Wetlands (2022) 42: 37
https://doi.org/10.1007/s13157-022-01558-2

COASTAL WETLANDS

Predicting greenhouse gas fluxes in coastal salt marshes using artificial neural networks

Mohammed T. Zaki[1] · **Omar I. Abdul-Aziz**[1]

Received: 28 November 2021 / Accepted: 15 April 2022 / Published online: 27 April 2022

Abstract

Prediction of wetland greenhouse gas (GHG) fluxes has been a challenging undertaking. Machine learning techniques such as the artificial neural network (ANN) has a strong potential to provide high quality predictions of the wetland GHG fluxes. We developed eight different ANN models and investigated their suitability to predict the major GHG fluxes (CO_2 and CH_4) in coastal salt marshes (dominated by *Spartina alterniflora*) of Waquoit Bay, Massachusetts, USA. Based on the dominant environmental drivers, the daytime net uptake fluxes of CO_2 were predicted as a function of photosynthetically active radiation, soil temperature (ST), and porewater salinity (SS). The net emission fluxes of CH_4 were predicted as a function of ST and SS. Our models with the radial basis function neural network (RBNN) provided the most accurate and least-biased predictions of the net CO_2 uptake (Nash-Sutcliffe Efficiency, NSE = 0.98) and CH_4 emission (NSE = 0.90-0.92). The linear layer neural network generated the least successful and most biased predictions of the GHG fluxes (NSE = 0.48-0.80). Other ANNs, including the commonly-used feed forward neural network (FFNN), provided less accurate and more biased predictions of the CO_2 (NSE = 0.86-0.97) and CH_4 (NSE = 0.73-0.89) fluxes than the RBNN. We, therefore, recommend using RBNN as the first choice and FFNN (or its variant) as the second choice for predicting the GHG fluxes in coastal salt marshes. Our findings and tools would help derive plausible scenarios and guidelines for restoration, monitoring, and maintenance of coastal salt marshes in the U.S. and beyond.

Keywords Artificial neural network (ANN) · Coastal wetlands · Greenhouse gases · Machine learning · Predictions · Salt marshes

Introduction

Coastal salt marshes are intertidal wetlands, which exhibit the highest rates of ecosystem carbon accumulations (Ouyang and Lee 2014; Theuerkauf et al. 2015). The net ecosystem exchange (NEE) of carbon dioxide (CO_2) and methane (CH_4) are the major greenhouse gas (GHG) fluxes in tidal salt marshes (Mozdzer and Megonigal 2013; Abdul-Aziz et al. 2018). The NEE of CO_2 is the difference between gross primary productivity and ecosystem respiration (Reichstein et al. 2005). The net uptake typically dominates the daytime CO_2 fluxes due to photosynthesis in the presence of sunlight and favorable temperature (Juszczak et al. 2012; Schäfer et al. 2014). During nighttime, net emission fluxes of CO_2 result from ecosystem respiration, which is mainly driven by soil temperature (Lloyd and Taylor 1994; Qi et al. 2002; Smith et al. 2019).

The NEE of CH_4 is often dominated by CH_4 emission as the outcome of soil microbial processes and gas transport (Conrad 1989). CH_4 is produced by methanogenic bacteria under anaerobic conditions, whereas CH_4 is oxidized by methanotrophic bacteria mostly under aerobic conditions. These processes of CH_4 production and oxidation are primarily controlled by soil temperature and moisture content (or water table) in freshwater wetlands (Walter and Heimann 2000; Nahlik and Mitsch 2011). However, subject to the regular cycles of tidal flooding, other soil characteristics such as porewater salinity can also impact the underlying processes of both CO_2 and CH_4 fluxes in coastal salt marshes (Poffenbarger et al. 2011; Abdul-Aziz et al. 2018).

Subject to the complex biogeochemical processes and interactions, it has been challenging to predict the GHG fluxes and carbon storage in coastal wetlands (Oikawa et al.

✉ Omar I. Abdul-Aziz
oiabdulaziz@mail.wvu.edu; omariaaziz@gmail.com

[1] West Virginia University, Po Box 6103, Morgantown, WV 26506-6103, USA

2017; Abdul-Aziz et al. 2018). However, accurate modeling and prediction tools are essential to guide the restoration, monitoring, and maintenance activities of tidal salt marshes. Previous research (e.g., Walter and Heimann 2000; Zhang et al. 2002; St-Hilaire et al. 2010) attempted to develop process-based models of the GHG fluxes mainly for inland freshwater wetlands. In contrast, Abdul-Aziz et al. (2018) determined the environmental controls of CO_2 and CH_4 fluxes in coastal salt marshes and developed emergent power law scaling models to acceptably predict the GHG fluxes. Their success inspires further developments into the empirical and data-driven predictive modeling of the GHG fluxes with the necessary mechanistic underpinnings.

Machine learning offers data-driven tools that could be utilized to develop models for accurate predictions of the GHG fluxes in coastal salt marshes. Relationships between the GHG fluxes and environmental drivers in salt marshes are highly non-linear (Moseman-Valtierra et al. 2016; Abdul-Aziz et al. 2018). Machine learning techniques such as the artificial neural network (ANN) has a strong potential to represent highly non-linear relations and provide high quality predictions of ecosystem carbon fluxes (Kordowski and Kuttler 2010; Jammet et al. 2017; Tramontana et al. 2020). However, ANN has considerable variants, representing different levels of accuracy and uncertainty. Therefore, evaluation of various ANNs is important for developing the most accurate and consistent models to predict the GHG fluxes in coastal salt marshes.

In general, ANNs are known as 'black-box' models that identify complex non-linear relationships between the predictors and response variable and reproduce the data of the response by flexibly adjusting assigned weights and biases for different predictors (Dengel et al. 2013). For the last several decades, scientists working with terrestrial carbon dynamics have been extensively utilizing ANN models to predict CO_2 fluxes based on the environmental drivers such as radiation, temperature, soil water content, vapor pressure deficit, and vegetation indices. The most commonly-used ANN for this purpose has been the feed forward neural network (FFNN). For example, many studies (e.g., Moffat et al. 2010; Knox et al. 2018; Schäfer et al. 2019) developed FFNN models, often involving multiple hidden layers in the network structure, to predict the CO_2 fluxes in forests, croplands, and wetlands from the environmental drivers.

Apart from the FFNN models, the radial basis function neural network (RBNN) and the generalized regression neural network (GRNN) models have also been employed to predict ecosystem CO_2 fluxes, especially for sites representing complex environmental settings. The RBNN model has mostly been utilized to predict CO_2 fluxes of urban or peri-urban ecosystems, involving the aerodynamic variables (e.g., wind speed, wind direction) alongside the typical environmental drivers as the key predictors (Schmidt et al. 2008; Kordowski and Kuttler 2010). Archibald et al. (2009) developed a GRNN model to predict CO_2 fluxes in an African savanna using climate variables, vegetation indices, and various water stress indicators. Zhu et al. (2013) employed GRNN to predict CO_2 fluxes in high latitude wetlands using climatic variables, water table depth, soil organic carbon, and porosity. CO_2 fluxes in rainfed croplands were predicted by Safa et al. (2021) with RBNN based on the amounts of irrigation and precipitation. Further, recent studies (e.g., Safa et al. 2019; Tramontana et al. 2020) employed a multilayer perceptron neural network (MLNN) to achieve strong predictions of CO_2 fluxes in terrestrial ecosystems based on environmental and vegetation indices, with a coefficient of determination of up to 0.95.

Overall, the ANN models, particularly FFNN and MLNN, have been employed to predict the GHG fluxes for various wetland ecosystems, mostly focusing on the northern peatlands (e.g., Dengel et al. 2013), temperate and boreal freshwater wetlands (e.g., Goodrich et al. 2015; Rey-Sanchez et al. 2018; Delwiche et al. 2021; Irvin et al. 2021), and agricultural peatlands (e.g., Hatala et al. 2012; Knox et al. 2015). However, the existing ANN-based studies did not focus much on predicting the GHG fluxes in coastal salt marshes.

The main objective of our study is to develop various types of ANN models and investigate their suitability to predict the major GHG fluxes (CO_2 and CH_4) in coastal salt marshes. Our hypothesis is that the GHG fluxes in coastal salt marshes are predominantly driven by a small set of environmental variables, which can be utilized to develop accurate predictive models. The dominant predictors of CO_2 and CH_4 fluxes were first identified by developing univariate (i.e., single-predictor) models with eight types of ANN. Optimal ANN models were then developed based on the dominant predictors through a forward selection modeling approach, and the most successful ANN models were identified to predict the major GHG fluxes in coastal salt marshes.

Materials and Methods

Study sites and dataset

The salt marshes of our study are located in Waquoit Bay, Cape Code, Massachusetts, USA, with opening into the North Atlantic Ocean (Fig. 1). The four salt marsh sites (Sage Lot Pond, Eel Pond, Great Pond, and Hamblin Pond) exhibited varying ranges of nitrogen (N) loading (~5 to 126 $kg \cdot ha^{-1} \cdot year^{-1}$), although no significant differences in the measured GHG fluxes were observed with these variations in N loading (Abdul-Aziz et al. 2018). The salt marsh sites were dominated by *Spartina*

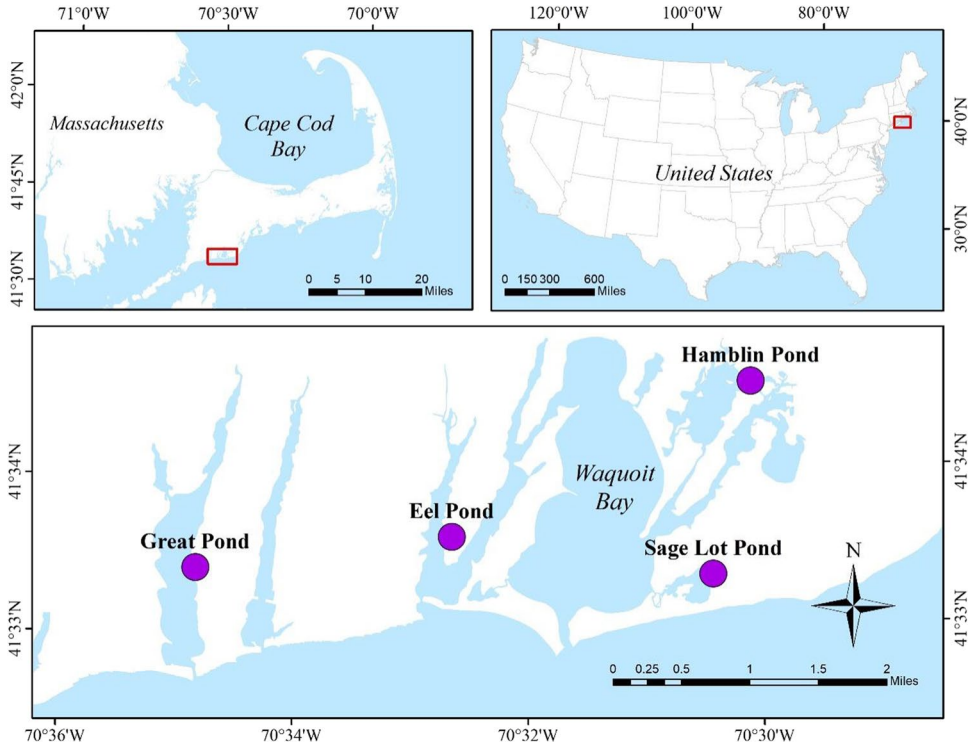

Fig. 1 Locations of the four salt marsh sites used for evaluation of the different types of ANN models to predict GHG fluxes

alterniflora, a C_4 halophyte plant, which is prevalent in majority of the salt marshes along the Atlantic coast of USA (Boyd et al. 2017).

The dataset used in this study was originally published by Abdul-Aziz et al. (2018). The data included chamber-based instantaneous measurements of the NEE of CO_2 and CH_4, as well as the corresponding environmental variables during different days of May-October in 2013 (Table S1 in the supplementary information). The environmental variables included photosynthetically active radiation (PAR), soil temperature (ST), porewater salinity (SS), pH, water level relative to the soil surface (WL), and soil moisture content (SM). The NEE of CO_2 represented the net uptake fluxes of CO_2 in daytime (measured during 8:00 a.m. to 4.30 p.m. Eastern Standard Time; sample size, n = 137) and NEE of CH_4 represented the day and nighttime net emission fluxes of CH_4 (n = 107). Negative and positive values of NEE indicated, respectively, the net uptake and emission fluxes. The different sample sizes between the datasets for the two GHG fluxes led to differences in the summary statistics (e.g., mean, standard deviation, minimum, maximum, and 25^{th} to 75^{th} percentiles) of the associated environmental variables (Table 1 and Table S1). Further details into the data collections and processing can be found in Abdul-Aziz et al. (2018). The complete dataset is available online at no-cost in the figshare data repository (Abdul-Aziz et al., 2021).

Framework for modeling and analysis

Eight types of conventional ANN models were developed to predict the GHG fluxes in the coastal salt marshes. We first employed the linear layer neural network (LLNN), which uses a linear function (Fig. 2a) and is arguably the simplest ANN model (Zhang et al. 2008). Then, we developed models with the feed forward neural network (FFNN), followed by its two variants: multi-layer perceptron neural network (MLNN) and cascade forward neural network (CFNN). The three models are unidirectional (i.e., involve no feedback loops) and utilize a sigmoid (e.g., hyperbolic tangent-sigmoid, and log-sigmoid) or piecewise-linear (e.g., rectified linear unit, and saturated linear) transfer function in hidden layer(s) for non-linear transformations of data (Fig. 2b-d; also see Wang et al. 2019; Stursa and Dolezel 2019). MLNN utilizes multiple hidden layers to compensate for the lost efficiency of predictions typically seen in FFNN (Stathakis 2009). CFNN uses weighted connections from the input to both hidden and linear layers to learn the underlying relationships quicker than FFNN (Saeedi et al. 2016). Then, we developed models with the layer-recurrent neural network (LRNN) and the non-linear autoregressive neural network with exogenous inputs (NARX). These models use a sigmoid or piecewise-linear function in their hidden layers, similar to the FFNN (Kun et al. 2018; Korprasertsak and Leephakpreeda 2019). However, both models further involve a single

Table 1: Summary of the predictive modeling data of the GHG fluxes and the environmental predictors collected during May-October of 2013 across the four tidal salt marshes of Waquoit Bay, Massachusetts, USA.

Dataset	Variables	Mean	Standard deviation	Minimum	25th percentile	50th percentile	75th percentile	Maximum
Modeling of daytime net CO_2 uptake (n = 137)	Net CO_2 uptake ($\mu molCO_2 \cdot m^{-2} \cdot s^{-1}$)	-5.33	4.72	-17.10	-9.27	-3.47	-1.04	-0.05
	PAR ($\mu mol \cdot m^{-2} \cdot s^{-1}$)	1395.53	519.73	303.70	1053.18	1514.95	1867.03	2093.08
	ST (°C)	17.57	4.15	8.89	14.79	17.22	20.59	26.10
	SS (ppt)	30.50	4.70	10.00	28.00	32.00	33.00	40.00
Modeling of day and nighttime net CH_4 emission (n = 107)	Net CH_4 emission ($nmolCH_4 \cdot m^{-2} \cdot s^{-1}$)	0.62	0.55	0.10	0.24	0.35	0.94	2.35
	ST (°C)	17.21	4.39	8.75	14.75	16.92	20.40	26.35
	SS (ppt)	31.22	3.97	20.00	29.00	32.00	34.00	40.00

*PAR, ST, and SS refer to photosynthetically active radiation, soil temperature, and porewater salinity. *n* represents sample sizes of the modeling datasets. The different sample sizes (*n*) between the datasets for the two GHG fluxes ($n = 137$ for CO_2 and 107 for CH_4) led to the differences in summary statistics (e.g., mean, standard deviation, minimum, maximum, and 25th to 75th percentiles) of the associated environmental variables.

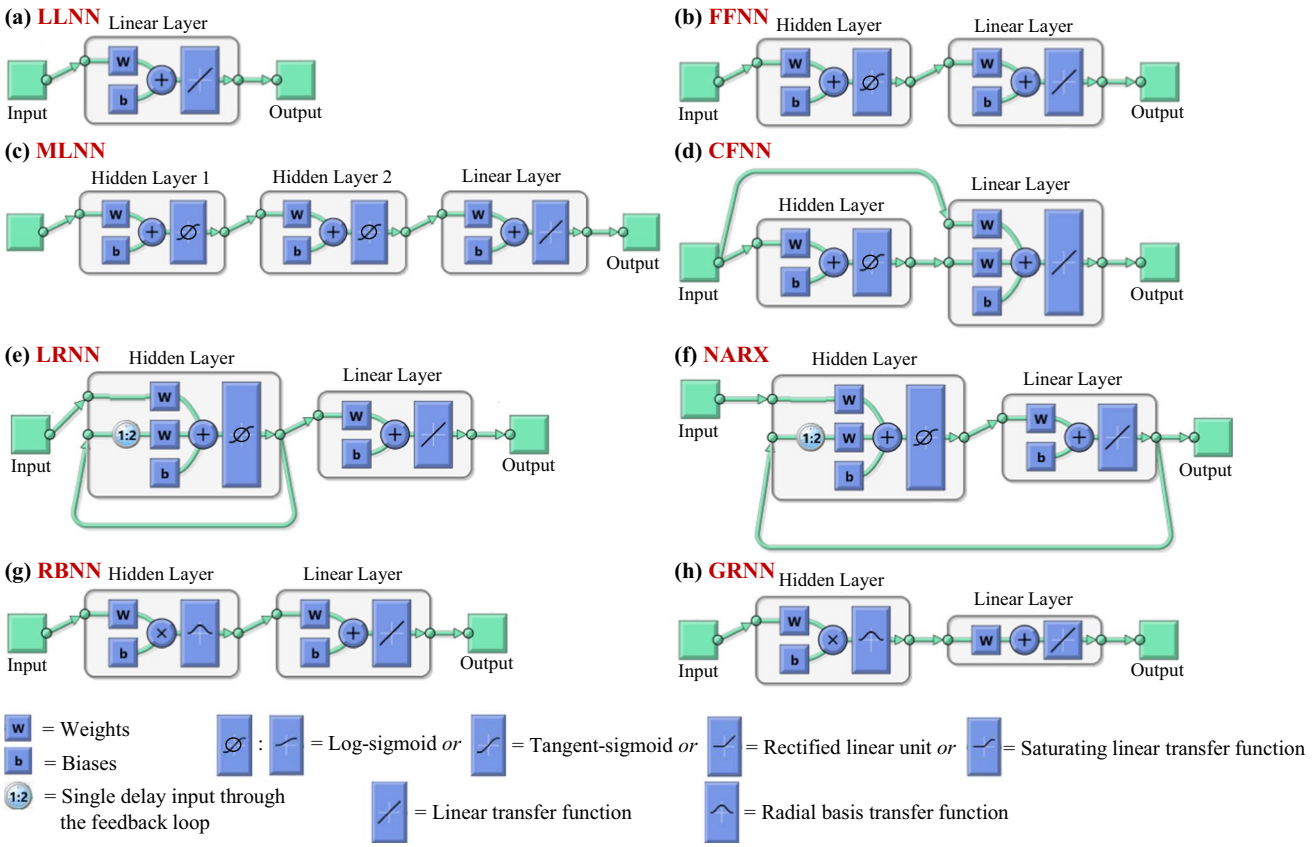

Fig. 2 Configurations of the eight ANN models developed to predict the GHG fluxes in coastal salt marshes based on MATLAB2020a. LLNN = linear layer neural network, FFNN = feed forward neural network, MLNN = multi-layer perceptron neural network, CFNN = cascade forward neural network, LRNN = layer-recurrent neural network, NARX = nonlinear autoregressive neural network with exogenous inputs, RBNN = radial basis function neural network, and GRNN = generalized regression neural network

delay feedback loop to also incorporate latent structures from the previous time-step and improve the network for the current time (Hussain et al. 2015). In LRNN, the feedback loop is designed to work within the non-linear hidden layer, whereas in NARX the feedback loop is implemented between the non-linear hidden layer and the linear output layer (Fig. 2e-f).

Finally, we employed the radial basis function neural network (RBNN) and its variant called the generalized regression neural network (GRNN). Instead of using a sigmoid or piecewise linear function for data transformations, RBNN and GRNN involves locally-tuned networks, which helps overcome the issue of obtaining a suboptimal solution (e.g., a local minima) in the learning process (Fig. 2g-h; also see Zirkohi et al. 2010). Each hidden neuron adjusts the weights of predictors using Euclidean distances and then transfers their summation to a radial basis transfer function within the hidden layer. Results of the hidden neurons for all ANNs using either the sigmoid or the piecewise linear or the radial basis function are transferred to the output layer, which involves a linear transfer function. However, contrary to the rest of the ANNs, the output layer in GRNN utilizes a linear transfer function without the bias term to facilitate a rapid training of the model (Fig. 2h; also see Ozyildirim and Avci 2013). Further details into the eight types of ANN with the necessary mathematical formulations are given in the supplementary information (Text S1).

We used MATLAB2020a for all coding, analysis, and visualization related to the model developments and evaluations. Built-in MATLAB functions were utilized to develop the eight ANN models (Text S1 in the supplementary information). Prior to model development, data for all variables were standardized via a Z-transformation (i.e., $Z = \left(X - \overline{X}\right)/S_X$, where X = original variable, $\overline{X}$ = mean of X, and S_X = standard deviation of X). This standardization of the data brings different units and magnitudes of variables to a comparable range and improves performance of the ANN models (Ladlani et al., 2012; Irvin et al. 2021). We estimated each ANN model 100 times by randomly resampling data (maintaining the original sample size in each iteration) for the GHG fluxes and the concurrent environmental variables using a boot-strap method. This helped achieve robust estimations of the ANN models and quantify the uncertainty associated with the predictions. For each iteration of model estimations, cross-validation was performed by randomly splitting the resampled dataset into the training, validation, and testing ratios of 80:10:10 (respectively), 70:15:15, 60:20:20, 50:25:25, 40:30:30, 30:35:35, and 20:40:40. These ratios represented the most desirable to the least desirable data-splits, given the sample sizes (n = 107 to 137) in the study. Our approach helped incorporations of prediction performance from the various combinations of data-splitting for model training, validation, and testing. Following previous studies (e.g., Kawamoto and Kabashima 2017; Willard et al., 2020), we reported the average prediction performance along with associated uncertainty from the 100 sets of cross-validations.

The quality of predictions from an ANN model may depend on the underlying training algorithm. We evaluated four algorithms to train the eight ANN models: Levenberg-Marquardt (LM) backpropagation, Bayesian regularization (BR) backpropagation, Broyden-Fletcher-Goldfarb-Shannon (BFGS) quasi-Newton backpropagation, and scaled conjugate gradient (SCG) backpropagation (see Text S2 in the supplementary information for details). During the learning procedure in the MATLAB platform, each algorithm iteratively updated the network weights and biases using the training dataset until (1) the best possible performance with the validation dataset was achieved or (2) algorithm's requirements to avoid getting stuck at local minima was met or (3) 1000 epochs were completed. The testing dataset was then used to independently evaluate the performance of the final ANN models.

An essential part of developing ANN models is optimizing the training dataset for network hyper-parameters. Important hyper-parameters for ANNs are typically the number of hidden neurons, selection of transfer function, learning rate, and the number of training epochs (Smithson et al. 2016; Jo et al. 2019). Neurons are the basic units of an ANN, whereas the transfer functions perform a non-linear transformation of the input data. The learning rate controls how slowly or quickly the network updates the weights and biases, and the training epochs refer to the number of network updates towards obtaining the final models. The FFNN, MLNN, CFNN, LRNN, and NARX models were optimized for 1 to 20 hidden neurons, four transfer functions (hyperbolic tangent-sigmoid, log-sigmoid, rectified linear unit, saturating linear function; see Text S1 in the supplementary information for details), and four learning rates (0.1, 0.01, 0.001, and 0.0001). We optimized the LLNN model only for the learning rates as it utilizes one hidden neuron and a linear transfer function to train the input data. For RBNN and GRNN models, the incorporated number of hidden neurons are the same as the length (i.e., sample size) of the input data (Sudheer and Jain 2003). However, the spread parameter, which represents the distribution density of the radial basis function in both models, has to be optimized (Benoudjit and Verleysen 2003). The RBNN and GRNN models were optimized for a spread of 1 to 20 and the four learning rates. We evaluated all possible combinations of the hyper-parameters following a grid search approach and selected the combination with the best cross-validation performance by minimizing the root-mean-square error (RMSE) as the

objective function and as a prediction error metric (Text S3 in the supplementary information; also see Chittaragi et al. 2019; Lorencin et al. 2021). The training epochs for all ANNs were optimized in the MATLAB environment through the convergence of training and validation errors (RMSE).

We initially developed univariate (single predictor) models to determine the strength of individual environmental variables (PAR, ST, SS, pH, WL, and SM) to predict the net fluxes of CO_2 uptake and CH_4 emission (Table S1 in the supplementary information). The prediction efficiency of the ANN models was assessed by computing the Nash-Sutcliffe efficiency (NSE), which can range from negative infinity ($-\infty$) to unity (1.0). The square of Pearson's correlation coefficient (r^2, which can range from 0 to 1.0) was also computed as an indicator of linear correspondence between the observed and predicted values of the GHG fluxes. The model prediction accuracy was measured by computing the RMSE and the ratio of RMSE to the standard deviation of observations (RSR). RMSE (range: 0 to ∞) presents the overall error in model predictions in actual units (e.g., $\mu molCO_2 \cdot m^2 \cdot s^{-1}$ and $nmolCH_4 \cdot m^2 \cdot s^{-1}$), which is easy to understand but can widely vary based on the magnitude or scale of the response variable (i.e., CO_2 or CH_4 fluxes). RSR (range: 0 to ∞) provides a normalized index of model accuracy, which is helpful for comparison of prediction errors across the various scales of the response variables.

The variables exhibiting the strongest predictions were used to obtain the optimum set of predictors for the corresponding GHG fluxes through a forward selection approach. The Akaike information criterion (AIC; Akaike 1974) was used to obtain the sets of optimum predictors for the CO_2 and CH_4 fluxes, alongside the prediction efficiency (NSE), correspondence (r^2), and accuracy (RMSE and RSR) metrics (Text S3 in the supplementary information). The optimal predictor sets were then used to predict the respective GHG fluxes by developing multivariate ANN models and utilizing all available data of the selected environmental predictors (Table 1).

The means of NSE, r^2, RMSE, and RSR obtained from the 100 resampled estimates were used to assess the overall prediction efficiency, correspondence, and accuracy of the ANN models. Similar to r^2, NSE = 1.0 refers to a perfectly predictive model. However, unlike r^2 (which cannot be negative), NSE < 0 (i.e., negative) suggests that the average of observations represents a better model than the proposed model (Nash and Sutcliffe 1970). In contrast, a perfect model has an RSR of 0; an RSR between 0 and 0.50 (and an NSE between 0.75 and 1.00) indicates a very good model, whereas a satisfactory model has an RSR between 0.50 and 0.70 (and an NSE between 0.50 and 0.75) (Moriasi et al. 2007). Although these criteria were originally developed for watershed hydrologic models and may appear stringent for ecological models, we still set the thresholds as reference metrics to achieve good quality predictions of the GHG fluxes in coastal salt marshes.

Results and Discussion

Dominant predictors of the GHG fluxes based on univariate ANN models

The univariate ANN models suggested PAR, ST, and SS as moderate to strong predictors (as indicated by NSE, r^2, RMSE, and RSR) of the GHG fluxes in the coastal salt marshes, whereas pH, WL, and SM were weak predictors (Fig. 3, 4; Fig. S1-S6 in the supplementary information). The predictive control of PAR on the net uptake fluxes of CO_2 mainly reflected the photosynthetic activities of the salt marsh plants during the daytime. However, the mechanistic link of PAR to CH_4 fluxes through primary productivity is confounding and unclear (Gomez-Casanovas et al. 2020). In fact, by using the dataset of the current study, Abdul-Aziz et al. (2018) found PAR to be a statistically insignificant predictor ($p\ value = 0.54$) of the CH_4 emission fluxes. They further demonstrated that the apparent control of PAR on the emission fluxes of CH_4 in coastal salt marshes had been spurious, representing mostly a surrogate effect of ST on methanogenesis. However, weakened relationship between ST and CH_4 emission in deep freshwater wetlands (Zhu et al. 2021) possibly might indicate PAR as a predictor that indirectly influences emission of CH_4 by triggering plant-mediated transport (Koebsch et al. 2015).

The strong linkage of ST on the net CO_2 uptake reiterated the strong influence of high temperature on the photosynthesis in coastal wetlands (Guo et al. 2009; Inglett et al. 2012). Specifically, temperature controls the activation process of the primary photosynthetic enzyme (RuBisCO) in the plants (C_4) of salt marshes (Sage and Kubien 2007). The link of SS indicated the adverse impacts of high salinity on the productivity of halophytic plants in salt marshes (Vasquez et al. 2006; Mateos-Naranjo et al. 2010; Pierfelice et al. 2017). High salinity leads to the production and accumulation of phytotoxic substances (e.g., hydrogen sulfide) in the anaerobic marsh sediments (Bradley and Morris 1990; Lamers et al. 2013). The phytotoxins impact leaf chlorophyll content, protein synthesis, and lipid metabolism of marsh plants —— leading to a reduced primary productivity (Mateos-Naranjo et al. 2010; Pierfelice et al. 2017).

The strong linkage of ST with the net CH_4 emission fluxes indicated temperature as the dominant driver of methanogenesis in the presence of adequate organic substrate in the wetland soil (Martin and Moseman-Valtierra 2017; Abdul-Aziz et al. 2018). ST drives the microbial activities involving both methanogenesis (CH_4 production) and methanotrophy (CH_4 oxidation). However, methanogenesis can be more

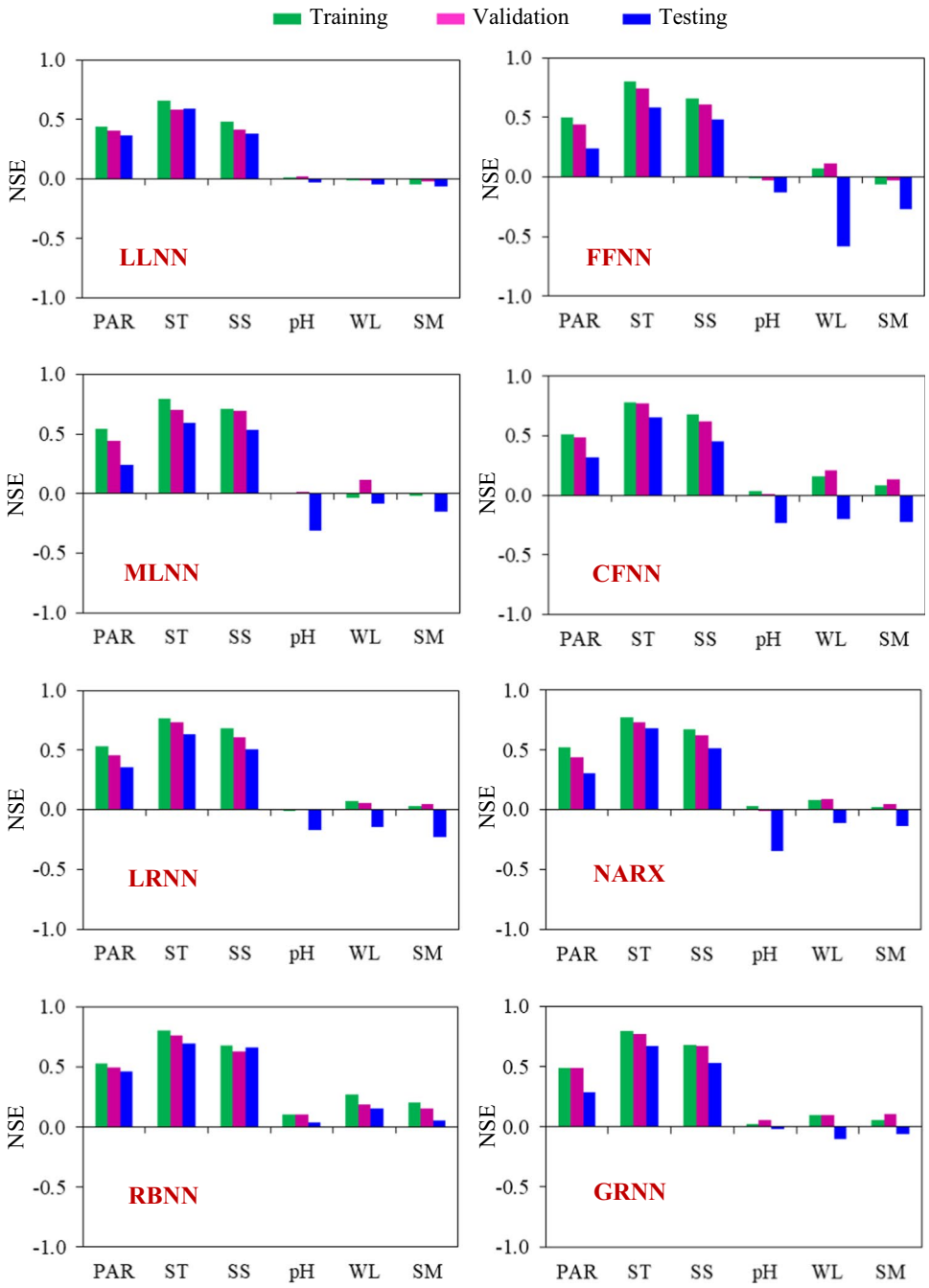

Fig. 3 Average Nash–Sutcliffe efficiency (NSE) of the univariate models of net CO_2 uptake fluxes for the eight ANN models. PAR, ST, SS, WL, and SM refer to photosynthetically active radiation, soil temperature, porewater salinity, water level with respect to soil surface, and soil moisture content.

sensitive to high ST than that of methanotrophy —— which ultimately results in increased emission of CH_4 to the atmosphere (Walter and Heimann 2000). The control of SS on the net CH_4 emission can be attributed to the highly sulfate-rich soil in salt marshes (Poffenbarger et al. 2011; Vivanco et al. 2015; Abdul-Aziz et al. 2018). Total anaerobic decomposition in sediments is typically dominated by sulfate reduction, which can hinder methanogenesis by outcompeting methanogens (Bartlett et al. 1987; Poffenbarger et al. 2011; Weston et al. 2014). Further, CH_4 can also be oxidized by sulfate-reducers (Bartlett et al. 1987; Segers 1998). Together, these processes limit CH_4 production (while enhancing oxidation) and reduce emission to the atmosphere in highly saline coastal salt marshes.

Previous studies (e.g., Wilson and Morris 2012; Abdul-Aziz et al. 2018) reported higher fluxes of both CO_2 uptake and CH_4 emission during high tides than low tides, subject to the higher flushing of salt accumulated in the marsh soil by the high tides. However, we found a weak relationship of WL with the GHG fluxes; this might have been caused by the inherent time-lag between the well (where water levels were measured) and tidal water levels. In contrast, the

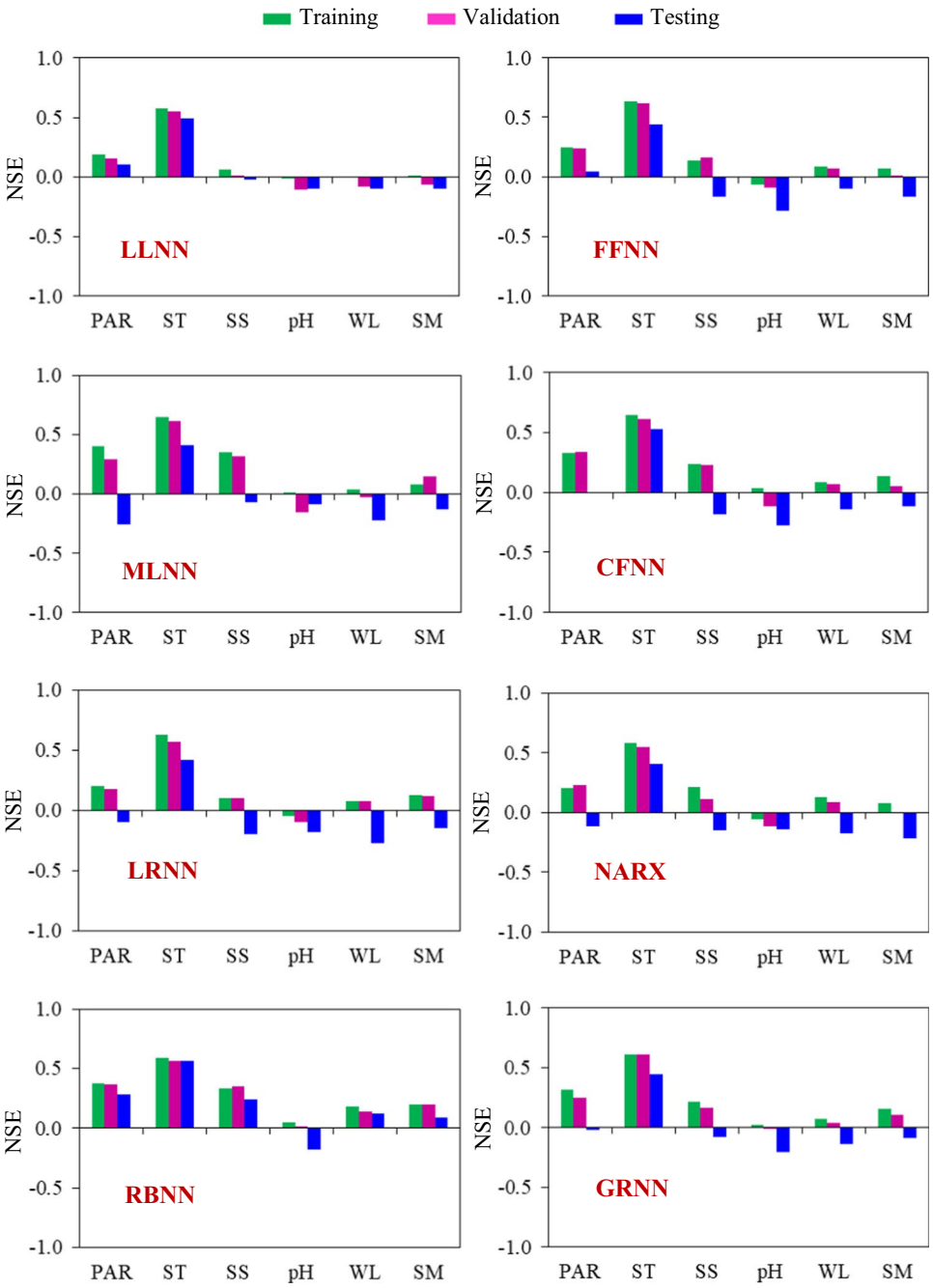

Fig. 4 Average Nash–Sutcliffe efficiency (NSE) of the univariate models of net CH_4 emission fluxes for the eight ANN models. PAR, ST, SS, WL, and SM refer to photosynthetically active radiation, soil temperature, porewater salinity, water level with respect to soil surface, and soil moisture content.

weak relation of SM with the GHG fluxes may be attributed to the predominant soil saturation in our study area, which represented low marshes (Table S1 in the supplementary information). Further, the near neutral pH of soil porewater in our salt marshes (mean ~7 and standard deviation ~ 0.30; Table S1 in the supplementary information) resulted in its weak relationships with the GHG fluxes. However, given that tidal hydrology contributes to the variation of soil temperature and porewater salinity (Wang et al. 2007), the controls of ST and SS might have reflected the overall effects of tidal hydrology on the GHG fluxes (Abdul-Aziz et al. 2018).

Optimal multivariate ANN models to predict the GHG fluxes

The AIC plots for model training suggested that the net uptake fluxes of CO_2 were optimally predicted by using up to three predictors (i.e., ST, SS, and PAR) for most ANNs, whereas two variables (ST and SS) represented the optimal predictor set for net CH_4 emission (Fig. S7 in the supplementary information). However, some ANNs (e.g., RBNN and GRNN) exhibited potential for including additional predictors to predict the GHG fluxes. This could be attributed

to the over-fitting tendency of the RBF-based ANNs considering that they use the same number of hidden neurons as the sample size of the input data (Skolthanarat et al. 2014). The NSE, r^2, RMSE, and RSR of the corresponding models provided complementary results to that for AIC (Fig. S8-S11 in the supplementary information).

Based on the identified dominant predictors of the GHG fluxes, we developed eight multivariate ANN models to predict the daytime net uptake fluxes of CO_2 as a function of PAR, ST, and SS. The net emission fluxes of CH_4 were predicted as a function of ST and SS. All models were optimized for network hyper-parameters to obtain the maximum accuracy in cross-validations (minimum RMSE) with the predicted GHG fluxes using the best performing training algorithm. Although the different training algorithms (LM, BR, BFGS, SCG) provided similar performance in predictions, the LM algorithm provided the most accurate predictions from the various ANN models across the training, validation, and testing phases (Table S2 and S3 in the supplementary information). In cross-validations with the LM algorithm, prediction performance metrics were nearly robust with various amounts of data-splitting among the model training, validation, and independent testing phases (Table S4 and S5 in the supplementary information).

The optimal models with FFNN, MLNN, CFNN, LRNN, and NARX mostly utilized the sigmoid transfer functions for the non-linear transformations of data (Table S6 in the supplementary information). The optimal FFNN and MLNN models did not require more than 10 hidden neurons, whereas CFNN, LRNN, and NARX typically used up to 15 hidden neurons. The value of spread parameter for the optimal models of RBNN were between 10 and 15, whereas the spread values ranged from 5 to 15 for GRNN. The different learning rates (0.1, 0.01, 0.001, and 0.0001) did not have much effect during optimization of the ANN models. Further, the optimal models for the different ANNs were typically achieved within 50 epochs of model estimations. The low standard deviations of the performance metrics (NSE, r^2, RSR, and RMSE) in predicting the GHG fluxes further indicated robust estimations of the final models for all eight ANNs (Table S7 in the supplementary information).

Among the multivariate ANN models of net CO_2 uptake, RBNN provided the most accurate and least biased predictions across training, validation, and testing (NSE = 0.98, r^2 = 0.99, RSR = 0.11 to 0.12, RMSE = 0.53 to 0.59 $\mu molCO_2 \cdot m^{-2} \cdot s^{-1}$) (Table 2 and Fig. 5). The FFNN, MLNN, CFNN, LRNN, NARX, and GRNN models provided very good predictions (NSE = 0.86 to 0.97, r^2 = 0.92 to 0.99, RSR = 0.14 to 0.33, RMSE = 0.67 to 1.55 $\mu molCO_2 \cdot m^{-2} \cdot s^{-1}$), although the associated uncertainties were quite high, specifically in model validation and testing. However, LLNN produced a biased model that yielded the least successful

Table 2: Average performance of the optimal ANN models in predicting the GHG fluxes in the coastal salt marshes.

Response	ANN	Training				Validation				Testing			
		NSE	r^2	RMSE†	RSR	*NSE*	*r^2*	*RMSE†*	*RSR*	NSE	r^2	RMSE†	RSR
Daytime net uptake fluxes of CO_2 *(n = 137)*	LLNN	0.80	0.90	2.08	0.44	*0.80*	*0.90*	*2.07*	*0.44*	0.79	0.90	2.10	0.45
	FFNN	0.92	0.96	1.25	0.26	*0.91*	*0.96*	*1.31*	*0.28*	0.89	0.95	1.50	0.32
	MLNN	0.89	0.95	1.45	0.31	*0.86*	*0.96*	*1.38*	*0.29*	0.88	0.92	1.55	0.33
	CFNN	0.95	0.97	1.03	0.22	*0.91*	*0.97*	*1.23*	*0.26*	0.90	0.95	1.49	0.32
	LRNN	0.94	0.97	1.08	0.23	*0.92*	*0.97*	*1.20*	*0.25*	0.90	0.95	1.45	0.31
	NARX	0.94	0.97	1.07	0.23	*0.92*	*0.96*	*1.26*	*0.27*	0.89	0.95	1.50	0.32
	RBNN	0.98	0.99	0.58	0.12	*0.98*	*0.99*	*0.53*	*0.11*	0.98	0.99	0.59	0.12
	GRNN	0.97	0.99	0.67	0.14	*0.93*	*0.97*	*1.15*	*0.24*	0.92	0.96	1.34	0.28
Day and nighttime net emission fluxes of CH_4 *(n = 107)*	LLNN	0.66	0.82	0.32	0.58	*0.59*	*0.83*	*0.32*	*0.58*	0.48	0.82	0.34	0.61
	FFNN	0.80	0.90	0.22	0.41	*0.78*	*0.93*	*0.21*	*0.39*	0.77	0.90	0.25	0.46
	MLNN	0.82	0.91	0.22	0.40	*0.77*	*0.92*	*0.22*	*0.40*	0.79	0.91	0.24	0.45
	CFNN	0.86	0.93	0.20	0.36	*0.78*	*0.93*	*0.21*	*0.39*	0.76	0.89	0.26	0.48
	LRNN	0.84	0.93	0.21	0.38	*0.82*	*0.93*	*0.21*	*0.38*	0.78	0.90	0.25	0.45
	NARX	0.86	0.93	0.20	0.36	*0.84*	*0.93*	*0.21*	*0.38*	0.73	0.88	0.27	0.49
	RBNN	0.91	0.96	0.15	0.28	*0.92*	*0.97*	*0.14*	*0.26*	0.90	0.96	0.16	0.29
	GRNN	0.89	0.95	0.17	0.31	*0.84*	*0.93*	*0.21*	*0.37*	0.77	0.90	0.26	0.47

*LLNN = linear layer neural network, FFNN = feed forward neural network, MLNN = multi-layer perceptron neural network, CFNN = cascade forward neural network, LRNN = layer-recurrent neural network, NARX = nonlinear autoregressive neural network with exogenous inputs, RBNN = radial basis function neural network, and GRNN = generalized regression neural network. *n* represents sample size of the modeling dataset. NSE = Nash–Sutcliffe efficiency, r^2 = Pearson's correlation coefficient, RMSE = root-mean-square error, and RSR = ratio of RMSE to the standard deviation of observations.

†units of RMSE is $\mu molCO_2 \cdot m^{-2} \cdot s^{-1}$ for CO_2 fluxes and $nmolCH_4 \cdot m^{-2} \cdot s^{-1}$ for CH_4 fluxes.

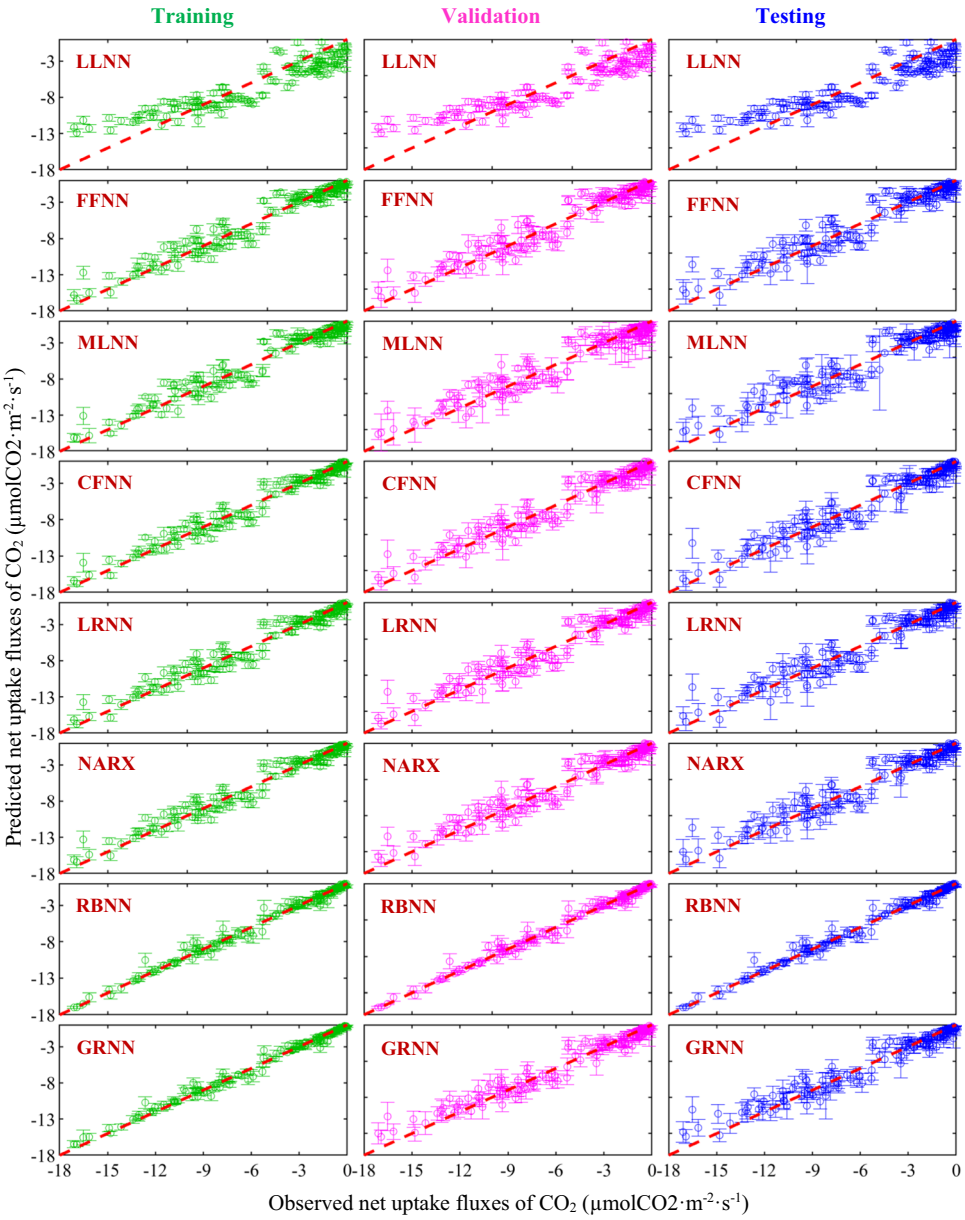

Fig. 5 Observed versus predicted net uptake fluxes of CO_2 from the eight ANN models for training, validation, and testing phases. Circles referred to the mean, and error bars denoted the standard deviations of the predicted net CO_2 fluxes. The red hashed line represented the 1:1 line.

predictions (NSE = 0.79 to 0.80, r^2 = 0.90, RSR = 0.44 to 0.45, RMSE = 2.07 to 2.10 $\mu molCO_2 \cdot m^2 \cdot s^{-1}$). It provided under-predictions of the high fluxes and over-predictions in the medium range (Fig. 5).

For the net emission fluxes of CH_4, the multivariate RBNN model provided the strongest predictions (NSE = 0.90 to 0.92, r^2 = 0.96 to 0.97, RSR = 0.26 to 0.29, RMSE = 0.14 to 0.16 $nmolCH_4 \cdot m^2 \cdot s^{-1}$) across the training, validation, and testing phases (Table 2 and Fig. 6). The LLNN model was biased and generated the least successful predictions (NSE = 0.48 to 0.66, r^2 = 0.82 to 0.83, RSR = 0.58 to 0.61, RMSE = 0.32 to 0.34 $nmolCH_4 \cdot m^2 \cdot s^{-1}$), with remarkable under-predictions in the higher fluxes of CH_4 emission. The remaining ANN models (FFNN, MLNN, CFNN, LRNN, NARX, and GRNN) performed stronger (NSE = 0.73 to 0.89, r^2 = 0.88 to 0.95, RSR = 0.31 to 0.49, RMSE = 0.17 to 0.27 $nmolCH_4 \cdot m^2 \cdot s^{-1}$) than LLNN, but weaker than RBNN. These models also under-predicted the higher fluxes of net CH_4 emission and had in general higher uncertainties in predictions (Fig. 6).

Insights into the varying prediction performance of different ANNs

To investigate the different quality of predictions from the various ANNs, we explored the relationships of the GHG fluxes with the environmental drivers through scatter-plots (Fig. S12 and S13 in the supplementary information). As apparent, the expected mechanistic trends of the fluxes with the predictors were associated with a remarkable noise. For

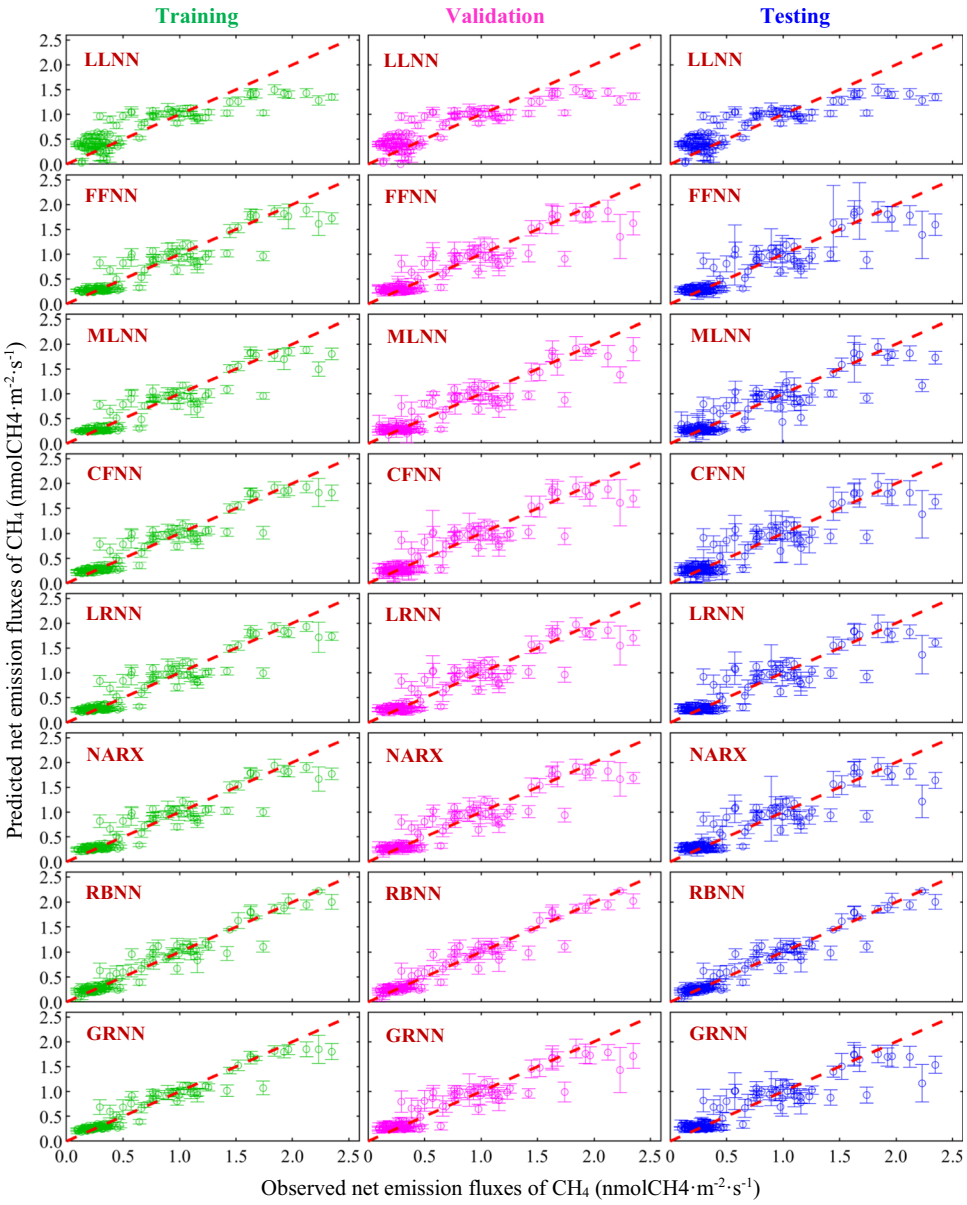

Fig. 6 Observed versus predicted net emission fluxes of CH_4 from the eight ANN models for training, validation, and testing phases. Circles referred to the mean, and error bars denoted the standard deviations of the predicted net CH_4 fluxes. The red hashed line represented the 1:1 line.

example, a seasonal hysteresis was apparent in the plot of PAR versus net CO_2 uptake —— with lower fluxes during September-October than June-August at the similar PAR values, which could be attributed to the lower temperatures in the fall than in the summer. Although no hysteresis was visible in the plots of net CO_2 uptake with ST, SS or other environmental drivers, the underlying signals (e.g., increasing uptake with increasing ST or decreasing uptake with increasing SS) were masked with much noise (Fig. S12 in the supplementary information).

We posit that the radial basis transfer function (RBF) in the hidden layers of RBNN, in concert with a bias term in the output layer (Fig. 2g), learnt signals from the noisy patterns in the data more efficiently and accurately than the remaining ANN models (Sudheer and Jain 2003; Xie et al. 2011). This resulted in the strongest predictions of the GHG fluxes by RBNN among all ANNs. Although RBNN and GRNN used the same network structure in the hidden layer (Fig. 2g-h), the lack of a bias term in the output layer of GRNN contributed to its weaker predictions and higher uncertainties. In contrast, the least successful and biased predictions of the net CO_2 uptake and CH_4 emission fluxes from LLNN suggested that the model with a simple linear network structure had not been successful to learn the non-linear and noisy relationships in the respective datasets (Fig. 2a).

Comparison with the previously developed emergent scaling models

The emergent power law scaling models of Abdul-Aziz et al. (2018) successfully predicted the salt marsh GHG fluxes using the datasets of the current study: net uptake fluxes of CO_2 (NSE = 0.90 to 0.91, RSR = 0.30 to 0.31) and net emission fluxes of CH_4 (NSE = 0.80 to 0.83, RSR = 0.42 to 0.49). The emergent scaling models had a simple structure and explicit mathematical formulations, reflecting the ubiquitous power law in nature (Enquist et al. 2003; Schwefel et al. 2017). A comparison of the overall performance (with average NSE and RSR across training, validation, and testing; see Table 2) of the ANN models in the current study with the emergent scaling models led to interesting observations. The emergent scaling models provided much better predictions than the LLNN models for both GHG fluxes. In contrast, the RBNN models provided more accurate predictions of the GHG fluxes than the emergent scaling models. However, based on the averaged metrics, the emergent scaling models predicted the GHG fluxes as well as that from the remaining ANN models (FFNN, MLNN, CFNN, LRNN, NARX, and GRNN). This is remarkable given that the scaling models were parametric in type with a simple explicit structure (emergent power law) and few parameters.

Limitations, challenges, and recommendations

The prevalence of noise and uncertainty in ecological and environmental measurements can often lead to challenges such as overfitting (specifically for small datasets) in the training phase of an ANN model (Choi and Park 2001). Over-training of ANNs can contribute to a decreased generalization, resulting in large inaccuracies of model predictions at variable time and space (Maher and Eyre 2011). We employed ANN algorithms (on MATLAB 2020a platform) that were equipped with a stopping criterion as soon as the optimal (i.e., most accurate) model was estimated in the training phase (see Materials and methods). This stopping criterion helped avoid a substantial over-training of our models for the net uptake fluxes of CO_2 and the emission fluxes of CH_4. However, an important recommendation from the current study would be to collect more data over longer periods in coastal salt marshes to develop more reliable ANN models.

In general, most ANN-based studies of ecosystem carbon fluxes developed models by employing FFNN with multiple hidden layers to enhance accuracy in predictions (e.g., Safa et al. 2019; Tramontana et al. 2020). In contrast, our final ANN models involved a single hidden layer —— except for MLNN that had two hidden layers by default and LLNN that did not accommodate any hidden layer (Fig. 2). Our RBNN and GRNN models used RBF for non-linear transformation and strictly required one hidden layer for a global convergence (Park and Sandberg 1991; Mahato and Paul 2019). However, we explored the performance of multivariate FFNN, CFNN, LRNN, and NARX models for the GHG fluxes in salt marshes by involving multiple hidden layers (up to five). No consistent improvement in predictions of the GHG fluxes was apparent across model training, validation, and testing with the increasing number of hidden layers (Fig. S14-S15 in the supplementary information). Further, the simulation time required on a Dell workstation (Processor: Intel Xeon E5 @2.40GHz, RAM: 32GB, and operating system: 64-bit Windows 10) increased in an exponential manner due to the inclusion of additional hidden layers in the model structure (Fig. S14-S15 in the supplementary information). Overall, the analysis indicated that one hidden layer was adequate to obtain an optimal prediction of the GHG fluxes in salt marshes from the FFNN, CFNN, LRNN, and NARX models.

In the backdrop of insufficient data and decreasing resources for environmental management (Wenger and Olden 2012; Yates et al. 2018), transferability of ecological models across various scales in space (e.g., regional, global), time (e.g., diurnal, seasonal, annual), and ecological units (e.g., ecosystem, species) is of paramount importance. Although some ANN models could be employed as powerful tools to obtain highly accurate predictions of ecological variables and indicators, it is quite difficult to interpret how different predictors are used through the hyper-parameterized structure of neural networks to simulate a response variable (Zhang et al. 2018). Being such a 'black-box' model, ANNs trained for a particular species, site or ecosystem are often not transferable to their counterparts. However, the lack of transferability is not unique with ANN-based modeling; it is also prevalent (although potentially at a reduced level) in process-based and simple empirical models. Future research should investigate ways to make the presented ANN models of GHG fluxes (particularly RBNN) generalizable using appropriately selected dominant predictors across salt marshes representing gradients in time, space, species, and associated processes. We further recommend evaluations of these ANNs in predicting the lateral carbon fluxes (e.g., dissolved organic and inorganic carbon) alongside the vertical fluxes (i.e., GHG fluxes) in the future research. This could help obtain an accurate and complete estimations of carbon budget in coastal salt marshes (Bogard et al. 2020).

Conclusions

We evaluated the hypothesis that the GHG fluxes in coastal salt marshes are predominantly driven by a small set of environmental variables, which can be utilized to develop accurate predictive models. The eight different ANN models

indicated sunlight (PAR), soil temperature (ST), and porewater salinity (SS) as the dominant predictors for the daytime fluxes of net CO_2 uptake in coastal salt marshes. ST and SS were the strongest predictors for the net emission fluxes of CH_4. Our overall finding is that the neural network-based machine learning models can be useful tools to accurately predict the major GHG fluxes in coastal salt marshes. However, transferability of these models should be evaluated with additional data from other salt marshes representing diverse hydroclimatic, vegetation, and salinity regimes. Another key finding is that use of a machine learning technique does not inherently guarantee a highly accurate prediction. The quality of predictions depends on the algorithms of individual ANNs to represent the nonlinear patterns in data. For example, our models with the radial basis function neural network (RBNN) provided the most accurate and least-biased predictions of the net CO_2 uptake (NSE = 0.98) and CH_4 emission (NSE = 0.90-0.92). In contrast, the linear layer neural network (LLNN) resulted in the least successful and most biased predictions of the GHG fluxes (NSE = 0.48-0.80). Other ANNs, including the feed forward neural network (FFNN), provided less accurate and more biased predictions of the CO_2 (NSE = 0.86-0.97) and CH_4 (NSE = 0.73-0.89) fluxes than RBNN. We, therefore, recommend using RBNN as the first choice and FFNN (or any of its variants) as the second choice for predicting the GHG fluxes in coastal salt marshes with a high accuracy and consistency. The models (except for LLNNs) would be useful tools to derive plausible scenarios and guidelines for restoration, monitoring, and maintenance of salt marshes (dominated by *Spartina alterniflora*) along the U.S. Atlantic coast and around the world.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13157-022-01558-2.

Acknowledgements This research was funded by a grant from the U.S. National Science Foundation (NSF) awarded to Abdul-Aziz (NSF CBET Environmental Sustainability Award No. 1705941). The datasets used in this study were collected through a project from the National Oceanic and Estuarine Administration (NOAA)'s National Estuarine Research Reserve Science Collaborative (NOAA Project No. NA09NOS4190153), awarded to Abdul-Aziz. All data were adequately described in the main text, figures, tables, and in the supporting information. The complete dataset is available in the figshare data repository at https://doi.org/10.6084/m9.figshare.15125148.v1.

Availability of data and material The dataset was deposited in figshare under the following reference: Abdul-Aziz, Omar I.; Tang, Jianwu; Moseman-Valtierra, Serena (2021): GHG flux dataset of Waquoit Bay, MA, USA saltmarshes (May-October 2013). figshare. Dataset. https://doi.org/10.6084/m9.figshare.15125148.v1.

Code availability Built-in functions of MATLAB2020a were used for all coding, analysis, and visualization related to the model developments and evaluations. The MATLAB functions used in this study are described in Text S1 in supplementary information.

Author contributions Abdul-Aziz conceptualized the research idea. Abdul-Aziz and Zaki designed the methodology, conducted the analyses, and summarized the results. Both authors contributed to the writing. Abdul-Aziz administered the projects funding the research and supervised Zaki. All authors have read and agreed to the published version of the manuscript.

Funding U.S. National Science Foundation (NSF) (NSF CBET Environmental Sustainability Award No. 1705941) and National Oceanic and Estuarine Administration (NOAA)'s National Estuarine Research Reserve Science Collaborative (NOAA Project No. NA09NOS4190153), awarded to Abdul-Aziz.

Declarations

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.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Abdul-Aziz OI, Ishtiaq KS, Tang J, Moseman-Valtierra S, Kroeger KD, Gonneea ME, Mora J, Morkeski K (2018) Environmental controls, emergent scaling, and predictions of greenhouse gas (GHG) fluxes in coastal salt marshes. J Geophys Res-Biogeo 123:2234–2256. https://doi.org/10.1029/2018JG004556

Abdul-Aziz OI, Tang J, Moseman-Valtierra S (2021): GHG flux dataset of Waquoit Bay, MA, USA saltmarshes (May-October 2013). figshare. Dataset. https://doi.org/10.6084/m9.figshare.15125148.v1

Akaike H (1974) A new look at the statistical model identification. IEEE T Automat Contr 19:716–723. https://doi.org/10.1109/TAC.1974.1100705

Archibald SA, Kirton A, Van der Merwe MR, Scholes RJ, Williams CA, Hanan N (2009) Drivers of inter-annual variability in Net Ecosystem Exchange in a semi-arid savanna ecosystem, South Africa. Biogeosciences 6:251–266. https://doi.org/10.5194/bg-6-251-2009

Bartlett KB, Bartlett DS, Harriss RC, Sebacher DI (1987) Methane emissions along a salt marsh salinity gradient. Biogeochemistry 4:183–202. https://doi.org/10.1007/BF02187365

Benoudjit N, Verleysen M (2003) On the kernel widths in radial-basis function networks. Neural Process Lett 18:139–154. https://doi.org/10.1023/A:1026289910256

Bogard MJ, Bergamaschi BA, Butman DE, Anderson F, Knox SH, Windham-Myers L (2020) Hydrologic export is a major component of coastal wetland carbon budgets. Global Biogeochem Cy 34:e2019GB006430. https://doi.org/10.1029/2019GB006430

Boyd BM, Sommerfield CK, Elsey-Quirk T (2017) Hydrogeomorphic influences on salt marsh sediment accumulation and accretion in two estuaries of the US Mid-Atlantic coast. Mar Geol 383:132–145. https://doi.org/10.1016/j.margeo.2016.11.008

Bradley PM, Morris JT (1990) Influence of oxygen and sulfide concentration on nitrogen uptake kinetics in *Spartina alterniflora*. Ecology 71:282–287. https://doi.org/10.2307/1940267

Chittaragi NB, Limaye A, Chandana NT, Annappa B, Koolagudi SG (2019) Automatic Text-Independent Kannada Dialect

Identification System. In: Satapathy S, Bhateja V, Somanah R, Yang XS, Senkerik R (eds) Adv Intell Syst. Springer, Singapore, pp 79–87. https://doi.org/10.1007/978-981-13-3338-5_8

Choi DJ, Park H (2001) A hybrid artificial neural network as a software sensor for optimal control of a wastewater treatment process. Water Res 35:3959–3967. https://doi.org/10.1016/S0043-1354(01)00134-8

Conrad R (1989) Control of methane production in terrestrial ecosystems. In: Andreae MO, Schimel DS (eds) Exchange of trace gases between terrestrial ecosystems and the atmosphere. Dahlem Workshop Reports, Wiley Chichester, UK, pp 39–58

Delwiche KB, Knox SH, Malhotra A, Fluet-Chouinard E, McNicol G, Feron S, Ouyang Z, Papale D, Trotta C, Canfora E, Cheah YW (2021) FLUXNET-CH_4: a global, multi-ecosystem dataset and analysis of methane seasonality from freshwater wetlands. Earth Syst Sci Data 13:3607–3689. https://doi.org/10.5194/essd-13-3607-2021

Dengel S, Zona D, Sachs T, Aurela M, Jammet M, Parmentier FJ, Oechel W, Vesala T (2013) Testing the applicability of neural networks as a gap-filling method using CH_4 flux data from high latitude wetlands. Biogeosciences 10:8185–8200. https://doi.org/10.5194/bg-10-8185-2013

Dunfield P, Dumont R, Moore TR (1993) Methane production and consumption in temperate and subarctic peat soils: response to temperature and pH. Soil Biol Biochem 25:321–326. https://doi.org/10.1016/0038-0717(93)90130-4

Enquist BJ, Economo EP, Huxman TE, Allen AP, Ignace DD, Gillooly JF (2003) Scaling metabolism from organisms to ecosystems. Nature 423:639–642. https://doi.org/10.1038/nature01671

Goodrich JP, Campbell DI, Roulet NT, Clearwater MJ, Schipper LA (2015) Overriding control of methane flux temporal variability by water table dynamics in a Southern Hemisphere, raised bog. J Geophys Res-Biogeo 120:819–831. https://doi.org/10.1002/2014JG002844

Gomez-Casanovas N, DeLucia NJ, DeLucia EH, Blanc-Betes E, Boughton EH, Sparks J, Bernacchi CJ (2020) Seasonal controls of CO_2 and CH_4 dynamics in a temporarily flooded subtropical wetland. J Geophys Res-Biogeo 125:e2019JG005257. https://doi.org/10.1029/2019JG005257

Guo H, Noormets A, Zhao B, Chen J, Sun G, Gu Y, Li B, Chen J (2009) Tidal effects on net ecosystem exchange of carbon in an estuarine wetland. Agr Forest Meteorol 149:1820–1828. https://doi.org/10.1016/j.agrformet.2009.06.010

Hatala JA, Detto M, Sonnentag O, Deverel SJ, Verfaillie J, Baldocchi DD (2012) Greenhouse gas (CO_2, CH_4, H_2O) fluxes from drained and flooded agricultural peatlands in the Sacramento-San Joaquin Delta. Agr Ecosyst Environ 150:1–18. https://doi.org/10.1016/j.agee.2012.01.009

Hussain AJ, Fergus P, Al-Jumeily D, Alaskar H, Radi N (2015) The Utilisation of Dynamic Neural Networks for Medical Data Classifications-Survey with Case Study. In: Huang DS, Han K (ed) Proceedings of the 2015 International Conference on Intelligent Computing, Springer, Fuzhou, China, pp 752-758. https://doi.org/10.1007/978-3-319-22053-6_80

Inglett KS, Inglett PW, Reddy KR, Osborne TZ (2012) Temperature sensitivity of greenhouse gas production in wetland soils of different vegetation. Biogeochemistry 108:77–90. https://doi.org/10.1007/s10533-011-9573-3

Irvin J, Zhou S, McNicol G, Lu F, Liu V, Fluet-Chouinard E, Ouyang Z, Knox SH, Lucas-Moffat A, Trotta C, Papale D (2021) Gap-filling eddy covariance methane fluxes: Comparison of machine learning model predictions and uncertainties at FLUXNET-CH_4 wetlands. Agr Forest Meteorol 308:108528. https://doi.org/10.1016/j.agrformet.2021.108528

Jo Y, Min K, Jung D, Sunwoo M, Han M (2019) Comparative study of the artificial neural network with three hyper-parameter optimization methods for the precise LP-EGR estimation using in-cylinder pressure in a turbocharged GDI engine. Appl Therm Eng 149:1324–1334. https://doi.org/10.1016/j.applthermaleng.2018.12.139

Juszczak R, Acosta M, Olejnik J (2012) Comparison of Daytime and Nighttime Ecosystem Respiration Measured by the Closed Chamber Technique on a Temperate Mire in Poland. Pol J Environ Stud 21:643–658

Kavaklioglu K, Ceylan H, Ozturk HK, Canyurt OE (2009) Modeling and prediction of Turkey's electricity consumption using artificial neural networks. Energ Convers Manage 50:2719–2727. https://doi.org/10.1016/j.enconman.2009.06.016

Kawamoto T, Kabashima Y (2017) Cross-validation estimate of the number of clusters in a network. Sci. Rep. 7:1–17. https://doi.org/10.1038/s41598-017-03623-x

Knox SH, Sturtevant C, Matthes JH, Koteen L, Verfaillie J, Baldocchi D (2015) Agricultural peatland restoration: effects of land-use change on greenhouse gas (CO_2 and CH_4) fluxes in the Sacramento-San Joaquin Delta. Glob Change Biol 21:750–765. https://doi.org/10.1111/gcb.12745

Knox SH, Windham-Myers L, Anderson F, Sturtevant C, Bergamaschi B (2018) Direct and indirect effects of tides on ecosystem-scale CO_2 exchange in a brackish tidal marsh in Northern California. J Geophys Res-Biogeo 123:787–806. https://doi.org/10.1002/2017JG004048

Koebsch F, Jurasinski G, Koch M, Hofmann J, Glatzel S (2015) Controls for multi-scale temporal variation in ecosystem methane exchange during the growing season of a permanently inundated fen. Agr Forest Meteorol 204:94–105. https://doi.org/10.1016/j.agrformet.2015.02.002

Kordowski K, Kuttler W (2010) Carbon dioxide fluxes over an urban park area. Atmos Environ 44:2722–2730. https://doi.org/10.1016/j.atmosenv.2010.04.039

Korprasertsak N, Leephakpreeda T (2019) Robust short-term prediction of wind power generation under uncertainty via statistical interpretation of multiple forecasting models. Energy 180:387–397. https://doi.org/10.1016/j.energy.2019.05.101

Kun C, Zhiwei M, Yuehua L, Zhinong J, Jinjie Z (2018) Lithium-ion battery state of charge estimation based on dynamic neural network and Kalman filter. In: 2018 IEEE International Conference on Prognostics and Health Management (ICPHM). IEEE, Seattle, Washington, USA, pp 1-6. https://doi.org/10.1109/ICPHM.2018.8448734

Ladlani I, Houichi L, Djemili L, Heddam S, Belouz K (2012) Modeling daily reference evapotranspiration (ET_0) in the north of Algeria using generalized regression neural networks (GRNN) and radial basis function neural networks (RBFNN): a comparative study. Meteorol Atmos Phys 118:163–178. https://doi.org/10.1007/s00703-012-0205-9

Lamers LP, Govers LL, Janssen IC, Geurts JJ, Van der Welle ME, Van Katwijk MM, Van der Heide T, Roelofs JG (2013) Sulfide as a soil phytotoxin—a review. Front Plant Sci 4:268. https://doi.org/10.3389/fpls.2013.00268

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

Lorencin I, Anđelić N, Šegota SB, Musulin J, Štifanić D, Mrzljak V, Španjol J, Car Z (2021) Edge detector-based hybrid artificial neural network models for urinary bladder cancer diagnosis. In: Hassanien AE, Taha MHN, Khalifa NEM (eds) Enabling AI Applications in Data Science. Springer, Cham, pp 225–245. https://doi.org/10.1007/978-3-030-52067-0_10

Mahato S, Paul S (2019) Detection of major depressive disorder using linear and non-linear features from EEG signals. Microsyst Technol 25:1065–1076. https://doi.org/10.1007/s00542-018-4075-z

Maher D, Eyre BD (2011) Benthic carbon metabolism in southeast Australian estuaries: Habitat importance, driving forces, and

application of artificial neural network models. Mar Ecol Prog Ser 439:97–115. https://doi.org/10.3354/meps09336
Martin RM, Moseman-Valtierra S (2017) Different short-term responses of greenhouse gas fluxes from salt marsh mesocosms to simulated global change drivers. Hydrobiologia 802:71–83. https://doi.org/10.1007/s10750-017-3240-1
Mateos-Naranjo E, Redondo-Gómez S, Álvarez R, Cambrollé J, Gandullo J, Figueroa ME (2010) Synergic effect of salinity and CO_2 enrichment on growth and photosynthetic responses of the invasive cordgrass *Spartina densiflora*. J Exp Bot 61:1643–1654. https://doi.org/10.1093/jxb/erq029
Moffat AM, Beckstein C, Churkina G, Mund M, Heimann M (2010) Characterization of ecosystem responses to climatic controls using artificial neural networks. Glob Change Biol 16:2737–2749. https://doi.org/10.1111/j.1365-2486.2010.02171.x
Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. T ASABE 50:885–900. https://doi.org/10.13031/2013.23153
Moseman-Valtierra S, Abdul-Aziz OI, Tang J, Ishtiaq KS, Morkeski K, Mora J, Quinn RK, Martin RM, Egan K, Brannon EQ, Carey J (2016) Carbon dioxide fluxes reflect plant zonation and belowground biomass in a coastal marsh. Ecosphere 7:e01560. https://doi.org/10.1002/ecs2.1560
Mozdzer TJ, Megonigal JP (2013) Increased methane emissions by an introduced *Phragmite australis* lineage under global change. Wetlands 33:609–615. https://doi.org/10.1007/s13157-013-0417-x
Nahlik AM, Mitsch WJ (2011) Methane emissions from tropical freshwater wetlands located in different climatic zones of Costa Rica. Glob Change Biol 17:1321–1334. https://doi.org/10.1111/j.1365-2486.2010.02190.x
Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—A discussion of principles. J Hydrol 10:282–290. https://doi.org/10.1016/0022-1694(70)90255-6
Oikawa PY, Jenerette GD, Knox SH, Sturtevant C, Verfaillie J, Dronova I, Poindexter CM, Eichelmann E, Baldocchi DD (2017) Evaluation of a hierarchy of models reveals importance of substrate limitation for predicting carbon dioxide and methane exchange in restored wetlands. J Geophys Res-Biogeo 122:145–167. https://doi.org/10.1002/2016JG003438
Ouyang X, Lee SY (2014) Updated estimates of carbon accumulation rates in coastal marsh sediments. Biogeosciences 11:5057–5071. https://doi.org/10.5194/bg-11-5057-2014
Ozyildirim BM, Avci M (2013) Generalized classifier neural network. Neural Networks 39:8–26. https://doi.org/10.1016/j.neunet.2012.12.001
Park J, Sandberg IW (1991) Universal approximation using radial-basis-function networks. Neural Comput 3:246–257. https://doi.org/10.1162/neco.1991.3.2.246
Pierfelice KN, Graeme Lockaby B, Krauss KW, Conner WH, Noe GB, Ricker MC (2017) Salinity influences on aboveground and belowground net primary productivity in tidal wetlands. J Hydrol Eng 22:D5015002. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001223
Poffenbarger HJ, Needelman BA, Megonigal JP (2011) Salinity influence on methane emissions from tidal marshes. Wetlands 31:831–842. https://doi.org/10.1007/s13157-011-0197-0
Qi Y, Xu M, Wu J (2002) Temperature sensitivity of soil respiration and its effects on ecosystem carbon budget: nonlinearity begets surprises. Ecol Model 153:131–142. https://doi.org/10.1016/S0304-3800(01)00506-3
Reichstein M, Falge E, Baldocchi D, Papale D, Aubinet M, Berbigier P, Bernhofer C, Buchmann N, Gilmanov T, Granier A, Grünwald T (2005) On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob Change Biol 11:1424–1439. https://doi.org/10.1111/j.1365-2486.2005.001002.x
Rey-Sanchez AC, Morin TH, Stefanik KC, Wrighton K, Bohrer G (2018) Determining total emissions and environmental drivers of methane flux in a Lake Erie estuarine marsh. Ecol Eng 114:7–15. https://doi.org/10.1016/j.ecoleng.2017.06.042
Riegel JB, Bernhardt E, Swenson J (2013) Estimating above-ground carbon biomass in a newly restored coastal plain wetland using remote sensing. Plos one 8:e68251. https://doi.org/10.1371/journal.pone.0068251
Saeedi E, Hossain MS, Kong Y (2016) Side-channel information characterisation based on cascade-forward back-propagation neural network. J Electron Test 32:345–356. https://doi.org/10.1007/s10836-016-5590-4
Safa B, Arkebauer TJ, Zhu Q, Suyker A, Irmak S (2019) Net Ecosystem Exchange (NEE) simulation in maize using artificial neural networks. IFAC J Syst Contr 7:100036. https://doi.org/10.1016/j.ifacsc.2019.100036
Safa B, Arkebauer TJ, Zhu Q, Suyker A, Irmak S (2021) Gap Filling of Net Ecosystem CO_2 Exchange (NEE) above Rain-Fed Maize Using Artificial Neural Networks (ANNs). J Softw Eng Appl 14:150–171. https://doi.org/10.4236/jsea.2021.145010
Sage RF, Kubien DS (2007) The temperature response of C3 and C4 photosynthesis. Plant Cell Environ 30:1086–1106. https://doi.org/10.1111/j.1365-3040.2007.01682.x
Schäfer KV, Duman T, Tomasicchio K, Tripathee R, Sturtevant C (2019) Carbon dioxide fluxes of temperate urban wetlands with different restoration history. Agr Forest Meteorol 275:223–232. https://doi.org/10.1111/j.1365-3040.2007.01682.x
Schäfer KV, Tripathee R, Artigas F, Morin TH, Bohrer G (2014) Carbon dioxide fluxes of an urban tidal marsh in the Hudson-Raritan estuary. J Geophys Res-Biogeo 119:2065–2081. https://doi.org/10.1002/2014JG002703
Schmidt A, Wrzesinsky T, Klemm O (2008) Gap filling and quality assessment of CO_2 and water vapour fluxes above an urban area with radial basis function neural networks. Bound-Lay Meteorol 126:389–413. https://doi.org/10.1007/s10546-007-9249-7
Schwefel R, Hondzo M, Wüest A, Bouffard D (2017) Scaling oxygen microprofiles at the sediment interface of deep stratified waters. Geophys Res Lett 44:1340–1349. https://doi.org/10.1002/2016GL072079
Segers R (1998) Methane production and methane consumption: a review of processes underlying wetland methane fluxes. Biogeochemistry 41:23–51. https://doi.org/10.1023/A:1005929032764
Skolthanarat S, Lewlomphaisarl U, Tungpimolrut K (2014) Short-term load forecasting algorithm and optimization in smart grid operations and planning. In: 2014 IEEE Conference on Technologies for Sustainability (SusTech), IEEE, Portland, OR, USA, pp 165-171. https://doi.org/10.1109/SusTech.2014.7046238
Smith IA, Hutyra LR, Reinmann AB, Thompson JR, Allen DW (2019) Evidence for edge enhancements of soil respiration in temperate forests. Geophys Res Lett 46:4278–4287. https://doi.org/10.1029/2019GL082459
Smithson SC, Yang G, Gross WJ, Meyer BH (2016) Neural networks designing neural networks: multi-objective hyper-parameter optimization. In: Frank L (ed) Proceedings of the 35th International Conference on Computer-Aided Design, Austin, Texas, USA, pp 1-8. https://doi.org/10.1145/2966986.2967058
Stathakis D (2009) How many hidden layers and nodes? Int J Remote Sens 30:2133–2147. https://doi.org/10.1080/01431160802549278
St-Hilaire F, Wu J, Roulet NT, Frolking S, Lafleur PM, Humphreys ER, Arora V (2010) McGill wetland model: evaluation of a peatland carbon simulator developed for global assessments. Biogeosciences 7:3517–3530. https://doi.org/10.5194/bg-7-3517-2010
Stursa D, Dolezel P (2019) Comparison of ReLU and linear saturated activation functions in neural network for universal

approximation. In 22nd International Conference on Process Control (PC19). IEEE, Strbske Pleso, Slovakia, pp 146-151. https://doi.org/10.1109/PC.2019.8815057

Sudheer KP, Jain SK (2003) Radial basis function neural network for modeling rating curves. J Hydrol Eng 8:161–164. https://doi.org/10.1061/(ASCE)1084-0699(2003)8:3(161)

Theuerkauf EJ, Stephens JD, Ridge JT, Fodrie FJ, Rodriguez AB (2015) Carbon export from fringing saltmarsh shoreline erosion overwhelms carbon storage across a critical width threshold. Estuar Coast Shelf S 164:367–378. https://doi.org/10.1016/j.ecss.2015.08.001

Tramontana G, Migliavacca M, Jung M, Reichstein M, Keenan TF, Camps-Valls G, Ogee J, Verrelst J, Papale D (2020) Partitioning net carbon dioxide fluxes into photosynthesis and respiration using neural networks. Glob Change Biol 26:5235–5253. https://doi.org/10.1111/gcb.15203

Vasquez EA, Glenn EP, Guntenspergen GR, Brown JJ, Nelson SG (2006) Salt tolerance and osmotic adjustment of *Spartina alterniflora (Poaceae)* and the invasive M haplotype of *Phragmites australis (Poaceae)* along a salinity gradient. Am J Bot 93:1784–1790. https://doi.org/10.3732/ajb.93.12.1784

Vivanco L, Irvine IC, Martiny JB (2015) Nonlinear responses in salt marsh functioning to increased nitrogen addition. Ecology 96:936–947. https://doi.org/10.1890/13-1983.1

Walter BP, Heimann M (2000) A process-based, climate-sensitive model to derive methane emissions from natural wetlands: Application to five wetland sites, sensitivity to model parameters, and climate. Global Biogeochem Cy 14:745–765. https://doi.org/10.1029/1999GB001204

Wang H, Hsieh YP, Harwell MA, Huang W (2007) Modeling soil salinity distribution along topographic gradients in tidal salt marshes in Atlantic and Gulf coastal regions. Ecol Model 201:429–439. https://doi.org/10.1016/j.ecolmodel.2006.10.013

Wang X, Qin Y, Wang Y, Xiang S, Chen H (2019) ReLTanh: An activation function with vanishing gradient resistance for SAE-based DNNs and its application to rotating machinery fault diagnosis. Neurocomputing 363:88–98. https://doi.org/10.1016/j.neucom.2019.07.017

Wenger SJ, Olden JD (2012) Assessing transferability of ecological models: an underappreciated aspect of statistical validation. Methods Ecol Evol 3:260–267. https://doi.org/10.1111/j.2041-210X.2011.00170.x

Weston NB, Neubauer SC, Velinsky DJ, Vile MA (2014) Net ecosystem carbon exchange and the greenhouse gas balance of tidal marshes along an estuarine salinity gradient. Biogeochemistry 120:163–189. https://doi.org/10.1007/s10533-014-9989-7

Willard JD, Read JS, Appling AP, Oliver SK, Jia X, Kumar V (2020) Predicting water temperature dynamics of unmonitored lakes with meta transfer learning. Water Resour Res 57:e2021WR029579. https://doi.org/10.1029/2021WR029579

Wilson AM, Morris JT (2012) The influence of tidal forcing on groundwater flow and nutrient exchange in a salt marsh-dominated estuary. Biogeochemistry 108:27–38. https://doi.org/10.1007/s10533-010-9570-y

Xie T, Yu H, Wilamowski B (2011) Comparison between traditional neural networks and radial basis function networks. In Proceedings of the 2011 IEEE International Symposium on Industrial Electronics, IEEE, Gdansk, Poland, pp 1194-1199. https://doi.org/10.1109/ISIE.2011.5984328

Yates KL, Bouchet PJ, Caley MJ, Mengersen K, Randin CF, Parnell S, Fielding AH, Bamford AJ, Ban S, Barbosa AM, Dormann CF (2018) Outstanding challenges in the transferability of ecological models. Trends Ecol Evol 33:790–802. https://doi.org/10.1016/j.tree.2018.08.001

Zhang Z, Beck MW, Winkler DA, Huang B, Sibanda W, Goyal H (2018) Opening the black box of neural networks: methods for interpreting neural network models in clinical applications. Ann Transl Med 6:216. https://doi.org/10.21037/atm.2018.05.32

Zhang Y, Li C, Trettin CC, Li H, Sun G (2002) An integrated model of soil, hydrology, and vegetation for carbon dynamics in wetland ecosystems. Global Biogeochem Cy 16:9. https://doi.org/10.1029/2001GB001838

Zhang W, Zhong X, Liu G (2008) Recognizing spatial distribution patterns of grassland insects: neural network approaches. Stoch Env Res Risk A 22:207–216. https://doi.org/10.1007/s00477-007-0108-3

Zhu D, Wu N, Bhattarai N, Oli KP, Chen H, Rawat GS, Rashid I, Dhakal M, Joshi S, Tian J, Zhu QA (2021) Methane emissions respond to soil temperature in convergent patterns but divergent sensitivities across wetlands along altitude. Glob Change Biol 27:941–955. https://doi.org/10.1111/gcb.15454

Zhu X, Zhuang Q, Qin Z, Glagolev M, Song L (2013) Estimating wetland methane emissions from the northern high latitudes from 1990 to 2009 using artificial neural networks. Global Biogeochem Cy 27:592–604. https://doi.org/10.1002/gbc.20052

Zirkohi MM, Fateh MM, Akbarzade A (2010) Design of Radial Basis Function Network Using Adaptive Particle Swarm Optimization and Orthogonal Least Squares. J Softw Engineer Appl 3:704–708. https://doi.org/10.4236/jsea.2010.37080