MDPI

Communication

Comparative Assessment of Sap Flow Modeling Techniques in European Beech Trees: Can Linear Models Compete with Random Forest, Extreme Gradient Boosting, and Neural Networks?

Paulína Nalevanková [1,*], Peter Fleischer, Jr. [1,2,3], Mohammad Mukarram [1], Zuzana Sitková [4] and Katarína Střelcová [1]

1   Faculty of Forestry, Technical University in Zvolen, T.G. Masaryka 24, 96001 Zvolen, Slovakia
2   Department of Plant Ecophysiology, Institute of Forest Ecology, Slovak Academy of Sciences, Štúrova 2, 96053 Zvolen, Slovakia
3   Administration of Tatra National Park, Tatranská Lomnica, 05960 Vysoké Tatry, Slovakia
4   National Forest Centre, Forest Research Institute, T.G. Masaryka 22, 96001 Zvolen, Slovakia
*   Correspondence: xnalevankova@tuzvo.sk; Tel.: +421-45-5206490

Abstract: Transpiration and sap flow are physiologically interconnected processes that regulate nutrient and water uptake, controlling major aspects of tree life. They hold special relevance during drought, where wrecked sap flow can undermine overall tree growth and development. The present study encompasses five-year (2012–2015 and 2017) sap flow datasets on European beech (*Fagus sylvatica*). Four different techniques were used for sap flow modeling, namely, a linear model (LM), random forest (RF), extreme gradient boosting machine (XGBM), and neural networks (NN). We used six variants (Variants 1–6) differing in the captured conditions and the dataset size. The 'prediction power' was the ratio of the predicted and observed sap flow. We found the LM had the maximum prediction power for the overall sap flow in beech trees with 1 h shift of global radiation. In the reaming variants, the LM provided comparable prediction power to RF and XGBM. At the same time, NN exhibited relatively poor prediction power over other machine learning models. The study supports an easier-to-apply and computationally simpler approach (LM) to assess sap flow over more sophisticated machine learning approaches (RF, XGBM, and NN).

Keywords: sap flow; *Fagus sylvatica*; machine learning; drought; modeling

1. Introduction

Transpiration creates a negative pressure gradient for water movement through a cohesion-tension approach in xylem. On the other hand, sap flow is the movement of water, nutrients, sugars, hormones, and other dissolved substances in the plant. Thus, sap flow has critical relevance in photosynthesis and plant development considering the nutrient and photosynthates movement. Sap flow is not limited to xylem vessels, unlike transpiration, yet they are tightly interconnected processes. Therefore, any disruption in tree–water relations compromises the transpiration–sap flow dialogue.

An extensive repercussion of climate change is exponentially worsening global drought events. Trees in particular are more susceptible to drought as they have higher size-driven water demands when compared with tiny plants. Tree roots cannot compensate for such water demand during drought scenarios [1]. Transpiration is essential to the hydrological cycle because it shifts substantial water from the soil to the atmosphere. Its precise depiction within Land Surface Schemes in climate models is vital to generate accurate and trustworthy climate projections [2]. Accurate measurement or estimation of plant transpiration or sap flow, respectively, is significant for understanding a plant's water use efficiency. It is estimated that more than half of terrestrial precipitation is transpired by forests and thus constitutes a crucial component of global water cycle [3]. Transpiration is

primarily controlled by stomata, internal plant resistances, local hydrological conditions, and climatic variations, especially evaporative demands [4–6]. Although it is crucial to quantify plant transpiration, accurate and direct measuring is technically demanding, time and money consuming, and labor intensive [7].

Several approaches, which resulted in empirical or physical models, have been widely used to estimate transpiration based on relevant predictors [8,9]. Among these, the SIM-DualKc, Shuttleworth–Wallace, and modified Jarvis–Stewart models are widely used to estimate transpiration rates [8]. However, the application of these mathematical models is still restricted due to their complicated parameterization and requirement of a vast amount of observational data. Machine learning models have been suggested as an alternative to conventional models because they do not require the knowledge of internal factors and can capture non-linear relationships [10]. Furthermore, such models can render an enriched sap flow understanding with precise predictions and optimization of the tree–water relation in the agricultural and ecological context.

The manuscript focuses on comparing various approaches for modeling sap flow, with an emphasis on different soil moisture conditions. The study employs machine learning methods, including random forest, gradient boosting machine, deep neural networks, and a multiple linear regression model to predict European beech (*Fagus sylvatica*) sap flow.

2. Materials and Methods

2.1. Study Site

The study was conducted in the central region of Slovakia (Kremnické vrchy Mts.) at an altitude of 450 m above sea level (48°36′43″ N; 19°03′59″ E). The experimental plot, Bienská dolina, was established in the spring of 2012 in a 65-year-old European beech forest. The study area is situated in an oak–beech forest zone (according to the classification of Zlatník et al. [11]), with soil classified as Haplic Cambisol (Humic, Eutric, Endoskeletic, Siltic) derived from volcanic parent material. The experimental plot is located on a slight east-facing slope at the lower edge of the distribution of beech in Slovakia, which is potentially more susceptible to drought. The study site has a slightly warm and moderately humid climate, with a mean air temperature of 16 °C in July. The long-term mean annual air temperature and total annual precipitation from 1961 to 1990 were 7.3 °C and 690 mm, respectively. The sample plot covered an area of 608 m^2, and the European beech trees within it had a diameter at breast height (DBH) ranging from 5.7 to 42.3 cm and tree heights ranging from 8.4 to 29.3 m. The selected trees were representative, with an average height of 26.3 with a standard deviation of 1.3 m (ranging from 24.7 to 29.1 m) and an average DBH of 32.4 cm with a deviation of 4.8 cm (ranging from 27.1 to 42.3 cm). Sap flow data from the growing seasons of 2012, 2013, 2014, 2015, and 2017 were utilized in this study. The data from 2016 were not included in the analysis due to incompleteness as a result of technical reasons. A detailed description of the Bienská dolina study site was provided by Sitková et al. [12].

2.2. Sap Flow Measurements

Sap flow systems (model EMS51A) connected to a 16-channel data logger RailBox V16 were installed on the selected 12 sample trees (*Fagus sylvatica* L.). The system utilized a tissue heat balance method (THB) based on the volume (three-dimensional) heating of the stem segment to measure volumetric sap flow directly in units of kg of water per a specific period and per one centimeter of stem circumference [13,14]. The sap flow was measured at 5 min intervals and recorded as 20 min averages in the data logger. For further analysis, hourly data were used. A more detailed description of the method and equipment used is given in the article by Sitková et al. [12]. This study used average sap flow values from 12 measured beech trees.

2.3. Monitoring of Environmental Conditions

The meteorological variables used for the models were measured in an open grass area using an automatic weather station. The station was equipped with sensors for air temperature (*AT*, °C), relative humidity (*RH*, %), and global radiation (*Rs*, W·m^{-2}) (EMS33 and EMS11; Environmental Measuring System (EMS Brno) Ltd., Brno, Czech Republic) placed at the height of 2 m above the ground (low cut grass). Precipitation (*P*, mm) was measured using a rain gauge type 370 with a collecting area of 320 cm^2 (1 m above ground, Met One Instruments Inc., Grants Pass, OR, USA). The measurements were recorded at intervals of 5 min, and data were stored every 20 min in the data logger EdgeBox V12 (EMS Brno Ltd., Brno, Czech Republic), powered by a 12 V solar-charged battery.

Soil water potential (*SWP*, MPa) was monitored using three calibrated gypsum blocs (Delmhorst Inc., Towaco, NJ, USA) and data logger MicroLog SP3 (EMS Brno Ltd., Czech Republic). The lowest measurable limit of the equipment was −1.5 MPa. The *SWP* measurements were taken at 15, 30, and 50 cm soil depths using six different soil probes across the experimental plot. In this study, we utilized the average SWP values of the research plot.

Based on the measured data, we calculated the derived variables that represent hypothetical atmospheric evaporative demands, i.e., vapor pressure deficit (*VPD*, kPa [15]) and potential evapotranspiration (*PET*, mm), according to Penman [16].

2.4. Model Development and Machine Learning

The statistical analyses and visualizations were conducted using the R programming language (version 4.1.3, R Core Team, Vienna, Austria) and the following packages: TensorFlow (version 2.11.0), Keras (version 2.11.1), Ranger (version 0.14.1), and XGBoost (version 1.7.3.1).

Before developing the models, the input data, which consisted of all measured meteorological variables (*Rs*, *AT*, *RH*, *P*, *PET*, *VPD*, and *SWP*), were centered and scaled. All data were randomly divided into training (60%) and validation (40%) subsets, with the validation subset used to evaluate model performance after training on the training dataset. Model performance was assessed using three metrics: the coefficient of determination (R^2), root-mean-square error (RMSE), and mean absolute deviation (MAD). The coefficient of determination is a statistical metric employed to evaluate how well a regression model fits the data. It quantifies the portion of the variation in the dependent variable that can be accounted for by the independent variables within the regression model. Ranging between 0 and 1, a higher value of R^2 suggests a stronger fit, with values closer to 1 indicating a more favorable fit. Root-mean-square error is commonly used to measure the average magnitude of the errors between predicted and actual values. It is calculated by taking the square root of the average of the squared differences between the predictions and actual values. Mean absolute deviation measures the average absolute difference between each predicted value and the actual value in a dataset. It provides a measure of the average absolute error of the predictions [17].

The modeling process utilized six variants, each incorporating different input data ranges to assess their impact on sap flow estimation. These variants aimed to explore the effect of varying conditions on plant transpiration.

The threshold of global radiation value was used to mitigate the effects of nighttime values on sap flow estimation. Additionally, wetter and drier periods were defined based on *SWP* values. At the same time, according to previous research, the effect of a one-hour shift in global radiation was introduced [15].

A description of all used variants is given in Table 1. Variant 1 contained all available data, while Variant 2 included all data with the global radiation data shifted by one hour. Variant 3 represented periods characterized by reduced water availability in the soil (*SWP* < −0.8 MPa) with a simultaneous displacement of *Rs* by one hour. Variant 4 utilized data from Variant 3 but only included data with *Rs* values higher than 200 W m^{-2} (daylight data). Variant 5 contained only daylight data from the wet period (*SWP* from 0 to

-0.4 MPa) with Rs shifted by one hour. Variant 6 utilized all daylight data ($Rs > 200$ Wm^2) with no shift in Rs.

Table 1. Data size, data manipulation, and filtering applied in Variants 1–6 along with the corresponding measured sap flow sums.

| Variants | Data Manipulation and Filtering | | | | |
	Rs **Shifted 1 h**	Rs **above 200 W m^{-2}**	*SWP* **Values**	**Data Size** n	**Sum of Sap Flow (kg cm^{-1})**
Variant 1	-	-	from 0 to -1.45 MPa	10318	192.2
Variant 2	yes	-	from 0 to -1.45 MPa	10318	192.2
Variant 3	yes	-	from -0.8 to -1.45 MPa (drier soil conditions)	3323	58.7
Variant 4	yes	yes	from -0.8 to -1.45 MPa (drier soil conditions)	1158	22.0
Variant 5	yes	yes	from 0 to -0.4 MPa (wetter soil conditions)	2332	50.1
Variant 6	-	yes	from 0 to -1.45 MPa	3488	160.3

For each dataset (Variants 1–6), four different models were used to predict sap flow: neural network, random forest, gradient boosting machine, and linear models. All applied models belong to supervised learning algorithms.

Before modeling sap flow by machine learning models, so-called hyperparameters were estimated. Hyperparameters for each model were tuned following the methodology used in our previous study [18]. We must be aware of the fact that different hyperparameters result in different model parameter estimates.

The neural network (NN) is a collection of algorithms that aims to identify underlying links in a data structure using a method that imitates how the human brain functions. For the analysis, the Keras package (version 2.9.0) and TensorFlow package (version 2.9.0) were used. The neural network architecture consisted of three hidden layers with 16, 32, and 16 neurons, respectively. Stochastic gradient descent was chosen as the optimization algorithm, with a batch size of 40 determining the number of training samples processed before updating internal parameters. The model was trained for 500 epochs and incorporated a dropout rate of 0.8 to prevent overfitting. A learning rate of 0.01 was employed to control parameter updates, and the activation function ReLU was chosen to introduce non-linearity and capture complex input representations.

Random forest (RF) is an ensemble of decision trees trained using the bagging approach, making up the "forest". The bagging method's central premise is that combining learning models improves the end outcome. The accuracy of the result grows as the number of trees increases. The random forest model was built using the Ranger method from the caret package (version 6.0-93). It consisted of 1000 decision trees, forming an ensemble to make predictions collectively. The mtry parameter was set to 9, indicating that 9 features were randomly sampled at each tree split. A minimal node size of 7 was specified, requiring a minimum number of observations to create a terminal node. The splitting rule employed was variance, a common criterion for assessing split quality during tree construction.

The extreme gradient boosting machine (XGBM) also utilizes decision trees as its basis, similar to random forest (RF). However, the key difference between the two is that RF utilizes averaging, while XGBM uses additive (ensemble) modeling. Furthermore, RF combines findings at the end of the process, while XGBM combines the results along the way [18]. For the XGBoost algorithm, the xgboost package (version 1.7.4.9) was utilized. A colsample_value of 0.9 was chosen, which randomly sampled 90% of features for each tree to promote diversity within the ensemble. An 'eta' value of 0.3, known as the learning rate, controlled the contribution of each tree to the final prediction. The 'max_depth' parameter

was set to 3, restricting the complexity of individual trees. A 'min_child_weight' of 5 was used, ensuring a minimum sum of instance weight in each child node. The 'n_estimators' parameter was set to 150, determining the total number of boosting rounds and allowing the model to learn complex relationships through the ensemble effect. A subsample value of 0.8 was employed, randomly sampling 80% of observations for each tree.

In addition, a multiple linear regression (LM) model was utilized as a benchmark in the study, alongside the RF, XGBM, and NN models. To evaluate the performance of these models in estimating daily *Fagus sylvatica* transpiration, three commonly used statistical indices, the coefficient of determination, root-mean-square error, and mean absolute deviation, were employed.

3. Results and Discussion

In this article, we compared the model performance and prediction power of the XGBM, NN, RF, and LM methods for predicting sap flow based on 5-year observations. We used metrics R^2, RMSE, and MAD to compare the performance of the various models (Variants 1–6). The overall prediction power is presented by the ratio between the modeled (predicted) and measured *SF* integrated over whole time intervals specific for each variant.

When inspecting sap flow in drier conditions, the performance of machine learning models (Figure 1: B2, C2, and D2; and Table 2: RMSE, R^2, and MAD in Variants 3 and 4) was superior to the linear model (Figure 1: A2), while, in wetter conditions, the overall performance of the linear model metrics was substantially increased (Figure 1: A1, B1, C1, and D1; Figure 2: Variant 5; Table 2: Variant 5).
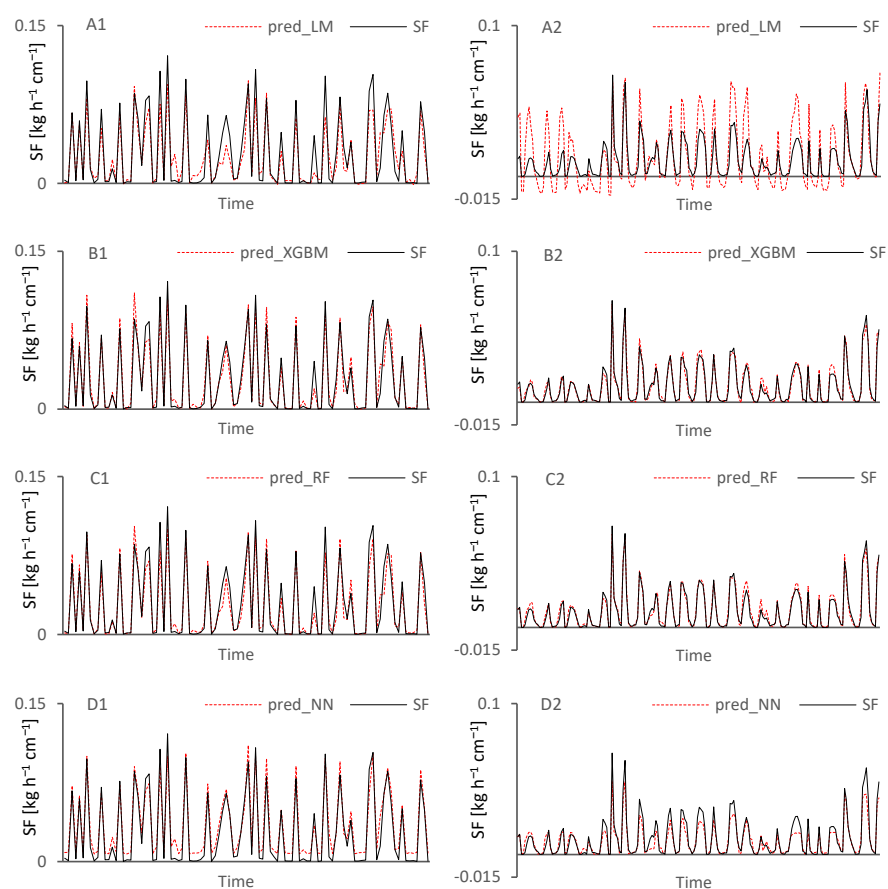

Figure 1. Comparison of hourly sap flow data modeled by linear model (*pred_LM*), extreme gradient boosting machine (*pred_XGBM*), random forest (*pred_RF*), and neural network (*pred_NN*) to measured sap flow (*SF*) in two different soil water potential ranges: from 0 to −0.4 MPa (left; **A1**, **B1**, **C1**, **D1**) and from −0.8 to −1.45 MPa (right; **A2**, **B2**, **C2**, **D2**).

Table 2. The performance of individual models for Variants 1, 2, 3, 4, 5, and 6 differed in *Rs* shift, *SWP* threshold, and/or light conditions. Real measured (*real SF*) and predicted sap flow (*pred. SF*) are expressed in kg cm^{-1} per measured/modeled period. Model performance was assessed by coefficient of determination (R^2), root-mean-square error (RMSE), and mean absolute deviation (MAD). Additionally, the prediction power is expressed as the ratio of predicted and measured sums of sap flow. *n* = data size.

| Variant 1			Method				
				NN	RF	XGBM	LM
model description	*n*	10318	RMSE	0.007	0.006	0.005	0.014
all available data used	*real SF*	192.2	R^2	0.937	0.954	0.970	0.764
Rs non-shifted			MAD	0.002	0.001	0.001	0.005
			pred. SF	211.0	192.8	192.9	194.8
			pred. SF/SF	1.1	1.0	1.0	1.0
Variant 2			Method				
				NN	RF	XGBM	LM
model description	*n*	10318	RMSE	0.010	0.006	0.005	0.014
all available data used	real *SF*	192.2	R^2	0.894	0.961	0.973	0.762
Rs 1 h shifted			MAD	0.005	0.001	0.001	0.006
			pred. SF	253.6	194.9	194.7	193.2
			pred. SF/SF	1.3	1.0	1.0	1.0
Variant 3			Method				
				NN	RF	XGBM	LM
model description	*n*	3323	RMSE	0.008	0.005	0.005	0.013
SWP below −0.8 MPa	*real SF*	58.7	R^2	0.887	0.954	0.960	0.713
Rs 1 h shifted			MAD	0.004	0.001	0.001	0.007
			pred. SF	59.0	58.5	58.2	59.2
			pred. SF/SF	1.0	1.0	1.0	1.0
Variant 4			Method				
				NN	RF	XGBM	LM
model description	*n*	1158	RMSE	0.010	0.007	0.006	0.014
SWP below −0.8 Mpa	*real SF*	22.0	R^2	0.834	0.917	0.937	0.683
Rs > 200 W m^{-2}			MAD	0.008	0.002	0.002	0.008
Rs 1 h shifted			*pred. SF*	21.0	22.1	22.2	22.1
			pred. SF/SF	1.0	1.0	1.0	1.0
Variant 5			Method				
				NN	RF	XGBM	LM
model description	*n*	2332	RMSE	0.010	0.007	0.006	0.014
SWP above −0.4 Mpa	*real SF*	50.1	R^2	0.901	0.956	0.963	0.815
Rs > 200 W m^{-2}			MAD	0.006	0.002	0.001	0.006
Rs 1 h shifted			*pred. SF*	59.2	49.7	49.8	49.7
			pred. SF/SF	1.2	1.0	1.0	1.0
Variant 6			Method				
				NN	RF	XGBM	LM
model description	*n*	3488	RMSE	0.018	0.009	0.007	0.021
Rs > 200 W m^{-2}	*real SF*	160.3	R^2	0.709	0.922	0.954	0.626
Rs non-shifted			MAD	0.015	0.005	0.004	0.014
			pred. SF	207.4	160.5	160.7	162.7
			pred. SF/SF	1.3	1.0	1.0	1.0
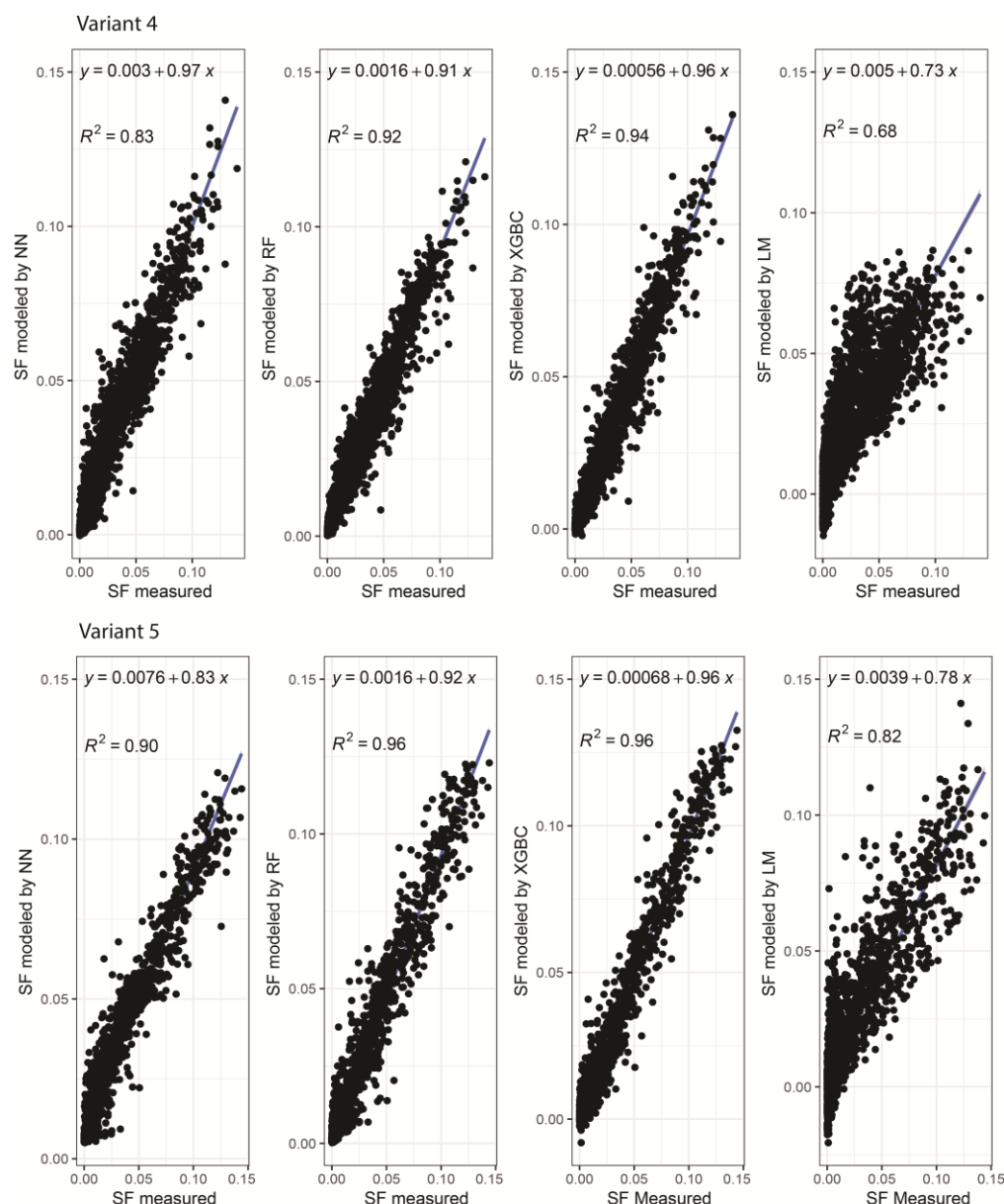
Figure 2. Measured vs. predicted sap flow (kg h^{-1} cm^{-1}) in dry (Variant 4, upper part) and wet conditions (Variant 5, lower part).

In our study, we observed relatively poor prediction power of the NN compared to the remaining algorithms (RF, XGBM, and LM), as shown in Table 2's predicted versus real measured sap flow values (*pred. SF/SF*). Despite the lower prediction power of the NN, the model performance metrics were comparable to those of RF and XGBM, mainly within Variants 3 and 4, hinting at a systematic bias when assessing sap flow values using the NN in wetter conditions (Variants 1, 2, 5, and 6). This contrasts with the fact that a NN easily outperforms other machine learning models such as XGBM and RF on datasets with distributed representation, such as pictures, voice, and text [19]. Sagawa et al. [20] pointed out that models that are very accurate on average can still perform poorly on rare and unusual examples.

Furthermore, we noted that all the models had a better performance when comparing between wet variants (*SWP* > −0.4 MPa) and dry variants (*SWP* < −0.8 MPa). The performance metrics were substantially better in Variant 5 (wet conditions) for LM (R^2 = 0.82) than in Variant 4 (dry conditions) (R^2 = 0.68) (Figure 2). The difference was also pronounced for the NN, R^2 = 0.90 and 0.83 for wet and dry conditions, respectively.

Better performance metrics, such as R^2, were also observed for RF and XGBM in Variant 5. The graph comparing the measured and modeled sap flow indicates minimized bias and more precise values with RF and XGBM (Figure 2). However, in Variant 5, we noticed that the *NN* tended to overestimate low sap flow values, particularly in the range of 0.0 to 0.5 kg h^{-1} cm^{-1}. While the model metrices were quite good, there was about 20% overestimation *pred. SF/SF* (Table 2). Overall, during wet conditions (Variant 5), it seemed that RF and XGBM were precise and unbiased in predicting sap flow and NN precise and biased, while the LM was imprecise and unbiased (Figure 2). Similar model behavior was noticed during the dry conditions of Variant 4 except the NN was relatively less biased.

We observed comparably good prediction power of the LM over RF and XGBM. Although the NN had better performance metrics in terms of R^2, RMSE, and MAD, it overestimated the sap flow for the projected period (Table 2). Therefore, it exhibited overall poor prediction power in most variants. Alternatively, the LM exhibited inferior values for the same metrics (i.e., lower R^2 and higher RMSE and MAD). Yet, it more appropriately predicted the amount of transpired water. The LM consistently overestimated the high values and underestimated the low values (e.g., Figure 1: A2), resulting in better or comparable prediction power better than or comparable to the NN and RF and XGBM models. Furthermore, when the global radiation was 1 h shifted, the LM surpassed the RF, XGBM, and NN models (Table 2: Variant 2). This was unexpected since these RF, XGBM, and NN models can cover non-linear relations; however, the LM is not capable of that. So, it seems that by incorporating systematic biases (over- and underestimation), the LM produces a more precise overall prediction.

A NN uses pattern matching to solve the regression problem. We noticed a conclusive systematic overestimation of about 10–30% in the predicted sap flow of Variants 1, 2, 5, and 6, as suggested by *pred. SF/SF* (Table 2). This could have occurred due to a positive skew of sap flow data and the fact that the NN was optimized by the squared loss function (mean squared error). This led to large estimating errors in the region of higher values, which were consequently weighted more strongly than many minor estimation errors in the region of low values, resulting in overall overestimation in presented models. When the data were non-skewed, the NN seemed to perform comparably to other algorithms (Variants 3 and 4).

Our results agree with the 'no free lunch' theorem, stating that no "perfect" machine learning algorithm can solve any problem successfully. Every issue has a specific solution that works well, while other solutions can fall short significantly. In our case, model performance substantially varied with the portion and nature of the data used for model creation.

It has long been known that certain ecosystems, species, or geographic areas exhibit a time lag between *SF* and environmental variables [20,21]. Generally, sap flow is delayed relative to solar radiation [12,22]. The shifting of climatic variables against sap flow (1 h) did not yield any pronounced enhancement in our models. Nonetheless, global radiation shifting significantly downgraded the NN's prediction power and performance metrics when we compared Variants 1 and 2 (Table 1). We also recorded a slight increase in the tree-based approaches (XGBM and RF).

In the study of Li et al. [23], the performance of six models for simulating sap flow in *Agathis australis* using a sizeable dataset was assessed. From these models, a linear model, a back-propagation neural network (BPNN), and a convolutional neural network (CNN) were comparable with our linear model (LM) and neural network (NN) model. Our models were parameterized using the Variant 1 dataset, where there was no data manipulation (non-shifted). The performance of our models displayed good agreement with those used in that study. Specifically, the R^2 value of our linear regression model was 0.764, closely matching the R^2 value of 0.796 reported by Li et al. [23]. Similarly, our neural network model displayed an R^2 value of 0.939, comparable to the *BPNN* and *CNN* models in this study, with R^2 values of 0.928 and 0.936, respectively. This comparison implies that the performance of our models aligns well with those used in above mentioned study.

When applied to describe transpiration in *Fagus sylvatica* (European Beech), the performance of our NN and RF models was significantly better than that of the NN and RF models used in the Wu et al. [24] study on maize. In Amir et al.'s [25] research on tomatoes, a variety of models was applied, including a gradient boosting model, which reached a maximum R^2 of 0.8. Tu et al.'s study [26], which employed a three-layer back-propagation NN, demonstrated that varying the number of predictors led to an R^2 range from 0.80 up to 0.95, aligning with our findings.

It appears that models perform better when applied to trees compared to low-growing plants. This observation is supported by Xing et al. [7] and Liu et al. [27]. Similar to the approach proposed by Tu et al. in 2019 [28], our model could potentially be improved by introducing a phenological function.

4. Conclusions

In our investigation, we documented a remarkably high prediction power exhibited by the linear model when applied to the prediction of long-term cumulative sap flow. This particular model demonstrated its competitiveness when compared to more intricate machine learning techniques, including random forest (RF), extreme gradient boosting (XGBM), and neural networks (NN) techniques. It seems that by incorporating systematic biases (over- and underestimation), the linear model produced a more precise overall prediction then the NN and one comparable to RF and XGBM. We can conclude, when predicting integrated sap flow over longer time periods by using the linear model, that we do not obtain a markedly biased estimate. When modeling the amounts of transpired water, it is not considered incorrect to use a much easier-to-apply and computationally simpler linear model. By predicting the courses of sap flow (hourly steps), the advanced machine learning techniques (NN, RF, and XGBM) are more suitable.

The generalization of our results should not go beyond *Fagus sylvatica* species in oak–beech altitudinal forest zones (classification according to Zlatník et al. [11]). These promising results should be used with a certain amount of caution until they are confirmed in different conditions or on other species of trees.

Author Contributions: Conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, P.N. and P.F.J.; writing—review and editing, M.M.; supervision, project administration, funding acquisition, Z.S. and K.S. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mukarram, M.; Choudhary, S.; Kurjak, D.; Petek, A.; Khan, M.M.A. Drought: Sensing, Signalling, Effects and Tolerance in Higher Plants. *Physiol. Plant.* **2021**, *172*, 1291–1300. [CrossRef] [PubMed]
2. Oogathoo, S.; Houle, D.; Duchesne, L.; Kneeshaw, D. Vapour Pressure Deficit and Solar Radiation Are the Major Drivers of Transpiration of Balsam Fir and Black Spruce Tree Species in Humid Boreal Regions, Even during a Short-Term Drought. *Agric. For. Meteorol.* **2020**, *291*, 108063. [CrossRef]
3. Schlesinger, W.H.; Jasechko, S. Transpiration in the Global Water Cycle. *Agric. For. Meteorol.* **2014**, *189*, 115–117. [CrossRef]
4. Eamus, D.; Boulain, N.; Cleverly, J.; Breshears, D.D. Global Change-Type Drought-Induced Tree Mortality: Vapor Pressure Deficit Is More Important than Temperature per Se in Causing Decline in Tree Health. *Ecol. Evol.* **2013**, *3*, 2711–2729. [CrossRef] [PubMed]

5. Lüttschwager, D.; Jochheim, H. Drought Primarily Reduces Canopy Transpiration of Exposed Beech Trees and Decreases the Share of Water Uptake from Deeper Soil Layers. *Forests* **2020**, *11*, 537. [CrossRef]

6. Zavadilová, I.; Szatniewska, J.; Petrík, P.; Mauer, O.; Pokorný, R.; Stojanović, M. Sap Flow and Growth Response of Norway Spruce under Long-Term Partial Rainfall Exclusion at Low Altitude. *Front. Plant Sci.* **2023**, *14*, 1089706. [CrossRef]

7. Xing, L.; Cui, N.; Liu, C.; Zhao, L.; Guo, L.; Du, T.; Zhan, C.; Wu, Z.; Wen, S.; Jiang, S. Estimation of Daily Apple Tree Transpiration in the Loess Plateau Region of China Using Deep Learning Models. *Agric. Water Manag.* **2022**, *273*, 107889. [CrossRef]

8. Fan, J.; Zheng, J.; Wu, L.; Zhang, F. Estimation of Daily Maize Transpiration Using Support Vector Machines, Extreme Gradient Boosting, Artificial and Deep Neural Networks Models. *Agric. Water Manag.* **2021**, *245*, 106547. [CrossRef]

9. Wang, H.; Guan, H.; Simmons, C.T. Modeling the Environmental Controls on Tree Water Use at Different Temporal Scales. *Agric. For. Meteorol.* **2016**, *225*, 24–35. [CrossRef]

10. Fan, J.; Yue, W.; Wu, L.; Zhang, F.; Cai, H.; Wang, X.; Lu, X.; Xiang, Y. Evaluation of SVM, ELM and Four Tree-Based Ensemble Models for Predicting Daily Reference Evapotranspiration Using Limited Meteorological Data in Different Climates of China. *Agric. For. Meteorol.* **2018**, *263*, 225–241. [CrossRef]

11. Zlatník, A. Overview of Groups of Types of Geobiocoenes Originally Forest and Shrubby. *Zprávy Geogr. Čsav Brno* **1976**, *13*, 55–56.

12. Sitková, Z.; Strelcová, K.; Jezík, M.; Sitko, R.; Pavlenda, P.; Hlásny, T. How Does Soil Water Potential Limit the Seasonal Dynamics of Sap Flow and Circumference Changes in European Beech? *Lesn. Cas.* **2014**, *60*, 19. [CrossRef]

13. Čermák, J.; Kučera, J.; Nadezhdina, N. Sap Flow Measurements with Some Thermodynamic Methods, Flow Integration within Trees and Scaling up from Sample Trees to Entire Forest Stands. *Trees* **2004**, *18*, 529–546. [CrossRef]

14. Kučera, J.; Čermák, J.; Penka, M. Improved Thermal Method of Continual Recording the Transpiration Flow Rate Dynamics. *Biol Plant* **1977**, *19*, 413–420. [CrossRef]

15. Nalevanková, P.; Sitková, Z.; Kučera, J.; Střelcová, K. Impact of Water Deficit on Seasonal and Diurnal Dynamics of European Beech Transpiration and Time-Lag Effect between Stand Transpiration and Environmental Drivers. *Water* **2020**, *12*, 3437. [CrossRef]

16. Penman, H.L.; Keen, B.A. Natural Evaporation from Open Water, Bare Soil and Grass. *Proc. R. Soc. London. Ser. A. Math. Phys. Sci.* **1948**, *193*, 120–145. [CrossRef]

17. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 112.

18. Leštianska, A.; Fleischer, P.; Merganičová, K.; Fleischer, P.; Nalevanková, P.; Střelcová, K. Effect of Provenance and Environmental Factors on Tree Growth and Tree Water Status of Norway Spruce. *Forests* **2023**, *14*, 156. [CrossRef]

19. Shavitt, I.; Segal, E. Regularization Learning Networks: Deep Learning for Tabular Datasets. In *Proceedings of the Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2018; Volume 31.

20. Sagawa, S.; Koh, P.W.; Hashimoto, T.B.; Liang, P. Distributionally Robust Neural Networks for Group Shifts: On the Importance of Regularization for Worst-Case Generalization. *arXiv* **2019**, arXiv:1911.08731.

21. Hayat, M.; Zha, T.; Jia, X.; Iqbal, S.; Qian, D.; Bourque, C.P.-A.; Khan, A.; Tian, Y.; Bai, Y.; Liu, P.; et al. A Multiple-Temporal Scale Analysis of Biophysical Control of Sap Flow in Salix Psammophila Growing in a Semiarid Shrubland Ecosystem of Northwest China. *Agric. For. Meteorol.* **2020**, *288–289*, 107985. [CrossRef]

22. Zhang, R.; Xu, X.; Liu, M.; Zhang, Y.; Xu, C.; Yi, R.; Luo, W.; Soulsby, C. Hysteresis in Sap Flow and Its Controlling Mechanisms for a Deciduous Broad-Leaved Tree Species in a Humid Karst Region. *Sci. China Earth Sci.* **2019**, *62*, 1744–1755. [CrossRef]

23. Li, Y.; Ye, J.; Xu, D.; Zhou, G.; Feng, H. Prediction of Sap Flow with Historical Environmental Factors Based on Deep Learning Technology. *Comput. Electron. Agric.* **2022**, *202*, 107400. [CrossRef]

24. Wu, Z.; Cui, N.; Gong, D.; Zhu, F.; Xing, L.; Zhu, B.; Chen, X.; Wen, S.; Liu, Q. Simulation of Daily Maize Evapotranspiration at Different Growth Stages Using Four Machine Learning Models in Semi-Humid Regions of Northwest China. *J. Hydrol.* **2023**, *617*, 128947. [CrossRef]

25. Amir, A.; Butt, M.; Van Kooten, O. Using Machine Learning Algorithms to Forecast the Sap Flow of Cherry Tomatoes in a Greenhouse. *IEEE Access.* **2021**, *9*, 154183–154193. [CrossRef]

26. Tu, J.; Liu, Q.; Wu, J. Recognition of Dominant Driving Factors behind Sap Flow of Liquidambar Formosana Based on Back-Propagation Neural Network Method. *Ann. Forest Sci.* **2021**, *78*, 95. [CrossRef]

27. Liu, X.; Kang, S.; Li, F. Simulation of Artificial Neural Network Model for Trunk Sap Flow of Pyrus Pyrifolia and Its Comparison with Multiple-Linear Regression. *Agric. Water Manag.* **2009**, *96*, 939–945. [CrossRef]

28. Tu, J.; Wei, X.; Huang, B.; Fan, H.; Jian, M.; Li, W. Improvement of Sap Flow Estimation by Including Phenological Index and Time-Lag Effect in Back-Propagation Neural Network Models. *Agric. Forest Meteorol.* **2019**, *276–277*, 107608. [CrossRef]