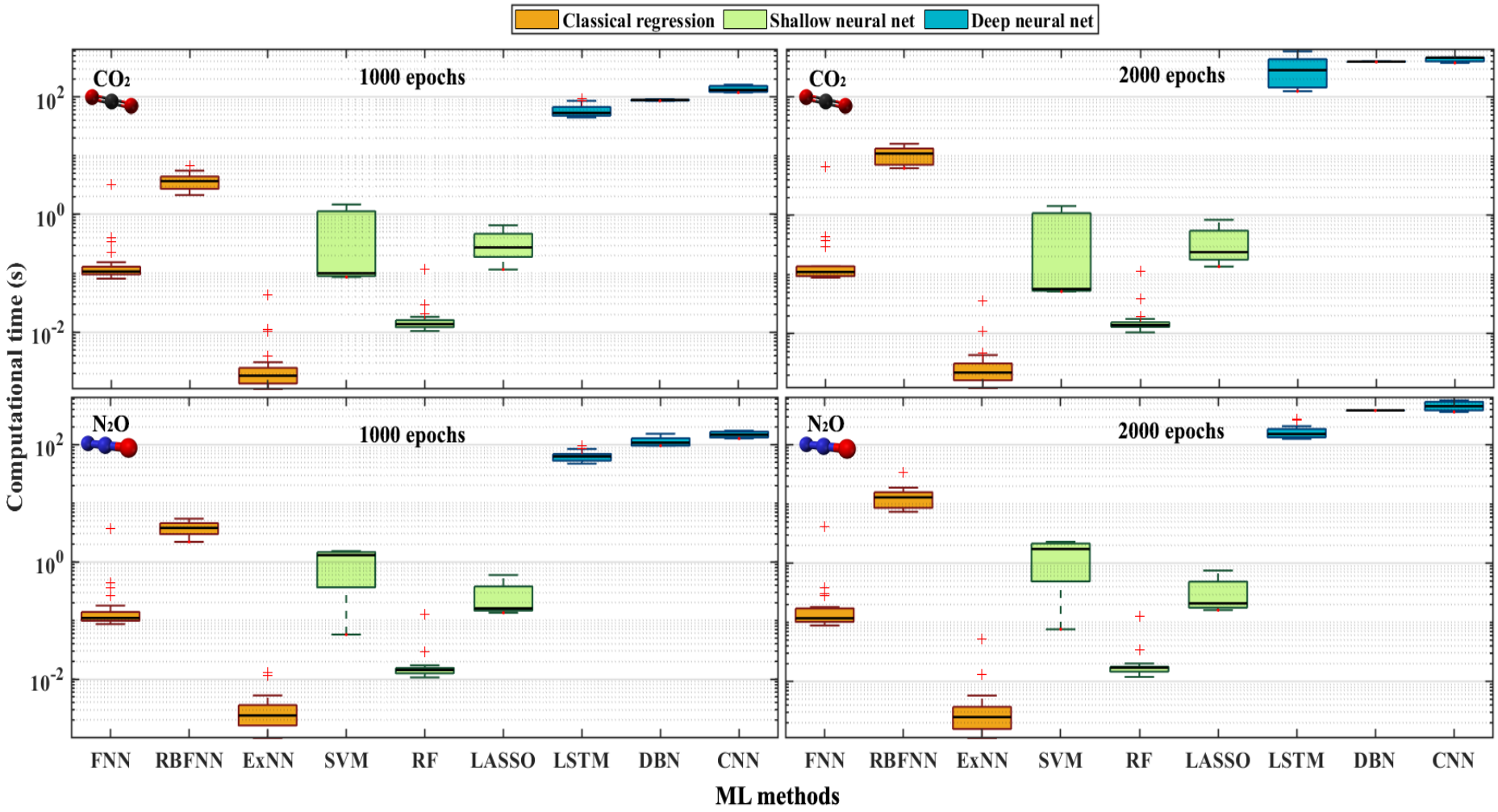

Figure 11. Comparative assessment of the ML models regarding their computation time (scaled on Log_{10}).

Computation time. The calculation time of the ML models during the training stage is presented in Figure 11. The numbers, in seconds, have been rounded up to the nearest integer. The computations were performed using a computer equipped with Intel core i5, CPU 2 GHz and 8 GB RAM. The computation time for training deep and shallow neural networks increases drastically by increasing the number of neurons and epochs. Among all ML models, DNNs were the most computationally time-intensive (at least 10 times higher than other ML models), due to the use of a multitude of hidden layers and complex neural cells, i.e., memory cell, recurrent cell, restricted Boltzmann machines, convolution and pooling. Despite its training time, LSTM remains the most efficient and the less computationally time-intensive DNN for the prediction of GHG emissions. ExNN and RF were the most computationally rapid models, achieving computation time of the order of milliseconds and less than 0.1 second, respectively. This performance in time execution, is due to the lighter least square optimization process employed in ExNN, resulting in an optimized number of neurons in the hidden layer. However, its prediction accuracy remains weak compared to other ML models. On the other hand, the RF model with its simple yet effective decision tree concept is a good choice for CO_2 predictions, based on both the prediction accuracy (with R=0.75) and computational time. The slowest ML models are: CNN for deep ML (caused by the processing time consumed during the convolution and pooling tasks), RBFNN for shallow ML (the result of the building of RBF interpolation during the training phase) and SVM for classical regression, CPU time remains acceptable (less than 5 seconds) and is due to the choice of the kernel (linear is the fastest) and the optimization problem involved as a quadratic program.

Nevertheless, the advantage of these ML approaches lies in the fact that the "expensive" computational cost is only required in the training stage. Trained ML models; classical and shallow as well as deep learning, could make these predictions within a few fractions of a

second.

CONCLUSIONS

In this study, various ML models, ranging from classical regression to shallow and deep learning were investigated for predicting GHG emissions from agricultural soils. An extensive comparison of the models was performed by evaluating and comparing several statistical metrics. Input data, i.e., air and soil temperatures, soil volumetric water content, precipitation, wind speed, humidity, surface pressure and crop N uptake were used in the ML models. The key findings of this study can be summarized as follows: (1) The LSTM model with recurrent networks containing memory cells as building blocks is the best suited for predicting future CO_2 and N_2O fluxes, and if necessary for filling gaps in data. The comparison of LSTM results with those of the Root Zone Water Quality Model (RZWQM2) showed that LSTM gave higher predictive performance. However, like most DL models, the training of LSTM remains computationally time intensive; (2) CNN and DBN are better suited to predict peak N_2O fluxes; (3) Shallow ML models are the most sensitive to tuning the hyperparameters and the optimal choice of these parameters (e.g. number of hidden units, epochs, choice of activation functions, etc.) is very crucial for accurate predictions; (4) The computation time of the shallow and deep ML models is strongly linked to the number of neurons and epochs; (5) The RF model could be considered as a fast, effective ML model and a good choice for predicting cyclical and seasonal trends of CO_2 fluxes; and (6) Optimization of the ML configuration is essential for reliable predictions with a reasonable computational time.

In future studies, further developments can be conducted to improve the predictive accuracy by combining data-driven ML and biophysical-based approaches (e.g. developing hybrid models) and optimizing their configuration and parameters in an ensemble learning approach.

Acknowledgments

This study was financially supported by the Agriculture and Agri-Food Canada Greenhouse Gases Program (AGGP2-020). We are grateful to Mr. Guy Vincent and his family, for the use of their farm to conduct the GHG studies. We also wish to thank the many undergraduate and graduate students associated with the McGill Water Innovation Lab, who helped to collect the field data.

References

Adeyemi, O., Grove, I., Peets, S., Domun, Y., Norton, T., 2018. Dynamic Neural Network Modelling of Soil Moisture Content for Predictive Irrigation Scheduling. Sensors 18, 3408. https://doi.org/10.3390/s18103408

Aneja, V.P., Schlesinger, W.H., Erisman, J.W., 2009. Effects of agriculture upon the air quality and climate: Research, policy, and regulations. Environ. Sci. Technol. 43, 4234–4240. https://doi.org/10.1021/es8024403

Bengio, Y., 2000. Gradient-based optimization of hyperparameters. Neural Comput. 12, 1889–1900. https://doi.org/10.1162/089976600300015187

Berglund, Ö., Berglund, K., Klemedtsson, L., 2010. A lysimeter study on the effect of temperature on CO2 emission from cultivated peat soils. Geoderma 154, 211–218. https://doi.org/10.1016/j.geoderma.2008.09.007

Bergstra, J., Ca, J.B., Ca, Y.B., 2012. Random Search for Hyper-Parameter Optimization Yoshua Bengio, Journal of Machine Learning Research.

Breiman L., 2001. Machine Learning, 45(1), 5–32. Stat. Dep. Univ. California, Berkeley, CA

94720. https://doi.org/10.1023/A:1010933404324

Chang, C.C., Lin, C.J., 2011. LIBSVM: A Library for support vector machines. ACM Trans. Intell. Syst. Technol. 2. https://doi.org/10.1145/1961189.1961199

Chlingaryan, A., Sukkarieh, S., Whelan, B., 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Comput. Electron. Agric. 151, 61–69. https://doi.org/10.1016/J.COMPAG.2018.05.012

Cortes, C., Vapnik, V., Saitta, L., 1995. Support-Vector Networks Editor, Machine Leaming. Kluwer Academic Publishers. https://doi.org/https://doi.org/10.1007/BF00994018

Crézé, C.M., Madramootoo, C.A., 2019. Water table management and fertilizer application impacts on CO2, N2O and CH4 fluxes in a corn agro-ecosystem. Sci. Rep. 9, 2692. https://doi.org/10.1038/s41598-019-39046-z

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., van Aardenne, J.A., Monni, S., Doering, U., Olivier, J.G.J., Pagliari, V., Janssens-Maenhout, G., 2018. Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2. Earth Syst. Sci. Data 10, 1987–2013. https://doi.org/10.5194/essd-10-1987-2018

Daw, N.W., 2009. The foundations of development and deprivation in the visual system. J. Physiol. 587, 2769–73. https://doi.org/10.1113/jphysiol.2009.170001

Dou, J., Yunus, A.P., Merghadi, A., Shirzadi, A., Nguyen, H., Hussain, Y., Avtar, R., Chen, Y.L., Pham, B.T., Yamagishi, H., 2020. Different sampling strategies for predicting landslide susceptibilities are deemed less consequential with deep learning. Sci. Total Environ. 137320. https://doi.org/10.1016/j.scitotenv.2020.137320

Dou, J., Yunus, A.P., Tien Bui, D., Merghadi, A., Sahana, M., Zhu, Z., Chen, C.W., Khosravi, K., Yang, Y., Pham, B.T., 2019. Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. Sci. Total Environ. 662, 332–346. https://doi.org/10.1016/j.scitotenv.2019.01.221

Drury, C.F., Reynolds, W.D., Tan, C.S., Mclaughlin, N.B., Yang, X.M., Calder, W., Oloya, T.O., Yang, J.Y., 2014. Impacts of 49-51 years of fertilization and crop rotation on growing season nitrous oxide emissions, nitrogen uptake and corn yields. Can. J. Soil Sci. 94, 421–433. https://doi.org/10.4141/CJSS2013-101

Duan, K., Keerthi, S.S., Poo, A.N., 2003. Evaluation of simple performance measures for tuning SVM hyperparameters. Neurocomputing 51, 41–59. https://doi.org/10.1016/S0925-2312(02)00601-X

Edwards, K.P., Madramootoo, C.A., Whalen, J.K., Adamchuk, V.I., Mat Su, A.S., Benslim, H., 2018. Nitrous oxide and carbon dioxide emissions from surface and subsurface drip irrigated tomato fields. Can. J. Soil Sci. 98, 389–398. https://doi.org/10.1139/cjss-2017-0001

Fang, K., Shen, C., Kifer, D., Yang, X., 2017. Prolongation of SMAP to Spatiotemporally Seamless Coverage of Continental U.S. Using a Deep Learning Neural Network. Geophys. Res. Lett. 44, 11,030-11,039. https://doi.org/10.1002/2017GL075619

Fischer, T., Krauss, C., 2018. Deep learning with long short-term memory networks for financial market predictions. Eur. J. Oper. Res. 270, 654–669. https://doi.org/10.1016/J.EJOR.2017.11.054

Ganzenmüller, R., Pradhan, P., Kropp, J.P., 2019. Sectoral performance analysis of national greenhouse gas emission inventories by means of neural networks. Sci. Total Environ. 656, 80–89. https://doi.org/10.1016/j.scitotenv.2018.11.311

Gilhespy, S.L., Anthony, S., Cardenas, L., Chadwick, D., del Prado, A., Li, C., Misselbrook, T., Rees, R.M., Salas, W., Sanz-Cobena, A., Smith, P., Tilston, E.L., Topp, C.F.E., Vetter, S., Yeluripati, J.B., 2014. First 20 years of DNDC (DeNitrification DeComposition): Model evolution. Ecol. Modell. 292, 51–62. https://doi.org/10.1016/J.ECOLMODEL.2014.09.004

González Costa, J.J., Reigosa, M.J., Matías, J.M., Covelo, E.F., 2017. Soil Cd, Cr, Cu, Ni, Pb and Zn sorption and retention models using SVM: Variable selection and competitive model. Sci. Total Environ. 593–594, 508–522. https://doi.org/10.1016/j.scitotenv.2017.03.195

Grinblat, G.L., Uzal, L.C., Larese, M.G., Granitto, P.M., 2016. Deep learning for plant identification using vein morphological patterns. Comput. Electron. Agric. 127, 418–424. https://doi.org/10.1016/J.COMPAG.2016.07.003

Heard, J., Hay, D., 2006. Typical nutrient content, uptake pattern and carbon: nitrogen ratios of prairie crops, in: Designing Cropping Systems That Prosper in Variable Weather: Proceedings of the 7th Manitoba Agronomists Conference, Winnipeg. Manitoba.

Hinton, G.E., Osindero, S., Teh, Y.-W., 2006. A Fast Learning Algorithm for Deep Belief Nets. Neural Comput. 18, 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527

Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. Science 313, 504–7. https://doi.org/10.1126/science.1127647

Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. Neural Comput. 9, 1735–

1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hong, Y., Shen, R., Cheng, H., Chen, Y., Zhang, Y., Liu, Yaolin, Zhou, M., Yu, L., Liu, Yi, Liu, Yanfang, 2019. Estimating lead and zinc concentrations in peri-urban agricultural soils through reflectance spectroscopy: Effects of fractional-order derivative and random forest. Sci. Total Environ. 651, 1969–1982. https://doi.org/10.1016/j.scitotenv.2018.09.391

Hu, X., Belle, J.H., Meng, X., Wildani, A., Waller, L.A., Strickland, M.J., Liu, Y., 2017. Estimating PM 2.5 Concentrations in the Conterminous United States Using the Random Forest Approach. Environ. Sci. Technol. 51, 6936–6944. https://doi.org/10.1021/acs.est.7b01210

Huang, G.-B., Chen, L., 2007. Convex incremental extreme learning machine. Neurocomputing 70, 3056–3062. https://doi.org/10.1016/J.NEUCOM.2007.02.009

Huang, G., Huang, G.-B., Song, S., You, K., 2015. Trends in extreme learning machines: A review. Neural Networks 61, 32–48. https://doi.org/10.1016/J.NEUNET.2014.10.001

Hubel, D.H., Wiesel, T.N., 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. 160, 106–54. https://doi.org/https://doi.org/10.1113/jphysiol.1962.sp006837

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J.G.J., Peters, J.A.H.W., van Aardenne, J.A., Monni, S., Doering, U., Petrescu, A.M.R., 2017. EDGAR v4.3.2 Global Atlas of the three major Greenhouse Gas Emissions for the period 1970–2012. Earth Syst. Sci. Data Discuss. 1–55. https://doi.org/10.5194/essd-2017-79

Jiang, He, W., Zhou, W., Hou, Y., Yang, J.Y., He, P., 2019a. Exploring management strategies

to improve maize yield and nitrogen use efficiency in northeast China using the DNDC and DSSAT models. Comput. Electron. Agric. 166, 104988. https://doi.org/10.1016/J.COMPAG.2019.104988

Jiang, Q., Qi, Z., Madramootoo, C.A., Smith, W., Abbasi, N.A., Zhang, T.-Q., 2018. Comparison of RZWQM2 and DNDC model in simulating greenhouse gas emission, crop yield and subsurface drainage, in: 2018 Detroit, Michigan July 29 - August 1, 2018. American Society of Agricultural and Biological Engineers, St. Joseph, MI, p. 1. https://doi.org/10.13031/aim.201801588

Jiang, Qi, Z., Madramootoo, C.A., Crézé, C., 2019b. Mitigating greenhouse gas emissions in subsurface-drained field using RZWQM2. Sci. Total Environ. 646, 377–389. https://doi.org/10.1016/J.SCITOTENV.2018.07.285

Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. Comput. Electron. Agric. 147, 70–90. https://doi.org/10.1016/J.COMPAG.2018.02.016

Keerthi, S.S., 2002. Efficient tuning of SVM hyperparameters using radius/margin bound and iterative algorithms. IEEE Trans. Neural Networks 13, 1225–1229. https://doi.org/10.1109/TNN.2002.1031955

Khaledian, Y., Miller, B.A., 2020. Selecting appropriate machine learning methods for digital soil mapping. Appl. Math. Model. 81, 401–418. https://doi.org/10.1016/j.apm.2019.12.016

Kubat, M., 1998. Decision trees can initialize radial-basis function networks. IEEE Trans. Neural Networks 9, 813–821. https://doi.org/10.1109/72.712154

Kulkarni, V.Y., Sinha, P.K., 2012. Pruning of random forest classifiers: A survey and future directions, in: Proceedings - 2012 International Conference on Data Science and

Engineering, ICDSE 2012. pp. 64–68. https://doi.org/10.1109/ICDSE.2012.6282329

Kuremoto, T., Kimura, S., Kobayashi, K., Obayashi, M., 2014. Time series forecasting using a deep belief network with restricted Boltzmann machines. Neurocomputing 137, 47–56. https://doi.org/10.1016/j.neucom.2013.03.047

Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A., 2017. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. IEEE Geosci. Remote Sens. Lett. 14, 778–782. https://doi.org/10.1109/LGRS.2017.2681128

Längkvist, M., Karlsson, L., Loutfi, A., 2014. A review of unsupervised feature learning and deep learning for time-series modeling. Pattern Recognit. Lett. 42, 11–24. https://doi.org/10.1016/j.patrec.2014.01.008

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

Li, X., Xiao, J., He, B., Altaf Arain, M., Beringer, J., Desai, A.R., Emmel, C., Hollinger, D.Y., Krasnova, A., Mammarella, I., Noe, S.M., Ortiz, P.S., Rey-Sanchez, A.C., Rocha, A. V., Varlagin, A., 2018. Solar-induced chlorophyll fluorescence is strongly correlated with terrestrial photosynthesis for a wide variety of biomes: First global analysis based on OCO-2 and flux tower observations. Glob. Chang. Biol. 24, 3990–4008. https://doi.org/10.1111/gcb.14297

Linquist, B., Van Groenigen, K.J., Adviento-Borbe, M.A., Pittelkow, C., Van Kessel, C., 2012. An agronomic assessment of greenhouse gas emissions from major cereal crops. Glob. Chang. Biol. 18, 194–209. https://doi.org/10.1111/j.1365-2486.2011.02502.x

Liu, D., Mishra, A.K., Yu, Z., Yang, C., Konapala, G., Vu, T., 2017. Performance of SMAP,

AMSR-E and LAI for weekly agricultural drought forecasting over continental United States. J. Hydrol. 553, 88–104. https://doi.org/10.1016/J.JHYDROL.2017.07.049

Liu, Y., Zhang, Q., Song, L., Chen, Y., 2019. Attention-based recurrent neural networks for accurate short-term and long-term dissolved oxygen prediction. Comput. Electron. Agric. 165, 104964. https://doi.org/10.1016/j.compag.2019.104964

Lloyd, K., Madramootoo, C.A., Edwards, K.P., Grant, A., 2019. Greenhouse gas emissions from selected horticultural production systems in a cold temperate climate. Geoderma 349, 45–55. https://doi.org/10.1016/j.geoderma.2019.04.030

Luo, G., 2016. A review of automatic selection methods for machine learning algorithms and hyper-parameter values. Netw. Model. Anal. Heal. Informatics Bioinforma. 5, 1–16. https://doi.org/10.1007/s13721-016-0125-6

MacKay, D.J.C., 1992a. A Practical Bayesian Framework for Backpropagation Networks. Neural Comput. 4, 448–472. https://doi.org/10.1162/neco.1992.4.3.448

MacKay, D.J.C., 1992b. Bayesian Interpolation. Neural Comput. 4, 415–447. https://doi.org/10.1162/neco.1992.4.3.415

Maldonado, W., Barbosa, J.C., 2016. Automatic green fruit counting in orange trees using digital images. Comput. Electron. Agric. 127, 572–581. https://doi.org/10.1016/j.compag.2016.07.023

Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P.R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J.B.R., Chen, Y., Zhou, X., Gomis, M.I., Lonnoy, E., Maycock, T., Tignor, M., Waterfield, T., 2018. Global warming of 1.5°C. An IPCC Spec. Rep. impacts Glob. Warm. 1.

McSwiney, C.P., Robertson, G.P., 2005. Nonlinear response of N2O flux to incremental fertilizer addition in a continuous maize (Zea mays L.) cropping system. Glob. Chang. Biol. 11, 1712–1719. https://doi.org/10.1111/j.1365-2486.2005.01040.x

Møller, M.F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. Neural Networks 6, 525–533. https://doi.org/10.1016/S0893-6080(05)80056-5

Necpálová, M., Anex, R.P., Fienen, M.N., Del Grosso, S.J., Castellano, M.J., Sawyer, J.E., Iqbal, J., Pantoja, J.L., Barker, D.W., 2015. Understanding the DayCent model: Calibration, sensitivity, and identifiability through inverse modeling. Environ. Model. Softw. 66, 110–130. https://doi.org/10.1016/J.ENVSOFT.2014.12.011

Ng, W., Minasny, B., McBratney, A., 2020. Convolutional neural network for soil microplastic contamination screening using infrared spectroscopy. Sci. Total Environ. 702, 134723. https://doi.org/10.1016/j.scitotenv.2019.134723

Ni, X., Liao, S., Wu, F., Groffman, P.M., 2019. Short-term precipitation pulses stimulate soil CO 2 emission but do not alter CH 4 and N 2 O fluxes in a northern hardwood forest. Soil Biol. Biochem. 130, 8–11. https://doi.org/10.1016/j.soilbio.2018.11.021

Oertel, C., Matschullat, J., Zurba, K., Zimmermann, F., Erasmi, S., 2016. Greenhouse gas emissions from soils—A review. Chemie der Erde. https://doi.org/10.1016/j.chemer.2016.04.002

Pachauri, R., Allen, M., Barros, V., Broome, J., 2014. Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate.

Pham, B.T., Nguyen, M.D., Dao, D. Van, Prakash, I., Ly, H.B., Le, T.T., Ho, L.S., Nguyen,

K.T., Ngo, T.Q., Hoang, V., Soh, L.H., Ngo, H.T.T., Tran, H.T., Do, N.M., Van Le, H., Ho, H.L., Tien Bui, D., 2019. Development of artificial intelligence models for the prediction of Compression Coefficient of soil: An application of Monte Carlo sensitivity analysis. Sci. Total Environ. 679, 172–184. https://doi.org/10.1016/j.scitotenv.2019.05.061

Pradhan, P., Tingsanchali, T., Shrestha, S., 2020. Evaluation of Soil and Water Assessment Tool and Artificial Neural Network models for hydrologic simulation in different climatic regions of Asia. Sci. Total Environ. 701, 134308. https://doi.org/10.1016/j.scitotenv.2019.134308

Prasad, R., Deo, R.C., Li, Y., Maraseni, T., 2018. Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition. Geoderma 330, 136–161. https://doi.org/10.1016/j.geoderma.2018.05.035

Probst, P., Wright, M.N., Boulesteix, A.L., 2019. Hyperparameters and tuning strategies for random forest. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. https://doi.org/10.1002/widm.1301

Qi, Y., Li, Q., Karimian, H., Liu, D., 2019. A hybrid model for spatiotemporal forecasting of PM 2.5 based on graph convolutional neural network and long short-term memory. Sci. Total Environ. 664, 1–10. https://doi.org/10.1016/j.scitotenv.2019.01.333

Rahnemoonfar, M., Sheppard, C., 2017. Deep Count: Fruit Counting Based on Deep Simulated Learning. Sensors 17, 905. https://doi.org/10.3390/s17040905

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

Rodríguez-Galiano, V., Mendes, M.P., García-Soldado, M.J., Chica-Olmo, M., Ribeiro, L., 2014. Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain). Sci. Total Environ. 476–477, 189–206. https://doi.org/10.1016/j.scitotenv.2014.01.001

Rodríguez, J.D., Pérez, A., Lozano, J.A., 2010. Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation. IEEE Trans. Pattern Anal. Mach. Intell. 32, 569–575. https://doi.org/10.1109/TPAMI.2009.187

Rosenzweig, C., Horton, R.M., Bader, D.A., Brown, M.E., DeYoung, R., Dominguez, O., Fellows, M., Friedl, L., Graham, W., Hall, C., Higuchi, S., Iraci, L., Jedlovec, G., Kaye, J., Loewenstein, M., Mace, T., Milesi, C., Patzert, W., Stackhouse, P.W., Toufectis, K., 2014. Enhancing climate resilience at nasa centers: A collaboration between science and stewardship. Bull. Am. Meteorol. Soc. 95, 1351–1363. https://doi.org/10.1175/BAMS-D-12-00169.1

Ross R. Bender, J.W.H.M.L.R. and F.E.B., 2013. Modern Corn Hybrids' Nutrient Uptake Patterns. Better Crop. with Plant Food, Int. Plant Nutr. Inst. 97, 7–10.

Sainath, T.N., Vinyals, O., Senior, A., Sak, H., 2015. Convolutional, Long Short-Term Memory, fully connected Deep Neural Networks, in: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 4580–4584. https://doi.org/10.1109/ICASSP.2015.7178838

Scher, S., 2018. Toward Data- Driven Weather and Climate Forecasting: Approximating a Simple General Circulation Model With Deep Learning. Geophys. Res. Lett. 45, 12,616-12,622. https://doi.org/10.1029/2018GL080704

Schmidbauer, A., 2004. Effects of soil moisture and temperature on NO, NO_2, and N_2O emissions from European forest soils. J. Geophys. Res. 109, D17302. https://doi.org/10.1029/2004JD004590

Schmidt, A., Creason, W., Law, B.E., 2018. Estimating regional effects of climate change and altered land use on biosphere carbon fluxes using distributed time delay neural networks with Bayesian regularized learning. Neural Networks 108, 97–113. https://doi.org/10.1016/J.NEUNET.2018.08.004

Sharma, A., Vans, E., Shigemizu, D., Boroevich, K.A., Tsunoda, T., 2019. DeepInsight: A methodology to transform a non-image data to an image for convolution neural network architecture. Sci. Rep. 9, 1–7. https://doi.org/10.1038/s41598-019-47765-6

Shcherbak, I., Millar, N., Robertson, G.P., 2014. Global metaanalysis of the nonlinear response of soil nitrous oxide (N 2O) emissions to fertilizer nitrogen. Proc. Natl. Acad. Sci. U. S. A. 111, 9199–9204. https://doi.org/10.1073/pnas.1322434111

Sing, J.K., Basu, D.K., Nasipuri, M., Kundu, M., 2003. Improved k-means algorithm in the design of RBF neural networks, in: TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region. Allied Publishers Pvt. Ltd, pp. 841–845. https://doi.org/10.1109/TENCON.2003.1273297

Singh, A., Ganapathysubramanian, B., Singh, A.K., Sarkar, S., 2016. Machine Learning for High-Throughput Stress Phenotyping in Plants. Trends Plant Sci. 21, 110–124. https://doi.org/10.1016/J.TPLANTS.2015.10.015

Smith, W., Qi, Z., Grant, B., VanderZaag, A., Desjardins, R., 2019. Comparing hydrological frameworks for simulating crop biomass, water and nitrogen dynamics in a tile drained

soybean-corn system: Cascade vs computational approach. J. Hydrol. X 2, 100015. https://doi.org/10.1016/J.HYDROA.2018.100015

Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Stackhouse, P.W., Zhang, T., Westberg, D., Barnett, A.J., Bristow, T., Macpherson, B., Hoell, J.M., 2018. POWER Release 8.0.1 (with GIS Applications) Methodology (Data Parameters, Sources, & Validation).

Tibshirani, R., 2011. Regression shrinkage and selection via the lasso: a retrospective. J. R. Stat. Soc. Ser. B (Statistical Methodol. 73, 273–282. https://doi.org/10.1111/J.1467-9868.2011.00771.X@10.1111/(ISSN)1467-9868.TOP_SERIES_B_RESEARCH

Tibshirani, R., 1996. Regression Shrinkage and Selection Via the Lasso. J. R. Stat. Soc. Ser. B 58, 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

van Groenigen, J.W., Velthof, G.L., Oenema, O., Van Groenigen, K.J., Van Kessel, C., 2010. Towards an agronomic assessment of N2O emissions: A case study for arable crops. Eur. J. Soil Sci. 61, 903–913. https://doi.org/10.1111/j.1365-2389.2009.01217.x

van Kessel, C., Venterea, R., Six, J., Adviento-Borbe, M.A., Linquist, B., van Groenigen, K.J., 2013. Climate, duration, and N placement determine N2O emissions in reduced tillage systems: A meta-analysis. Glob. Chang. Biol. https://doi.org/10.1111/j.1365-2486.2012.02779.x

Walker, A.L., 2016. Machine learning in geosciences and remote sensing. Geosci. Front. 7, 3–10. https://doi.org/10.1016/J.GSF.2015.07.003

Wu, C., Chen, J.M., Desai, A.R., Lafleur, P.M., Verma, S.B., 2013. Positive impacts of precipitation intensity on monthly CO2 fluxes in North America. Glob. Planet. Change 100, 204–214. https://doi.org/10.1016/j.gloplacha.2012.10.019

Wu, Z., Zhou, Y., Wang, H., Jiang, Z., 2020. Depth prediction of urban flood under different rainfall return periods based on deep learning and data warehouse. Sci. Total Environ. 716, 137077. https://doi.org/10.1016/j.scitotenv.2020.137077

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

Yan, G., Mu, C., Xing, Y., Wang, Q., 2018. Responses and mechanisms of soil greenhouse gas fluxes to changes in precipitation intensity and duration: A meta-analysis for a global perspective. Can. J. Soil Sci. https://doi.org/10.1139/cjss-2018-0002

Yan, H., Zhang, J., Rahman, S.S., Zhou, N., Suo, Y., 2020. Predicting permeability changes with injecting CO2 in coal seams during CO2 geological sequestration: A comparative study among six SVM-based hybrid models. Sci. Total Environ. 705, 135941. https://doi.org/10.1016/j.scitotenv.2019.135941

Yang, M., Mou, Y., Meng, Y., Liu, S., Peng, C., Zhou, X., 2020. Modeling the effects of precipitation and temperature patterns on agricultural drought in China from 1949 to 2015. Sci. Total Environ. 711, 135139. https://doi.org/10.1016/j.scitotenv.2019.135139

Yu, H., Xie, T., Paszczyñski, S., Wilamowski, B.M., 2011. Advantages of radial basis function networks for dynamic system design, in: IEEE Transactions on Industrial Electronics. pp.

5458–5459. https://doi.org/10.1109/TIE.2011.2164773

Yu, K., Hiscox, A., DeLaune, Ronald D., DeLaune, R.D., Reddy, K.R., Richardson, C.J., Megonigal, J.P., 2013. Greenhouse Gas Emission by Static Chamber and Eddy Flux Methods. https://doi.org/10.2136/sssabookser10.c22

Yue, Y., Li, J.H., Fan, L.F., Zhang, L.L., Zhao, P.F., Zhou, Q., Wang, N., Wang, Z.Y., Huang, L., Dong, X.H., 2020. Prediction of maize growth stages based on deep learning. Comput. Electron. Agric. 172, 105351. https://doi.org/10.1016/j.compag.2020.105351

Zhang, C., Ma, Y., 2012. Ensemble machine learning: methods and applications. Springer Science & Business Media.

Zhang, H., Wu, P., Yin, A., Yang, X., Zhang, M., Gao, C., 2017. Prediction of soil organic carbon in an intensively managed reclamation zone of eastern China: A comparison of multiple linear regressions and the random forest model. Sci. Total Environ. 592, 704–713. https://doi.org/10.1016/j.scitotenv.2017.02.146

Zhang, J., Zhu, Y., Zhang, X., Ye, M., Yang, J., 2018. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. J. Hydrol. 561, 918–929. https://doi.org/10.1016/j.jhydrol.2018.04.065

Zhang, Y., Niu, H., 2016. The development of the DNDC plant growth sub-model and the application of DNDC in agriculture: A review. Agric. Ecosyst. Environ. 230, 271–282. https://doi.org/10.1016/J.AGEE.2016.06.017

Declaration of competing interests

☒ 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.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Graphical abstract

Highlights

- Machine learning was used for predicting agricultural soil greenhouse gas emissions.
- The performance of nine alternative machine learning models were compared.
- Deep learning LSTM revealed the best performance in predicting both N_2O and CO_2 fluxes.
- Classical RF found to be a fast and effective data-driven model for predicting CO_2 fluxes.