Dendrochronologia 40 (2016) 102–109

Contents lists available at ScienceDirect

Dendrochronologia

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

Original article

Should artificial neural networks replace linear models in tree ring based climate reconstructions?

Jernej Jevšenak[a], Tom Levanič[b,*]

[a] PhD candidate at University of Ljubljana, Biotechnical Faculty, Department of Forestry and Renewable Forest Resources, Večna pot 83, 1000 Ljubljana, Slovenia
[b] Slovenian Forestry Institute, Večna pot 2, 1000 Ljubljana, Slovenia

ARTICLE INFO

Article history:
Received 4 April 2016
Received in revised form 26 July 2016
Accepted 8 August 2016
Available online 26 August 2016

Keywords:
Artificial neural network
Dendroclimatology
Levenberg-Marquardt algorithm
Bayesian regularization
Transfer function
Nonlinear models

ABSTRACT

Studies focused on tree ring—climate relationships usually use linear methods to find the optimal transfer function. In our study, three sites with three different tree species from the Western Balkan region were selected to compare linear and artificial neural network (ANN) nonlinear models and to see whether linear models can be potentially replaced with ANN in climate reconstruction. For each site, one linear and two different ANN models were calculated. For all analysed sites, we were able to find a better fit using the advanced technique of ANN. All calibration and verification statistics were in favour of ANN models. A climate variable was reconstructed for a selected site using linear and nonlinear ANN methods. We demonstrated that ANN is always a more effective method, which always produce better results than linear models. The key to success is a properly selected training algorithm, which prevents overfitting and is able to find the optimal transfer function, also linear, if that is the case.

1. Introduction

Studying tree rings provides us with an insight into the history of forests. Tree rings are directly influenced by climate and, when this relationship is well understood, reliable climate reconstructions can be calculated. In addition to climate reconstructions, understanding climate—growth relationships is very important for predicting the effects of environmental changes on tree growth and vitality.

Linear methods have traditionally been used to study the relationship between climate and a proxy variable, such as tree-ring width, early- and latewood width, wood-anatomical parameters (for a review see Fonti et al., 2010), wood density and stable carbon, oxygen and hydrogen isotopes ratio (for a review see McCarrol and Loader, 2004). The most common linear methods are linear regression, multiple linear regression, canonical correlation analysis, regression of principal component analysis (Fritts, 1976) and non-metric multidimensional scaling (Bunn et al., 2005; Patón and García, 2010).

From a theoretical point of view, the relationship between proxy variable and climate variable should be more or less nonlinear because 1) there is a limited interval of climate conditions in which trees can grow and 2) trees respond differently to optimal and suboptimal growth conditions. Nonlinearity has been considered in many previous tree-ring studies (e.g. Graumlich and Brubaker, 1986; Robinson et al., 1990; Vaganov et al., 2006; Lloyd et al., 2013; Breitenmoser et al., 2014). However, although the linear function is often a very good fit for climate and proxy records, this might just be a case of short calibration interval and not take into account the whole ecological response of the studied tree species. With current increasing temperature, the chance of capturing a nonlinear relationship in the calibration period increases. This phenomenon is probable also related to the divergence problem mentioned in various dendrochronological studies (e.g. D'Arrigo et al., 2008; Büntgen et al., 2009).

Artificial neural networks (ANN) are very suitable for modelling nonlinear functions between input and output data and therefore very suitable for analysing climate—growth relationships. ANN are widely used in various fields, such as economics (e.g. McNelis, 2005) and artificial intelligence (e.g. Jones, 2008) but are also gaining importance in environmental research, such as in forestry and tree-ring studies. Castro et al. (2013) used ANN for modelling the growth and yield of *Eucalyptus* stands located in northern Brazil. Chon et al. (2000) predicted the population dynamics of *Thecodiplosis japonensis* with ANN models. Fernandes et al. (2013) compared ANN and partial least squares for intra-growth ring wood density measure-

* Corresponding author.
E-mail addresses: jernej.jevsenak@gmail.com (J. Jevšenak), tom.levanic@gozdis.si (T. Levanič).

http://dx.doi.org/10.1016/j.dendro.2016.08.002
1125-7865/

ment with hyperspectral imaging. For a review of ANN applications in forestry, see Imada (2014).

The ANN method has already been compared with the linear approach in the field of dendroclimatology. Woodhouse (1999) performed one of the first studies in which ANNs were compared to linear models and she thus opened a new chapter in the methodological development of dendrochronological studies. Even though Woodhouse (1999) failed to construct a model that performed better than the linear model, most subsequent papers have reported the opposite results, in which the ANN method outperformed the linear approach (D'Odorico et al., 2000; Zhang et al., 2000; Ni et al., 2002; Balybina, 2010). Helama et al. (2009) compared ANN, linear scaling and multiple linear regression (MLR). Reconstructions calculated using MLR and ANN showed similar statistical reliability. Despite all the examples of better performance with ANN, this method is still not widely accepted by the relevant research community.

The main goal of this study was to compare the performance of linear regression models and nonlinear ANN models based on two different training algorithms. We hypothesized that ANN with proper configurations will always outperform linear regression. Our hypothesis was tested on three different tree species from three different sites, with different levels of climate response in TRW. One of the aims was also to answer the question of whether ANN, as an advanced technique, could potentially replace traditional linear models in tree-ring width based climate reconstructions.

2. Methods

2.1. Sample sites and meteorological data

We compared linear regression and nonlinear ANN on three datasets (three chronologies) from the Western Balkan region. Samples from black pine (*Pinus nigra*) were collected in Albania at five locations dispersed from north to south of the country (Levanič et al., 2015). TRW chronologies were compared to the mean temperature from June to July. Climate data for Albania was extracted from the KNMI Climate Explorer web page (http://climexp.knmi.nl/). Another conifer tree species included was European larch (*Larix decidua*) from the SE Alps in Slovenia (Hafner et al., 2014). TRW indices were compared with mean temperature from May to July (meteorological station Villacher Alpe, Auer et al., 2007). Samples from pedunculate oak (*Quercus robur*) were collected in the Srem region in northern Serbia (Stojanović et al., 2015) and compared to the mean temperature from April to July (meteorological station Ljubljana; ARSO, 2015). For all locations, chosen temperature data are based on preliminary correlation analysis and information from the original papers—see Table 1.

2.2. Artificial neural networks

The term *artificial neural network* (*neural net* or ANN) refers to a large class of models and learning methods suitable for regression and classification. ANN models can solve not only linear but also nonlinear problems and are therefore not restricted to a linear relationship between dependent and independent variables, which makes them potentially a better predictor than linear models. Since the beginning of research on artificial intelligence, great strides have been made in recognizing that different designs of ANN respond differently depending on the type of problem being analysed. In the field of data analysis, the best results are often obtained by a multilayer perceptron with backpropagation algorithm (Alsmadi et al., 2009; Arora and Suman, 2012; Skrobanski et al., 2012), which was also used in our study.

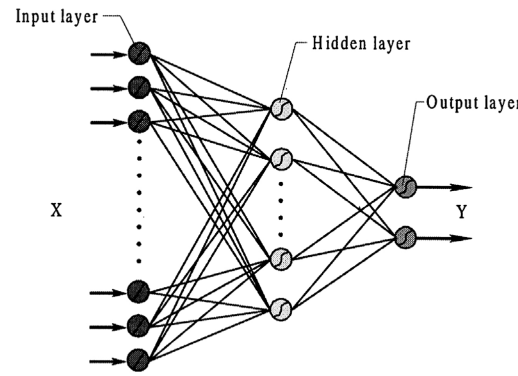

Fig. 1. Example of a simple feed-forward artificial neural network that learns via an iterative back-propagation procedure (from Lek and Guégan, 1999).

A neural network can be very complex but in our case it consisted of one input layer (TRW indices), one output layer (climate variable) and one hidden layer with different numbers of neurons. Each neuron could be thought of as a processing unit connected to weighted inputs. The summed value in a neuron is then applied to an activation function and sent forward to output. Outputs and inputs are then compared and the error is calculated. This process is called training—an iterative procedure in which weights are adjusted in a way that input data fits outputs well. For regression, the sum of squared errors as a measure of fit is used. With iteration, error function is minimized (Fig. 1).

There are various training algorithms available for ANN; in our study we used Levenberg-Marquardt (LM) and Bayesian regularization (BR). The Levenberg-Marquardt algorithm provides a numerical solution to the problem of minimizing a nonlinear function. It blends the steepest descent method and the Gauss–Newton algorithm. It is fast and has stable convergence. In the ANN field, this algorithm is suitable for training small- and medium-sized problems (for algorithm derivation and detailed explanation see Wilamowski and Irwin (2011)). Bayesian regularization is a more robust training algorithm than standard back-propagation and can reduce or eliminate the need for lengthy cross-validation. It is difficult to overfit a model with this algorithm, because it calculates and trains on a number of effective network parameters or weights, effectively turning off those that are not relevant (Burden and Winkler, 2008). LM and BR algorithms are implemented in Matlab R2015b and Neural Network Toolbox 8.4 (The MathWorks, Inc., Natick, Massachusetts, United States), which were also used in our study. More about the theory of artificial neural networks can be found elsewhere (e.g. Bishop, 1996; Hastie et al., 2009).

2.3. Model strategy and evaluation

To evaluate the performance of linear and nonlinear models, all three datasets were split into half—one half was used to calibrate our models (calibration period); the other was used as an independent dataset (verification period) to test the performance of the models. Calibration and verification datasets were then switched and we tested the performance of our models once again. One linear and two ANN models were constructed for each calibration period. One ANN was based on the Levenberg-Marquardt training algorithm (ANN_LM) and one on Bayesian regularization (ANN_BR).

The quality of each model was evaluated using explained variance (R^2), root mean squared error (*RMSE;* Willmott, 1981), reduction of error (*RE*), coefficient of error (*CE*) and index of agreement (*d*; Willmott, 1981). RE and CE (Cook et al., 1999) are commonly used in dendroclimatic studies (Lorenz, 1956; Fritts, 1976; Briffa et al., 1988; Robinson et al., 1990). Their maximum value is 1, indicating a perfect fit, any positive value indicates the