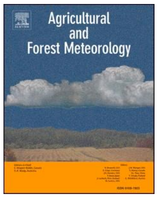
Agricultural and Forest Meteorology 350 (2024) 109955

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

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



Explainable machine learning for predicting stomatal conductance across multiple plant functional types

Srishti Gaur[a], Darren T. Drewry[a,b,c,*]

[a] Department of Food, Agricultural, and Biological Engineering, Ohio State University, Columbus, OH 43210, USA
[b] Department of Horticulture and Crop Science, Ohio State University, Columbus, OH 43210, USA
[c] Translational Data Analytics Institute, Ohio State University, Columbus, OH 43210, USA

ARTICLE INFO

Keywords:
Stomatal conductance
Ecohydrology
Explainable machine learning
Plant functional types
Uncertainty

ABSTRACT

Stomatal conductance (g_s) is a key leaf-level function controlling water, carbon, and energy exchange between vegetation and the surrounding environment. Conventionally, semi-empirical models have been used to model g_s, but these models require re-parameterization as ecosystems undergo phenological changes over the growing season. In contrast, machine learning (ML) models offer a potential path to overcome this problem but are less interpretable than process-based models. This study explores ML as an approach to develop flexible and robust models of g_s for a range of plant functional types (PFTs), including C3 crops, C3 grasses, shrubs, and tree species across different continents. An explainable machine-learning approach (eXML) was used here to provide novel interpretations and insights into the ML model formulations and relative predictor importance. We contrast the performance of three ML architectures: extreme gradient boosting, random forests, and neural networks. Models were developed and examined using many combinations of environmental and physiological predictors. The results demonstrated that ML models significantly outperform conventional semi-empirical models in predicting g_s responses to the environment, while not requiring re-parameterization as is required in the semi-empirical paradigm. Particular focus is placed on models formulated around predictor sets that are: (a) relevant to g_s estimation in modern terrestrial biophysical simulation models, and (b) composed of variables describing environmental and physiological drivers that can be remotely sensed non-invasively. "Generalized" models developed using data from all four PFTs demonstrated strong predictive performance using only three predictor variables, capturing 63–80 % of the variability in stomatal conductance across all ML architectures. Four predictor variables resulted in models capturing 79–83 % of g_s variability, and models developed using all five predictor variables examined here were able to capture as much as 87 % of g_s variability across all PFTs. Uncertainty in g_s predictions was quantified using quantile regression. Shapley additive explanations was applied to unravel instance-based positive and negative contributions of environmental and physiological predictors to g_s modeling, while illustrating that the models are consistent with the underlying ecophysiology. This work demonstrates the power of ML to introduce a new paradigm in the simulation of highly dynamic ecophysiological processes critical to environmental prediction.

1. Introduction

Plants control the energy and water budget of the vegetated land surface by balancing the need for photosynthetic carbon dioxide (CO_2) uptake while limiting the water lost during transpiration through pores located on leaf surfaces known as stomata (Buckley and Mott, 2013; Miner et al., 2017; Bonan, 2019). As the ecophysiological control on plant water use, stomata play a central role in regulating leaf and plant water status and transport, photosynthesis, and drought sensitivity (Hetherington and Woodward, 2003; Damour et al., 2010; Buckley, 2017; Liao et al., 2022). Stomatal activity is often measured and quantified in models as stomatal conductance (g_s) and provides a central link between simulated processes of energy balance and photosynthesis (Drewry et al., 2010a; Buckley, 2017; Vialet-Chabrand and Lawson, 2019).

Stomatal responses to the local environment (e.g., incident radiation, relative humidity, temperature, and CO_2 concentration) have been widely studied (Jarvis, 1976; Ball et al., 1987; Leuning, 1990; Leuning

---

* Corresponding author at: Department of Food, Agricultural, and Biological Engineering, Ohio State University, Columbus, OH 43201, USA.
E-mail address: drewry.19@osu.edu (D.T. Drewry).

https://doi.org/10.1016/j.agrformet.2024.109955
Received 5 July 2023; Received in revised form 20 January 2024; Accepted 1 March 2024

et al., 1995; Loranty et al., 2010). This has led to the development of semi-empirical formulations to estimate g_s as a function of environmental conditions at the leaf surface. Specifically, the Ball–Berry model (Ball et al., 1987; Collatz et al., 1991) relates g_s to leaf surface relative humidity, CO_2 concentration, and photosynthetic activity. The Leuning model (Leuning, 1990) relates g_s to vapor pressure deficit, CO_2 concentration, and photosynthetic activity. These g_s models are widely utilized in land surface models that require g_s estimation to resolve energy balance and biochemical processes related to the land-atmosphere exchange of carbon, water, and energy (Drewry et al., 2010a, 2010b; De Kauwe et al., 2015; Franks et al., 2018; Saunders et al., 2021; Chitsaz et al., 2023). A key consideration when using either of these semi-empirical formulations is the fixed parameters that need to be determined prior to applying the model. A challenge has been the need to modify / re-estimate model parameter values as vegetation phenological, physiological, and biochemical status change over the course of a growing season. Often these models are parameterized from leaf gas exchange measurements (Kim and Leith, 2003; Ko and Piccinni, 2009; Saunders et al., 2021; Garen et al., 2022) which are often time prohibitive to make with regular frequency. This points to the need for new approaches to estimate g_s that rely on non-parametric methods.

Machine learning (ML) frameworks have recently emerged as powerful non-parametric prediction tools across a wide range of scientific and engineering fields (Reichstein et al., 2019; Sarker, 2016; Hsieh, 2022). ML offers a potential path to overcome this parameterization issue in g_s modeling, as well as providing a flexible approach to explore the value of information spanning the leaf/canopy environment, biochemical activity, and other variables that can be derived from modern remote sensing and field observing systems. However, ML models do require retraining if the distribution of input data changes over time or the relationship between predictors and response changes over time, a phenomenon known as model drift (Vela et al., 2022).

The development of ML models requires care as these approaches can be plagued by over-fitting to limited data (Schmidt et al., 2019). However, this issue can be mitigated by tuning hyperparameters and performing robust cross-validation during the training process on extensive datasets that capture the underlying range of variability of the predicted variables (Vincent and Jidesh, 2023; Tabe-Bordbar et al., 2018). ML models are less transparent and interpretable in comparison to mechanistic or process-based models (Lundberg and Lee, 2017). Explainable ML (eXML) methods have received significant attention in advancing reliability, interpretability and providing insights into opaque ML modeling frameworks (Lundberg et al., 2020; Newman and Furbank, 2021).

Of the few studies that have utilized ML for g_s estimation most of these studies have restricted focus either to a single plant species or a specific geographic/climatic region (Ellsäßer et al., 2020; Vitrack-Tamam et al., 2020; Houshmandfar et al., 2021). Ellsäßer et al. (2020) utilized meteorological and drone-recorded data to predict g_s using multiple ML models. Their study discovered that predicting stomatal conductance from remotely sensed data was less successful and needed further study. Houshmandfar et al. (2021) compared the performance of a Jarvis-type model with ML models for predicting g_s in wheat. This study suggested that the ML models they produced exhibit high prediction accuracy when trained on larger datasets but have limited capability to extrapolate to unseen regions of data. Saunders et al. (2021) developed ML models for g_s using environmental predictors for 36 tree species from 5 forest biomes across 6 continents and suggested that future work evaluate the ability of ML to predict g_s across contrasting climate regimes and vegetation types. Here we develop a range of ML models of g_s for multiple plant functional types (PFTs) to evaluate the extent to which ML can provide non-parametric models capable of accurately simulating g_s within and across PFTs.

Gas exchange datasets are widely used in eco-hydrology to model leaf-level functions and their response to environmental and physiological changes. These datasets serve as the foundation for scaling leaf-level functions from leaf to canopy scale in land surface models (Yang et al., 2020; Ely et al., 2021). However, the collection of these datasets is weather-dependent, time-consuming and requires specific training, logistics and equipment (Ellsworth et al., 2012; Weerasinghe et al., 2014; Ely et al., 2021). Because of these constraints, such measurements tend to be clustered in time and space in areas where data collection campaigns have been focused. This may introduce uncertainty in models developed using these datasets, and accounting for these uncertainties will improve model extensibility.

In this paper we explored the use of three different ML architectures for model development: random forests (RF), extreme gradient boosting (XGBoost), and multilayer perceptron neural networks (MLPNN). Each of these architectures has inherent features that make them unique from each other. RF and XGBoost are tree-based ensemble algorithms with fundamentally distinct architectures. RF is an evolution of the bagging technique that combines multiple decision trees that each contribute to the predictions, that has been successfully applied to a wide range of problems over its history (Breiman, 2001; Fawagreh et al., 2014; Tyralis et al., 2019). RF is an efficient ensemble learning model that ensures high predictive precision, immediacy, and flexibility in improving predictions compared to other supervised approaches like decision trees (Aria et al., 2021; Dinh et al., 2023; Gaur et al., 2023). RF is also known to be robust to outliers and missing data (Tang and Ishwaran, 2017). XGBoost is a boosting strategy based on sequential learning in which each decision tree is built in turn so that each tree improves relative to the previous instance (Chen and Guestrin, 2016). XGBoost has gained popularity due to its scalability, ability to deal with noisy data, and ability to handle sparse data (Chen and Guestrin, 2016; Sahoo et al., 2021; Narbeav et al., 2023). MLPNN contains embedded hierarchical architectures to obtain higher-level features from input datasets through representation learning (Ivakhnenko and Lapa, 1965; Schmidhuber, 2015). MLPNN has been demonstrated to robustly predict, learn, and classify non-linear data, making it an ideal candidate for modeling plant physiological responses to environment (Gaur et al., 2020; Isabona et al., 2022).

The previous studies applying ML to g_s utilized meteorological, remote sensing, and phenological predictors to predict g_s, however, consideration of plant physiology (e.g., biochemical aspects like net assimilation) have not been taken into account. The datasets used here to develop and test ML models of g_s (including environmental and physiological variables) span four PFTs from data points collected across different continents. In order to understand the importance of model complexity (number of predictors used to develop a model) in predictive performance and trade-offs between specific predictors, we systematically develop models using each single predictor, and all possible combinations of predictor variables. Particular focus is placed on ML models that utilize the same predictor variables as the conventional semi-empirical Ball–Berry and Leuning models (i.e., Ball–Berry and Leuning proxies), as well as predictor sets that can leverage modern remote sensing capabilities towards non-invasive and continuous assessments of stomatal conductance. We have applied a robust cross-validation scheme to overcome the issues associated with the overfitting of ML models.

To our knowledge no previous study has leveraged explainability metrics to provide insights into the relative importance of environmental and physiological predictors, and the unique synergistic combinations of predictors that emerge for modeling this process. Here we use these techniques to not only explain the outcomes of ML models (global explanations) but also to provide underlying reasoning on how particular predictions were made (instance-based explanations), which explains the order of importance of different environmental and physiological predictors in predicting g_s using a popular exML tool SHapley Additive exPlanations (SHAP). This study also advances the use of ML for g_s prediction by quantifying the predictive uncertainty of these models.

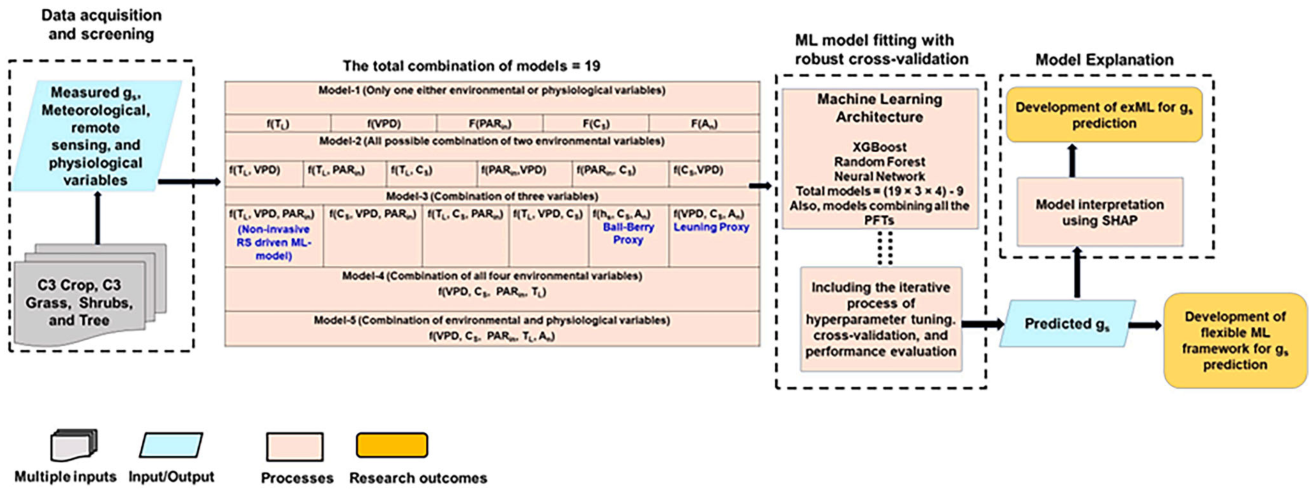


Fig. 1. Methodology schematic spanning data acquisition and screening, ML model fitting with robust cross-validation for a wide range of models defined by unique predictor variable sets, and the application of analytical techniques to provide explanatory power for each of the ML models.

2. Methodology

Stomatal conductance is a central variable in many models of terrestrial ecohydrology and land surface biophysics. The two most popular semi-empirical models, those of Ball–Berry (Ball et al., 1987; Leuning, 1990; Leuning et al., 1995), calculate g_s as a function of net photosynthesis, carbon dioxide concentration at the leaf surface, and either relative humidity or vapor pressure deficit at the leaf surface (Anderegg et al., 2017). The parameters of these semi-empirical models are often determined through data fitting and require re-parameterization to be accurate as vegetation phenology and physiology change over the course of a growing season (Saunders et al., 2021; Chen et al., 2020). The focus of this study was to examine the ability of ML to non-parametrically simulate this critical biophysical process for multiple distinct PFTs, and to assess the ability ML models to have generalized predictive capabilities across multiple PFTs. To this end, we use three machine learning architectures for model development to understand if there are unique advantages to using a specific ML architecture for this ecophysiological process. The sub-sections below provide details on the data used in this study, methodologies for ML development and evaluation, and interpretation of model results. A schematic of the over-arching methodology for this study is presented in Fig. 1, including data screening and pre-processing, how the full set of ML models evaluated in this study are formulated, model training and cross-validation, and the evaluation of explainability indicators.

2.1. Datasets and predictor variables

Here we use existing global datasets of stomatal conductance along with contemporaneous measurements of environmental and physiological variables as the basis for developing ML models of g_s. The datasets are found in Anderegg et al. (2018) and Lin et al. (2015). Anderegg et al. (2018) synthesized usable data from peer-reviewed literature for approximately 34 species spanning wide geographical and taxonomic coverage of global forest biomes collected with gas exchange instruments. These datasets have been utilized by Saunders et al. (2021) examine the utility of ML for predicting g_s using meteorological forcings. Lin et al. (2015) gathered leaf-level gas exchange datasets spanning a wide range of PFTs and biomes including 314 species from 56 experimental sites around the globe (17 sites from Australasia, 15 sites from Europe, 14 sites from North America, 6 sites from Asia, 3 sites from South America and 1 site from Africa). These measurements include instantaneous measurements under ambient conditions using gas exchange instruments. Most of these measurements are made on upper canopy leaves during the growing season.

The predictor variables used here include leaf temperature (T_L [°C]), incident photosynthetically active radiation (PAR_{in} [μmol m^{-2} s^{-1}]), net photosynthetic rate (A_n [mol m^{-2} s^{-1}]), and ambient environmental conditions at the leaf surface including carbon dioxide concentration (C_s [ppm]), relative humidity (h_s [%]) and vapor pressure deficit (VPD [kPa]). Integrating these datasets allows us to examine model performance across four PFTs: C3 crops, C3 grasses, shrubs, and trees. The tree dataset was obtained from Anderegg et al. (2018), while the datasets for C3 crops, C3 grasses, and shrubs were obtained from Lin et al. (2015). These datasets span collection campaigns across several continents, but a spatial bias can be seen to exist as some locations had intensive data collection activities while other locations were not well sampled. Based on data availability for all predictors, the number of measurements for C3 crops, C3 grasses, shrubs, and trees is 818, 768, 630, and 2614, respectively. This suggests that based on the relative number of measurements, the dataset used in this study is biased towards tree PFT. We develop ML models here that utilize each PFT individually, as well as models that utilize all data collectively, as well as performing statistical analysis of uncertainty on the modeling predictions. Relative humidity measurements are available only for the tree species and so h_s is utilized for g_s prediction for only this PFT.

Fig. 2 shows the histograms of all predictor variables and g_s data used in this study (i.e., from across all PFTs). Notably, these data were all collected under a narrow range of C_s between 330 and 450 ppm, which is likely to minimize the importance of this environmental condition in our modeling results. Other predictor variables show a wider range of variability, spanning values typical of many field conditions. The distribution of PAR_{in} is left skewed toward saturating light conditions, but a wide range of incident light levels is represented in the full dataset used here. Fig. S1 presents similar plots for each of the individual PFTs. It is important to note that the models developed here are trained and validated around the data used in this study and are not necessarily valid for so-called “unseen” conditions (i.e., conditions outside the ranges used for model development).

Our focus in this study is to evaluate the performance of ML models formulated with the variables used in the two primary semi-mechanistic formulations described above, and to understand how model performance could be improved using additional predictors related to energy balance, a process tightly coupled to stomatal conductance (Drewry et al., 2010a, 2010b). To this end, we included PAR_{in} as a variable describing energy available for biochemical processes and T_L as a

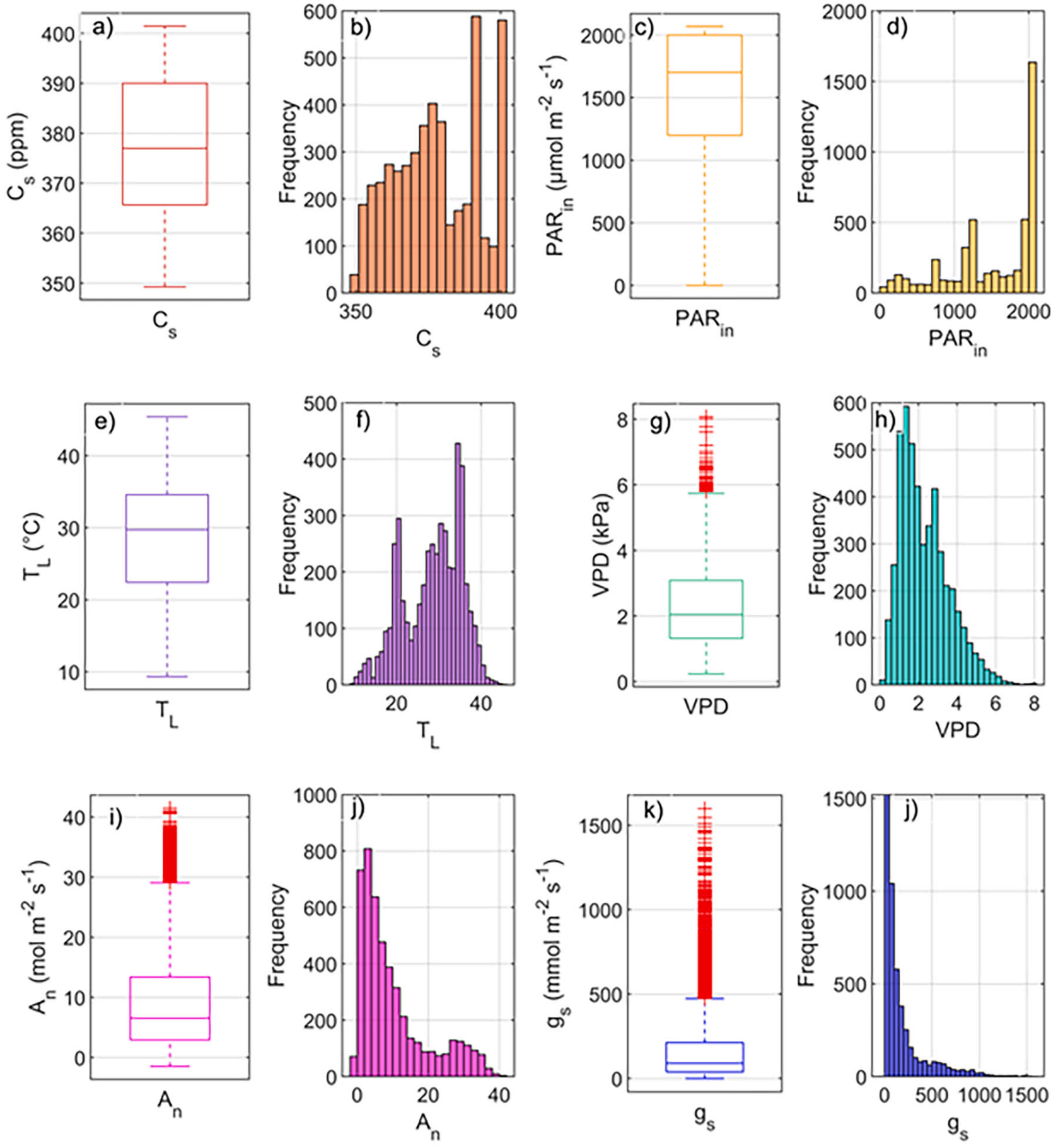


Fig. 2. The range of environmental and physiological variables, across all datasets (PFTs), used to develop and validate ML models in this study. This includes surface CO_2 concentration (a,b), downwelling PAR (c,d), leaf temperature (e,f), vapor pressure deficit (g,h), net photosynthesis (i,j) and stomatal conductance (k,j).

variable central to the leaf energy balance and that can be obtained through remote sensing (Chen and Liu, 2020; Sobejano-Paz et al., 2023).

2.2. Model formulation and uncertainty estimation

Fig. 1 presents the different combinations of predictors used to construct each of the 19 ML model formulations developed and evaluated here. Model-1 refers to the collection of model formulations that utilize only one environmental or physiological variable as a predictor (five in total: Model 1.1 to Model 1.5).

Model-2 refers to all model formulations that use combinations of two environmental variables (six in total: Model 2.1 (T_L, VPD), Model 2.2 (T_L, PAR_{in}), Model 2.3 (T_L, C_s), Model 2.4 (PAR_{in}, VPD), Model 2.5 (PAR_{in}, C_s), and Model 2.6 (C_s, VPD). Model-3 refers to the set of models that utilize three predictor variables: Model 3.1 (T_L, VPD, PAR_{in}), Model 3.2 (C_s, VPD, PAR_{in},), Model 3.3 (T_L, C_s, PAR_{in}), Model-3.4 (T_L, VPD, C_s), Model-3.5 (h_s, C_s, A_n) and Model-3.6 (VPD, C_s, A_n). Model-4 refers to the model formulation that does not require a measurement of biochemical activity (C_s, VPD, T_L and PAR_{in}) and Model-5 refers to the model that uses all five predictor variables.

The nineteen unique combinations of forcing variables specified above are used for each of the three ML model architectures examined here, resulting in 57 models developed for each PFT. Additionally, we develop these 57 models for the aggregate dataset containing data from all PFTs. This results in a total of 228 ML models evaluated here.

We will focus our analysis on a few of these formulations as they have particular significance:

- Ball–Berry Proxy (Model-3.5): this model utilizes the same set of predictors as used in the semi-empirical Ball–Berry model, and so will be contrasted directly with results of that semi-empirical formulation.

- Leuning Proxy (Model-3.6): this model utilizes the same set of predictors as used in the semi-empirical Leuning model, and so will be contrasted directly with results of that semi-empirical formulation.
- Non-invasive RS (Model 3.1): represents a potential non-invasive / non-contact remote sensing and environmental basis for predicting g_s.
- Model using all environmental predictors (Model 4): this model utilizes the complete set of environmental predictors.
- Model using all predictors (Model 5): this model utilizes all environmental and physiological predictor variables for g_s estimation.

The performance of the Ball–Berry Proxy is compared directly to predictions made with the semi-empirical Ball-Barry model for the tree PFT only, as that is the only PFT with associated h_s values in the dataset. This reduces the number of models evaluated in this study to 219. Likewise, the performance of the Leuning Proxy models is compared with semi-empirical Leuning models for all PFTs. The parameters of the Ball–Berry models, i.e., slope (*m*) and intercept (*b*), were optimized through a grid search to minimize the mean absolute error between the measured and predicted g_s. Likewise, the parameters of the Leuning model, slope (*m**) and intercept (*b**), are also optimized for all data as well as for individual PFTs. All the models described here are developed for each individual PFT and ML architecture combination. In the process of model development, these models are trained individually for each PFT. In addition, we develop a set of global ML models (terminology applies to this study only) by training them on the data from all four PFTs employing the best-performing ML architecture as determined from the individual PFT model testing. The best-performing ML architecture is chosen based on the performance metrics discussed in Section 2.4.

Uncertainty in the predictions can arise from uncertainties in measurement data, parameter values, and model structures (Solomatine and Shrestha, 2009; McMillen et al. 2018). Quantile regression (QR) is a popular method that accounts for uncertainty due to all of the aforementioned sources (Solomatine and Shrestha, 2009; Weerts et al., 2011; Rahmati et al., 2019; Gaur et al., 2021) regardless of the other methods that estimate uncertainty due to individual sources. Monte Carlo methods analyze uncertainty due to model parameters and Bayesian methods assess uncertainty due to input data (Solomatine and Shrestha, 2009). QR has gained popularity in estimating prediction uncertainty for ML models due to its versatility in incorporating multiple sources of uncertainty (Rahmati et al., 2019; Kasraei et al., 2021). In this study, QR is used to estimate the combined predictive uncertainty in g_s predictions due to predictors, parameters, and the structure of ML models.

QR estimates uncertainty between training datasets (ML predictors and response) and measured data. Parameters obtained from well-calibrated QR models are used to predict uncertainty over the validated datasets. A QR function is used to calculate the residuals (difference between measured and predicted data) over the user-defined quantiles (5 %, 50 %, 75 %, and 95 % in this study) where the model is individually calibrated for each quantile (Rahmati et al., 2019). QR utilizes normal quantile transformation to convert the data into the Gaussian domain. The details of the QR methods can be obtained from Koenkar and Hallock (2001), Koenkar (2005), Weerts et al. (2011), and Gaur et al. (2021).

2.3. Semi-empirical stomatal conductance models

The Ball–Berry (Ball et al., 1987; Ball, 1988) and Leuning et al. (1995) models estimate stomatal conductance as functions of leaf surface environmental conditions and net photosynthesis. These two models are widely utilized in ecophysiological and land surface models spanning scales of a single leaf, homogeneous plant canopy to that of the globe. Here we define these models that we use to evaluate and contrast ML proxies that we develop below.

The Ball–Berry (Ball et al., 1987) model defines g_s as a function of net photosynthesis, CO_2 at the leaf surface, a proportionality constant, and a linear offset.

$$g_s = mA_n\frac{h_s}{C_s} + b \tag{1}$$

The parameters *m* and *b* are also called Ball–Berry slope and offset, respectively.

Leuning et al. (1995) provided an alternative to the Ball–Berry model to address two issues: (i) the Ball–Berry model does not simulate g_s and A_n accurately when C_s approaches the CO_2 compensation point (Γ) and (ii) stomata respond to VPD rather than h_s. This resulted in the Leuning model for g_s:

$$g_s = m^*\frac{A_n}{\left(1+\frac{V}{D_0}\right)(C_s-\Gamma)} + b^* \tag{2}$$

The parameters *m** and *b** are the Leuning slope and offset. D_0 is an empirically determined parameter.

2.4. ML model development

ML models are flexible in nature due to their ability to automatically learn relationships between predictors and response variables (de Hond et al., 2022). This gives ML models the potential to more efficiently utilize information in predictor variables, as well as find value in predictor variables not used in conventional biophysical formulations. Here we train each of the model formulations described in Section 2.2 using three ML architectures: Random Forests, XGBoost and Neural Networks.

Each ML architecture contains two kinds of parameters, a set that is determined during model training (model parameters) and a set that controls the architecture and the learning processes (hyperparameters) (Vincent and Jidesh, 2023). Accurate hyperparameter values are crucial for model performance and are the first stage of model development, determined prior to model training (Kuhn and Johnson, 2013; Yang and Chui, 2021). This process is called hyperparameter tuning (HPT) and is often performed by grid search or other optimization strategies (Bergstra et al., 2012; Fuadah et al., 2022). In the case of MLP the hyperparameters are the number of hidden layers, the learning rate and the number of neurons. These parameters essentially define the structure of the NN. Likewise, with XGBoost and RF, the number of boosting iterations and n-estimator (number of trees in a forest) are hyperparameters associated with the structure. Following the specification of this structure, the training process is performed to determine the weights of the final model. Hyperparameter values may be unique to each dataset, and so HPT is performed uniquely for each architecture and PFT here. Table S1 lists the hyperparameters of different ML architectures used in this study together with the information on their physical significance and fitted ranges obtained after hyperparameter tuning for each PFT.

Another important factor in the development of rigorous ML models is cross-validation (CV) (King et al., 2021), which determines how well a ML architecture will perform on unseen data from the same distribution (Tabe-Bordbar et al., 2018; Brodeur et al., 2020). The fundamental idea behind CV is to divide the dataset into training and validation fractions. Following training the validation fraction is used to quantify model error on the fraction of data not used in training processes. This is often performed iteratively so that this process can be performed on many random breaks of the dataset. CV attempts to minimize common issues associated with ML such as over-fitting and poor generalizability (Tabe-Bordbar et al., 2018; Brodeur et al., 2020).

As presented in schematic in Fig. S2, in the ML development work presented here we use an iterative model development and testing framework based around a robust CV procedure. We use an iterative approach ($n = 500$) to randomly split the data into training (60 %) and validation (40 %) fractions. As per previous literature, the training data fraction can range from 60 to 95 % depending on factors such as the size of the underlying dataset (Deo and Şahin, 2015; Vu et al., 2016; Hou

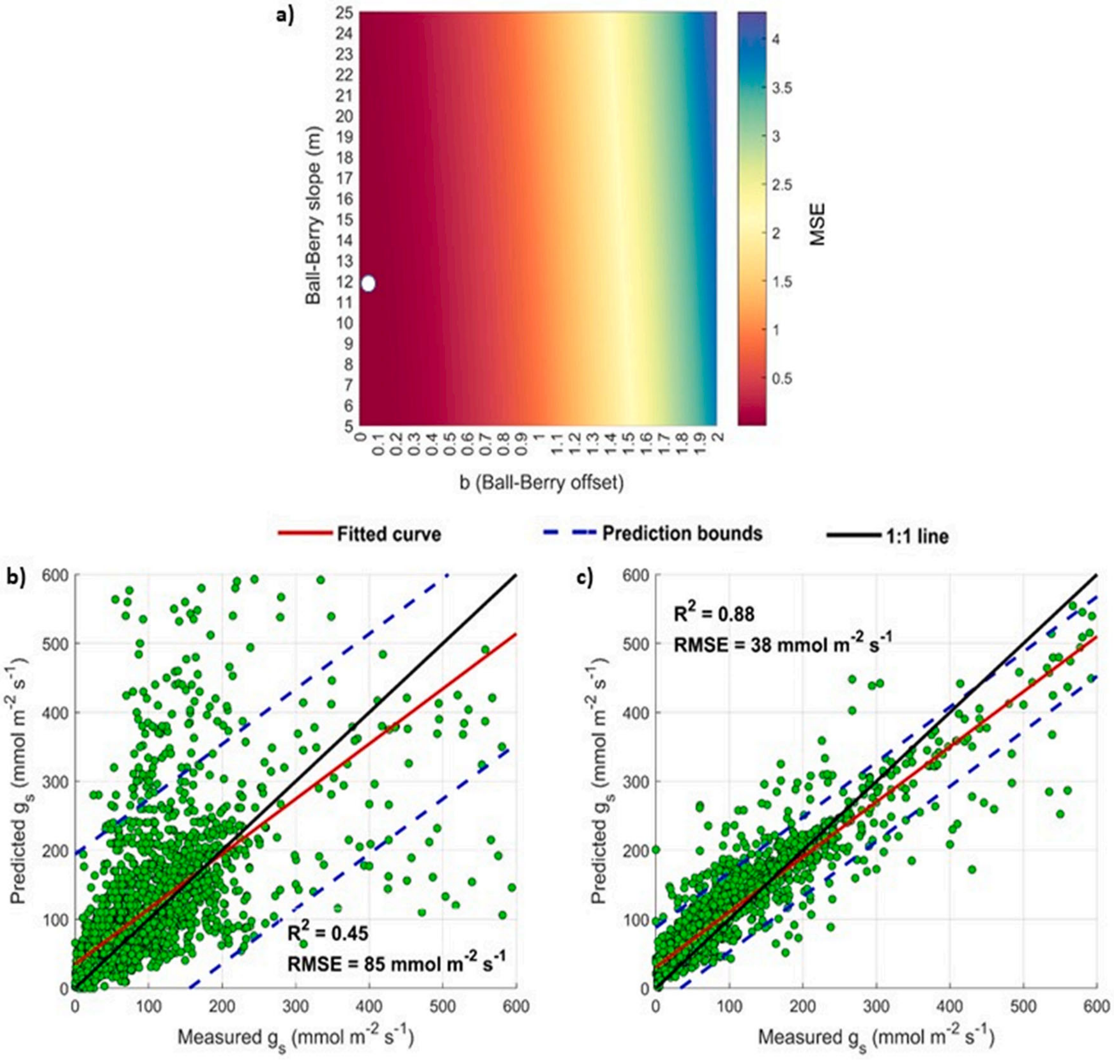


Fig. 3. (a) Range of values of Ball–Berry slope and offset parameters explored using grid search to find the optimal parameter values (minimized MSE) for the tree PFT (optimal combination is the white dot), (b) Performance of Ball–Berry model for the tree PFT and (c) the performance of ML Ball–Berry proxy for the tree PFT.

et al., 2017; Wang et al., 2018). Here we used a 60 / 40 split (training / validation) so as to ensure that a sufficient amount of data is left out for model validation. For each model developed we performed 500 random data splits to ensure that no bias from a biased data split impacted model performance. A unique optimum set of hyperparameters for each of the n data splits is determined by grid search HPT with a nested k-fold CV technique (Step-2 in Fig. S2). In k-fold CV the entire dataset is divided into k disjoint sets, the model trained on k-1 folds and validated on the remaining fold of data. The procedure is repeated k times, and overall accuracy is calculated as the average of accuracies across each fold (Yang and Chui, 2021). A 10-fold cross-validation is used in this study.

Once the n ML best performing hyperparameter sets have been specified using HPT, a model is trained for each of these hyperparameter sets using the training fraction corresponding to that unique data split. Following this training step, the trained model is used to predict g_s for the unique left out validation fraction. Mean square error (MSE) is used as the loss function to train ML models. The model's performance over training and validation fractions is then computed by averaging the performance over the n sets (Step-5 in Fig. S2). Two-performance metrics, the coefficient of determination (R^2) and root mean square error (RMSE), are used to quantify the performance of each ML architecture here.

At the last step, each ML model is trained independently using the n ML hyperparameters over the entire original training fraction (60 %). The left-out validation fraction (40 %) is used to perform the final validation of the trained model. The final model performances over training and validation fractions are determined by averaging the performance for the n sets.

2.5. Interpretations of ML models outcomes

ML models offer a path for improving the utilization of information in predictor variables for improved predictive accuracy but are challenging with respect to interpretation and developing insights into system dynamics as compared to mechanistic models (Schmidt et al., 2020). To overcome these problems, a popular eXML algorithm, SHapley Additive exPlanations (SHAP), was developed by Lundberg and Lee (2017) to interpret ML model outputs. The idea of SHAP originated from the Shapley value of game theory (Shapley, 1953), which defines the average marginal contribution of a player across all the coalitions to which the player belongs (Biecek and Burzykowski, 2020). In ML the SHAP value represents the contribution of individual predictor variables to the final outcome of the model. Apart from other eXML techniques, SHAP provides global as well as local explanations (Lundberg and Lee, 2017).

A SHAP value represents the expected marginal contribution of each

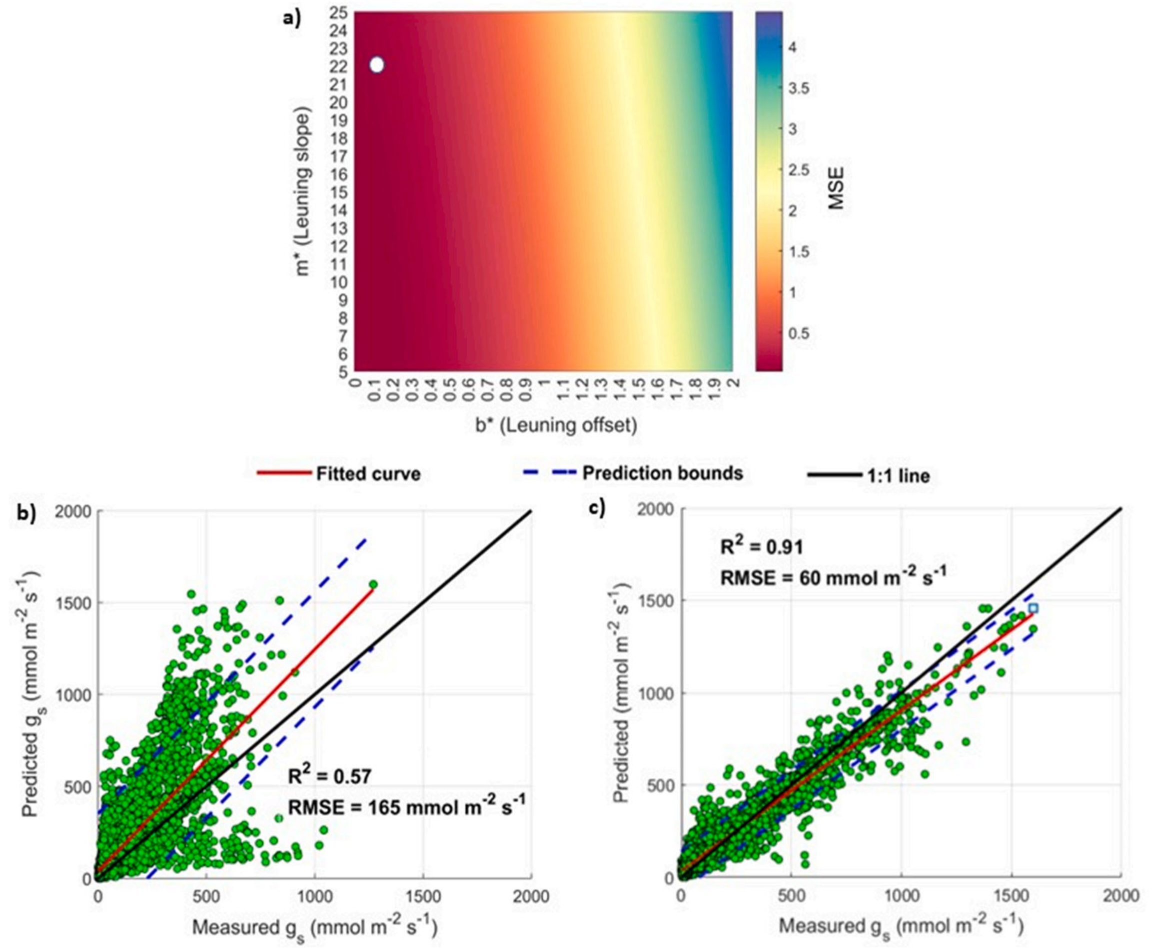


Fig. 4. (a) Range of values of Leuning slope and offset parameters explored using grid search to find the optimal parameter values (minimized MSE) for all PFTs (optimal combination is the white dot), (b) Performance of Leuning model for all PFTs and (c) the performance of ML Leuning proxy for all PFTs.

predictor which is calculated by the weighted average of a predictor's contribution (Lundberg and Lee, 2017). A detailed explanation of SHAP can be found in Lundberg and Lee (2017). Local explanation represents the contribution of each predictor for individual instances, whereas global explanations summarize the contribution of an individual predictor on the model as a whole, i.e., the aggregate of all instances (Lundberg et al., 2020). The global contribution is calculated by taking the mean of absolute SHAP values for individual instances.

Local interpretation focuses on the unique characteristics of each individual instance and offers justifications that may help in better comprehension of the predictor contribution which is disregarded by global interpretation techniques.

Here we use beeswarm plots to summarize SHAP values and interactions. Beeswarm plots illustrate the local explanations by depicting the magnitude, prevalence, and direction of the prediction's effect on model output (g_s predictions in our case) (Lundberg et al., 2020). Each dot in the beeswarm plot corresponds to a single data point evaluated by the model. When multiple points fall around the same position the SHAP plot appears denser (i.e., bulges in points in that region).

The SHAP dependence plot depicts how the value of a predictor influences the prediction for each datapoint in the dataset. These plots also account for the interaction effect of the main predictor with another. The x-axis of the dependence plot reflects the value of the main predictor, the y-axis represents the SHAP values of the respective main predictor, and the color variation represents the value of another interacting predictor.

3. Results and discussion

3.1. Stomatal conductance prediction using conventional approaches

Two widely utilized semi-empirical models of g_s, the Ball–Berry (Ball et al., 1987; Ball, 1988) and Leuning (1990) model, have been used here as the reference models for g_s predictions. Fig. 3(a) illustrates the range of Ball–Berry parameters explored in the grid search optimization. The Ball–Berry parameter values used are obtained by minimizing the mean square error (MSE) using grid search optimization. The Ball–Berry slope is the coefficient of the linear regression fit between g_s and A_nh_s/C_s (Eq. (1)) and represents the strength of the relationship between stomatal conductance and photosynthetic activity in the context of leaf-level humidity and CO_2 concentration (Ball, 1988; Miner et al., 2017). The Ball–Berry intercept represents the cuticular and residual g_s and is defined as the intercept of the linear regression fit between g_s and A_nh_s/C_s (Ball, 1988; Collatz et al., 1991). The specific values of *m* and *b* used here were found by determining the optimal values across the ranges of 5–20 and 0–2 respectively, based on known variability in these parameters for the PFTs used in this study (Leuning, 2002; Dewar, 2002 Miner et al., 2017; Saunders et al., 2021). The numerical resolution of 1 × 0.1 (*m* x *b*) was used for this grid search exercise.

The optimum values of *m* and *b* were determined to be 12 and 0 respectively and are represented by a white dot in Fig. 3a. Fig. 3b presents the performance of the Ball–Berry model for trees. The R^2 and RMSE values are found to be 0.45 and 85 mmol m^{-2} s^{-1} for the Ball–Berry model. Likewise, a white dot in Fig. 4a indicates the optimum

Table 1

Training and validation performance of the best-performing “generalized” model for each level of model complexity for data across all PFTs for in terms of R^2 and RMSE (mmol m^{-2} s^{-1}).

ML Models	XGBoost				RF				NN			
	R^2		RMSE		R^2		RMSE		R^2		RMSE	
	T_r	V_a	T_r	V_a	T_r	V_a	T_r	V_a	T_r	V_a	T_r	V_a
M5 $f(A_n, VPD, C_s, PAR_{in}, T_L)$	0.94	0.87	98	105	0.92	0.84	105	112	0.90	0.81	204	255
M4 $f(VPD, C_s, PAR_{in}, T_L)$	0.92	0.83	115	125	0.90	0.79	121	131	0.88	0.79	198	205
M3.1 $f(VPD, PAR_{in}, T_L)$	0.86	0.72	189	194	0.84	0.69	201	221	0.80	0.70	220	245
M3.2 $f(VPD, C_s, PAR_{in})$	0.88	0.73	192	201	0.86	0.71	199	210	0.82	0.68	221	241
M3.3 $f(C_s, PAR_{in}, T_L)$	0.85	0.70	178	185	0.86	0.73	181	189	0.81	0.69	191	202
M3.4 $f(VPD, C_s, T_L)$	0.80	0.67	165	171	0.82	0.64	171	179	0.79	0.63	195	201
M3.5 $f(A_n, h_s, C_s)$	Insufficient data											
M3.6 $f(A_n, VPD, C_s)$	0.91	0.80	129	139	0.89	0.76	132	149	0.86	0.71	141	155
M2.1 $f(VPD, T_L)$	0.58	0.55	589	614	0.55	0.51	612	624	0.49	0.47	617	629
M2.2 $f(PAR_{in}, T_L)$	0.45	0.43	614	625	0.45	0.42	625	631	0.39	0.37	645	645
M2.3 $f(C_s, T_L)$	0.54	0.51	594	598	0.52	0.49	602	687	0.42	0.40	607	695
M2.4 $f(VPD, PAR_{in})$	0.59	0.55	602	689	0.49	0.46	610	701	0.41	0.38	615	725
M2.5 $f(C_s, PAR_{in})$	0.42	0.39	621	635	0.40	0.38	630	645	0.41	0.39	645	655
M2.6 $f(VPD, C_s)$	0.49	0.38	635	699	0.39	0.35	641	701	0.40	0.38	650	745
M1.1 $f(T_L)$	0.21	0.19	1245	1289	0.19	0.17	1251	1291	0.18	0.17	1260	1298
M1.2 $f(VPD)$	0.24	0.21	1212	1254	0.21	0.19	1231	1265	0.19	0.18	1240	1271
M1.3 $f(PAR_{in})$	0.20	0.18	1245	1265	0.19	0.17	1251	1271	0.17	0.16	1261	1275
M1.4 $f(C_s)$	0.19	0.15	1210	1235	0.18	0.14	1215	1241	0.19	0.16	1221	1251
M1.5 $f(A_n)$	0.35	0.32	1121	1148	0.32	0.29	1126	1154	0.31	0.29	1128	1159

values of *m** and *b**, which were determined to be 22 and 0.1, respectively. Fig. 4b presents the performance of the Leuning model by combining the data from all PFTs. The R^2 and RMSE values were found to be 0.57 and 165 mmol m^{-2} s^{-1} for the Leuning model. Likewise, the optimum values of *m** and *b** have also been determined for individual PFT. Fig. S3 presents the performance of Leuning model for individual PFTs.

3.2. Stomatal response predictions using ML models and uncertainty estimation

We use the ML development framework illustrated in Fig. 1 to examine the ability of ML to predict g_s. For each ML experiment the data is split using a 60/40 (training/validation) random percentage split, performed 500 times to avoid statistical artifacts that may be present in any single random split of the data. Table 1 presents the performance of generalized ML models, developed using data from all PFTs, obtained by averaging the 500 unique sets of models obtained from Step-5 of Fig. S2.

Predictive performance increases as the number of predictors increases from one to five.

Figs. 3c and 4c present the performance of the ML proxies for the Ball–Berry (for the tree PFT only) and Leuning models (across all PFTs) using XGBoost. The R^2 values of 0.88 and 0.91 and RMSE values of 38 mmol m^{-2} s^{-1} and 60 mmol m^{-2} s^{-1} are obtained from the Ball–Berry and Leuning proxies, respectively. For both conventional g_s models, the ML proxies (ML models that use the same predictor sets) performed significantly better.

Fig. 5 presents the performance of the final ML models (obtained from Step 6 in Fig. S2) during validation using the best performing ML architecture XGBoost. There is a clear trend of improving model performance when a greater number of predictor variables is used.

The 1-predictor models (Model-1) vary in predictive performance for all PFTs from $R^2 = 0.15$ to 0.32. The best single predictor is A_n, which performed almost 50 % better than any other single predictor model. The range of predictive performance of 2-predictor models increases dramatically from $R^2 = 0.38$ to 0.55. It is notable that some environmental parameters that performed poorly in single-predictor models, such as PAR_{in} ($R^2 = 0.18$) and C_s ($R^2 = 0.15$) demonstrated much stronger performance when combined in 2-predictor models (Model-2.5 [PAR_{in}, C_s] $R^2 = 0.39$). The performance of 2-predictor models (Model 2.2 [T_L, PAR_{in}] $R^2 = 0.51$, Model-2.3 [T_L, C_s] $R^2 = 0.43$ and Model-2.6 [C_s, VPD] $R^2 = 0.38$) show the synergistic effect of two predictors when combined with well-performing environmental predictors from single predictor models, like VPD ($R^2 = 0.22$) and T_L ($R^2 = 0.19$). These synergistic impacts are primarily driven by the unique contributions of different predictors to plant ecophysiological responses. The two-predictor model formulated by VPD and T_L ($R^2 = 0.55$) demonstrates the combination of two strong environmental predictors and is one of the best performing 2-predictor models. Another strong 2-predictor model (Model-2 [PAR_{in}, VPD], $R^2 = 0.55$) shows the synergy between two environmental predictors PAR_{in} and VPD. The performance improved dramatically again with the inclusion of a third predictor,

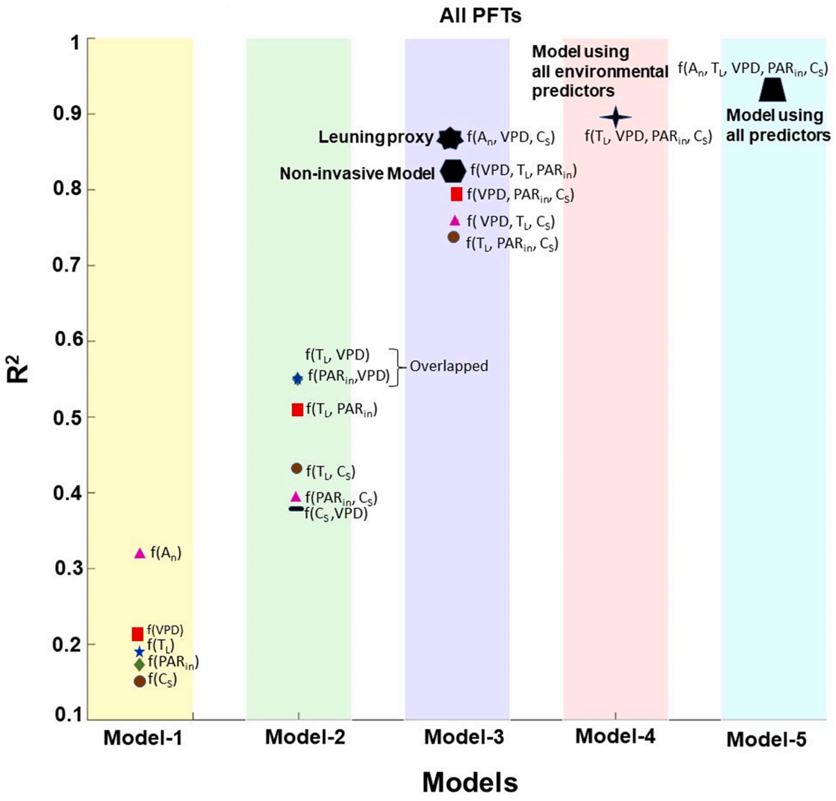


Fig. 5. Validation performance of ML model formulations. Each point represents the average performance of a single ML model developed for data aggregated over all PFTs.

with R^2 values from 0.74 to 0.87. The Leuning Proxy model in particular [A_n, VPD, C_s], demonstrated an R^2 performance of 0.87. Adding additional predictors beyond three resulted in nominal performance increases relative to three-predictor models, indicating that carefully chosen three-predictor models may be described as parsimonious, achieving near-optimal performance while minimizing model complexity. We also highlight an additional 3-variable model, the “non-invasive” model with a predictor set composed of environmental variables and leaf temperature which can be assessed remotely [PAR_{in}, T_L, VPD]. This model has an R^2 performance of 0.82.

A closer look at the performance of 1- to 3-predictor models reveals the importance of combining T_L and VPD on g_s prediction performance. Individually these two predictors account for approximately 20 % of the variability in g_s in single-predictor models. When combined they result in the strongest 2-predictor formulation, accounting for almost 60 % of g_s variability. Their combination in the 3-predictor “non-invasive” formulation described above accounts for greater than 80 % of the variability in g_s, making a strong case for the inclusion of these two variables in ML-based models of stomatal conductance. The performance of multiple predictor models is dependent on the unique information each predictor provides to plant ecophysiological responses.

The model with all four environmental variables outperformed all other models completely driven by environmental predictors with an R^2 of 0.89. The inclusion of all five predictors resulted in a nominal increase in performance, capturing 92 % of the variability in g_s.

Table 1 presents the performance of ML models developed using data combined across all four PFTs, i.e., “generalized” models. The performance of ML models differs in Table 1 and Fig. 5. This is because Table 1 presents the average performance over n-iterations for training and validation, whereas Fig. 5 presents the average performance of n ML hyperparameters over the original dataset during validation. Fig. 6 presents the validation performance of four of the models described in Section 2.2 above (Leuning, non-invasive, all environmental predictors and all predictors) obtained by Step-6 of Fig. S2, i.e., the average performance of models obtained using n-sets of optimum hyperparameters. Likewise, the predictive performance of the g_s models has been evaluated for the individual PFTs. In Tables S2–S5 we present the average validation performance of models developed specifically for each PFT. In general, similar model performance is seen for each individual PFT as is the case for the all-PFT model.

Fig. S4 presents the validation performance of the single-PFT g_s models. For single PFT models (Fig. S4), one-predictor and three-predictor models exhibit similar performance to the ‘generalized’ (all PFT combined) models, whereas the performance of two-predictor models varies across the different PFTs. We see similar trends in performance improvement as the number of predictors is increased, with the synergistic effects of specific predictor combinations resulting in significant performance improvements as two and three predictors are combined in ML model formulations. As with the all-PFT generalized models of Fig. 5, performance increases found as model complexity increases beyond three predictors are nominal and likely do not merit the associated increases in model complexity and additional data requirements.

To quantify uncertainty on the measured and predicted g_s responses, a QR model is calibrated on the g_s predictions and measurements for the training fraction, and calibrated parameters are subsequently used to check the performance of a QR model on the validation fraction. Fig. S5 presents the performance of calibrated QR models for the Leuning proxy, non-invasive model, model with all environmental predictors, and model with all predictors over the validation fraction. The performance of models over the validation fraction by fitting QR from the parameters obtained from calibration. Fig. S5 a, c, e, g present the scatter plot between NQT error and NQT g_s along with regression lines at 5 %, 25 % 75 % 95 % quantiles. Fig. S5 b, d, f, h show the coverage of measured g_s within the 90 % and 50 % confidence intervals obtained by predicted g_s. It is evident from Fig. 5 that most of the data lies within 90 % and 50 % confidence intervals, which suggests an acceptable amount of

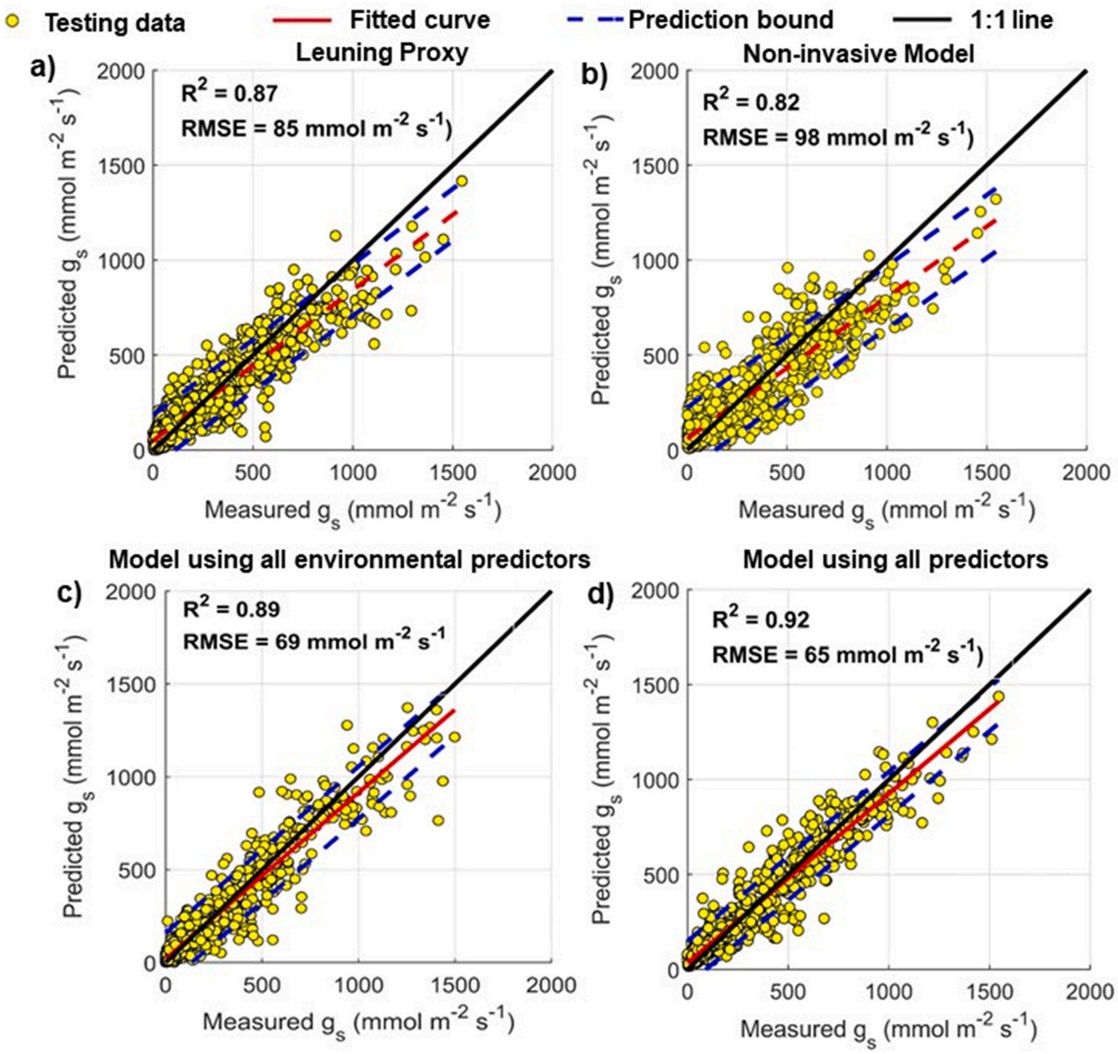


Fig. 6. Model validation performance across all PFTs for: (a) the ML Leuning proxy, (b) the non-invasive RS-based model, (c) the model using all environmental predictors, and (d) the model using all predictors.

uncertainty in the g_s predictions for all four models, given the breadth of data used in model development.

3.3. Global and local explanations of ML models

The SHAP technique is applied to four of the models of particular significance described in the Section 2.2: the Leuning proxy (Model-3.6), non-invasive model (Model-3.1), the model utilizing all environmental predictors (Model-4) and the model utilizing all sets of predictors (Model-5). Fig. 7 presents the SHAP summary plots that concisely illustrate the magnitude, prevalence,and direction of the predictor's effect. The plots in Fig. 7a, c, e, g are standard bar charts that show the average magnitude of SHAP values for each predictor. The global feature importance plot in Fig. 7a illustrates the importance of features in the order of $A_n > VPD > C_s$ as evidenced by their average SHAP values for the Leuning proxy. This order of importance remains the same when two additional predictors (PAR_{in} and T_L) are added (Fig. 7g). This could be due to relative individual performance of predictors PAR_{in} ($R^2 = 0.18$) and T_L ($R^2 = 0.19$) in comparison to other predictors such as A_n ($R^2 = 0.32$) in predicting g_s as illustrated from the performance of Model-1 in Fig. 5.

The plots in Fig. 7b, d, f, h are beeswarm plots as described in Section 2.4. The color variation from blue to red represents an increase in the magnitude of the predictor. For example, in Fig. 7b as A_n increases it results in higher predicted values of g_s, and vice versa. Likewise, lower magnitudes of VPD and C_s contribute to higher predicted values of g_s (i.e., positive impact on model output as evident from Fig. 7b), and vice versa. In Fig. 7b we also see the distribution of effect magnitudes, with low A_n values clustered around low predicted g_s magnitudes, but increasing A_n values having a larger increasing influence on predicted g_s magnitudes (extended long right tail of SHAP for A_n in Fig. 7b). The extended long tail shows the relative importance of a predictor. For example, lower magnitudes of VPD have larger positive impacts on g_s, relative to lower magnitudes of C_s. Similarly, in Fig. 7d, larger values of PAR_{in} are more important to the model than higher values of T_L.

The long tails also show that the predictors with low global importance can be important for specific situations (Lundberg et al., 2020). For example, in Fig. 7f T_L and the VPD have lower impacts on model predictions than the other variables, which is apparent in the beeswarm clustering around SHAP values of zero in Fig. 7f. But both of these variables have tails in either direction of these clusters that indicate that they can be influential on g_s model predictions in some circumstances.

The long tails for g_s model outcomes are mostly toward the right (greater positive impact on g_s predicted values), as circumstances such as A_n or PAR_{in} going to zero will cause stomatal closure. This results in the predictors examined here having the greatest influences on increases in g_s predictions. In the case of VPD and C_s lower values of these predictors positively impact the magnitude of predicted g_s. However, the

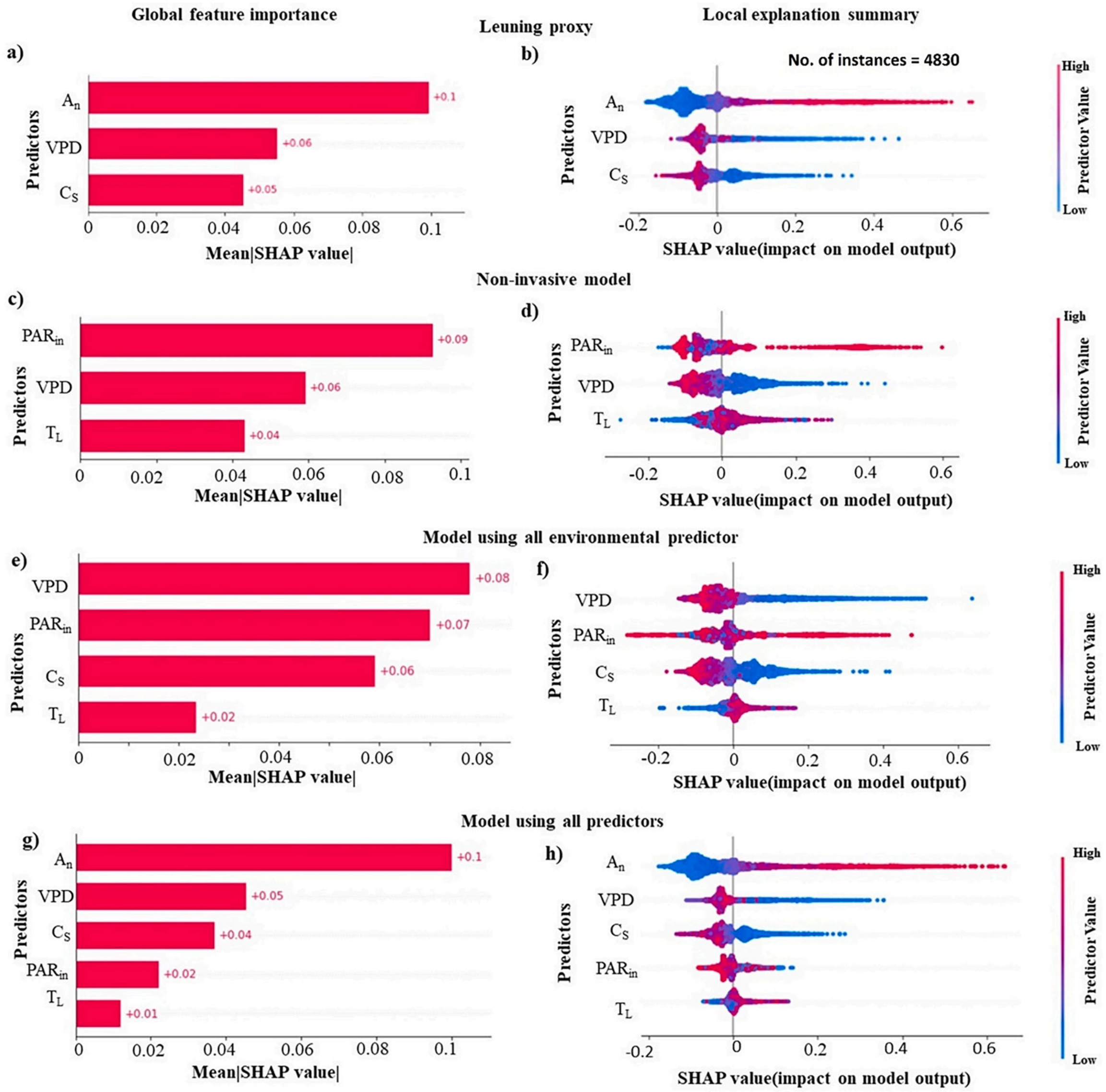


Fig. 7. SHAP summary plots for Leuning proxy [VPD, C_s, A_n], non-data invasive model [PAR_{in}, VPD, T_L], the model using all environmental predictors [VPD, C_s, PAR_{in}, T_L], and the model using all predictors [A_n, VPD, PAR_{in}, T_L, C_s] using SHAP global feature importance (a, c, e, f) and SHAP local explanation summaries (b, d, f, h).

extended right tail for VPD suggests that the lower values of VPD are more important to the model than the lower values of C_s, which is in part due to the narrow range of C_s values (approximately 330–450 ppm) evaluated here, relative to the significantly wider range of ambient VPD evaluated in these datasets (see Fig. 2).

The order of the global feature importance for the non-invasive model and the model using all environmental variables are $PAR_{in} > VPD > T_L$ and $VPD > PAR_{in} > C_s > T_L$, respectively. This is illustrated in Fig. 7c and e. The local explanation of the non-invasive model (Fig. 7d) show that higher values of PAR_{in} and T_L positively impact g_s predictions and vice versa. Nevertheless, the higher values of PAR_{in} are more important to the model than the higher values of T_L.

The model using all predictors shows the global feature importance of $A_n > VPD > C_s > PAR_{in} > T_L$. The higher importance of A_n, VPD, and C_s reaffirms the intuition of conventional semi-empirical models and physiological processes (Ball, 1988; Leuning et al., 1995). The local explanations in Fig. 7h indicate that greater magnitudes of A_n and T_L have a positive influence on g_s predictions and vice versa. In the case of VPD and C_s, lower magnitudes have more positive impacts on g_s predictions and vice versa. This could be due to the left-skewed distribution of PAR_{in} (Fig. 2), where most of the measurements have been captured under light-saturating conditions.

The results obtained from SHAP also highlight the importance of certain predictors or the combination of predictors (Fig. 7) which is a novel aspect introduced by this study. Fig. 7b and h suggest A_n is the most important predictor when all predictors or a combination of environmental predictors are used to predict g_s. Likewise, VPD is found to be the most important predictor when a certain combination of environmental predictors or all environmental predictors are used (Fig. 7d, f). This points to the utility of a combination of A_n and VPD for predicting g_s which is in line with the findings of Leuning et al. (1995) and Medlyn et al. (2011). A lower VPD can ameliorate transpiration and assimilation of CO_2 and thus, increase g_s (Ding et al., 2022). In the case of all environmental predictors or a set of them VPD and PAR_{in} are the two most important predictors in combination (Fig. 7d, f). Both VPD and PAR_{in} play an important role in plant water and energy relations, as atmospheric demand is characterized by VPD whereas PARin is a source of energy for plant assimilation. However, the combination of VPD and C_s is more important when both environmental and physiological predictors are considered.

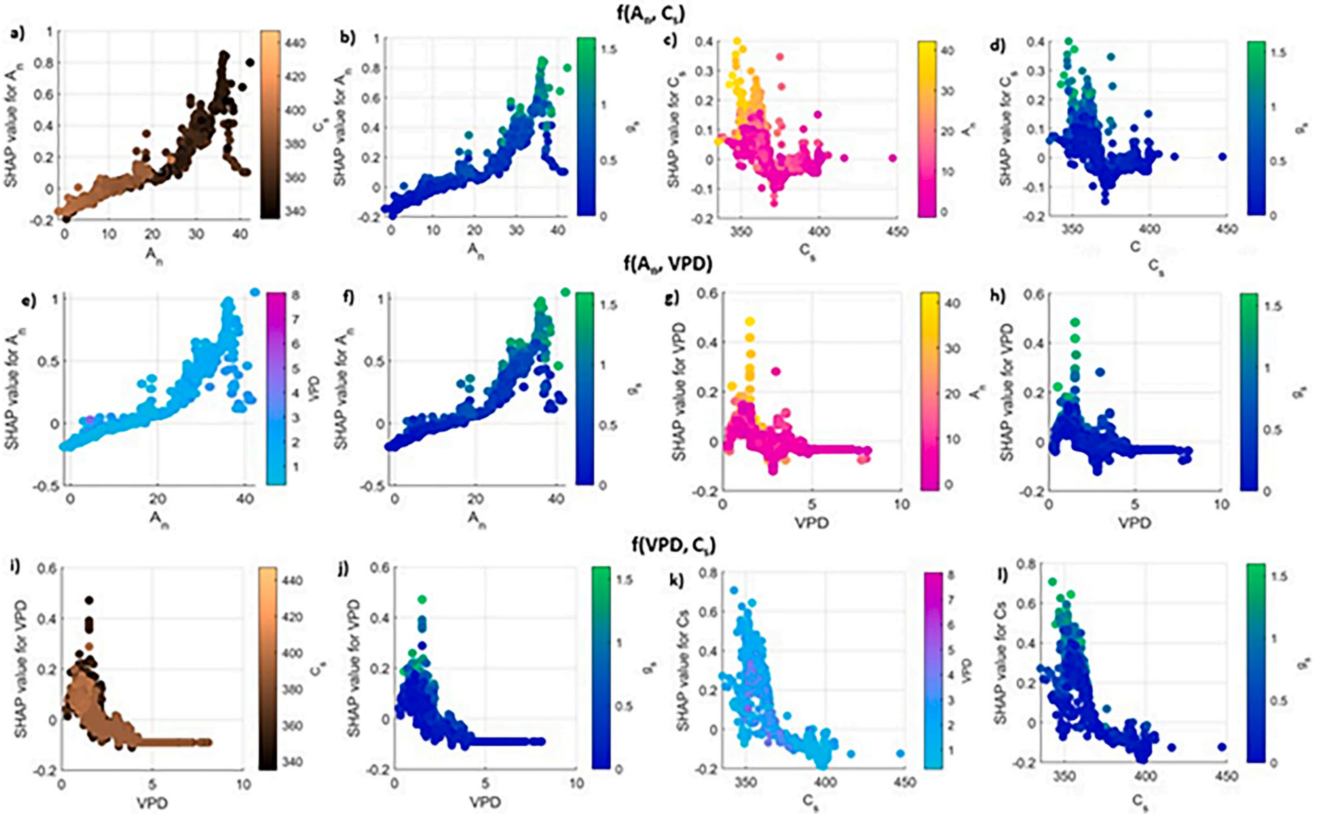


Fig. 8. SHAP dependence plots for models using two predictors of the Leuning Proxy at a time: (a, b, c, d) models developed from [A_n, C_s]; (e, f, g, h) models developed from [A_n, VPD], (i, j, k, l) [VPD, C_s]. Each panel is a 2-D scatter plot between the predictor and that predictors's SHAP value, with coloration to show the interaction with the other predictor variable, as well as g_s.

The findings obtained from SHAP suggest the relative importance of Leuning predictors, A_n, VPD, and C_s and their combinations in different models. Hence, the interaction in the different combinations of Leuning predictors is analyzed further through SHAP in the next section.

3.4. *Model explanation by individual predictors and corresponding interactions*

A SHAP dependence analysis is performed first by examining models based on all combinations of two predictors from the set of three Leuning predictors (Fig. 8). Likewise, Fig. 9 presents a SHAP dependence analysis for the Leuning proxy. These plots show how the SHAP value for a predictor varies as a function of the value of the predictor as a scatter plot, with coloration showing how these dynamics vary with the value of an additional predictor, or the predictand (g_s).

In Fig. 8a,b,e,f g_s increases with A_n, with the slope of the relationship increasing as A_n increases. This follows from well-understood stomatal responses under well-watered growth circumstances, where an increase in g_s enables plants to boost CO_2 uptake and hence enhance photosynthesis (Kusumi et al., 2012; Purcell et al., 2018). This relationship is maintained, and made somewhat stronger with less deviations and scatter, when an additional (third) predictor is added to the full Leuning proxy model (Fig. 9a–c). The coloration in Fig. 9b suggests a strong interaction between A_n and C_s,which appears to control the relationship between A_n and g_s as the slope of the A_n – SHAP value for A_n curve increases as C_s values decrease.

Plants control the degree of stomatal opening as a compromise between sustaining high photosynthesis rates and a low water loss rate. When plants are subjected to greater C_s levels, their stomata contract to avoid water loss, resulting in a decrease in g_s (Xu et al., 2016).

As with C_s, VPD also shows strong interactions with A_n as evident from the slope of Figs. 8e, g, and 9a, d. However, the extended longer right tail in the beeswarm plot (Fig. 7b) makes VPD a stronger predictor of g_s than C_s. This is also supported by the higher SHAP values of A_n when interacting with VPD (Fig. 8e), as compared to interacting with C_s (Fig. 8a).

The role of VPD is consistent across the 2-predictor models (Fig. 8g–j) and the Leuning proxy (Fig. 9d,e,f), with the largest SHAP magnitudes resulting in a positive influence on g_s prediction occurring when VPD is low. At values between 2 and 3 [kPa] the impact of VPD is to reduce g_s, with low magnitude associated SHAP values. These low magnitudes of the negative SHAP values associated with high VPD are clearly due to correlated changes in other predictors, such as decreases in A_n (Figs. 8g and 9d) and increases in C_s (Figs. 8i and 9e). Increases in VPD cause a corresponding rise in the rate of transpiration through the stomatal pore, forcing stomatal closure to prevent excessive water loss (Urban et al., 2017; Plantin and Blatt, 2018).

It is notable that the addition of a third predictor to the 2-predictor models results in a reduction in the range of SHAP values associated with each predictor, relative to the ranges of the 2-predictor models.

The strong interaction between the Leuning predictors [A_n, VPD, C_s] discussed above explains relative importance of the predictors in the model using all predictors, i.e., $A_n > \text{VPD} > C_s > PAR_{in} > T_{L.}$ In the model using all predictors, net assimilation is the most important, and the strong interaction of A_n and C_s causes C_s to climb in importance relative to importance in the environment-only model.

4. Conclusions

Plant responses to meteorological or physiological variables are not inherently linear (Cai et al., 2019; Saunders et al., 2021). Semi-empirical models center on the observed linear variation between g_s and A_n.

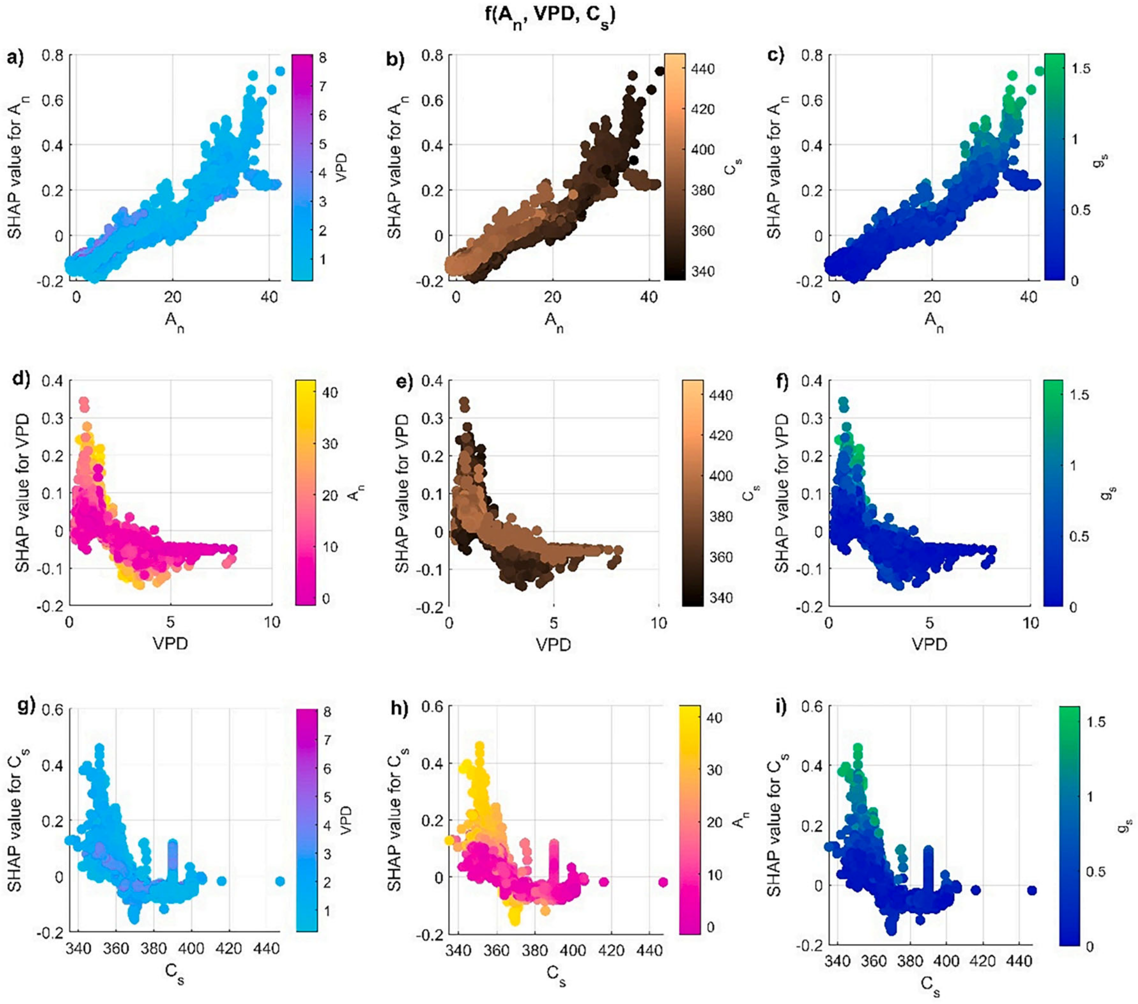


Fig. 9. SHAP dependence plots for the Leuning Proxy model [A_n, VPD, C_s]. Each panel is a 2-D scatter plot between the predictor and that predictor's SHAP value, with coloration to show the interaction with the other predictor variable, as well as g_s.

However, the relationship is not always linear at low photosynthetic flux density (PPFD), and g_s values measured at low PPFD can deviate significantly from g_s values measured at high PPFD (Collatz et al., 1991; Leuning et al., 1995; Ball, 1988; Barnard and Bauerle, 2013). To overcome this problem, we utilized an optimization technique to estimate the parameters of semi-empirical g_s models. ML proxies for conventional semi-empirical models outperformed the original semi-empirical models even after optimizing the parameters for the semi-empirical models (Figs. 3, 4 and S3). These findings suggest that ML may provide a powerful alternative for g_s prediction due to its non-parametric nature and capability to capture non-linear responses (Saunders et al., 2021; Gaur et al., 2020).

Earlier studies suggested the conventional semi-empirical models for g_s performed well when applied to single species at a specific site and subjected to reparameterization over different seasons, crop growth stages, and treatments (Liu et al., 2009; Miner and Bauerle, 2017; Wei et al., 2018; Qi et al., 2023). However, these models did not perform well when applied to different species as per the findings of Li et al. (2019) and Saunders et al. (2021). The relatively poor performance of the Ball–Berry and Leuning models in this study may be due in part to being applied to diverse collections of species at the PFT level, and for conditions that vary widely. This is not necessarily consistent with how they were originally formulated to be applied. That said these models are often applied in modern land surface models at the PFT level (Bonan et al., 2014; Jefferson et al., 2017), and the performance of these models that we have found in this study may be more demonstrative of the large-scale modeling context.

The Leuning ML proxy produces near-optimal predictive performance while constraining model complexity through the use of only three predictors, providing an example of a parsimonious formulation to model g_s.

We conducted a thorough evaluation of predictor sets and found several that provide excellent predictive skill while potentially being suitable for specific observational or simulation contexts, such as a predictor set that utilizes leaf temperature (which can be remotely sensed) and two environmental variables ([T_L, PAR_{in}, VPD]). Due to their non-parametric nature, these ML models overcome the parameterization / re-parameterization issues associated with standard semi-empirical models, and so offer a path forward for robust, flexible plant physiological modeling.

Evaluating the different sets of predictors provided an opportunity to develop some useful models utilizing non-conventional sets of predictors to predict g_s based on the synergy between two or more predictors. We note here the importance of considering potential multicollinearity between predictors when developing predictor sets for ML model formulations (Belsley et al., 1980). For example, foliar gas exchange experiments use transpiration rate to calculate g_s, and so transpiration rate should not be used as a predictor to model g_s.

By evaluating ML models for g_s across multiple PFTs, we demonstrated the flexibility of ML in modeling g_s responses across climate regions and different PFTs. The performance of ML models during validation (Figs. 5 and S3) suggests that model overfitting issues can be overcome to a considerable extent by performing robust cross-validation of ML models. This study also addressed the predictive uncertainty associated with g_s prediction (Fig. S5). Although, the uncertainty analysis cannot remove the uncertainty but certainly provides an idea about the credibility of modeling predictions in case of limited data availability.

While this study focuses on the application of machine learning to predict g_s, the general methodology for model development and robust CV can be used as a standard framework for developing and exploring non-parametric ML models of various environmental phenomena.

The findings of the study suggest that considering plant physiological function as a potential predictor to model g_s is important and can improve the performance of ML models. Finally, we note that the application of eXML (SHAP) here demonstrated the consistency of ML models with our understanding of vegetation ecophysiological function, providing confidence in and understanding of the operation of these ML models.

CRediT authorship contribution statement

Srishti Gaur: Conceptualization, Formal analysis, Investigation, Visualization, Methodology, Software, Validation, Writing – original draft. **Darren T. Drewry:** Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

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

Data availability

The data used in this work is publicly available at the time of this work.

Acknowledgments

The authors acknowledge support from the US National Science Foundation (Award #1954556). DTD also acknowledges support from the National Aeronautics and Space Administration (Award #80NSSC20K1789) as well the College of Food, Agricultural and Environmental Sciences at Ohio State University.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2024.109955.

References

Anderegg, W.R.L., Wolf, A., Arango-velez, A., Choat, B., Chmura, D.J., Jansen, S., et al., 2018. Woody plants optimise stomatal behaviour relative to hydraulic risk. Ecol. Lett. 21 (7), 968–977. https://doi.org/10.1111/ele.12962.

Anderegg, W.R.L., Wolf, A., Arango-Velez, A., Choat, B., Chmura, D.J., Jansen, S., et al., 2017. Plant water potential improves prediction of empirical stomatal models. PLOS One 12 (10), e0185481. https://doi.org/10.1371/journal.pone.0185481.

Aria, M., Cuccurullo, C., Gnasso, A., 2021. A comparison among interpretative proposals for Random Forests. Mach. Learn. Appl. 6, 100094 https://doi.org/10.1016/j.mlwa.2021.100094.

Ball, J.T., Woodrow, I.E., Berry, J.A., Biggins, J., 1987. A model predicting Stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions (Ed.). In: Progress in Photosynthesis Research: Volume 4 Proceedings of the VIIth International Congress on Photosynthesis. Springer Netherlands, Providence, Rhode Island, USA, pp. 221–224. August 10–15, 1986Dordrecht.

Ball, J.T., 1988. An Analysis of Stomatal Conductance. Stanford University.

Barnard, D.M., Bauerle, W.L., 2013. The implications of minimum stomatal conductance on modeling water flux in forest canopies. J. Geophys. Res. Biogeosci. 118, 1322–1333. https://doi.org/10.1002/jgrg.20112.

Belsley, D.A., Edwin, K., Roy, E.W., 1980. Regression diagnostics: identifying influential data and sources of collinearity. Wiley Series in Probability and Mathematical Statistics. Wiley. New York.

Bergstra, J.A., Klop, J.W., Nielsen, M., Schmidt, E.M., 2012. Algebraic specifications for parametrized data types with minimal parameter and target algebras. Automata, languages and programming. Algebraic specifications for parametrized data types with minimal parameter and target algebras (eds). In: Proceedings of the Automata, Languages and Programming. ICALP 1982. Lecture Notes in Computer Science, 140. Springer. https://doi.org/10.1007/BFb0012754. Berlin, Heidelberg.

Biecek, P., Burzykowski, T., 2020. Explanatory model analysis explore, explain and examine predictive models 2020. [online] Available from: https://pbiecek.github.io/ema/.

Bonan, G., 2019. Stomatal conductance. Climate Change and Terrestrial Ecosystem Modeling. Cambridge University Press, pp. 189–212.

Bonan, G.B., Williams, M., Fisher, R.A., Oleson, K.W., 2014. Modeling stomatal conductance in the earth system: linking leaf water-use efficiency and water transport along the soil–plant–atmosphere continuum. Geosci. Model Dev. 7, 2193–2222. https://doi.org/10.5194/gmd-7-2193-2014.

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

Brodeur, Z.P., Herman, J.D., Steinschneider, S., 2020. Bootstrap aggregation and cross-validation methods to reduce overfitting in reservoir control policy search. Water Resour. Res. 56 (8), e2020WR027184 https://doi.org/10.1029/2020WR027184.

Buckley, T.N., 2017. Modeling stomatal conductance. Plant Physiol. 174 (2), 572–582. https://doi.org/10.1104/pp.16.01772.

Buckley, T.N., Mott, K.A., 2013. Modelling stomatal conductance in response to environmental factors. Plant Cell Environ. 36 (9), 1691–1699. https://doi.org/10.1111/pce.12140.

Cai, Y., Guan, K., Lobell, D., Potgieter, A.B., Wang, S., Peng, J., et al., 2019. Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. Agric. For. Meteorol. 274, 144–159. https://doi.org/10.1016/j.agrformet.2019.03.010.

Chen, T., Guestrin, C., 2016. Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd Association for Computing Machinery Special Interest Group on Knowledge Discovery in Data (ACM SIGKDD) international conference, pp. 785–794.

Chen, J.M., Liu, J., 2020. Evolution of evapotranspiration models using thermal and shortwave remote sensing data. Remote Sens. Environ. 237, 111594 https://doi.org/10.1016/j.rse.2019.111594.

Chitsaz, N., Guan, H., Shanafield, M., Batelaan, O., 2023. Evaluating CO_2 effects on semi-empirical and empirical stomatal conductance simulation in land surface models. J. Hydrol. 620, 129385 https://doi.org/10.1016/j.jhydrol.2023.129385.

Collatz, G.J., Ball, J.T., Grivet, C., Berry, J.A., 1991. Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. Agric. For. Meteorol. 54 (2–4), 107–136. https://doi.org/10.1016/0168-1923(91)90002-8.

Damour, G., Simonneau, T., Cochard, H., Urban, L., 2010. An overview of models of stomatal conductance at the leaf level. Plant Cell Environ. 33, 1419–1438. https://doi.org/10.1111/j.1365-3040.2010.02181.x.

de Hond, A.A.H., Leeuwenberg, A.M., Hooft, L., et al., 2022. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. npj Digit. Med. 5, 2. https://doi.org/10.1038/s41746-021-00549-7.

De Kauwe, M.G., Kala, J., Lin, Y., Pitman, A.J., Medlyn, B.E., Duursma, R.A., et al., 2015. A test of an optimal stomatal conductance scheme within the CABLE land surface model. Geosci. Model Dev. 8, 431–452. https://doi.org/10.5194/gmd-8-431-2015, 2015.

Deo, R.C., Şahin, M., 2015. Application of the extreme learning machine algorithm for the prediction of monthly effective drought index in eastern Australia. Atmos. Res. 153, 512–525.

Dewar, R.C., 2002. The Ball–Berry–Leuning and Tardieu–Davies stomatal models: synthesis and extension within a spatially aggregated picture of guard cell function. Plant Cell Environ 25, 1383–1398. https://doi.org/10.1046/j.1365-3040.2002.00909.x.

Ding, J., Jiao, X., Bai, P., Hu, Y., Zhang, J., Li, J., 2022. Effect of vapor pressure deficit on the photosynthesis, growth, and nutrient absorption of tomato seedlings. Sci. Hortic. 293, 110736. https://doi.org/10.1016/j.scienta.2021.110736.

Dinh, P., Pham-Quoc, C., Thinh, T.N., Nguyen, B.K.D, Kha, P.C., 2023. A flexible and efficient FPGA-based random forest architecture for IoT applications. Internet Things 22. https://doi.org/10.1016/j.iot.2023.100813.

Drewry, D.T., Kumar, P., Long, S., Bernacchi, C., Liang, X.Z., Sivapalan, M., 2010a. Ecohydrological responses of dense canopies to environmental variability: 1. Interplay between vertical structure and photosynthetic pathway. J. Geophys. Res. 115, G4. https://doi.org/10.1029/2010JG001340.

Drewry, D.T., Kumar, P., Long, S., Bernacchi, C., Liang, X., Sivapalan, M., 2010b. Ecohydrological responses of dense canopies to environmental variability: 2. Role of acclimation under elevated CO_2. J. Geophys. Res. 115, G4. https://doi.org/10.1029/2010JG001341.

Ellsäßer, F., Röll, A., Ahongshangbam, J., Waite, P., Hendrayanto, B., et al., 2020. Predicting tree sap flux and stomatal conductance from drone-recorded surface temperatures in a mixed agroforestry system—A machine learning approach. Remote Sens. 12, 4070. https://doi.org/10.3390/rs12244070.

Ellsworth, D.S., Thomas, R., Crous, K.Y., Palmroth, S., Ward, E., Maier, C., DeLucia, E., Oren, R., 2012. Elevated CO_2 affects photosynthetic responses in canopy pine and subcanopy deciduous trees over 10 years: a synthesis from Duke FACE. Glob. Change Biol. 18, 223–242. https://doi.org/10.1111/j.1365-2486.2011.02505.x.

Ely, K.S., Rogers, A., Agarwal, D.A., Ainsworth, E.A., Albert, L.P., et al., 2021. A reporting format for leaf-level gas exchange data and metadata. Ecol. Inform. 61, 101232 https://doi.org/10.1016/j.ecoinf.2021.101232.

Isabona, J., Imoize, A.L., Ojo, S., Karunwi, O., Kim, Y., Lee, C.C., Li, C.T., 2022. Development of a multilayer perceptron neural network for optimal predictive modeling in urban microcellular radio environments. Appl. Sci. 12 (11), 5713. https://doi.org/10.3390/app12115713.

Fawagreh, K., Gaber, M.M., Elyan, E., 2014. Random forests: from early developments to recent advancements. Syst. Sci. Control. Eng. 2 (1), 602–609. https://doi.org/10.1080/21642583.2014.956265.

Franks, P.J., Bonan, G.B., Berry, J.A., Lombardozzi, D.L., Holbrook, N.M., Herold, N., et al., 2018. Comparing optimal and empirical stomatal conductance models for application in Earth system models. Glob Chang Biol. 24, 5708–5723. https://doi.org/10.1111/gcb.14445.

Fuadah, Y.N., Pramudito, M.A., Lim, K.M., 2022. An optimal approach for heart sound classification using grid search in hyperparameter optimization of machine learning. Bioengineering 10 (1), 45. https://doi.org/10.3390/bioengineering10010045.

Garen, J.C., Branch, H.A., Borrego, I., Blonder, B., Stinziano, J.R., Michaletz, S.T., 2022. Gas exchange analysers exhibit large measurement error driven by internal thermal gradients. New Phytol. 236, 369–384. https://doi.org/10.1111/nph.18347.

Gaur, S., Singh, R., Bandyopadhyay, A., Singh, R., 2023. Diagnosis of GCM-RCM-driven rainfall patterns under changing climate through the robust selection of multi-model ensemble and sub-ensembles. Clim. Change 176, 13. https://doi.org/10.1007/s10584-022-03475-z.

Gaur, S., Bandyopadhyay, A., Singh, R., 2021. From changing environment to changing extremes: exploring the future streamflow and associated uncertainties through integrated modelling system. Water Resour. Manag. 35, 1889–1911. https://doi.org/10.1007/s11269-021-02817-3.

Gaur, S., Mittal, A., Bandyopadhyay, A., Holman, I., Singh, R., 2020. Spatio-temporal analysis of land use and land cover change: a systematic model inter-comparison driven by integrated modelling techniques. Int. J. Remote Sens. 41 (23), 9229–9255. https://doi.org/10.1080/01431161.2020.1815890.

Hetherington, A.M., Woodward, F.I., 2003. The role of stomata in sensing and driving environmental change. Nature 424, 901–908. https://doi.org/10.1038/nature01843.

Hou, Y.-K., Chen, H., Xu, C.-Y., Chen, J., Guo, S.-L., 2017. Coupling a Markov chain and support vector machine for at-site downscaling of daily precipitation. J. Hydrometeorol. 18, 2385–2406.

Houshmandfar, A., Leary, G., Fitzgerald, G.J., Chen, Y., Tausz-Posch, S., Benke, K., et al., 2021. Machine learning produces higher prediction accuracy than the Jarvis-type model of climatic control on stomatal conductance in a dryland wheat agro-ecosystem. Agric. For. Meteorol. 304-305, 108423 https://doi.org/10.1016/j.agrformet.2021.108423.

Hsieh, W.W., 2022. Evolution of machine learning in environmental science—A perspective. Environ. Data Sci. 1, E3.

Ivakhnenko, A.G., Lapa, V.G., 1965. Cybernetic Predicting Devices. CCM Information Corporation.

Jarvis, P.G., 1976. The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. Phil. Trans. R. Soc. Lond. 610, B273593 https://doi.org/10.1098/rstb.1976.0035.

Jefferson, J.L., Maxwell, R.M., Constantine, P.G., 2017. Exploring the sensitivity of photosynthesis and stomatal resistance parameters in land surface model. J. Hydrometeor. 19, 897–915.

Kasraei, B., Heung, B., Saurette, D.D., Schmidt, M.G., Bulmer, C.E., Bethel, W., 2021. Quantile regression as a generic approach for estimating uncertainty of digital soil maps produced from machine-learning. Environ. Model. Softw. 144, 105139 https://doi.org/10.1016/j.envsoft.2021.105139.

Kim, S.H., Lieth, J.H., 2003. Coupled model of photosynthesis, stomatal conductance and transpiration for a rose leaf (Rosa hybrida L.). Ann. Bot. 91, 771–781.

King, R.D., Orhobor, O.I., Taylor, C.C., 2021. Cross-validation is safe to use. Nature Machine Intelligence, 3, p. 276. 10.1038/s42256-021-00332-z.

Ko, J., Piccinni, G., 2009. Characterizing leaf gas exchange responses of cotton to full and limited irrigation conditions. Field Crops Res. 112, 77–89.

Koenker, R., 2005. Quantile Regression. Cambridge University Press.

Koenker, R., Hallock, K.F., 2001. Quantile regression. J. Econ. Perspect. 15 (4), 143–156.

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer, New York. https://doi.org/10.1007/978-1-4614-6849-3.

Kusumi, K., Hirotsuka, S., Kumamaru, T., Iba, K., 2012. Increased leaf photosynthesis caused by elevated stomatal conductance in a rice mutant deficient in SLAC1, a guard cell anion channel protein. J. Exp. Bot. 63, 5635–5644. https://doi.org/10.1093/jxb/ers216.

Leuning, R., 1990. Modelling stomatal behaviour and photosynthesis of Eucalyptus grandis. Aust. J. Plant Physiol. 17, 159–175. https://doi.org/10.1071/PP9900159.

Leuning, R., Kelliher, F.M., de Pury, D.G.G., Schulze, E.D., 1995. Leaf nitrogen, photosynthesis, conductance and transpiration: scaling from leaves to canopies. Plant Cell Environ. 18, 1183–1200. https://doi.org/10.1111/j.1365-3040.1995.tb00628.x.

Leuning, R., 2002. Temperature dependence of two parameters in a photosynthesis model. Plant Cell Environ. 25, 1205–1210. https://doi.org/10.1046/j.1365-3040.2002.00898.x.

Li, J., Zhang, G., Li, X., Wang, Y., Wang, F., Li, X., 2019. Seasonal change in response of stomatal conductance to vapor pressure deficit and three phytohormones in three tree species. Plant Signal. Behav. 14, 12. https://doi.org/10.1080/15592324.2019.1682341.

Liao, Q., Gu, S., Kang, S., Du, T., Tong, L., Wood, J.D., et al., 2022. Mild water and salt stress improve water use efficiency by decreasing stomatal conductance *via* an osmotic adjustment in field maize. Sci. Total Environ. 805, 150364 https://doi.org/10.1016/j.scitotenv.2021.150364.

Lin, Y., Medlyn, B.E., Duursma, R.A., Prentice, I.C., Wang, H., Baig, S., et al., 2015. Optimal stomatal behaviour around the world. Nat. Clim. Chang. 5, 459–464.

Liu, F., Andersen, M.N., Jensen, C.R., 2009. Capability of the 'Ball–Berry' model for predicting stomatal conductance and water use efficiency of potato leaves under different irrigation regimes. Sci. Hortic. 122 (3), 346–354. https://doi.org/10.1016/j.scienta.2009.05.026.

Loranty, M.M., Mackay, D.S., Ewers, B.E., Traver, E., Gruger, E.L., 2010. Contribution of competition for light to within-species variability in stomatal conductance. Water Resour. Res. 46, W05516. https://doi.org/10.1029/2009WR008125.

Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. In: Proceedings of the Advances in Neural Information Processing Systems, pp. 4765–4774.

Lundberg, S.M., Erion, G., Chen, H., et al., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2, 56–67. https://doi.org/10.1038/s42256-019-0138-9.

McMillan, H.K., Westerberg, I.K., Krueger, T., 2018. Hydrological data uncertainty and its implications. WIREs Water. 5, e1319. https://doi.org/10.1002/wat2.1319.

Medlyn, B.E., Duursma, R.A., Eamus, D., Ellsworth, D.S., Prentice, I.C., Barton, C.V.M., Crous, K.Y., De Angelis, P., Freeman, M., Wingate, L., 2011. Reconciling the optimal and empirical approaches to modelling stomatal conductance. Glob. Change Biol. 17, 2134–2144. https://doi.org/10.1111/j.1365-2486.2010.02375.x.

Miner, G.L., Bauerle, W.L., Baldocchi, D.D., 2017. Estimating the sensitivity of stomatal conductance to photosynthesis: a review. Plant Cell Environ. 40 (7), 1214–1238.

Miner, G.L., Bauerle, W.L., 2017. Seasonal variability of the parameters of the Ball–Berry model of stomatal conductance in maize (*Zea mays* L.) and sunflower (*Helianthus annuus* L.) under well-watered and water-stressed conditions. Plant Cell Environ. 40, 1874–1886. https://doi.org/10.1111/pce.12990.

Narbaev, T., Hazir, Ö., Khamitova, B., Talgat, S., 2023. A machine learning study to improve the reliability of project cost estimates. Int. J. Prod. Res. https://doi.org/10.1080/00207543.2023.2262051.

Newman, S.J., Furbank, R.T., 2021. Explainable machine learning models of major crop traits from satellite monitored continent-wide field trial data. Nat. Plants 7, 1354–1363. https://doi.org/10.1038/s41477-021-01001-0.

Plantin, F., Blatt, M.R., 2018. Stomatal response to humidity: blurring the boundary between active and passive movement. Plant Physiol. 176, 485–488.

Purcell, C., Batke, S.P., Yiotis, C., Caballero, R., Soh, W.K., Murray, M., et al., 2018. Increasing stomatal conductance in response to rising atmospheric CO_2. Ann. Bot. 121 (6), 1137–1149.

Qi, Y., Zhang, Q., Hu, S., Wang, R., Wang, H., Zhang, K., Zhao, H., Zhao, F., Chen, F., et al., 2023. Applicability of stomatal conductance models comparison for persistent water stress processes of spring maize in water resources limited environmental zone. Agric. Water Manag. 277, 108090 https://doi.org/10.1016/j.agwat.2022.108090.

Rahmati, O., Choubin, B., Fathabadi, A., Coulon, F., Soltani, E., Shahabi, H., et al., 2019. Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and UNEEC methods. Sci. Total Environ. 688, 855–866. https://doi.org/10.1016/j.scitotenv.2019.06.320.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., et al., 2019. Deep learning and process understanding for data-driven Earth system science. Nature 566, 195.

Sahoo, R., Pasayat, A.K., Bhowmick, B., Fernandes, K., Tiwari, M.K., 2021. A hybrid ensemble learning-based prediction model to minimise delay in air cargo transport using bagging and stacking. Int. J. Prod. Res. 60 (2), 644–660. https://doi.org/10.1080/00207543.2021.2013563.

Saunders, A., Drew, D.M., Brink, W., 2021. Machine learning models perform better than traditional empirical models for stomatal conductance when applied to multiple tree species across different forest biomes. Trees For. People 6, 100139. https://doi.org/10.1016/j.tfp.2021.100139.

Sarker, I.H., 2016. Machine learning: algorithms, real-world applications and research directions. SN Comput. Sci. 160.

Schmidhuber, J., 2015. Deep learning in neural networks: an overview. Neural Netw. 61, 85–117.

Schmidt, L., Heße, F., Attinger, S., Kumar, R., 2020. Challenges in applying machine learning models for hydrological inference: a case study for flooding events across Germany. Water Resour. 56, e2019WR025924 https://doi.org/10.1029/2019WR025924.

Schmidt, J., Marques, M.R.G., Botti, S., et al., 2019. Recent advances and applications of machine learning in solid-state materials science. npj Comput. Mater. 5, 83. https://doi.org/10.1038/s41524-019-0221-0.

Shapley, L., Kuhn, H., Tucker, A., 1953. A value for n-Person games. Eds. Contributions to the Theory of Games II. Princeton University Press, pp. 307–317. https://doi.org/10.1515/9781400881970-018.

Sobejano-Paz, V., Mo, X., Liu, S., Mikkelsen, T.N., He, L., Jin, H., García, M., 2023. Heat dissipation from photosynthesis contributes to maize thermoregulation under suboptimal temperature conditions. Cold Spring Harbor Laboratory bioRxiv. 10.1101/2023.01.27.525868.

Solomatine, D.P., Shrestha, D.L., 2009. A novel method to estimate model uncertainty using machine learning techniques. Water Resour. Res. 45 (2009), 12.

Tabe-Bordbar, S., Emad, A., Zhao, S.D., Sinha, S., 2018. A closer look at cross-validation for assessing the accuracy of gene regulatory networks and models. Sci. Rep. 8, 6620.

Tang, F., Ishwaran, H., 2017. Random forest missing data algorithms. Stat. Anal. Data Min ASA Data Sci. J. 10, 363–377. https://doi.org/10.1002/sam.11348.

Tyralis, H., Papacharalampous, G., Langousis, A., 2019. A brief review of random forests for water scientists and practitioners and their recent history in water resources. Water 11 (5), 910. https://doi.org/10.3390/w11050910 (Basel).

Urban, J., Ingwers, M.W., Mcguire, M.A., Teskey, R.O., 2017. Increase in leaf temperature opens stomata and decouples net photosynthesis from stomatal conductance in Pinus taeda and Populus deltoides x nigra. J. Exp. Bot. 68, 1757–1767.

Vela, D., Sharp, A., Zhang, R., et al., 2022. Temporal quality degradation in AI models. Sci. Rep. 12, 11654. https://doi.org/10.1038/s41598-022-15245-z.

Vialet-Chabrand, S., Lawson, T., 2019. Dynamic leaf energy balance: deriving stomatal conductance from thermal imaging in a dynamic environment. J. Exp. Bot. 70, 2839–2855.

Vincent, A.M., Jidesh, P., 2023. An improved hyperparameter optimization framework for AutoML systems using evolutionary algorithms. Sci. Rep. 13, 4737.

Vitrack-Tamam, S., Holtzman, L., Dagan, R., Levi, S., Tadmor, Y., Azizi, T., et al., 2020. Random forest algorithm improves detection of physiological activity embedded within reflectance spectra using stomatal conductance as a test case. Remote Sens. 12 (14), 2213. https://doi.org/10.3390/rs12142213.

Vu, M.T., Aribarg, T., Supratid, S., Raghavan, S.V., Liong, S.Y., 2016. Statistical downscaling rainfall using artificial neural network: significantly wetter Bangkok? Theor. Appl. Climatol. 126, 453–467.

Wang, B., Zheng, L., Liu, D.L., Ji, F., Clark, A., Yu, Q., 2018. Using multi-model ensembles of CMIP5 global climate models to reproduce observed monthly rainfall and temperature with machine learning methods in Australia. Int. J. Climatol. 38, 4891–4902. https://doi.org/10.1002/joc.5705.

Weerts, A.H., Winsemius, H.C., Verkade, J.S., 2011. Estimation of predictive hydrological uncertainty using quantile regression: examples from the National Flood Forecasting System (England and Wales). Hydrol. Earth Syst. Sci. 15, 255–265. https://doi.org/10.5194/hess-15-255-2011.

Weerasinghe, L.K., Creek, D., Crous, K.Y., Xiang, S., Liddell, M.J., Turnbull, M.H., Atkin, O.K., 2014. Canopy position affects the relationships between leaf respiration and associated traits in a tropical rainforest in Far North Queensland. Tree Physiol. 34 (6), 564–584. https://doi.org/10.1093/treephys/tpu016.

Wei, Z., Du, T., Li, X., Fang, L., Fulai, L., 2018. Simulation of stomatal conductance and water use efficiency of tomato leaves exposed to different irrigation regimes and air CO_2 concentrations by a modified "Ball–Berry" model. Front. Plant Sci. 9 https://doi.org/10.3389/fpls.2018.00445.

Xu, Z., Jiang, Y., Jia, B., Zhou, G., 2016. Elevated-CO_2 response of stomata and its dependence on environmental factors. Front. Plant Sci. 7, 657. https://doi.org/10.3389/fpls.2016.00657.

Yang, J., Medlyn, B.E., De Kauwe, M.G., Duursma, R.A., Jiang, M., Kumarathunge, D., Crous, K.Y., Gimeno, T.E., Wujeska-Klause, A., Ellsworth, D.S., 2020. Low sensitivity of gross primary production to elevated CO_2 in a mature eucalypt woodland. Biogeosciences 17, 265–279. https://doi.org/10.5194/bg-17-265-2020, 2020.

Yang, Y., Chui, T.F.M., 2021. Modeling and interpreting hydrological responses of sustainable urban drainage systems with explainable machine learning methods. Hydrol. Earth Syst. Sci. 25, 5839–5858. https://doi.org/10.5194/hess-25-5839-2021.