PREDICTING THE VESSEL LUMEN AREA TREE-RING PARAMETER OF *QUERCUS ROBUR* WITH LINEAR AND NONLINEAR MACHINE LEARNING ALGORITHMS

JERNEJ JEVŠENAK[1], SAŠO DŽEROSKI[2, 3] and TOM LEVANIČ[1]

[1]*Slovenian Forestry Institute, Department of Forest Yield and Silviculture, Večna pot 2, 1000 Ljubljana, Slovenia*
[2]*Jožef Stefan Institute, Department of Knowledge Technologies, Jamova cesta 39, 1000 Ljubljana, Slovenia*
[3]*Jožef Stefan International Postgraduate School, Jamova cesta 39, 1000 Ljubljana, Slovenia*

Abstract: Climate-growth relationships in *Quercus robur* chronologies for vessel lumen area (VLA) from two oak stands (QURO-1 and QURO-2) showed a consistent temperature signal: VLA is highly correlated with mean April temperature and the temperature at the end of the previous growing season. QURO-1 showed significant negative correlations with winter sums of precipitation. Selected climate variables were used as predictors of VLA in a comparison of various linear and nonlinear machine learning methods: Artificial Neural Networks (ANN), Multiple Linear Regression (MLR), Model Trees (MT), Bagging of Model Trees (BMT) and Random Forests of Regression Trees (RF). ANN outperformed all the other regression algorithms at both sites. Good performance also characterised RF and BMT, while MLR, and especially MT, displayed weaker performance. Based on our results, advanced machine learning algorithms should be seriously considered in future climate reconstructions.

Keywords: dendroclimatology, artificial neural networks, multiple linear regression, machine learning, vessel lumen area, *Quercus robur*.

1. INTRODUCTION

Long systematic instrumental records of climate data are rare and spatially limited (for a review see Jones and Bradley, 1992). Natural archives are therefore used to extract information about the climate in the past. Tree-rings are among the most commonly used proxy records, especially in temperate zones, where seasonal growth results in annually resolved tree-rings.

The standard methodology which is usually used to study the relationship between climate (e.g., temperature, precipitation, moisture, cloudiness *etc.*) and tree-ring

proxies is to apply (multiple) linear regression (MLR; Cook and Kairiukstis, 1992). However, while linear transfer functions usually fit data well, there are some concerns about predictions for data points close to the edge of observed data and points out of the calibration data. Linear models assume that the dependency between tree-ring parameters and climate changes linearly from the most favourable to the most unfavourable conditions. This assumption is contradictory to the well-accepted concept of ecological amplitudes (Braak and Gremmen, 1987; Fritts, 1976), and trees can thus grow only in a limited range of growing conditions. The relationship between tree-ring parameters and environmental factors should, therefore, change when approaching the edges of ecological amplitudes.

There is growing evidence that nonlinear techniques are better at describing the relationship between tree-ring

Corresponding author: T. Levanič
e-mail: tom.levanic@gozdis.si

proxies and climate variables (Balybina, 2010; D'Odorico *et al.*, 2000; Helama *et al.*, 2009; Jevšenak and Levanič, 2016; Ni *et al.*, 2002; Zhang *et al.*, 2000). Sun *et al.* (2017) showed a significant improvement over linear regression by using linear spline regression and a likelihood-based model to explain nonlinearity in the relationship between tree rings and precipitation. One of the most accurate and tested process-based models, the Vaganov-Shaskin model (VS model, Shishov *et al.*, 2016; Vaganov *et al.*, 2011), uses a piecewise linear function to model the relationship between daily environmental data and tree-rings. A similar idea is implemented in a simplified version of the VS model – the VS-Lite model (Tolwinski-Ward *et al.*, 2011). The advantages of using nonlinear models have also been demonstrated in other studies using tree-ring data, such as predicting resilience to disturbance (Billings *et al.*, 2015), modelling tree growth response after partial cuttings (Montoro Girona *et al.*, 2017) and predicting the response of tree growth to climate change (Williams *et al.*, 2010).

In this article, we compare various regression methods. To this end, we used Vessel Lumen Area (VLA) tree-ring chronologies of pedunculate oak (*Quercus robur* L.) from two lowland sites in Slovenia. VLA is gaining importance in dendroclimatological studies, especially at mesic sites, where earlywood conduits often have a better climatic signal than traditional tree-ring width (Campelo *et al.*, 2010; Fonti and Garcia-Gonzalez, 2008; Jevšenak and Levanič, 2015).

The formation of earlywood vessels usually starts at the end of March or early April (Goršić, 2013; Gričar, 2010), when the temperature or moisture threshold, *i.e.*, the degree-day sum, is surpassed. This threshold is site- and species-specific. The period of most intense xylem cell production in oaks from nearby areas is April-May (Gričar, 2010), when the relationship between earlywood conduits and climate is more or less linear. Again, in the period from middle May-June onward, we expect saturation of the function that describes the relationship between earlywood vessels and climate. The two functions that could model such relationships more accurately (as compared to linear regression) are the sigmoid-shaped and piecewise linear functions. Both were applied in our study.

The goals of our study were 1) to analyse the relationship between VLA and climate, 2) to select predictors of VLA based on the results of the climate-growth relationship and 3) to apply those predictors to compare MLR and four different nonlinear modelling techniques (ANN, MT, BMT and RF).

2. MATERIAL AND METHODS

Study objects and wood-anatomical analysis

The two VLA chronologies used in our study were sampled from two lowland forest sites in Slovenia (**Fig. 1**). The selected sites are dominated by *Quercus robur*,

but differ in some site-specific conditions (see **Table 1**). The two sites have approximately equal temperature distributions, while the QURO-2 site (meteorological station Ljubljana, **Fig. 1C**), gets slightly more precipitation than the QURO-1 site (meteorological station Maribor; **Fig. 1D**). Climate data for both meteorological stations were provided by the Slovenian Environment Agency (ARSO, 2017).
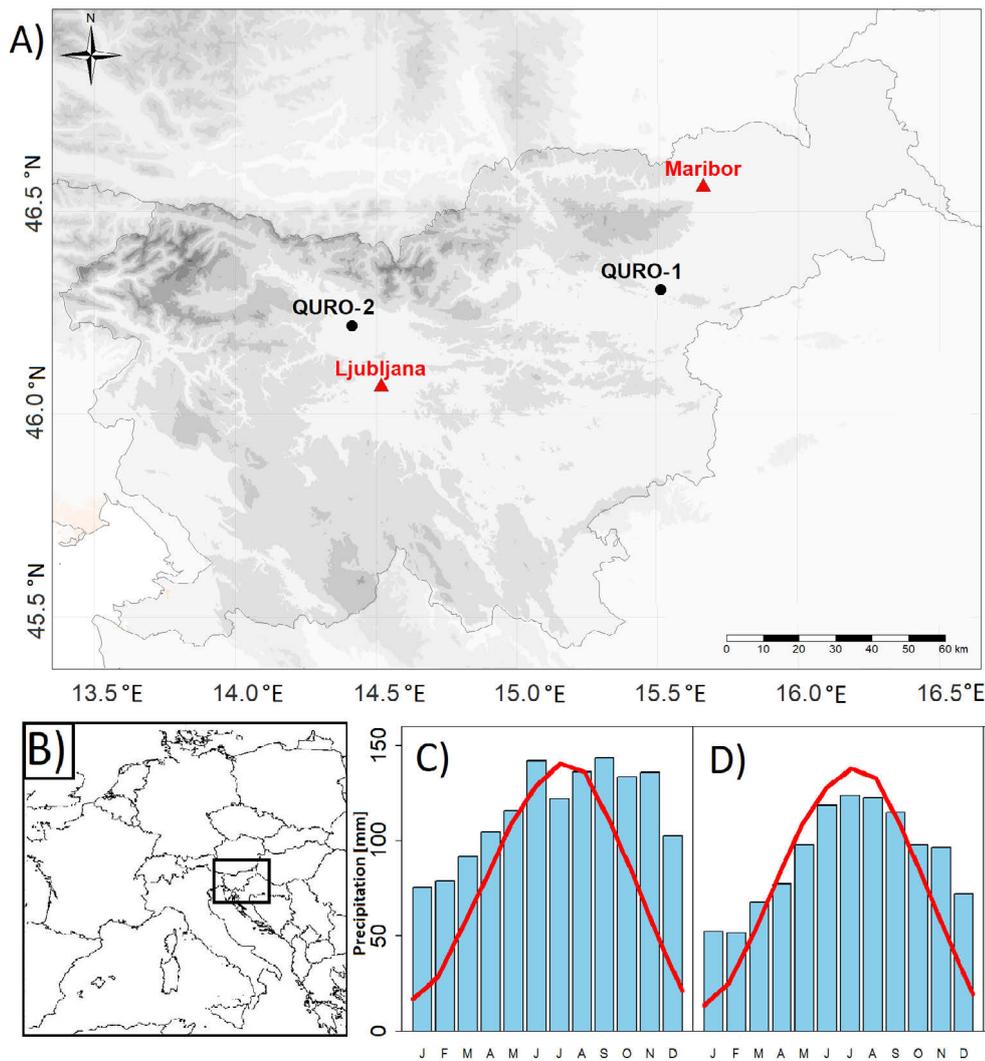
From each site, 6–8 mature and dominant trees (see **Table 2**, **Fig. 2**) were sampled for wood-anatomical analysis with a 10-mm increment borer. One core per tree was taken at breast height of 1.30 m. The extracted wooden cores were air dried, radially oriented and fitted in wooden holders. They were then sanded to a high polish in a laboratory, as suggested by Fonti *et al.* (2010). The sanding process starts with coarse 80-grit and ends with fine 1000-grit sanding paper. After sanding, vessels were filled with chalk, and high-quality images were taken with the ATRICS system (Levanič, 2007). ATRICS is an automated image acquisition system based on a motorised stage and a high-resolution microscope camera, which captures images and stitches them into a single, long image of the entire core. The captured images were analysed using the ImageJ program (Schindelin *et al.*, 2012), combined with macro EWVA (**Fig. 3**; Jevšenak and Levanič, 2014). First, the contrast between the vessels and the adjacent surface was enhanced, based on the selected black and white threshold. Each group of vessels was then recognised as a consecutive year and, finally, vessel lumens were measured. In each tree-ring, all vessel lumens were measured, and the parameter vessel lumen area (VLA) was subsequently calculated, considering only vessels with an area greater than 10,000 μm^2. The same threshold was used by Fonti and Garcia-Gonzalez (2008). VLA chronologies were not detrended, since we did not observe any age-related trend (see **Fig. 4**).

Nonlinear machine learning methods

The nonlinear methods used in our study belong to the field of machine learning. We used a single layer perceptron as an artificial neural network (ANN) (Bishop, 1995; Hastie *et al.*, 2009), which we trained with the Bayesian regularization training algorithm (Burden and Winkler, 2008). The described ANN method is available in the *brnn* R package (Pérez-Rodríguez and Gianola, 2016; Pérez-Rodríguez *et al.*, 2013) and usually fits a sigmoid-shaped transfer function.

Model trees (MT; Quinlan, 1992) are tree-structured models in which each terminal node in the tree contains a prediction in the form of a linear equation. A model tree is equivalent to a piecewise linear function and, when visualised, offers some interpretable features, such as the relative importance of predictors and the different linear equations used in the different terminal nodes.

In order to further improve the predictive capabilities of MT, bagged model trees (BMT) were used, which are derived by the ensemble method of bagging (Breiman,

Fig. 1. *A) Locations of the research sites (circles) and the meteorological stations (triangles), B) location of Slovenia in Europe and climagraphs of C) Ljubljana and D) Maribor. The lines of climagraphs represent mean monthly temperature and the blocks represent monthly amounts of precipitation.*

Table 1. *General description of the analysed sites.*

Location / Site denotation	Year of sampling	Latitude	Longitude	Elevation (m)	Bedrock	Forest soil type	Meteorological station	Average age of measured tree
Mlace / QURO-1	2012	N46°18'21"	E15°30'35"	280–315	Marl	Eutric brown soil	Maribor	150
Sorsko Plain / QURO-2	2015	N46°11'23"	E14°25'26"	360–365	Alluvial loams and clays	Dystric brown soil	Ljubljana	90

1996). Bagging involves the construction of several bootstrap samples of data obtained through random sampling with a replacement, and learning an MT from each bootstrap sample. Combining the predictions of individual MTs in the ensemble prevents overfitting the data.

Random Forests (RF; Breiman, 2001; Ho, 1995) of regression trees is an ensemble method that applies a randomised regression tree construction algorithm to each of the bootstrap samples. In regression trees, the predictions in the terminal nodes are real value constants, rather than linear models. The randomised algorithm considers only a small number of randomly selected predictors at each step of tree construction. We used the MT, BMT and RF algorithms for learning from the *RWeka* R package (Hornik *et al.*, 2009).

ANN, MT, BMT and RF are advanced algorithms used in the field of machine learning. These algorithms can also find linear relationships, in cases in which the dependency to be modelled is linear. We hypothesised that nonlinear techniques would provide better models,
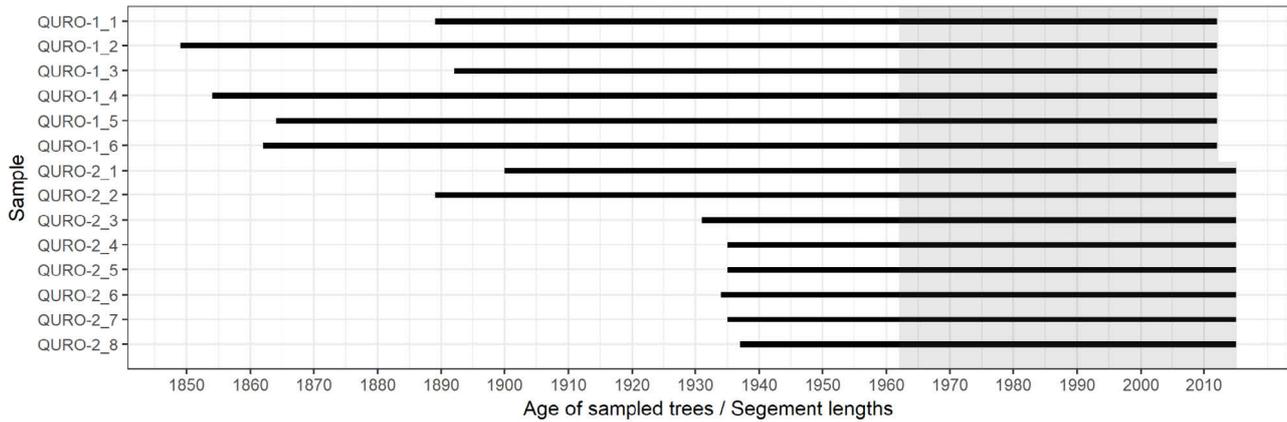
Fig. 2. *Individual sample lengths. The shaded area represents the analysed period used for model comparison for QURO-1 and QURO-2.*

Table 2. *Characteristics of site chronologies: number of samples (N), chronology span, mean and standard deviation (Std) of vessel areas, minimum and maximum range of vessel areas (Min – Max), rbar ($\bar{r}$) and autocorrelation with lag 1, 2 and 3 (AC).*

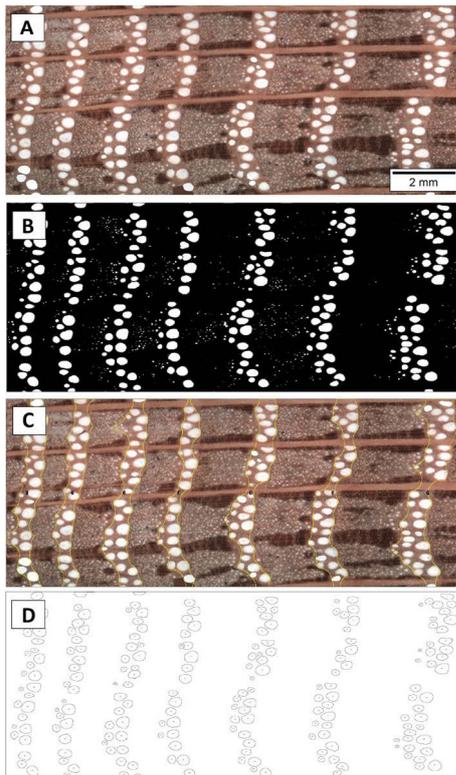| Site	N	Chronology span	Mean ± Std (μm^2 10^4)	Min – Max (μm^2 10^4)	$\bar{r}$	AC_1	AC_2	AC_3
QURO-1	6	2012–1961	6.259 ± 0.603	5.133–7.743	0.44	0.42	0.40	0.31
QURO-2	8	2015–1961	4.691 ± 0.395	4.072–5.756	0.32	0.61	0.52	0.46

Fig. 3. *An example of a workflow of earlywood vessel analysis with ImageJ and macro EWVA: A) original image, B) image converted to a black and white mask for segmentation of the original image, vessels below the threshold were excluded from the analysis, C) recognised groups of earlywood vessels belonging to different years overlaid over the original image and D) measured earlywood vessels.*

which means more explained variance and lower error on validation data.

Modelling strategy and comparison of nonlinear methods for VLA prediction

Pearson correlation coefficients were used to analyse the climate-growth relationship between climate data and VLA chronologies to get a first impression of patterns in the climate signal. Secondly, based on the results of the climate-growth relationship, climate variables were selected to be used as predictors for VLA. MLR models were then fitted by using these predictors, and the statistical properties of these models were assessed. To fit MLR models, stepwise selection of predictors with backward elimination was used. In addition, model trees (MT) were fitted to explore the understandable rules for VLA predictions. Finally, the performance of linear and nonlinear machine learning regression methods was evaluated by 3-fold cross-validation (Witten *et al.*, 2011), which is implemented in the *compare_methods()* function of the *dendroTools* R package (Jevšenak and Levanič, 2018).

In 3-fold cross-validation, the datasets were randomly split into 3 equal parts (folds) and models were systematically calibrated (learned) on 2 folds and validated on the remaining fold that had not been used for calibration (learning). To minimise coincidence due to specific train and test splits, cross-validation was repeated 100 times, and the mean and standard deviation of the performance metrics across all repeats of cross-validation were shown. To ensure objective comparison, the cross-validation
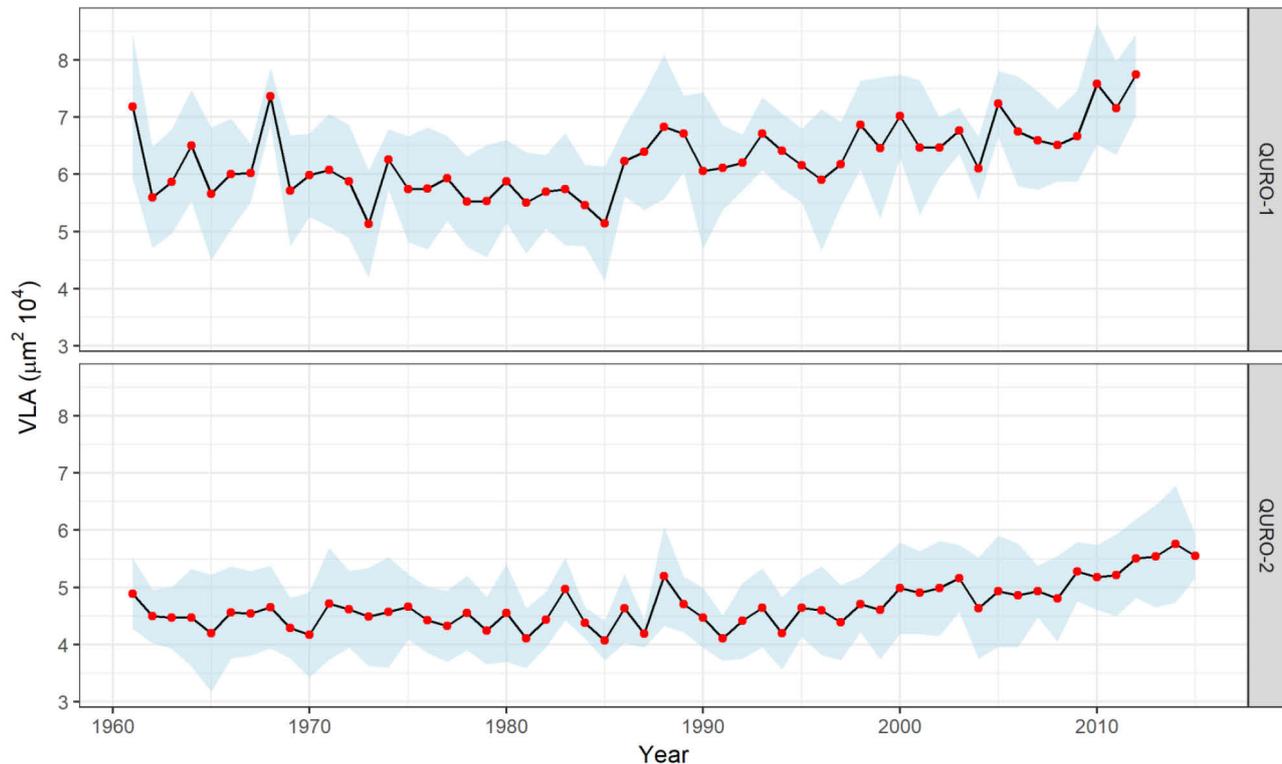
Fig. 4. *Chronologies of vessel lumen area (VLA) for QURO-1 and QURO-2 with 95% confidence interval of the mean.*

splits were identical for all nonlinear and MLR models, which was achieved by using the *set.seed()* function in R (R Core Team, 2018).

In addition, for each individual cross-validation split, the different models were ranked in terms of the calculated performance metrics, with rank 1 representing the best performance. The *ties.method* argument from the *rank()* function in R was set to '*min*'. Both models with the same performance metrics values were therefore given the lowest rank. The sum of shares of rank 1 consequently does not necessarily equal 1. The mean ranks and shares of rank 1 for each method are reported in the Results section.

For each cross-validation split, MLR was fitted by a stepwise selection of predictors with backward elimination. For each individual MLR model, therefore, only statistically significant and non-correlated predictors were used. One of the major strengths of machine learning algorithms is that they can be run with a large set of predictors with no *a priori* assumptions regarding which predictors are most important, or the underlying nature of the relationship between the response and predictor variables. ML models were therefore trained using all available predictors.

In the preliminary experimentation phase, different parameter settings of the four nonlinear algorithms were explored. We optimised our methods by experimentation and with the *caret* R package (Kuhn *et al.*, 2017). The *train* function from the *caret* R package sets up a grid of tuning parameter's values for a number of regression routines, fits each model and calculates a resampling-based performance measure. The final models were tuned for each dataset separately and used to compare the performance of different regression algorithms. A list with tuned parameter values is presented in Supplementary **Table S1**.

The measures of performance used in our study were the correlation coefficient (r), the root mean squared error (RMSE; Willmott, 1981), the root relative squared error (RRSE; Witten *et al.*, 2011), the index of agreement (d; Willmott, 1981), the reduction of error (RE; Fritts, 1976; Lorenz, 1956), the coefficient of efficiency (CE; Briffa *et al.*, 1988) and the detrended efficiency (DE; Helama *et al.*, 2018). Models with higher validation (testing) values of r, d, RE, CE and DE and lower values of RMSE and RRSE were considered to have better predictive capabilities. In addition, to address potential over- or under-prediction associated with a given approach, bias was calculated as the difference between observed and estimated mean response for the validation data. Bias is reported in the form of histograms for each method.

3. RESULTS

Chronology characteristics and the climate-growth relationship

The two site chronologies had a common increasing trend in VLA in the last 2–3 decades (**Fig. 4**). Even though the trees considered differed in terms of age (**Table 1**, **Fig. 2**) and were sampled from sites with different site characteristics, the increasing trend in VLA was obvious and consistent for both sites. The values of VLA in the QURO-1 chronologies ranged from 5.133–7.743 $\mu m^2 10^4$, while VLA values for QURO-2 ranged from 4.072–5.756 $\mu m^2 10^4$. Higher coherence between individual VLA chronologies was observed for QURO-1 ($\bar{r} = 0.44$), while QURO-2 had less coherent individual chronologies ($\bar{r} = 0.32$). Autocorrelation was significant in both analysed chronologies. On average, the vessels with the largest area were those from the QURO-1 site (**Table 2**).

VLA chronologies were compared to mean monthly temperatures and monthly sums of precipitation (**Table 3**). The two sites showed a very similar temperature pattern – VLA was significantly correlated with mean spring temperatures and mean temperatures from the previous growing season. Analysing only individual months of the current growing season, April had the strongest correlation coefficient at both sites. For the previous growing season, QURO-1 was correlated with temperatures from July to September, while QURO-2 correlated with temperatures from July to November. Only QURO-2 had significant correlations between the sum of precipitation and the VLA chronologies. Winter precipitation showed negative correlations with VLA from QURO-2.

Based on the correlation analysis, the predictors of VLA were selected for each site individually. All months with significant correlations from the previous growing season were averaged, so we obtained a single additional predictor representing the previous growing season temperatures. Precipitation values for the months from January to March were summed, so we obtained a single additional winter precipitation predictor for the QURO-1 site. All the monthly temperature variables from the current growing season with significant correlations were kept as individual predictors. The strongest linear relationship between VLA and monthly temperatures was observed for April, thus we expected to observe a more nonlinear relationship by adding temperature information for March and May. The final list of predictors for QURO-1 and QURO-2 is given in **Table 4**.

The predictors were fitted by the multiple regression method, using the stepwise selection of predictors (**Table 5**). None of the diagnostic plots of the linear regression models for the two sites showed any problematic patterns (see **Fig. S1**). As in the correlation analysis (**Table 3**), April temperatures were the most significant variables. Both models showed some similar patterns: 1) they excluded May temperatures as a significant predictor, 2) both models explained a similar amount of variance, 0.63 with QURO-1 and 0.62 with QURO-2 and 3) both models had a similar value for the coefficient for temperatures from the previous growing season.

For both sites, model trees (MT) were fitted to extract rules for predicting the VLA at each of the analysed sites (**Fig. 5**). For each instance (year), the values of its predictors determine which linear equation will be used to make the final prediction. For example, for QURO-1, if the mean April temperature (T_APR) is smaller than 9.97°C, linear equation 1 (LM1) is used. Otherwise, the January-March sum of precipitation ($P\_JAN-MAR$) is considered, with the threshold to select linear equation 2 (LM2) or 3 (LM3) set to 56.9 mm. This threshold indicates the point at which the relationship between VLA and the climate

Table 3. *Pearson correlation coefficients between vessel lumen area (VLA) and climate data: mean monthly temperatures (TEMP) and sum of monthly precipitation (PREC) for QURO-1 and QURO-2. Months with capital letters refer to the current growing season, while months with lowercase letters refer to the year of the previous growing season. Only correlation coefficients with p ≤ 0.01 are shown.*

	Month	QURO-1		QURO-2	
		TEMP	PREC	TEMP	PREC
Previous growing season	Jul	0.45		0.53	
	Aug	0.49		0.46	
	Sep	0.37		0.38	
	Oct			0.35	
	Nov			0.37	
	Dec				
	Jul–Sep	0.57			
	Jul–Nov			0.71	
Current growing season	JAN		−0.41	0.44	
	FEB				
	MAR	0.38	−0.33		
	APR	0.60		0.63	
	MAY	0.32		0.47	
	JUN	0.47		0.50	
	JAN–MAR		−0.43		

Table 4. *Description of predictors of the VLA for QURO-1 and QURO-2.*

	Predictor	Description
QURO-1	T_Jul_Sep	Average temperature from July to September, previous growing season
	T_MAR	Average March temperature, current growing season
	T_APR	Average April temperature, current growing season
	T_MAY	Average May temperature, current growing season
	T_JUN	Average June temperature, current growing season
	P_JAN-MAR	Sum of precipitation from January to March, current growing season
QURO-2	T_Jul-Nov	Average temperature from July to November, previous growing season
	T_JAN	Average January temperature, current growing season
	T_APR	Average April temperature, current growing season
	T_MAY	Average May temperature, current growing season
	T_JUN	Average June temperature, current growing season

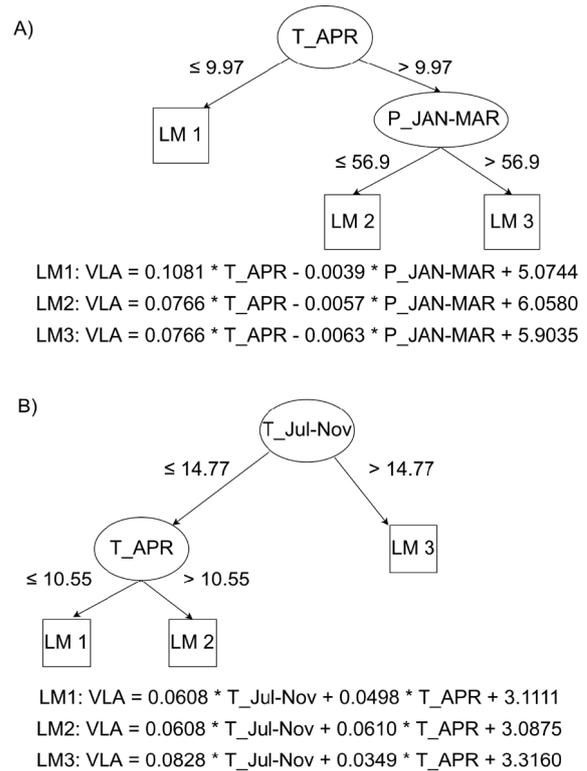**Table 5.** *Multiple linear regression summary statistics for QURO-1 and QURO-2.*

	Variable	Coefficients	*t* value	*p* value
QURO-1	Intercept	0.496164	0.418	0.67767
	T_Jul–Sep	0.165491	2.617	0.01196
	T_MAR	0.042491	1.437	0.15744
	T_APR	0.159552	3.474	0.00113
	T_JUN	0.064903	1.422	0.16164
	P_JAN–MAR	–0.00598	–2.294	0.02639
	R²	**Adj R²**		
	0.6304	0.5903		
QURO-2	Intercept	1.08324	1.792	0.07914
	T_Jul–Nov	0.12812	2.570	0.01319
	T_JAN	0.03792	2.501	0.01569
	T_APR	0.08273	2.879	0.00586
	T_JUN	0.04669	1.800	0.07782
	R²	**Adj R²**		
	0.6241	0.5941		

variables changes. While for QURO-1, *T_APR* is the most important predictor, for QURO-2 temperatures from the previous growing season (*T_Jul-Nov*) are in the root of MT. If *T_Jul-Nov* is greater than 14.77°C, linear equation 3 (LM3) is used; otherwise *T_APR* is considered. Finally, if *T_APR* is greater than 10.55 mm, MT will use LM1; otherwise LM2 is applied to make the final prediction.

Comparison of different ML algorithms for VLA prediction

MLR, ANN, MT, BMT and RF were used to model the relationship between VLA and the selected climate predictors (**Table 4**). The models were primarily evaluated in terms of their performance metrics on validation data (**Table 6**), which can be considered to be objective indicators. For QURO-1, ANN provided the best performance in 53% of individual cross-validation splits. ANN showed superior results for all mean validation statistics. RF, MLR and BMT showed similar performance at QURO-1, while MT had the worst validation performance. The statistical bias for validation data (**Fig. 6**) is reflected in the results from **Table 6**. Bias, with the highest count of values at 0, is an indication of no systematic over- or under-estimation of VLA. Histograms with a peak at 0 were obtained for BMT, ANN and RF, while for MLR and MT, the peak at 0 was not so obvious.

Similarly, as with QURO-1, ANN also had the best validation performance for QURO-2. The second-best performance was obtained by RF, followed by BMT. The worst validation results were recorded for MLR and MT, which performed the best only in 4% (MLR) and 3% (MT) of individual cross-validation splits (**Table 6**). Histograms of mean bias (**Fig. 6**) reflect the performance of models from **Table 6**. The histogram of mean bias for the ANN method shows the best performance, followed by RF and BMT, while the histograms of mean bias for

Fig. 5. *Examples of the Model Tree (MT) for A) QURO-1 and B) QURO-2 sites.*

LM1: VLA = 0.1081 * T_APR - 0.0039 * P_JAN-MAR + 5.0744
LM2: VLA = 0.0766 * T_APR - 0.0057 * P_JAN-MAR + 6.0580
LM3: VLA = 0.0766 * T_APR - 0.0063 * P_JAN-MAR + 5.9035

LM1: VLA = 0.0608 * T_Jul-Nov + 0.0498 * T_APR + 3.1111
LM2: VLA = 0.0608 * T_Jul-Nov + 0.0610 * T_APR + 3.0875
LM3: VLA = 0.0828 * T_Jul-Nov + 0.0349 * T_APR + 3.3160

MLR and MT display weaker performance. Based on the results from **Table 6** and **Fig. 6**, the final selected models for both sites would be those learned by ANN.
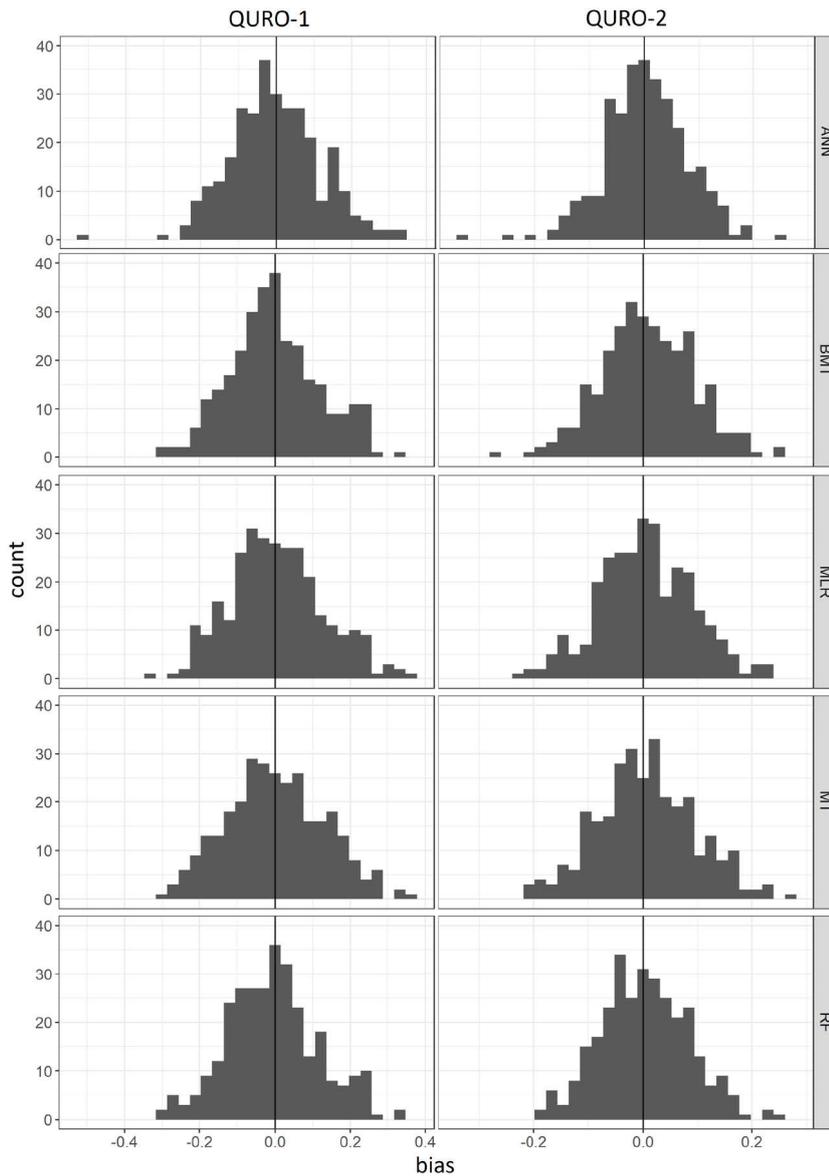
4. DISCUSSION

Climate signal in VLA

The main goal of our study was to compare various regression methods in the task of VLA prediction. To do so, we first analysed the climate signal of VLA chronologies to select predictors for different regression models. The correlation coefficients showed a similar pattern for both sites: the VLA tree-ring parameter was correlated with temperatures at the end of the previous growing season and temperatures at the time of earlywood formation. Many other studies have also reported positive correlations between mean temperatures and VLA (e.g., Fonti and Garcia-Gonzalez, 2008; Garcia-Gonzalez and Fonti, 2008; Matisons and Dauškane, 2009; Tumajer and Treml, 2016).

The April temperature signal is in accordance with xylogenetic studies from similar sites (e.g., Goršić, 2013; Gričar, 2010), which have reported an expansion of earlywood conduits in April. The spring temperature signal stored in vessel characteristics is related to the onset of cambial activity and the duration of the enlargement period before lignification of the vessel secondary wall

Table 6. *Comparison of the performance of five predictive modelling methods for A) QURO_1 and B) QURO_2 sites. Methods were evaluated by 3-fold cross-validation repeated 100 times. The five predictive modelling methods were Multiple Linear Regression (MLR), Artificial Neural Networks (ANN), Model Trees (MT), Bagging of Model Trees (BMT) and Random Forests of Regression Trees (RF). The performance measures were the correlation coefficient (r), root relative squared error (RRSE), root mean squared error (RMSE), index of agreement (d), reduction of error (RE), coefficient of efficiency (CE) and detrended efficiency (DE), calculated on the training (train) and testing (test) data of cross-validation splits. Across the 100 runs of cross-validation, we present the mean, standard deviations (Std), mean rank and share of rank 1 for each performance measure. The best values of each performance measure on the test set are highlighted in bold.*

	MLR		ANN		MT		BMT		RF	
	mean	std	mean	std	mean	std	mean	std	mean	std
(A) QURO-1										
r_{train}	0.797	0.037	0.815	0.043	0.789	0.059	0.810	0.030	0.898	0.018
r_{test}	0.694	0.103	**0.726**	0.095	0.633	0.134	0.692	0.112	0.689	0.097
$RMSE_{train}$	0.356	0.027	0.343	0.042	0.359	0.047	0.360	0.025	0.302	0.022
$RMSE_{test}$	0.453	0.063	**0.428**	0.066	0.488	0.079	0.447	0.065	0.453	0.066
$RRSE_{train}$	0.601	0.048	0.580	0.077	0.606	0.080	0.608	0.041	0.510	0.035
$RRSE_{test}$	0.786	0.134	**0.741**	0.123	0.845	0.148	0.771	0.106	0.778	0.083
d_{train}	0.877	0.026	0.878	0.059	0.871	0.042	0.851	0.030	0.898	0.019
d_{test}	0.794	0.066	**0.808**	0.073	0.753	0.081	0.757	0.067	0.723	0.069
RE_{test}	0.405	0.227	**0.473**	0.184	0.312	0.257	0.435	0.158	0.429	0.110
CE_{test}	0.364	0.237	**0.436**	0.202	0.265	0.273	0.394	0.174	0.387	0.132
DE_{test}	0.317	0.249	**0.394**	0.215	0.210	0.289	0.348	0.188	0.339	0.155
	rank	rank1	rank	rank1	rank	rank1	rank	rank1	rank	rank1
r_{train}	3.92	0.00	2.67	0.02	4.16	0.03	2.95	0.00	1.05	0.95
r_{test}	2.91	0.11	**1.85**	**0.50**	3.93	0.06	2.93	0.15	3.12	0.20
$RMSE_{train}$	3.45	0.00	2.33	0.05	3.93	0.07	3.90	0.00	1.13	0.88
$RMSE_{test}$	2.95	0.14	**1.89**	**0.49**	3.86	0.06	2.82	0.12	3.22	0.21
$RRSE_{train}$	3.45	0.00	2.33	0.05	3.93	0.07	3.90	0.00	1.13	0.88
$RRSE_{test}$	2.95	0.14	**1.89**	**0.49**	3.86	0.06	2.82	0.12	3.22	0.21
d_{train}	2.64	0.07	2.52	0.13	3.39	0.13	4.71	0.00	1.49	0.68
d_{test}	2.18	0.24	**1.50**	**0.65**	3.32	0.09	3.45	0.02	4.29	0.03
RE_{test}	2.96	0.14	**1.89**	**0.49**	3.86	0.06	2.82	0.12	3.22	0.21
CE_{test}	2.96	0.14	**1.89**	**0.49**	3.86	0.06	2.82	0.12	3.22	0.21
DE_{test}	2.95	0.14	**1.89**	**0.49**	3.86	0.06	2.82	0.12	3.22	0.21
$Overall_{train}$	3.37	0.02	2.46	0.06	3.85	0.08	3.86	0.00	1.20	0.85
$Overall_{test}$	2.75	0.16	**1.79**	**0.53**	3.74	0.07	3.00	0.11	3.46	0.16
(B) QURO-2										
r_{train}	0.796	0.044	0.843	0.035	0.800	0.039	0.839	0.028	0.907	0.020
r_{test}	0.717	0.093	**0.781**	0.084	0.687	0.104	0.740	0.087	0.747	0.094
$RMSE_{train}$	0.232	0.019	0.209	0.027	0.232	0.019	0.216	0.015	0.180	0.013
$RMSE_{test}$	0.293	0.041	**0.257**	0.042	0.297	0.040	0.278	0.040	0.274	0.043
$RRSE_{train}$	0.601	0.058	0.541	0.080	0.602	0.052	0.560	0.044	0.465	0.044
$RRSE_{test}$	0.792	0.164	**0.694**	0.150	0.801	0.147	0.745	0.126	0.733	0.107
d_{train}	0.874	0.032	0.899	0.059	0.866	0.030	0.883	0.024	0.921	0.018
d_{test}	0.793	0.057	**0.839**	0.077	0.765	0.068	0.789	0.052	0.782	0.058
RE_{test}	0.410	0.253	**0.546**	0.202	0.401	0.234	0.483	0.189	0.503	0.142
CE_{test}	0.346	0.299	**0.496**	0.248	0.337	0.283	0.428	0.230	0.451	0.179
DE_{test}	0.297	0.356	**0.457**	0.295	0.286	0.359	0.384	0.300	0.409	0.236
	rank	rank1	rank	rank1	rank	rank1	rank	rank1	rank	rank1
r_{train}	4.54	0.00	2.44	0.00	4.30	0.00	2.72	0.00	1.00	1.00
r_{test}	3.61	0.06	**1.70**	**0.61**	4.21	0.04	2.82	0.07	2.65	0.22
$RMSE_{train}$	4.42	0.00	2.22	0.01	4.26	0.00	3.08	0.00	1.01	0.99
$RMSE_{test}$	3.70	0.03	**1.75**	**0.64**	4.05	0.03	2.78	0.11	2.72	0.20
$RRSE_{train}$	4.42	0.00	2.22	0.01	4.26	0.00	3.08	0.00	1.01	0.99
$RRSE_{test}$	3.70	0.03	**1.75**	**0.64**	4.05	0.03	2.78	0.11	2.72	0.20
d_{train}	4.05	0.00	1.99	0.11	4.34	0.00	3.49	0.00	1.13	0.89
d_{test}	3.09	0.03	**1.30**	**0.85**	4.02	0.03	3.15	0.04	3.44	0.05
RE_{test}	3.70	0.03	**1.75**	**0.64**	4.05	0.03	2.78	0.11	2.72	0.20
CE_{test}	3.70	0.03	**1.75**	**0.64**	4.05	0.03	2.78	0.11	2.72	0.20
DE_{test}	3.70	0.03	**1.75**	**0.64**	4.05	0.03	2.78	0.11	2.72	0.20
$Overall_{train}$	4.36	0.00	2.22	0.03	4.29	0.00	3.09	0.00	1.04	0.97
$Overall_{test}$	3.53	0.04	**1.62**	**0.68**	4.08	0.03	2.88	0.08	2.88	0.17

Fig. 6. *Histograms of mean bias, calculated as the difference between observed and estimated mean responses for validation data.*

(Perez-de-Lis *et al.*, 2016). Rising temperatures before bud break increase the expression of genes involved in polar auxin transport (Schrader *et al.*, 2003), which induces radial growth. The onset of cambial activity starts approximately one month before bud swelling (Gonzalez-Gonzalez *et al.*, 2013; Sass-Klaassen *et al.*, 2011). Earlywood conduits are therefore mainly developed during the period with limited (or no) photosynthetic activity and trees use stored carbohydrates from the previous growing season, leading to the preservation of the temperature signal of the previous growing season in the vessel characteristics.

The correlation coefficients between the sum of monthly precipitation and VLA were negative and significant only for the QURO-1 site, but nevertheless lower

than temperature correlations. Negative correlations between VLA and winter precipitation are usually attributed to the excess of water accumulated in the soil (e.g.; González-González *et al.*, 2015). The QURO-1 site is in close proximity to small artificial lakes and water is, therefore, available throughout the year. Water saturation affects anatomical features, mainly the formation of smaller vessels (Ballesteros *et al.*, 2010).

Temperature is, therefore, the most important limiting growth factor at both sites. However, many other studies have reported that earlywood vessel characteristics are mostly related to water availability (Copini *et al.*, 2016; Garcia-Gonzalez and Eckstein, 2003; Garcia-Gonzalez and Fonti, 2008; Kames *et al.*, 2016; St George and Nielsen, 2000). Based on literature published to date,

there are no obvious site factors that would explain when VLA is more related to temperature and when to water availability.

Evaluation of various ML algorithms

To the best of our knowledge, our comparative study is the first in the dendroclimatological field to have used VLA tree-ring proxy and several climate predictors to compare different linear and nonlinear methods. In most previous comparative studies with tree-ring and climate data, ANN outperformed MLR (Balybina, 2010; D'Odorico *et al.*, 2000; Jevšenak and Levanič, 2016; Zhang *et al.*, 2000). We confirmed the high predictive abilities of ANN, which outperformed the other regression methods for both sites.

The high performance of nonlinear algorithms might be partially explained by the selection of predictors. By employing months with lower (linear) correlations, such as March and May, we allowed the nonlinear algorithms to use this information and explain the VLA more accurately. April had the strongest linear correlations with VLA at both sites, but temperatures in March and May probably cause nonlinear reactions in VLA. Advanced ML algorithms can use this information and model the relationship between climate and VLA more accurately. ML methods such as ANN should, therefore, become a more standard approach in dendroclimatology.

Despite the obvious outperformance of ANN over all the other regression methods used in our study, we recommend testing several different regression algorithms and selecting the optimal one for climate reconstruction. In a related study, Jevšenak *et al.* (2017) showed that different datasets favour different regression algorithms and that there is no single method that always outperforms the other regression methods. The methodological approach used in our study to compare different regression methods is implemented in the *compare_methods()* function from the *dendroTools* R package (Jevšenak and Levanič, 2018). This function could be used to find the optimal regression method for any dataset, prior to climate reconstruction.

The aggregate expression of the many interacting nonlinear, rate-limited processes within a tree is often effectively a linear response to local climate in its ring widths (Cook and Pederson, 2011). Linearity is, therefore, possible on a smaller part of the entire interval of climate-growth response but, when approaching the edge of this spectrum, the relationship between tree growth and climate is likely to become more or less nonlinear. In addition, in most climate reconstruction models, the reconstruction process estimates data points out of calibration data, *i.e.*, in the spectrum of the climate-tree response, which is expected to react nonlinearly. This is an additional argument in favour of nonlinear methods being used in dendroclimatology, since their predictions are likely to be more accurate.

Our comparative study showed that nonlinear algorithms can provide better models. Machine learning algorithms have several advantages, especially when a number of predictors, with a broader spectrum of climate and tree response data, are considered. In pursuit of more accurate climate reconstructions, it seems reasonable to switch from linear to advanced nonlinear algorithms.

5. CONCLUSIONS

We analysed the relationship between climate and the VLA parameter and used the results to select predictors to train and compare linear and nonlinear models. The two sites showed very similar results. ANN showed superior results for all mean validation statistics. The second-best models were RF and BMT, followed by MLR and MT. The question of whether it is reasonable to substitute the current *gold standard* arises. The current gold standard in the field of dendrochronological studies is MLR, but there are alternatives that might provide better predictions. The high performance of ANN was confirmed in our comparative study, but there is no reason to believe that ANN will always outperform other regression methods. We, therefore, suggest implementing our comparison strategy as a standard preliminary process, before using a selected method. To do so, there is a *compare_methods()* function available in the *dendroTools* R package.

SUPPLEMENTARY MATERIAL

Supplementary material, containing additional Table S1 and Figure S1, is available online at https://dx.doi.org/10.1515/geochr-2015-0097.

REFERENCES

ARSO, 2017. Archive – observed and measured meteorological data in Slovenia. WEB Site: http://meteo.arso.gov.si/met/sl/archive/. Accesed: 2017 August 8.

Ballesteros JA, Stoffel M, Bollschweiler M, Bodoque JM and Díez-Herrero A, 2010. Flash-flood impacts cause changes in wood anatomy of Alnus glutinosa, Fraxinus angustifolia and Quercus pyrenaica. *Tree Physiology* 30(6): 773–781, DOI 10.1093/treephys/tpq031.

Balybina AS, 2010. Reconstructing the air temperature from dendrochronological data from the Preolkhon area using the neural network method. *Geography and Natural Resources* 31(1): 30–33, DOI 10.1016/j.gnr.2010.03.006.

Billings SA, Glaser SM, Boone AS and Stephen FM, 2015. Nonlinear tree growth dynamics predict resilience to disturbance. *Ecosphere* 6(11): 1–13, DOI 10.1890/ES15-00176.1.

Bishop CM, 1995. *Neural Networks for Pattern Recognition*. Oxford, University Press, Inc.: 482 pp.

Braak CJFT and Gremmen NJM, 1987. Ecological Amplitudes of Plant Species and the Internal Consistency of Ellenberg's Indicator Values for Moisture. *Vegetatio* 69(1/3): 79–87, DOI 10.1007/BF00038689.

Breiman L, 1996. Bagging predictors. *Machine Learning* 24(2): 123–140, DOI 10.1023/A:1018054314350.

Breiman L, 2001. Random forests. *Machine Learning* 45(1): 5–32, DOI 10.1023/A:1010933404324.

Briffa KR, Jones PD, Pilcher JR and Hughes MK, 1988. Reconstructing summer temperatures in northern Fennoscandinavia back to A.D.1700 using tree ring data from Scots Pine. *Arctic and Alpine Research* 20: 385–394, DOI 10.2307/1551336.

Burden F and Winkler D, 2008. Bayesian regularization of neural networks. In: Livingstone DJ, ed., *Artificial Neural Networks: Methods and Applications*. Humana Press, Totowa, NJ, 458: 23–42.

Campelo F, Nabais C, Gutierrez E, Freitas H and Garcia-Gonzalez I, 2010. Vessel features of Quercus ilex L. growing under Mediterranean climate have a better climatic signal than tree-ring width. *Trees-Structure and Function* 24(3): 463–470, DOI 10.1007/s00468-010-0414-0.

Cook ER and Kairiukstis LA, 1992. *Methods of Dendrochronology : Applications in the Environmental Sciences*. Dordrecht, Kluwer Academic Publishers: 394 pp.

Cook ER and Pederson N, 2011. Uncertainty, Emergence, and Statistics in Dendrochronology. In: Hughes MK, TW Swetnam, HF Diaz, eds., *Dendroclimatology: Progress and Prospects*. Springer Netherlands, Dordrecht: 77–112.

Copini P, den Ouden J, Robert EMR, Tardif JC, Loesberg WA, Goudzwaard L and Sass-Klaassen U, 2016. Flood-Ring Formation and Root Development in Response to Experimental Flooding of Young Quercus robur Trees. *Frontiers in Plant Science* 7: 775, DOI 10.3389/fpls.2016.00775.

D'Odorico P, Revelli R and Ridolfi L, 2000. On the use of neural networks for dendroclimatic reconstructions. *Geophysical Research Letters* 27(6): 791–794, DOI 10.1029/1999GL011049.

Fonti P and Garcia-Gonzalez I, 2008. Earlywood vessel size of oak as a potential proxy for spring precipitation in mesic sites. *Journal of Biogeography* 35(12): 2249–2257, DOI 10.1111/j.1365-2699.2008.01961.x.

Fonti P, von Arx G, Garcia-Gonzalez I, Eilmann B, Sass-Klaassen U, Gartner H and Eckstein D, 2010. Studying global change through investigation of the plastic responses of xylem anatomy in tree rings. *New Phytologist* 185(1): 42–53, DOI 10.1111/j.1469-8137.2009.03030.x.

Fritts HC, 1976. *Tree Rings and Climate*. London, Academic Press: 567 pp.

Garcia-Gonzalez I and Eckstein D, 2003. Climatic signal of earlywood vessels of oak on a maritime site. *Tree Physiology* 23(7): 497–504, DOI 10.1093/treephys/23.7.497.

Garcia-Gonzalez I and Fonti P, 2008. Ensuring a representative sample of earlywood vessels for dendroecological studies: an example

from two ring-porous species. *Trees-Structure and Function* 22(2): 237–244, DOI 10.1007/s00468-007-0180-9.

Gonzalez-Gonzalez BD, Garcia-Gonzalez I and Vazquez-Ruiz RA, 2013. Comparative cambial dynamics and phenology of Quercus robur L. and Q-pyrenaica Willd. in an Atlantic forest of the northwestern Iberian Peninsula. *Trees-Structure and Function* 27(6): 1571–1585, DOI 10.1007/s00468-013-0905-x.

González-González BD, Vázquez-Ruiz RA and García-González I, 2015. Effects of climate on earlywood vessel formation of Quercus robur and Q. pyrenaica at a site in the northwestern Iberian Peninsula. *Canadian Journal of Forest Research* 45(6): 698–709, DOI 10.1139/cjfr-2014-0436.

Goršić E, 2013. Diameter increment dynamics of pedunculate oak (Quercus robur L.) in Croatia. Ph.D. dissertation, University of Zagreb, Zagreb: 154 pp.

Gričar J, 2010. Xylem and phloem formation in sessile oak from Slovenia in 2007. *Wood Research* 55(4): 15–22.

Hastie T, Tibshirani R and Friedman JH, 2009. The Elements of Statistical Learning : Data Mining, Inference, and Prediction. New York, Springer: 745 pp.

Helama S, Makarenko NG, Karimova LM, Kruglun OA, Timonen M, Holopainen J, Merilainen J and Eronen M, 2009. Dendroclimatic transfer functions revisited: Little Ice Age and Medieval Warm Period summer temperatures reconstructed using artificial neural networks and linear algorithms. *Annales Geophysicae* 27(3): 1097–1111, DOI 10.5194/angeo-27-1097-2009.

Helama S, Sohar K, Läänelaid A, Bijak S and Jaagus J, 2018. Reconstruction of precipitation variability in Estonia since the eighteenth century, inferred from oak and spruce tree rings. *Climate Dynamics* 50(11): 4083–4101, DOI 10.1007/s00382-017-3862-z.

Ho TK, 1995. *Random decision forests*. In Proceedings of the Third International Conference on Document Analysis and Recognition. Montreal, Canada: 278–282 pp.

Hornik K, Buchta C and Zeileis A, 2009. Open-source machine learning: R meets Weka. *Computational Statistics* 24(2): 225–232, DOI 10.1007/s00180-008-0119-7.

Jevšenak J, Džeroski S and Levanič T, 2017. On the use of machine learning methods to study the relationships between tree-ring characteristics and the environment. *Acta silvae et ligni* 114: 21–29, DOI 10.20315/ASetL.114.2.

Jevšenak J and Levanič T, 2014. Macro EWVA - an effective tool for analysis of earlywood conduits of ring porous species. *Acta silvae et ligni* 104: 15–24, DOI 10.20315/ASetL.104.2.

Jevšenak J and Levanič T, 2015. Dendrochronological and wood-anatomical features of differently vital pedunculate oak (*Quercus robur* L.) stands and their response to climate. *Topola* 195/196: 85–96.

Jevšenak J and Levanič T, 2016. Should artificial neural networks replace linear models in tree ring based climate reconstructions? *Dendrochronologia* 40: 102–109, DOI 10.1016/j.dendro.2016.08.002.

Jevšenak J and Levanič T, 2018. dendroTools: R package for studying linear and nonlinear responses between tree-rings and daily environmental data. *Dendrochronologia* 48: 32–39, DOI 10.1016/j.dendro.2018.01.005.

Jones PD and Bradley RS, 1992. Climatic variations in the longest instrumental records. In: Jones PD, RS Bradley, eds., *Climate Since A.D. 1500*. Routledge, London: 246–268.

Kames S, Tardif JC and Bergeron Y, 2016. Continuous earlywood vessels chronologies in floodplain ring-porous species can improve dendrohydrological reconstructions of spring high flows and flood levels. *Journal of Hydrology* 534: 377–389, DOI 10.1016/j.jhydrol.2016.01.002.

Kuhn M, Wing J, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z, Kenkel B, Benesty M, Lescarbeau R, Ziem A, Scrucca L, Tang Y, Candan C, Hunt T and the R Core Team, 2017. caret: Classification and Regression Training. R package version 6.0-76, WEB Site: https://CRAN.R-project.org/package=caret. Accesed: 2017 August 10.

Levanič T, 2007. ATRICS - A new system for image acquisition in dendrochronology. *Tree-Ring Research* 63(2): 117−122, DOI 10.3959/1536-1098-63.2.117.

Lorenz EN, 1956. *Empirical Orthogonal Functions and Statistical Weather Prediction*. Massachusetts, Massachusetts Institute of Technology, Department of Meteorology: 49 pp.

Matisons R and Dauškane I, 2009. Influence of climate on earlywood vessel formation of Quercus robur at its northern distribution range in central regions of Latvia. *Acta Universitatis Latviensis* 753: 49–58.

Montoro Girona M, Rossi S, Lussier J-M, Walsh D and Morin H, 2017. Understanding tree growth responses after partial cuttings: A new approach. *PLoS ONE* 12(2): e0172653, DOI 10.1371/journal.pone.0172653.

Ni FB, Cavazos T, Hughes MK, Comrie AC and Funkhouser G, 2002. Cool-season precipitation in the southwestern USA since AD 1000: Comparison of linear and nonlinear techniques for reconstruction. *International Journal of Climatology* 22(13): 1645−1662, DOI 10.1002/joc.804.

Perez-de-Lis G, Rossi S, Vazquez-Ruiz RA, Rozas V and Garcia-Gonzalez I, 2016. Do changes in spring phenology affect earlywood vessels? Perspective from the xylogenesis monitoring of two sympatric ring-porous oaks. *New Phytologist* 209(2): 521–530, DOI 10.1111/nph.13610.

Pérez-Rodríguez P and Gianola D, 2016. Brnn: Brnn (Bayesian Regularization forFeed-forward Neural Networks). R package version 0.6, WEB Site: http://CRAN.R-project.org/package=brnn. Accesed: 2017 April 20.

Pérez-Rodríguez P, Gianola D, Weigel KA, Rosa GJM and Crossa J, 2013. Technical Note: An R package for fitting Bayesian regularized neural networks with applications in animal breeding. *Journal of Animal Science* 91(8): 3522–3531, DOI 10.2527/jas.2012-6162.

Quinlan JR, 1992. *Learning with continuous classes*. In Proceedings of the 5th Australian Joint Conference on Artificial Intelligence (AI '92). Hobart: 343−348 pp.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, WEB Site: http://www.R-project.org/. Accesed: 2017 November 15.

Sass-Klaassen U, Sabajo CR and den Ouden J, 2011. Vessel formation in relation to leaf phenology in pedunculate oak and European ash. *Dendrochronologia* 29(3): 171–175, DOI 10.1016/j.dendro.2011.01.002.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P and Cardona A, 2012. Fiji: an open-source platform for biological-image analysis.

Nature Methods 9(7): 676−682, DOI 10.1038/nmeth.2019.

Schrader J, Baba K, May ST, Palme K, Bennett M, Bhalerao RP and Sandberg G, 2003. Polar auxin transport in the wood-forming tissues of hybrid aspen is under simultaneous control of developmental and environmental signals. *Proceedings of the National Academy of Sciences of the United States of America* 100(17): 10096–10101, DOI 10.1073/pnas.1633693100.

Shishov VV, Tychkov II, Popkova MI, Ilyin VA, Bryukhanova MV and Kirdyanov AV, 2016. VS-oscilloscope: A new tool to parameterize tree radial growth based on climate conditions. *Dendrochronologia* 39: 42–50, DOI 10.1016/j.dendro.2015.10.001.

St George S and Nielsen E, 2000. Signatures of high-magnitude 19th-century floods in Quercus macrocarpa tree rings along the Red River, Manitoba, Canada. *Geology* 28(10): 899–902, DOI 10.1130/0091-7613(2000)28<899:SOHTFI>2.0.CO;2.

Sun Y, Bekker MF, DeRose RJ, Kjelgren R and Wang SYS, 2017. Statistical treatment for the wet bias in tree-ring chronologies: A case study from the Interior West, USA. *Environmental and Ecological Statistics* 24(1): 131–150, DOI 10.1007/s10651-016-0363-x.

Tolwinski-Ward SE, Evans MN, Hughes MK and Anchukaitis KJ, 2011. An efficient forward model of the climate controls on inter-annual variation in tree-ring width. *Climate Dynamics* 36(11–12): 2419−2439, DOI 10.1007/s00382-010-0945-5.

Tumajer J and Treml V, 2016. Response of floodplain pedunculate oak (*Quercus robur* L.) tree-ring width and vessel anatomy to climatic trends and extreme hydroclimatic events. *Forest Ecology and Management* 379: 185–194, DOI 10.1016/j.foreco.2016.08.013.

Vaganov EA, Anchukaitis KJ and Evans MN, 2011. How Well Understood Are the Processes that Create Dendroclimatic Records? A Mechanistic Model of the Climatic Control on Conifer Tree-Ring Growth Dynamics. In: Hughes MK, TW Swetnam, HF Diaz, eds., *Dendroclimatology: Progress and Prospects. Developments in Paleoenvironmental Research*. Springer Netherlands, 11: 37–75.

Williams AP, Michaelsen J, Leavitt SW and Still CJ, 2010. Using Tree Rings to Predict the Response of Tree Growth to Climate Change in the Continental United States during the Twenty-First Century. *Earth Interactions* 14(19): 1–20, DOI 10.1175/2010EI362.1.

Willmott CJ, 1981. On the validation of models. *Physical Geography* 2(2): 184−194.

Witten IH, Frank E and Hall MA, 2011. *Data Mining: Practical Machine Learning Tools and Techniques*. Burlington, Morgan Kaufmann Publishers: 629 pp.

Zhang QB, Hebda RJ, Zhang QJ and Alfaro RI, 2000. Modeling tree-ring growth responses to climatic variables using artificial neural networks. *Forest Science* 46(2): 229−239.