Revised Title page

Use of Machine Learning Techniques to Model Wind Damage to Forests

Emma Hart[a], Kevin Sim[a], Kana Kamimura[b,c], Celine Meredieu[d], Dominique Guyon[b], Barry Gardiner[b,e1]

[a] *School of Computing, Edinburgh Napier University, Scotland, UK*

[b] *INRA UMR 1391 ISPA, F-33140 Villenave d'Ornon, France; Bordeaux Sciences Agro, UMR 1391 ISPA, F-33170 Gradignan, France.*

[c] *Institute of Mountain Science, Shinshu University, 8304 Minamiminowa, Kamiina, Nagano 399-4598, Japan.*

[d] *INRA, UE FP, 69 route d'Arcachon, F-33612 Cestas cedex, France.*

[e] *EFI Planted Forests Facility, 69 Route de Arcachon, F-33612 Cestas cedex, France.*

[*] Corresponding author. Tel.:+335 35 38 52 50; fax: +33 5 56 680 223.
E-mail address: barry.gardiner@efi.int.

Highlights

Use of Machine Learning Techniques to Model Wind Damage to Forests

1. Machine learning techniques were accurate in predicting wind damage to trees.
2. Random forests proved the most accurate and discriminating methodology.
3. Models were sensitive to removal of site and stand but not tree characteristics.
4. All models were able to accurately replicate a mechanistic wind risk model.
5. Machine learning techniques could help the management of wind damage to forests.

Abstract

This paper tested the ability of machine learning techniques, namely artificial neural networks and random forests, to predict the individual trees within a forest most at risk of damage in storms. Models based on these techniques were developed individually for both a small forest area containing a set of 29 permanent sample plots that were damaged in Storm Martin in December 1999, and from a much larger set of 235 forest inventory data damaged in Storm Klaus in January 2009. Both data sets are within the Landes de Gascogne Forest in Nouvelle Aquitaine, France. The models were tested both against the data from which they were developed, and against the data set from the other storm. For comparison with an earlier study using the same data, logistic regression models were also developed. In addition, the ability of machine learning techniques to substitute for a mechanistic wind damage risk model by training them with previous mechanistic model predictions was tested.

All models were accurate at identifying whether trees would be damaged or not damaged but the random forests models were more accurate, had higher discriminatory power, and were almost totally unaffected by the removal of any individual input variable. However, if all information relating to a stand was removed the random forests model lost accuracy and discriminatory power. The other models were similarly affected by the removal of all site information but none of the models were affected by removal of all tree information, suggesting that damage in the Landes de Gascogne Forest occurs at stand scale and is not controlled by individual tree characteristics. The models developed with the large comprehensive database were also accurate in identifying damaged trees when applied to the small forest data damaged in the earlier storm. However, none of the models developed with the smaller forest data set could successfully discriminate between damaged and undamaged trees when applied across the whole landscape. All models were very successful in replicating the predictions of the mechanistic wind risk model and using them as a substitute for the mechanistic model predictions of critical wind speed did not affect the damage model results.

Overall the results suggest that random forests provide a significant advantage over other statistical modelling techniques and the random forest models were found to be more robust in their predictions if all input variables were not available. In addition, the ability to replace the mechanistic wind damage model suggests that random forests could provide a powerful tool for damage risk assessment at the stand or single tree level over large regions and provide rapid assessment of the impact of different management strategies or be used in the development of optimised forest management with multiple objectives and constraints including the risk of wind damage.

Machine learning; forest damage; wind risk, risk models, GALES, forest planning

1. Introduction

Wind causes more than 50% by volume of all damage to European forests and is the major damage agent on the continent (Schelhaas et al., 2003). On average 2 storms each year cause major damage in some part of Europe, where major damage is defined as disrupting the normal harvesting and supply of timber in a region. In south-west France there have been two major storms in the recent past that have threatened the viability of the forest industry in the Nouvelle Aquitaine region. On 27 December 1999 Storm Martin caused a loss of 26 million m^3 of timber (equivalent to 3.5 years of normal harvest) in the north of the region and on 24 January 2009 Storm Klaus caused 41 million m^3 of timber loss further south. The damage was predominately (37 million m^3) to maritime pine (*Pinus pinaster* Ait.) and the damage from the two storms represented 15% and 32% of the maritime pine standing volume in the region respectively.

There are also now increasing concerns that wind damage in Europe and many other parts of the world may increase with the changing climate (Csilléry et al., 2017; Haarsma et al., 2013; Kunkel et al., 2013; Lindner et al., 2010) due to the increasing intensity of low pressure systems whether extra tropical or tropical (hurricanes and typhoons). Therefore, in order to plan for the future there is a need for accurate models predicting tree vulnerability to wind damage and the level of risk. Such wind risk models form part of the risk assessment process that is an integral part of forest management (Cucchi et al., 2005; Gardiner and Welten, 2013; Hanewinkel et al., 2010) and allow managers and planners to decide on choice of species, silvicultural/management approaches, and rotation lengths for forest stands as a function of the site conditions (e.g. soil type, slope, water table depth, wind climate, etc.).

A number of modelling approaches to wind risk in forests are available. These include mechanistic (Gardiner et al., 2008) and statistical approaches (Albrecht et al., 2010). Previous attempts to model the observed damage patterns in the Landes de Gascogne Forest in Nouvelle Aquitaine, France using these two very different approaches are described in Kamimura et al. (2016). The mechanistic approach used the GALES model (Hale et al., 2015) and the statistical approach was based on logistic regressions (e.g. Valinger and Fridman, 2011). The results showed mixed success. The models were first tested on a small forest area that had a detailed survey of tree characteristics and damage following the Martin storm. Both models made accurate predictions of which individual trees were damaged in the storm. However, when the models were applied across the whole forest at the regional scale the logistic regression model performed poorly and GALES only worked well in areas with similar soil conditions to those from previous tree pulling tests used in the model parameterisation (Cucchi et al., 2004).

In environmental science there has been an increased use of Artificial Intelligence (AI) techniques in modelling studies

(Chen et al., 2008). These techniques have also been increasingly used in forestry (e.g. Lagerquist et al., 2017) although the ideas of using AI in forestry have already been around for a long time (Kourtz, 1990). However, very little attention has been paid to the use of AI in modelling the risk of wind damage with the exception of the work of Hanewinkel (2005) and Hanewinkel et al. (2004) who investigated the use of artificial neural networks. They found that the use of artificial neural networks allowed enhanced identification of damaged trees compared to the more classic approach using a logistic regression model.

In this paper we present analysis of the data on wind damage at an individual tree level from the Landes de Gascogne Forest using two methods that are based on machine learning (ML) techniques (Alpaydin, 2014). This was to determine if such approaches can provide a better prediction of wind risk than was possible with more conventional approaches as reported by Kamimura et al. (2016). The approach we took were based on artificial neural networks (NN) (Patterson, 1996) and random forests (RF) (Breiman, 2001). We also developed logistic regression models (LOG) for comparison with the previous work (designated LR in Kamimura et al. (2016)). We analysed damage from the small Nezer Forest (~80 km^2) containing a set of 29 permanent sample plots that were damaged in Storm Martin in December 1999 and from a much larger set of 235 plots from the National Forest Inventory in the Landes de Gascogne Forest (~10,000 km^2) that were examined directly after damage from Storm Klaus in January 2009. The purpose was to evaluate the accuracy and discriminatory ability of the models using all available input data and to test the models both on the data set from which they were developed and the other independent data set to see how portable the models were. We wanted to test whether these new approaches provided an improvement in damage prediction and to determine which group of input parameters are most important for model performance. We do not attempt to directly identify the factors controlling the propensity of trees to damage, which has been the subject of numerous previous studies (e.g. Albrecht et al., 2010; Colin et al., 2009; Dobbertin, 2002; Nicoll et al., 2006; Valinger and Fridman, 2011)

We also tested whether such ML models could replace the mechanistic model GALES by "learning" how to predict the critical wind speed for tree damage from a large number of GALES runs on data representing the range of conditions found in the Landes de Gascogne Forest. The purpose was to determine the potential of providing a faster method of calculating the vulnerability of forests, and one that could be represented in a relatively simple equation. This could allow rapid calculation of risk over large areas and be extremely helpful in testing different management and planning scenarios with the consequences immediately available to the end-users. Such ML models could also be used in optimisation of forest planning when there are multiple objectives and constraints (e.g. risk of wind damage) as previously demonstrated by Zeng et al., (2007).

2. Materials and Methods

2.1. General Approach

The general modelling approach followed was similar to Kamimura et al. (2016) (see their Fig. 2). The main differences are that models were developed separately using the National Forest Inventory data (NFI data), collected after Storm Klaus (Inventaire Forestier National. 2009*), and the Nezer Forest data, collected after Storm Martin (Chehata et al., 2014). The models were developed from each data set using a balanced selection of trees (similar number of undamaged and damaged trees) selected from 90% of the data (see Section 2.3.5 below). The models were then tested against the remaining 10% of the data (Part 2 of Fig. 1). This was repeated 10 times with a different 10% of the data being used for testing each time. Finally, both sets of models were tested with the other independent data by creating 10 versions of each model using a different selection of balanced data and testing against the whole of the other data set. This was to check how transferable the models were and to check their ability to predict the damage from a different storm from the one used in their development. In this paper we did not consider the type of damage (breakage or overturning) but combined all trees known to have been damaged by a storm.

In addition a set of models was developed to predict critical wind speeds (CWS) using an artificially generated data set to see if it was possible to substitute for GALES (Part 1 of Fig. 1). CWS calculated both by GALES and by these GALES substitute models were subsequently used in the development of the damage models along with characteristics of the individual trees, stand, and site (Part 2 of Fig. 1).

In the model development and validation we focussed on the CWS and WAsP calculations at 29 m above the ground for the Nezer Forest and at 40 m above the ground for the NFI data. This was to help maintain the focus of the paper and to ensure direct compatibility with Kamimura et al. (2016). Results for other calculation heights are presented in Appendix A and indicated where appropriate.

* https://inventaire-forestier.ign.fr/spip.php?article610

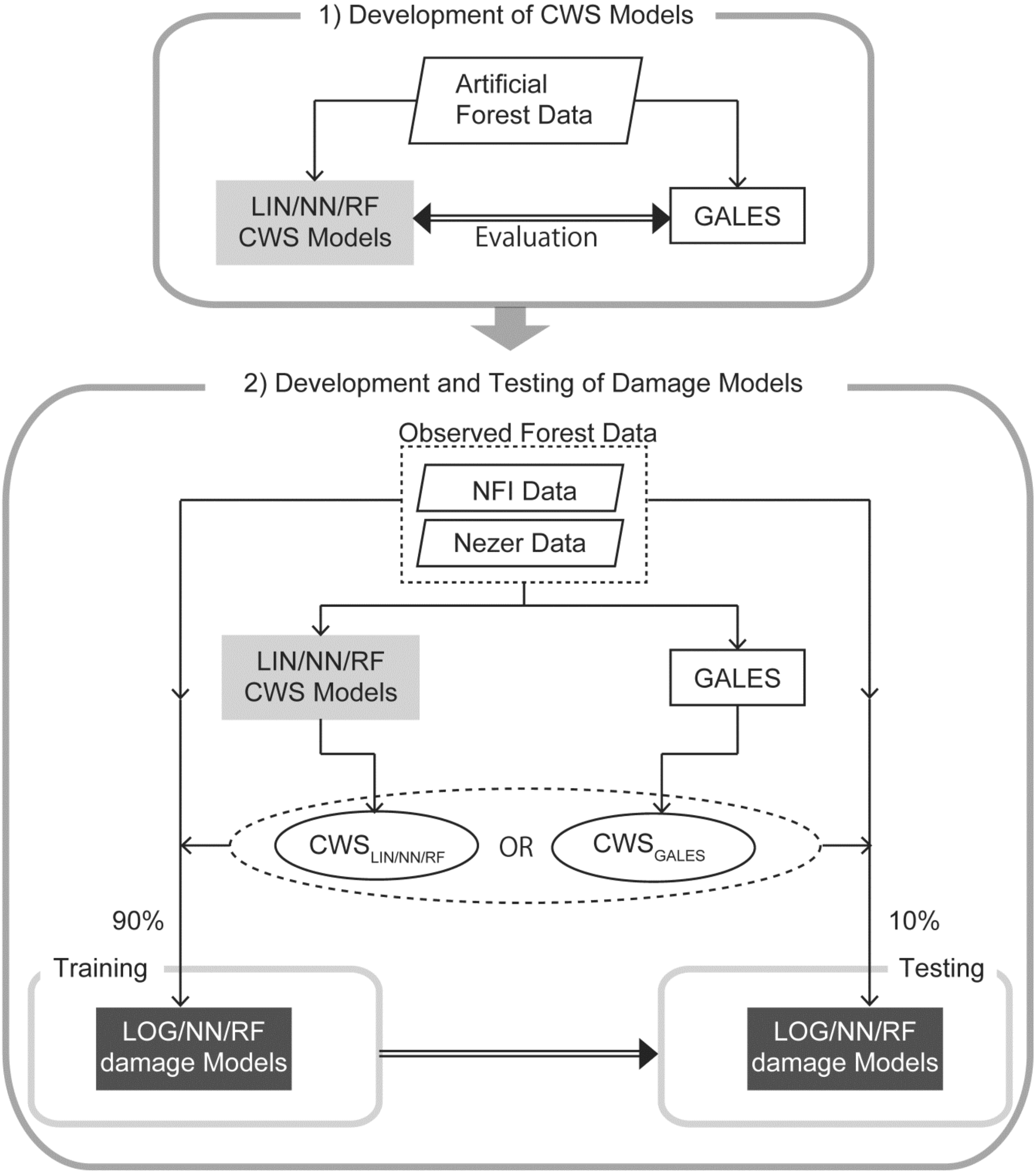


Fig. 1 Outline of modelling approach (LOG: logistic regression model, LIN: linear regression model, NN: artificial neural networks, RF: random forests; CWS: critical wind speed). In Part 1 (top) three modelling approaches (LIN, NN, RF) were trained to predict the CWS for damage based on a very large set (1970 individual trees) of previous simulations using GALES. In Part 2 (bottom) three modelling approaches (LOG, NN, RF) were trained (left-hand side) to predict damage using either the NFI or the Nezer Forest data (90% of data from each forest) together with either the GALES derived CWS, or the CWS values predicted using the models developed in Part 1. This produced a set of damage models (LOG/NN/RF) based on the Nezer Forest data and a set of damage models based on the NFI data. All damage models were then tested on the remaining 10% of the appropriate data set (right-hand side). The pattern of training and testing was repeated 10 times using 90% of the data for the training and a different remaining 10% of the data each time for validation. Compare with Fig. 2 in Kamimura et al. (2016).

2.2. Machine Learning Methods

Loosely inspired by biological neural networks, artificial neural networks (NN) are able to approximate a non-linear function to describe a mapping between a set of inputs and outputs. They are able to learn from incomplete and noisy datasets, making them particularly suitable for applications within forestry where data is hard to collect and likely to contain inaccuracies due to measurement difficulties. Previous applications of NNs in forestry have dealt with mortality estimation (Guan and Gertner, 1995; Hasenauer et al., 2001), and uncertainty assessment of forest growth models (Guan et al., 1997). However, a weakness in the neural network approach is that the learned function describing the non-linear mapping cannot be easily understood in terms of processes controlling behaviour, e.g. wind damage in forests. They are therefore tools that can be of practical use but do not easily provide scientific insight.

Random forests (RF) are a more recent technique (Breiman, 2001) that have also proved successful in developing models from noisy and unbalanced data. The RF algorithm builds a collection of independent decision trees whose results are combined to make a prediction for a given data record. The technique has the advantage of being very fast to train, and typically overcomes overfitting problems associated with decision tree methods. They are becoming extremely popular in many aspects of forest modelling (e.g. Seidl et al., 2011).

Logistic regression models (LOG) have been regularly used in assessing the risk of wind damage because their dependent variables are categorical and if the binary dependent variable is binary (0/1) they are ideal for wind damage prediction (damaged/undamaged). In particular, logistic regression models can be used to identify which factors are associated with wind damage. In this paper, a logistic regression model similar to those developed by Albrecht et al. (2012), Valinger and Fridman (2011) and Kamimura et al. (2016) was used.

2.3. Software and Methods

The WEKA software "workbench" (Waikato Environment for Knowledge Analysis) incorporates a large number of standard Machine Learning Techniques (ML) including the methods described above in a freely available tool (Frank et al., 2016). With it, a specialist in a particular field is able to use ML to derive useful knowledge from databases that are far too large to be analysed by hand. The workbench can either be used through a supplied Graphical User Interface, or incorporated directly in Java code using a supplied library. All experiments described here are conducted using Weka version 3-6-13. The three models used are described below. The NN and RF can be both be trained as classifiers, i.e. predicting a class value (damaged/no damage) or to undertake regression, i.e. output a continuous value. We did not attempt any model tuning in order to determine how well the WEKA software performed “off the shelf”.

2.3.1. Artificial Neural Network

The *artificial neural network* contains an input layer consisting of n neurons, each corresponding to one of the selected inputs variables. In classification mode, the output layer contains two neurons, one indicating the positive class, and the other the negative class. When used for regression, there is a single output neuron. In addition, there is a single hidden layer consisting of (inputs+outputs)/2 neurons. Each neuron receives a weighted sum of inputs $x = \sum_{i=1}^{i=k} w_i v_i$, where v_i = the value of the input and w_i the weight connecting the input to the neuron, and outputs a value $s(x)$ using a sigmoid activation function as defined in Eq. 1:

$$(x) = \frac{1}{1+e^{-cx}} \quad (1)$$

Weights are initialised at random and the *backpropagation* algorithm used to find a set of weights that minimizes the total error at the outputs, summed over all input records:

$$E = \frac{1}{2}\sum_{i=1}^{p} \parallel o_i - t_i \parallel^2 \quad (2)$$

Backpropagation is a gradient descent technique that modifies each weight in small steps based on the gradient of the error function with respect to the weight concerned, e.g.

$$w_n = w_n - \eta \frac{\delta E}{\delta w_n} \quad (3)$$

where w_n is the total error calculated at each step. The learning rate η is an adjustable parameter that modifies the step size, but was set to 0.3 in all our experiments. An additional *momentum* term is used that enables the gradient descent algorithm to escape from local minima, and is set to a default value of 0.2. Backpropagation is applied for a fixed number of 500 iterations for each model. These represent the default settings in the WEKA software.

2.3.2. Random Forests

The *Random Forests* algorithm uses a bagging approach, combined with a Random Tree learning algorithm. In bagging, multiple random subsets of the dataset are created by sampling n instances with replacement from the dataset. For each subset, a random tree classifier is grown: at each node, m variables are selected at random, from which the one that optimizes the information gain is chosen. We use the default Weka parameters: a forest of 100 random trees are created; each tree has unlimited depth and is grown without pruning; at each node m = log2(*number_of_attributes*) + 1 are randomly selected.

2.3.3. Logistic Regression

Logistic regression estimates the probability of a binary response variable based on the set of predictor inputs. The Weka implementation of the multinomial logistic regression model with a ridge estimator is loosely based on the description given by Le Cessie and Van Houwelingen (1992).

Given k classes, and n instances with m attributes, an m*(k-1) parameter matrix $\boldsymbol{\beta}$ is calculated. The probability of class i is given by Eq. 4 where Y_i are the mutually independent response variables (1,0), $p(X_i)$ is the probability that $Y_i = 1$, and X_i are the m-dimensional rows of covariates.

$$p(X_i) = \frac{\exp(X_i\beta)}{\{1+\exp(X_i\beta)\}} \qquad (4)$$

The log likelihood is given by Eq. 5. A ridge estimator is used to improve the parameter estimates and diminish the error made by further prediction. In order to find the matrix β for which l is minimised, a Quasi-Newton Method is used to search for the optimized values of the m*(k-1) variables. Before Weka runs the optimization procedure, the matrix β is compressed into a m*(k-1) vector. The default Weka parameter for the ridge estimator λ of $1\text{x}10^{-8}$ is used.

$$l = \sum_i \ [Y_i \log\left(p(X_i) + (1-Y_i)\log\{1-p(X_i)\}\right] + \lambda * \ \beta^2 \qquad (5)$$

2.3.4. Models

We evaluate the models above with respect to two functions:

- Damage Prediction: We adopted a dichotomous model which predicts damage at the level of individual trees in two categories, damaged or undamaged. A separate model was trained for each of the two data sets. For each of the three classification methods described, the default parameters supplied with Weka were used to train the model.
- Critical Wind Speed Prediction: A linear regression model (LIN) was used instead of the logistic regression model (LOG) because it is more appropriate for a variable output (non-dichotomous). All models (LIN, NN, RF) were trained to predict critical wind speeds for breakage and overturning at tree level using values obtained from running a GALES simulation as training data (see 2.4.1 below). The variables used to train the models are given in Table 1.

2.3.5. Training and Pre-Processing

Cross-validation is used to obtain an unbiased estimate of the performance of each model on unseen test data. For each model, the dataset is randomly divided into 10 subsets (folds) of equal size. 9 folds are combined to train a model, with the left-out fold used for testing the trained model. The procedure is repeated leaving each of the 10 folds out in turn. The final reported accuracy is the average of the accuracy value obtained on each of the 10 folds.

For damage prediction, given that the data is unbalanced in terms of the ratio of damaged/undamaged trees, it is preferable to bias the data used to train the models towards a uniform class distribution. The Weka *SpreadSubsample* filter is applied to the subset of data used in each training fold during cross-validation: this produces a new dataset twice the size of the minority class, by selecting all instances of the minority class (damaged tree in this case) and randomly sampling from the majority class (undamaged trees in this case). In order to eliminate variability due to the effects of random sampling in this way, 10 new data-sets were created as just described. All models are trained and tested as described above using each sub-sampled data-set, with mean results, standard deviations and/or boxplots used to report findings.

2.3.6. Outputs from each model

Damage-prediction models: For the NN, Weka returns a probability distribution based on the outputs from the network defining the probability of a tree being damaged, for each input vector. The discrimination threshold is set at 0.5, such that a probability of greater than or equal to 0.5 results in the tree being classified as damaged. The same threshold is used with the LOG and the RF models. No adjustment of this threshold was made in order to determine how well the models performed without any tuning.

Critical wind-speed models: the LIN, NN and RF models output a single real-valued number for the critical wind speed for breakage and a single real-valued number for the critical wind speed for overturning.

2.3.7. Performance Metrics

For the dichotomous models, we record *classification accuracy*, i.e. the proportion of true results (both true positives and true negatives) among the total number of cases examined. In addition, we report the area underneath the receiver-operating curve (AUC). This plots the false positive rate against the false negative rage: a perfect classifier would have an AUC of 1.0; an area of 0.5 is equivalent to random guessing. Typically, an AUC > 0.7 is considered to be *fair*, above 0.8 *good* and above 0.9 to be *excellent* (Hosmer and Lemeshow, 2000).

For prediction of numeric values (i.e. critical wind speed) the correlation coefficient is reported. All statistics were either calculated within the WEKA software or with Matlab 2016a (Mathworks, Natick MA, USA).

2.4. GALES

GALES is a hybrid mechanistic model for predicting the critical wind speeds (CWS) for damage to forest stands and trees due to overturning and breakage and is designated a CWS model in the convention adopted by Gardiner et al. (2008). If wind climate data is available then the probability of such wind speeds being exceeded and damage occurring is also calculated, and this version of the model is called ForestGALES and is designated a Wind Risk Management tool (WRM) using the same designation system. GALES requires information on the tree species, tree diameter at breast height (*DBH*), tree height, stand mean tree diameter at breast height (DBH_{mean}), stand mean tree height, mean stand spacing, soil type and rooting depth. Although GALES calculates the CWS for both stem breakage and overturning (uprooting), in this paper the CWS used in damage model development is always the minimum of the two, i.e. the most likely to occur and we did not attempt to discriminate between damage types.

Full details of the model and its validation can be found in Gardiner et al. (2000) and Hale et al. (2015). The parameters in GALES used for maritime pine stands are given in Cucchi et al. (2005).

2.4.1. GALES artificial training dataset

A large number of potential maritime pine stands with characteristics that covered the full range of possible characteristics (see Table 1 for details of the ranges sampled) were created as inputs to GALES. The stand characteristics were selected using Latin Hyper Cube Sampling to give uniform sampling. 10,000 stands were created, which after filtering for duplicates, constraining the ratio of stand mean tree height to stand mean *DBH* between 30 (very high taper) and 130 (very low taper), and constraining individual tree *DBH* and height to be within ±70% of the stand mean values, left 1970 simulations.

GALES was then run for the 1970 stands and the CWS values for tree overturning and stem breakage were calculated at 10 m above the zero-plane displacement (*d*+10m), which is the standard height for such measurements in Gardiner et al. (2000) and at 29m and 40m above the ground, which correspond to the maximum tree heights in the Nezer Forest and in the whole of the NFI data set respectively (Kamimura et al., 2016).

The outputs from the GALES runs were then used to train LIN, NN and RF models to predict CWS for overturning and breakage at *d*+10 m, 29 m and 40 m. The trained models were finally tested by comparing their predictions of CWS against GALES calculated CWS at *d*+10 m and 29 m for the Nezer Forest and at *d*+10 m and 40 m for the NFI data (see Part 1 in Fig. 1).

2.5. WAsP predicted wind speeds

The Wind Atlas Analysis and Application Program (WAsP) (Mortensen et al., 1993) was used to estimate the wind speeds above the forest during the Martin and Klaus storms. A land-use map (elevation range, 0 to 300 m; contour interval = 50 m) plus an aerodynamic roughness map (water = 0.003 m; unforested areas = 0.01 m; forest =1.0 m) was used in the simulations. The input wind speeds for WAsP were taken from the coastal meteorological station at Cap Ferret (approximately 25 km north-west of the Nezer Forest at 44°38'N, 1°15'W). Wind speeds were simulated at a horizontal resolution of 500 x 500 m, at a height of 29 m (just above height of tallest trees in the Nezer Forest) for storm

Martin, and at heights of 29 and 40 m (just above height of tallest trees in the NFI data) for storm Klaus. Full details are given in Kamimura et al. (2016).

2.6. Data

2.6.1. Study site and data

The field data used in this study are the same data as used in Kamimura et al. (2016). There are two groups of data. The first is from a field survey of 29 permanent plots ($400m^2 \cdot plot^{-1}$) in the Nezer Forest, located in Nouvelle-Aquitaine region (44°34'20''N, 1°2'20''W). Tree size was surveyed in 1998, and damaged trees were determined after storm Martin in 1999 (Table 2). Data consist of tree height, stem diameter at breast height (*DBH*, 1.3 m), tree location, and damage status for most trees. The data was not sub-divided as was the case in Kamimura et al. (2016). The second data set was from field surveys of the National Forest Inventory in France (Inventaire Forestier National; NFI, (Robert et al., 2009)) in the same region, which is predominately maritime pine stands. The annual survey plots (1 point for 10 km^2) are chosen in a systematic sub-sample of the 5-year sample covering the entire country. The forest field plots are composed of four concentric plots allowing the measurement of different tree diameter classes (Robert et al., 2009). We used data collected from 2007 to 2008 from a total of 235 plots chosen in two ecological regions of the Landes de Gascogne Forest, and wherever more than half of the trees in each plot were maritime pine. After storm Klaus in 2009, damaged trees in the NFI plots were identified by an additional follow up field survey to list damaged trees (Table 2). For each plot in the two data sets we added mean plot height, the mean plot *DBH* and the average stem spacing derived from the individual tree data. Spatial information included the distance of each tree from the windward stand edge (west) and the upwind gap size (distance in a westerly direction between the forest and the next forest block) were also estimated based on the position of the inventory plot (only accurate to within 500m). However, in this paper we assumed like Kamimura et al. (2016) that all the trees were effectively at a new edge because the best results were previously found with this assumption. This assumption is justified by the observation from aerial photography that damage propagated through stands during the storms and this led to new trees becoming exposed to an advancing damaged forest edge. The NFI plots were identified either within the Landes (main forest production area inland from the coast) or Dunes (forest along coastal dunes) areas based on the ecological region given in the NFI survey, whereas all the plots in the Nezer Foret were designated as Landes. Soil characteristics and hydrological status were derived from the French soils database (GISsol, 2011) and the ecological observations in the NFI plots (Bruno and Bartoli, 2001). Soils are mainly sandy podzols and arenosols, respectively in the Landes and in the Dunes areas. Gleys and brown soils are also present but only in the Landes area. In the Nezer Forest the soils are hydromorphic podzols, and their dominant hydrological status is "slightly wet". Soil depth is greater in the Dunes and Landes area with a dry hydrological status than in those Landes areas with a wetter hydrological status. An outline of the data used in the development of the models is provided in Table 3.

Full details of the data and the calculation of derived parameters is provided in Kamimura et al. (2016) and the location of the forests and the individual sample plots is given in Fig. 1 of Kamimura et al. (2016).

Table 1: Characteristics of the data set used to train the LIN, NN and RF models to simulate GALES critical wind speed predictions for maritime pine

Model Variable	*Mean Value*	*Range*	*Comment*
Soil	3	None	Fixed as *podzol*
Rooting	2	None	Fixed as Deep rooting ≥ 80 cm
Upwind gap width (m)	245.6	0-500	When gap = 0m then tree is effectively inside forest
Position relative to edge (m)	0	None	Fixed to always be at stand edge
Tree DBH (cm)	41.9	2.5-110	
Tree Height (m)	23.6	2.5-40	40 m is just above the maximum tree height of maritime pine in Landes de Gascogne Forest
Tree taper (m/m)	23.6	30-130	Constrained between 30 and 130 so trees not too thin or too tapered
Stand DBH (cm)	43.9	5-65	
Stand height (m)	24.8	2.5-35	
Stand taper (m/m)	60.9	30-130	Constrained between 30 and 130 so trees not too thin or too tapered
Tree DBH/Stand DBH	0.98	0.3-1.7	Constrained that tree size is within range ±70% of stand size
Tree height/Stand height	0.98	0.3-1.7	Constrained that tree size is within range ±70% of stand size
Stand density (trees/ha)	1840	30-3600	

Table 2: Levels of damage in the Nezer Forest and within the NFI database.

Data	*Number of Trees*	*% Damaged*	*% Undamaged*
Nezer Forest	1080	12% (134 trees)	88% (946 trees)
NFI	1705	33% (566 trees)	67% (1139 trees)

Table 3: Parameters and their range and standard deviation used in the model development for Nezer Forest and the NFI database. *DBH* is diameter at breast height (1.3m above ground) and *CI_BAL* is a competition index based on the basal area of all trees larger than the subject tree (Biging and Dobbertin, 1995)

Model Variable	*NFI: Range (Stdev)*	*Nezer Forest: Range (Stdev)*
Gap size (m)	41-328.2 (66.7)	28.4-262.5 (66.4)
Stand Mean DBH (cm)	8.0-65.1 (12.9)	3.9-43.4 (10.6)
Stand Mean Height (m)	4.1-32.8 (6.7)	2.8-26.3 (6.4)
Stand Density (ha)	28.3-2740.7 (399.7)	200-3594 (676.1)
Stand Mean CI_BAL	0.00-57.9 (9.7)	1.1-19.6 (6.6)
Tree DBH (cm)	7.6-111.00 (14.4)	2.5-61.0 (11.3)
Tree Height (m)	3.60-38.60 (6.9)	2.3-26.7 (6.6)
Tree CI_BAL	0.00-270.7 (18.1)	0.00-35.9 (9.7)
Distance from Edge (m)	0	0
CWS Breakage at d+10m GALES (ms^{-1})	10.9-45.4 (5.8)	12.7-46.2 (8.0)
CWS Overturning at d+10m GALES (ms^{-1})	10.0-32.5 (5.2)	11.3-40.0 (7.2)
CWS Breakage at 29m GALES (ms^{-1})	16.0-58.8 (5.5)	24.3-60.8 (7.6)
CWS Overturning at 29m GALES (ms^{-1})	13.7-48.2 (5.1)	25.0-53.7 (6.7)
CWS Breakage at 40m GALES (ms^{-1})	20.3-63.6 (5.8)	Not calculated
CWS Overturning at 40m GALES (ms^{-1})	18.8-52.2 (5.3)	Not calculated
WAsP predicted wind speeds at 29m (ms^{-1})	21-42 (4.5)	26.2-31.8 (1.8)
WAsP predicted wind speeds at 40m (ms^{-1})	24-43 (4.4)	Not calculated
Soil (1=Arenosol, 2=brown soils, 3=podzol, 4=gleys)	1-4	3
Hydro (1=very wet, 2=slightly wet, 3=dry)	1-3	2
Dune (1=Dune area, 0 = Landes area)	0-1	0

3. Results

3.1. Predicting CWS

The LIN, NN and RF model simulations of CWS were compared to the actual CWS produced by GALES for the Nezer and NFI data at 29 m and 40 m above the ground respectively, and are displayed in Table 4. Information for predictions at *d*+10 m can be found in Table A1 in Appendix A.

Table 4: Results of comparison of predictions from the trained LIN/NN/RF models and GALES for Nezer at 29 m and NFI data at 40 m. Numbers are correlation coefficient between trained model results and GALES predictions and root-mean square (RMS) error is given in brackets in ms^{-1}.

Training Set	*Test Set*	*Output*	*LIN*	*NN*	*RF*
GALES 29 m predictions from artificial data	Nezer	CWS for breakage	0.8836 (6.4165)	0.9251 (10.2185)	0.9137 (6.5713)
GALES 29 m predictions from artificial data	Nezer	CWS for overturning	0.9131 (3.0748)	0.9516 (3.838)	0.9394 (4.6022)
GALES 40 m predictions from artificial data	NFI	CWS for breakage	0.7659 (6.0699)	0.8565 (4.8805)	0.8437 (4.6879)
GALES 40 m predictions from artificial data	NFI	CWS for overturning	0.8264 (3.6150)	0.9347 (3.398)	0.9004 (2.8682)

The results show a high level of correlation between the predictions of GALES and those of the models. In all cases the models are correlated to the GALES predictions with r^2 values greater than 0.77 and in most cases above 0.9. In all cases the predictions of breakage are slightly less well correlated than the predictions of overturning. This might be a reflection of the fact that only approximately 15% of trees were damaged by breakage during the two storms (trees in the Landes de Gascogne Forest are more susceptible to overturning), and the models are consequently better trained to predict overturning than breakage (more examples of overturning). In all cases the LIN models perform least well, the RF second best and the NN performs best (average correlations of 0.847, 0.899 and 0.917 respectively). However, the RMS errors in the predictions are quite large with values ranging between 2.87 to 10.22 ms^{-1}, and with an average value of 5.02 ms^{-1}. This suggests that such models can be used for predictions for multiple trees and forest stands over large areas but not for precise predictions for a small number of trees or individual stands. Overall the models appear better at predicting the CWS at d+10 m rather than at fixed heights with r^2 values greater than 0.94 (see Table A1 in Appendix A). This is probably due to the fact that d+10 m is at a relatively consistent height above the modelled trees (<10 m), whereas with the fixed height values of 29 and 40 m the distance from the top of the trees to the calculation height is much more variable (22.5 to 37.5 m).

A large advantage was obtained in computational efficiency. The GALES model used in this paper required 0.37 ms to calculate the CWS for damage of a single tree using already known tree characteristics, whereas the LIN and NN derived models only required 0.013 ms per tree. This represents a 28 times increase in calculation speed. The RF derived CWS model required 0.065 ms per tree, a calculation speed more than 5.7 times faster than GALES. In the GALES version of Gardiner et al. (2000) there is an iterative solution for calculating the additional moment provided by the overhanging displaced mass of the canopy during a storm (Neild and Wood, 1999), whereas in in this paper we used a simple analytical bending equation (Gardiner, 1992). Additional simulations showed that a further computational efficiency of a factor of 2 would be obtained over the more complicated version of GALES. All calculations were based on 10 runs for all 1705 trees in the NFI data set using a MathCad program (PTC, Needham, United States) on a Dell Latitude© laptop (Dell, Round Rock, United States) running at 2.1 GHz (4 CPUs) with 16.0 GB of memory.

3.2. Wind damage to individual trees

3.2.1. Nezer Forest

In Fig. 2 the performance of the three damage modelling approaches (LOG/NN/RF) in predicting damage or no damage for the Nezer Forest is illustrated (LOG_Nez, NN_Nez, RF_Nez). All the parameters in Table 3 were used with the GALES CWS and WAsP wind speed calculated at 29 m. The accuracy and AUC values are given in the *All Variables* column (indicating all possible variables used) in Table 5 and Table 6 respectively. The accuracy of the three models are all reasonably good (≥ 67%) but the NN model has a significantly higher accuracy than the LOG model with a value of 68.7% and the RF model has a statistically significantly higher accuracy than both other models with a value of 72.5%. All three models have high values of AUC (≥ 0.8), which indicate *good* discrimination between damaged and undamaged trees (Hosmer and Lemeshow, 2000). The AUC values for all three models are higher than the value obtained by Kamimura et al. (2016) for the Nezer Forest using logistic regression models (AUC = 0.76). However, the accuracies are lower for the LOG and NN models in comparison to the earlier work, which had an accuracy of between 71.9-72.4% in the Nezer Forest. However, in Kamimura et al. (2016) the model accuracy was optimized by adjusting the cut points for the probability of damage between 0 and 1 until the true positive rate equalled the true negative rate (Hosmer and Lemeshow, 2000). As described earlier, in this paper no model optimisation was performed and the cut point was fixed at 0.5 in order to determine model performance with no tuning.

The accuracy and AUC of the models for the same data but using the calculated critical wind speeds at d+10 m above the ground are presented in Fig. A.1 and Tables A2 and A3 of Appendix A. The results are very similar to the results using the CWS at 29 m and suggest that the height of CWS calculation is not especially critical and the inclusion of the WAsP calculated wind speeds made little difference to the accuracy or discriminatory ability of the models.

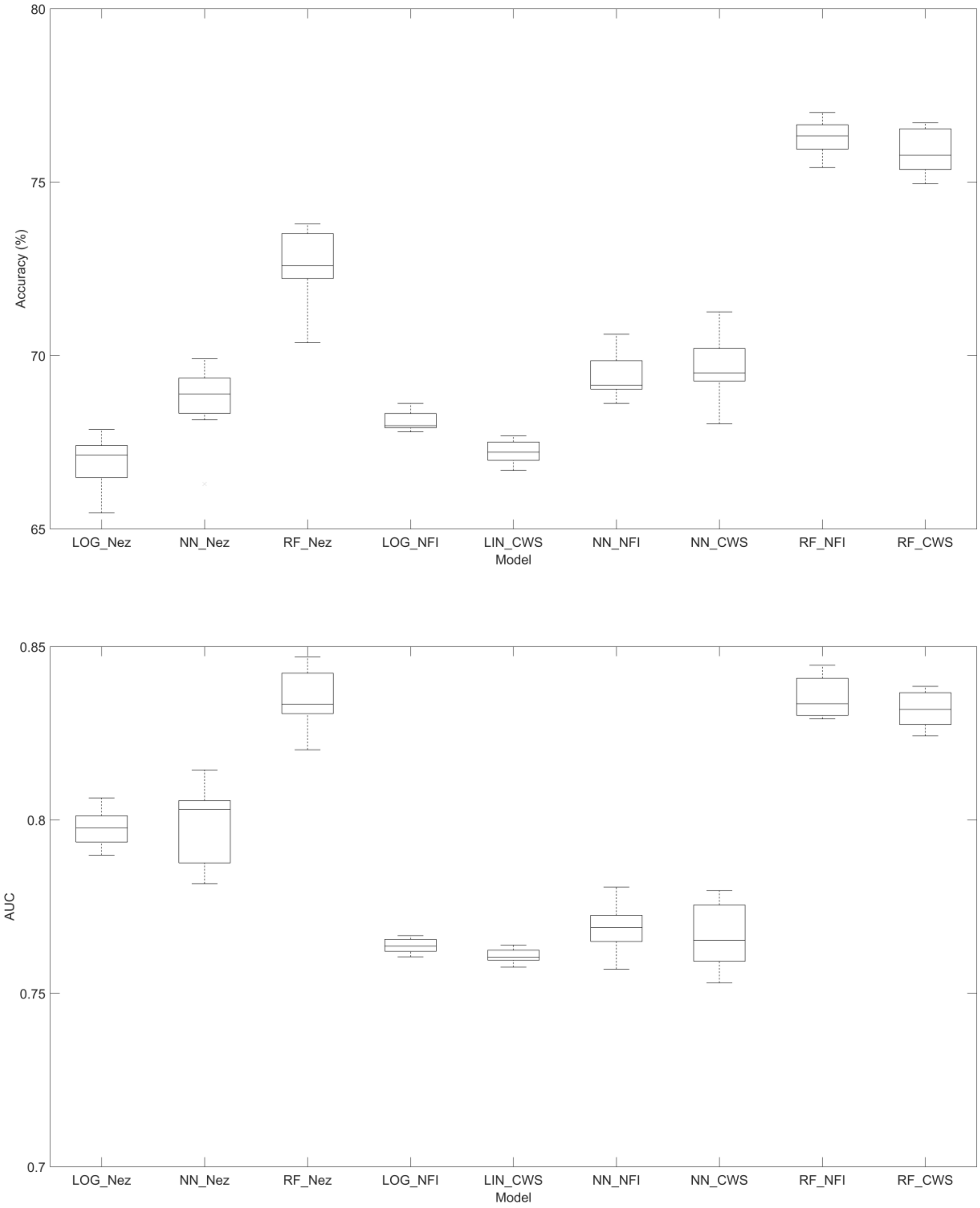


Fig. 2: Accuracy and AUC for the LOG, NN and RF damage model predictions using all data tree, stand and site data and the GALES predicted CWSs at 29 m against the Nezer Forest data (LOG_Nez, NN_Nez, RF_Nez) and the GALES predicted CWSs at 40 m against the NFI damage data (LOG_NFI, NN_NFI, RF_NFI). In addition a comparison is made for the NFI data (LIN_CWS, NN_CWS, RF_CWS) using the CWS values derived (see Part 1 of Fig. 1) from the three CWS models (LIN, NN, RF) instead of the GALES values.

3.2.2. NFI data (Landes de Gascogne Forest)

In Fig. 2 there is also the same analysis as presented for the Nezer Forest data but for the NFI data and using the GALES CWS and WAsP predicted wind speeds at 40 m (LOG_NFI, NN_NFI, RF_NFI). The values are tabulated in Table 5 and

Table 6. In addition the results using the model predicted CWSs calculated in Section 3.1 were also used (LIN_CWS, NN_CWS, RF_CWS) in place of the GALES derived CWS. The accuracies of the LOG and NN models are very similar to the logistic regression model of Kamimura et al. (2016) where the accuracy was 69.6% when the NFI data were used (see Table 8 in Kamimura et al., 2016), but the RF model is significantly more accurate (76.3%). The discriminatory behaviour of the LOG and NN models is also similar to the logistic regression model in Kamimura et al. (2016) with AUC values close to 0.77 compared to their value of 0.74. However, the RF model shows superior discriminatory power with an AUC value of 0.84. In the simulations using the model predicted CWSs in place of the GALES derived CWS (LIN_CWS, NN_CWS, RF_CWS) the AUC values are unaffected and only the accuracy of the simulations using the CWS derived from the linear regression model (LIN_CWS compared to LOG_NFI) showed a significant reduction ($p=0.0164$).

The results for the NFI data using calculations at d+10 m and 29 m and are shown in Fig. A.2 and Fig. A3, and Tables A2 and A3 in appendix A. They are very similar to the results presented here.

3.2.3. Model Sensitivity to Individual Parameters

The effects of leaving out one variable at a time on the accuracy and AUC value of the models for the Nezer Forest using the CWS and WAsP wind speed calculated at 29 m are given in Table 5 and Table 6 and plotted in Fig. A.4 of Appendix A. For each variable removal the model was always retrained with the remaining variables. The model performance using the CWS calculated at d+10 m are displayed in Fig. A.5 and tabulated in Tables A.2 and A.3 of Appendix A.

Variable removal only has an effect for the LOG model where the removal of stand density and mean stand *DBH* slightly reduce the accuracy and the removal of stand density slightly reduces the AUC (all significant at the $p=0.05$ level). However, for the NN and RF models the removal of no variable had a significant effect on either model accuracy or AUC. Note that in all the Nezer Forest simulations removing *Dune*, *Hydro* and *Soil* have no impact because they each only have a single value in this forest (Table 3).

The response of the models developed using the NFI data and the CWS and WAsP wind speed calculated at 40 m are also tabulated in Table 5 and Table 6 and plotted in Fig. A.6 of Appendix A. The results for the model performance using the CWS calculated at 29 m and d+10 m are displayed in Fig. A.7 and Fig. A.8 and Tables A.2 and A.3 of Appendix A. Removal of *Stand_density*, *Dune* and *Hydro* reduces the accuracy and AUC of the LOG model and additionally the removal of *Soil* and the WAsP calculated wind speed reduces the AUC of the LOG model. The NN model is only affected by the removal of *Hydro*, which reduces the AUC of the model. The RF model is not affected by the removal of any variable.

Overall there is relatively little impact of parameter removal on model performance. The LOG model is the most sensitive and the RF model almost completely insensitive. This is probably not surprising because of the way that the LOG and NN models utilise all the available variables, whereas the RF model creates nodes at each of which *m* variables are selected at random, from which the one that optimizes the information gain is chosen. Interestingly the removal of information on whether in the Dune or Landes area (*Dune*), the hydrological state of the soil, and to a lesser extent the soil type itself had an impact on the LOG and NN model developed using the NFI data. This suggests that this information provides an improvement in discrimination between damage and no damage but, because these variables are not strongly correlated to other variables, the models cannot create an equally effective alternative model when this information is missing.

Table 5: Mean accuracy of different models with each model variable removed in turn. Standard deviation is given in brackets. * indicates value significantly different ($p<0.05$) from the value with using all variables. The superscript letters against the values in the *All Variables* column (a, b, or c) indicate whether there are significant differences between the models for that particular height of CWS calculation at the $p=0.5$ level.

Data Set	*Model*	*CWS Height*	*All Variables*	*Average CI_BAL*	*CI_BAL*	*Tree DBH*	*Stand Density*	*Dune*	*Gap Size*	*Hydro*	*Stand DBH*	*Soil*	*Stand Height*	*Tree Height*	*CWS Break*	*CWS Overturn*	*WAsP Wind Speed*
Nezer	LOG	29 m	66.954[a] (0.76)	67.287 (0.801)	67.065 (0.929)	67.000 (0.688)	65.028* (0.772)	66.954 (0.76)	66.954 (0.76)	66.954 (0.76)	65.593* (0.581)	66.954 (0.76)	66.954 (0.76)	67.435 (1.053)	66.944 (1.206)	67.102 (0.795)	66.213 (1.042)
	NN		68.741[b] (1.028)	68.019 (1.329)	67.88 (0.961)	67.991 (1.573)	68.463 (1.407)	68.000 (1.279)	67.991 (1.176)	68.000 (1.279)	68.565 (1.107)	68.000 (1.279)	68.019 (1.414)	68.074 (1.621)	69.75 (2.046)	68.639 (1.162)	67.278 (1.054)
	RF		72.528[c] (1.02)	72.167 (1.164)	72.519 (0.801)	73.056 (1.011)	72.259 (0.83)	72.565 (0.903)	72.287 (0.952)	72.611 (0.704)	72.352 (0.783)	72.481 (0.877)	72.454 (0.836)	72.491 (0.918)	72.843 (0.95)	72.426 (0.864)	72.065 (0.924)
NFI	LOG	40 m	68.094[a] (0.283)	67.894 (0.282)	68.158 (0.277)	68.258 (0.321)	67.232* (0.212)	66.780* (0.324)	68.094 (0.283)	67.120* (0.373)	68.188 (0.458)	67.918 (0.322)	68.094 (0.283)	67.648 (0.215)	68.106 (0.269)	67.988 (0.335)	67.877 (0.303)
	NN		69.443[b] (0.679)	69.238 (0.672)	69.959 (0.990)	70.006 (0.643)	69.484 (0.665)	69.496 (0.957)	69.543 (0.345)	68.528 (0.684)	68.979 (0.911)	68.686 (0.548)	69.138 (1.045)	69.736 (1.382)	69.865 (0.725)	69.460 (0.858)	69.056 (0.499)
	RF		76.305 (0.466)[c]	75.701 (0.342)	76.587 (0.632)	76.493 (0.483)	75.900 (0.528)	76.534 (0.723)	76.076 (0.431)	75.742 (0.437)	76.082 (0.575)	76.328 (0.430)	76.100 (0.474)	76.211 (0.348)	76.217 (0.495)	76.416 (0.455)	75.672 (0.530)

Table 6: Mean AUC of different models with each model parameter removed in turn. Standard deviation is given in brackets. * indicates value significantly different ($p<0.05$) from the value with using all variables. The superscript letters against the values in the *All Variables* column (a, b, or c) indicate whether there are significant differences between the models for that particular height of CWS calculation at the $p=0.5$ level.

Data Set	*Model*	*CWS Height*	*All Variables*	*Average CI_BAL*	*CI_BAL*	*Tree DBH*	*Stand Density*	*Dune*	*Gap Size*	*Hydro*	*Stand DBH*	*Soil*	*Stand Height*	*Tree Height*	*CWS Break*	*CWS Overturn*	*WAsP Wind Speed*
Nezer	LOG	29 m	0.798[a] (0.005)	0.8 (0.005)	0.799 (0.005)	0.8 (0.006)	0.78* (0.005)	0.798 (0.005)	0.798 (0.005)	0.798 (0.005)	0.793 (0.006)	0.798 (0.005)	0.798 (0.005)	0.803 (0.004)	0.793 (0.006)	0.798 (0.005)	0.8 (0.005)
	NN		0.799[a] (0.011)	0.799 (0.012)	0.804 (0.01)	0.794 (0.011)	0.795 (0.015)	0.797 (0.013)	0.793 (0.012)	0.797 (0.013)	0.791 (0.021)	0.797 (0.013)	0.796 (0.011)	0.797 (0.011)	0.8 (0.01)	0.795 (0.011)	0.797 (0.013)
	RF		0.834[b] (0.009)	0.834 (0.006)	0.832 (0.008)	0.839 (0.008)	0.835 (0.008)	0.837 (0.008)	0.837 (0.007)	0.836 (0.01)	0.836 (0.009)	0.835 (0.009)	0.836 (0.008)	0.832 (0.011)	0.837 (0.008)	0.836 (0.009)	0.835 (0.008)
NFI	LOG	40 m	0.764[a] (0.002)	0.765 (0.002)	0.764 (0.002)	0.763 (0.002)	0.757* (0.002)	0.751* (0.002)	0.764 (0.002)	0.745* (0.002)	0.765 (0.002)	0.760* (0.002)	0.764 (0.002)	0.763 (0.002)	0.764 (0.002)	0.762 (0.002)	0.758* (0.002)
	NN		0.769[a] (0.007)	0.766 (0.008)	0.771 (0.007)	0.773 (0.006)	0.767 (0.004)	0.767 (0.011)	0.765 (0.005)	0.749* (0.008)	0.765 (0.006)	0.759 (0.008)	0.764 (0.006)	0.769 (0.009)	0.772 (0.009)	0.768 (0.006)	0.764 (0.006)
	RF		0.836[b] (0.006)	0.832 (0.006)	0.838 (0.005)	0.835 (0.005)	0.832 (0.005)	0.833 (0.006)	0.833 (0.005)	0.830 (0.005)	0.832 (0.007)	0.835 (0.005)	0.834 (0.006)	0.838 (0.006)	0.835 (0.006)	0.836 (0.006)	0.831 (0.005)

3.2.4. Model Sensitivity to Removal of Parameter Groups

The sensitivity of the models to the absence of groups of input variables was also tested. Four parameter groups were defined as *Stand* = {*Gap Size*, *Stand Mean DBH*, *Stand Mean Height*, *Stand Density*, *Stand Mean CI_BAL*}; *Tree* = {*Tree DBH*, *Tree Height*, *Tree CI_BAL*}, *Site* = {*WAsP 40m*, *Dune*, *Hydro*, *Soil*} and *CWS+WAsP* = {*CWS Breakage*, *CWS Overturn*, *WAsP 40m*}. The results are illustrated in Fig. 3.

There are clear differences in the behaviour of the three models. The LOG and NN models are badly affected by the removal of *Site* information and this was not compensated for by *Tree* or *Stand* information. *Site* information on its own reduced the performance of both the models by a large and significant amount and this reflects the findings from the single parameter removal in Section 3.2.3 that showed the LOG and NN models are sensitive to the removal of *Dune*, *Hydro*, or *Soil* information. Removal of *Stand* information had a small but significant influence on the LOG and NN models, but removal of just *Tree* information did not significantly affect the results. For the RF model the story is different and the loss of *Stand* information is the most important factor. In fact *Stand* information on its own is enough to produce high model accuracy and AUC values. In addition, the RF model results were slightly but significantly improved when *Tree* level information was <u>excluded</u>. The *CWS+WAsP* information on its own provided reduced but reasonable levels of accuracy and AUC for all models, and generally gave higher or equivalent results compared to any other single parameter group (except *Stand* with the RF model) suggesting that the GALES model does provide a reasonable assessment of damage risk in these forests.

In summary, all models benefit from *Stand* level information and results are improved in particular by *Site* information for the LOG and NN models. The LOG and NN models are unaffected and the RF model is slightly adversely affected by the inclusion of *Tree* information and all models performed reasonably, but with reduced accuracy and discrimination, when just the CWS values and the WAsP wind speed were used.

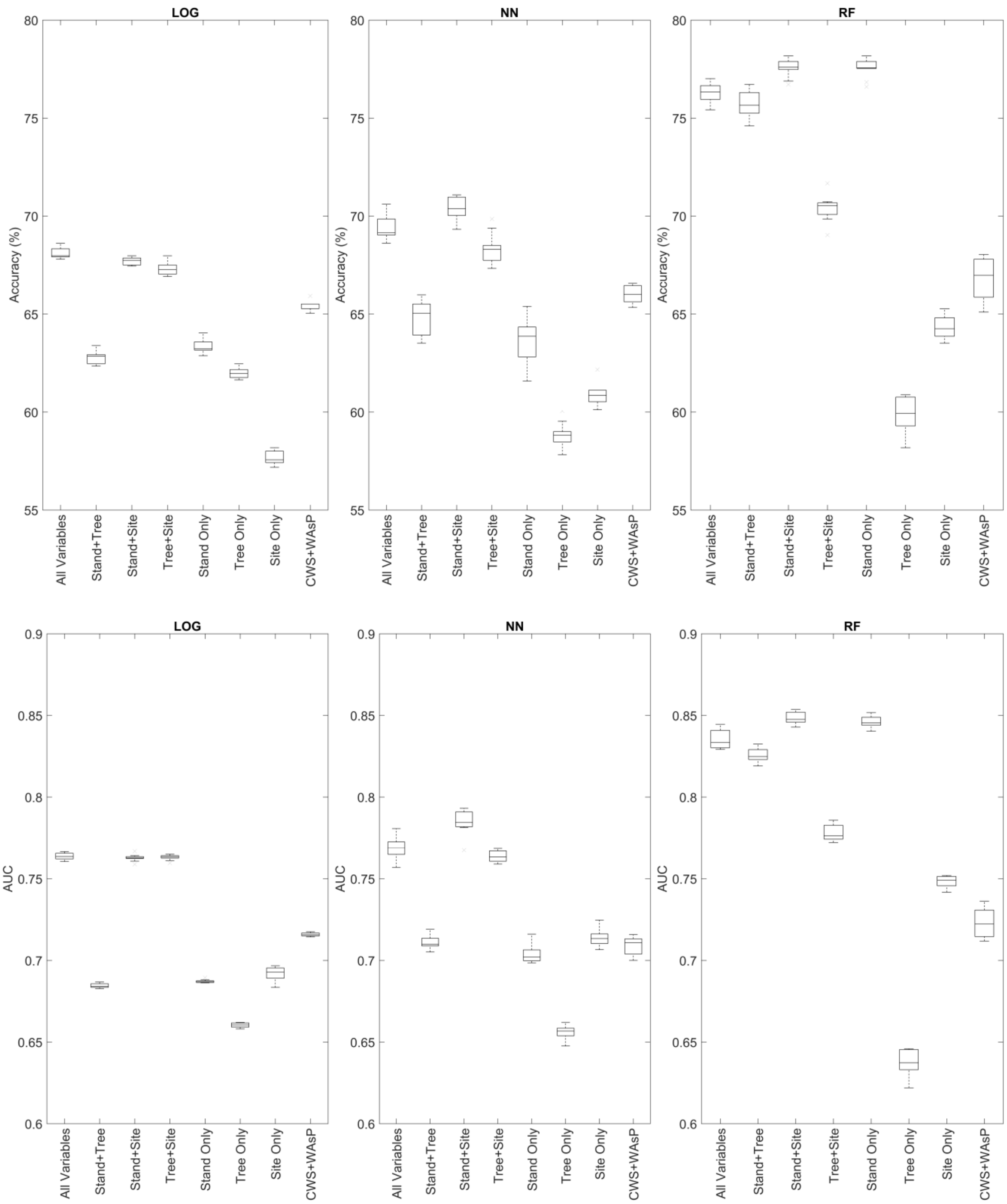
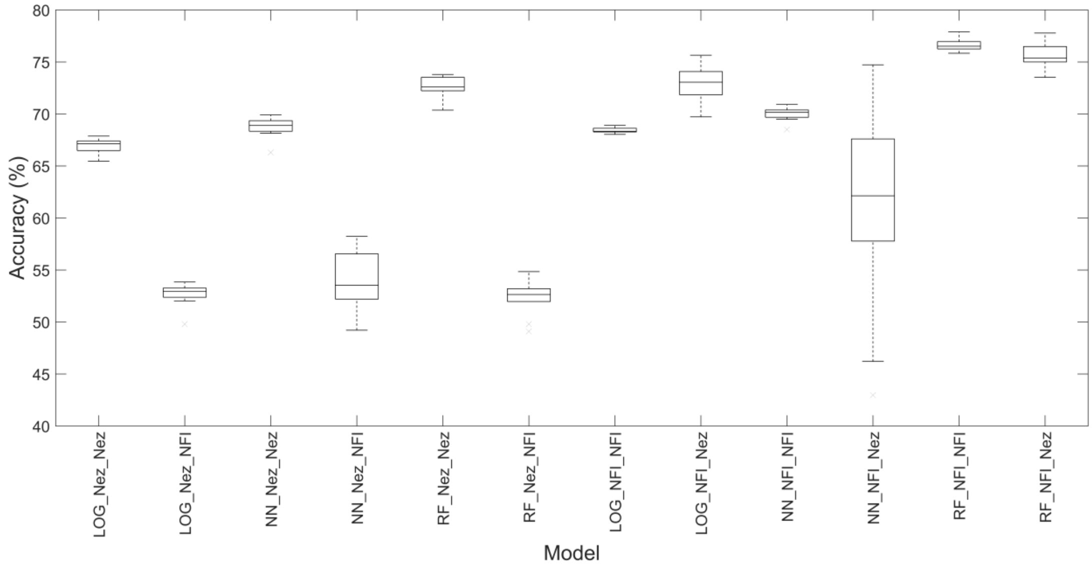


Fig. 3: Test of impact of leaving out different parameter groups in the damage models on the overall model accuracy and discriminatory ability (AUC) for the NFI data. *Stand* = {*Gap Size*, *Stand Mean DBH*, *Stand Mean Height*, *Stand Density*, *Stand Mean CI_BAL*}; *Tree* = {*Tree DBH*, *Tree Height*, *Tree CI_BAL*}, *Site* = {*WAsP 40m*, *Dune*, *Hydro*, *Soil*}, *CWS+WAsP* = {*CWS Breakage*, *CWS Overturn*, *WAsP 40m*}. Note change of scales on the y-axes compared to Fig. 2.

3.2.5. Portability of models

Model portability was tested by using the models developed from the Nezer Forest damage/no damage data and applying them to the NFI damage/no damage data in the same manner as Kamimura et al. (2016). But in addition we also tested the applicability of the NFI derived models on the smaller Nezer Forest data. In the same manner as discussed previously (Sections 3.2.1and 3.2.2) the test data was divided into 10 groups to allow 10 evaluations of model performance. Only calculations using the CSW calculated at d+10 m and 29 m were used because calculations at 40 m were not available in the Nezer Forest. The results are presented for the calculations at 29m in Fig. 4 and summarized for both heights in Table A4 in Appendix A. It is clear from the results that there is a severe reduction in model accuracy and discriminatory ability if the models developed on the Nezer Forest data (small forest area) are applied to the whole maritime pine forest estate in the Landes de Gascogne Forest (NFI data). In fact the models all fail to provide accurate predictions (all values between 50 and 55%) and have no discriminatory ability (AUC values close to 0.5). In the Nezer Forest there was a limited range of tree sizes, and there was no variation in soil or hydrological properties and the whole area was classified as a Landes ecological region. This meant there was no input data covering the larger range of conditions that exist in the NFI data. However, the models developed with the much larger data set from across the whole Landes de Gascogne Forest (NFI data) performed almost as well on the Nezer data set as when tested on the data from which it was originally developed. In the case of the LOG model the performance appeared to be actually enhanced in terms of accuracy (see Fig. 4 and compare LOG_NFI_NFI and LOG_NFI_Nez) although the difference was just not significant at the p=0.05 level (p=0.0592). The NN model had reduced accuracy and discriminatory ability (both significant at the p=0.05 level) and the accuracy was very variable between the 10 tests. The RF model had no loss of accuracy but a reduction in discriminatory ability (significant at p=0.05 level).

The results illustrate that the models developed from damage data in January 2009 (Storm Klaus) were able to successfully predict damage from a previous storm in December 1999 (Storm Martin) when the state of the soil and meteorological conditions were different. This suggests that such models, and especially the RF model, have the potential for predicting damage risk to individual trees for future storms if developed on a comprehensive enough data set. Unfortunately we have no other damage data sets with maritime pine on which to further test the models.



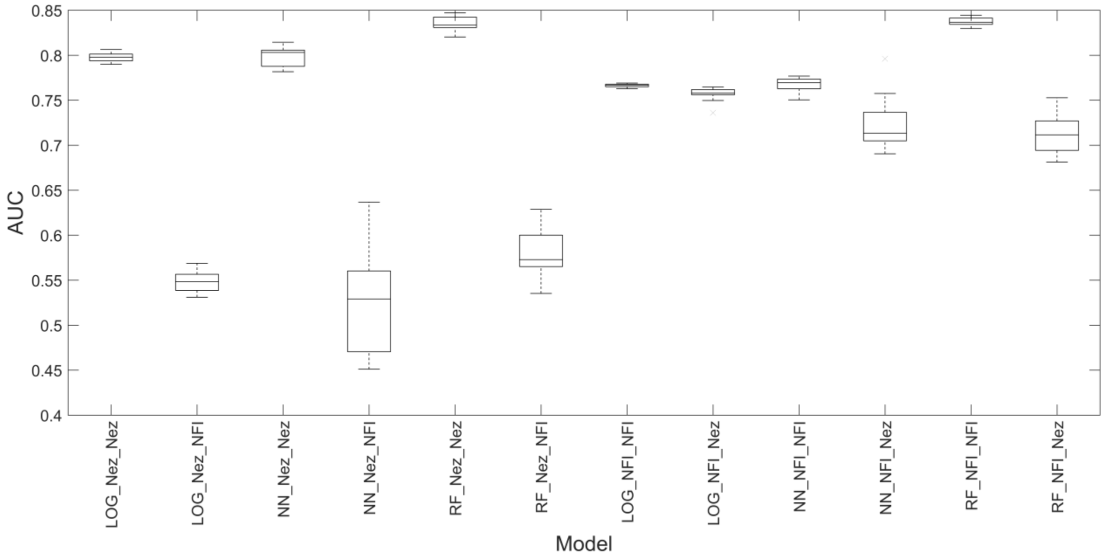


Fig. 4: Comparison of accuracy and AUC for predictions using the Nezer derived models on Nezer data (LOG_Nez_Nez, NN_Nez_Nez, RF_Nez_Nez), using the Nezer derived models on NFI data (LOG_Nez_NFI, NN_Nez_NFI, RF_Nez_NFI), NFI derived models on NFI data (LOG_NFI_NFI, NN_NFI_NFI, RF_NFI_NFI), and NFI derived models on Nezer data (LOG_NFI_Nez, NN_NFI_Nez, RF_NFI_Nez). All calculations used the CWSs calculated from GALES at 29m height.

4. Discussion

This paper follows on from the earlier work of Kamimura et al. (2016), which developed and tested the ability of logistic regression model and the hybrid mechanistic model GALES to calculate individual maritime pine trees at risk of wind damage in the Landes de Gascogne Forest of South-West France. That paper found good agreement of the predictions of the GALES model against observed damage for specific conditions of soil and soil hydrological status, specifically hydromorphic podzol, which was the only soil type on which tree pulling experiments in the region had been conducted and the values from which had been used to parameterise the model (Cucchi et al., 2005). However, when the soil and hydrological conditions changed the model had poor discrimination success between damaged and undamaged trees (typically AUC < 0.7). The logistic model was able to simulate well the damage in the Nezer Forest and the region represented by the NFI if the logistic model was calibrated for each forest area. However, the logistic model developed for the Nezer Forest had no discriminatory ability when applied to the NFI forest area with a much larger range of conditions. The logistic model was therefore not easily transferable even when the data from the NFI was filtered to only investigate soil and hydrological conditions similar to the ones in the Nezer Forest, where the model had been developed (Kamimura et al., 2016). This is a reflection of the fact that a model “trained” on a dataset with a limited range, and which tries to minimise errors with that dataset, fails to produce satisfactory results when used with a dataset with a wider range of characteristics (tree sizes, soil type, hydrological conditions, etc.)

In this paper we have attempted to determine whether other modelling approaches such as artificial neural networks and random forests are able to perform more accurately and with greater discrimination than a logistic regression model or the GALES model. In addition we wanted to determine if the models were more transferable from one area to another than was previously found in Kamimura et al. (2016). The same data sets were used in this paper and the parameterisation of the GALES model used in this paper to calculate critical wind speeds was identical to the previous work. In addition to developing artificial neural network and random forests models we again developed a logistic regression model for direct comparison with the previous work.

In addition, we wanted to determine if it was possible to substitute the hybrid-mechanistic model GALES by one of these modelling approaches if they were previously “trained” using outputs from the GALES model run over a large range of example stands. This could provide a very rapid method of calculating trees at risk over large areas such as the 790,000 ha of the Landes de Gascogne Forest or in computer simulations of different forest management scenarios such as have been conducted in Finland by Zeng et al. (2007). This would allow near rapid simulations of alternative management approaches for forest management planning and a very quick assessment of the impact of a plan on the current and future wind damage risk to the forest.

All the models in conjunction with regional predictions of wind speed during storms Martin and Klaus were successful at predicting individual tree damage within both the very well defined and measured Nezer Forest as well as across the whole of Landes de Gascogne Forest. However, overall there was little improvement in the accuracy or discriminatory ability of the artificial neural network model used in this study over the logistic regression model and results were similar to those obtained in the previous study both for the Nezer Forest and with the NFI data. This is in contrast to Hanewinkel et al. (2004) who found enhanced identification of damaged trees with the artificial neural network model compared to the logistic regression model. However, we did find that the random forests model produced enhanced accuracy and AUC values over all the other models for all circumstances (both forest test areas and for all heights of CWS calculation) and showed *good* discriminatory power (AUC between 0.827 and 0.837).

The random forests models were also found to be extremely insensitive to removing any individual variable but performance was adversely affected when all stand variables (*Gap Size*, *Stand Mean DBH*, *Stand Mean Height*, *Stand Density*, *Stand Mean CI_BAL*) were removed. In contrast both the logistic regression and artificial neural network models were more sensitive to the removal of individual variables and the logistic regression model particularly sensitive to the removal of the information on whether the stand was in the Dune or Landes area, the soil type and its hydrological status (*Dune*, *Soil* and *Hydro* variables). This was confirmed by the removal of groups of variables covering tree, stand and site conditions where the logistic regression and artificial neural network models were very sensitive to the removal of all site variables (*WAsP 40m*, *Dune*, *Hydro*, *Soil*), and performed best when site and stand information were available. These observations support the previous findings of Kamimura et al. (2016) where the logistic regression model lost discriminatory power if there was no information on whether the plot was in the Dune or Landes area, what the soil type was, and the hydrological status of the soil.

Interestingly the removal of either individual tree variables or all tree variables (*Tree DBH*, *Tree Height*, *Tree CI_BAL*) did not have a negative influence on any model performance and in fact there was a slight but significant improvement for the random forests model. This may be a reflection of the data distribution for tree variables that make it harder for the random forests method to find good unique values on which to split the data and build a good model. However, the fact that all models were not affected by the lack of tree data might suggest that for severe storms in forests similar to the Landes de Gascogne Forest the damage is controlled by stand and site characteristics and individual tree characteristics do not control the effective vulnerability to the wind. This would fit with the accepted view of the nature of damage within these forests, which is that it is triggered at vulnerable edges resulting from a recent clear-felling and then propagates through the stand damaging almost all trees regardless of their individual characteristics (Dupont et al., 2015; Kamimura et al., 2016).

All models were successful in replicating the outputs of the GALES model using the training data set with r^2 values, in almost all cases, greater than 0.9 between predicted critical wind speeds and the GALES derived critical wind speeds. This extremely strong correlation meant that substitution of model derived critical wind speeds for the GALES values in the damage model predictions of damage/no damage had almost no impact. However, the use of the critical wind speeds calculated by GALES or the CWS models as inputs for the damage models leads to concerns about error propagation. Therefore, because the performance of all the damage models was unaffected by the removal of critical wind speeds as inputs, it might be advisable to use damage models developed using only measured data. In addition, all the CWS models had a large standard deviation in their predictions indicating that the model derived critical wind speeds would only be appropriate for large areas and multiple simulations, such as investigating management options over a whole forest, rather than in calculations for individual trees or stands. Another use would be to provide a starting (seed) wind speed in the iterative calculations used in the GALES model itself (Hale et al., 2015).

The models developed with the large extensive data set across the whole of Landes de Gascogne Forest (NFI data) following damage caused by Storm Klaus in 2009 were successful in predicting the damaged trees in the smaller Nezer Forest for a completely different storm (Storm Martin in 1999). However, the models developed with the Nezer data showed no predictive ability for the storm damage in the larger NFI data set. This agrees with the findings of Kamimura et al. (2016), as discussed earlier, who were unable to successfully apply their logistic model developed with the Nezer data to predict damage in the whole Landes de Gascogne Forest and it is no surprise that models developed within a limited data set do not work in larger more complex areas.

Altogether the results suggest that the random forests modelling approach can very successfully predict the trees that will be damaged during a storm with an accuracy of up to 76% so long as good quality data are available to "train" the model. This data can be from any storm so long as there is a sufficient range of input conditions because the models were found to be transferable to other storms under such conditions. The random forests model could also be used in large-scale scenario testing to investigate different management options into the future. Such an approach would provide a powerful planning and public engagement tool because the models are fast and the impact of decisions could be visualised almost immediately.

5. Conclusions

The results from this investigation of new approaches to modelling forest wind damage suggest that artificial neural networks are no better than logistic regression models in their accuracy or discriminatory ability in determining which trees are likely to be damaged. However, no model tuning was employed with either approach so performance might be improved with adjustment of parameters such as the damage cut point. Even so, the models based on the random forests approach were found to be much more accurate and had higher discriminatory power than the logistic regression and neural network models in all circumstances and to give high accuracy (>75%) and good discrimination (AUC>0.8). In addition they were almost completely insensitive to the removal of any specific input variable and dependent on only stand level information to achieve good results. This would mean that they could be used successfully even if specific data were missing. Tree level information was found to be unimportant in all models suggesting that the dominant damage mechanism in these forests is propagation of damage from vulnerable forest edges, which affects all trees regardless of their size.

The random forests model along with the other approaches was also successfully able to predict the critical wind speeds (CWSs) predicted by the GALES model if trained on an extensive enough artificial data set. The models are much faster than GALES due to a lack of a requirement for iteration and so could be used for running large scale "what if" scenarios as part of scenario modelling and testing or planning exercises involving stakeholders.

The models that were developed all require extensive data sets of actual damage (large range of input variable values) for their development and could be transferred to other regions if the forest conditions in the new area are comprehensively covered within the model training data set. However, if the conditions are different and no detailed damage data from storms in the new area are available the models are unlikely to be transferable. In contrast, all the models can be trained to replace GALES if a large artificial data set covering the range of stand characteristics to be found in the new region is first used to “train” them and this could be extremely useful for large scale forest planning in any region that has its specific conditions and species incorporated in the GALES model.

Appendix A.

Supplementary data can be found in Appendix A.

Acknowledgements

We are grateful to Thierry Bélouard (IGN) who kindly provided the national forest inventory data in the Landes de Gascogne Forest (Projet CRA 2012 IGN-IDF-INRA), as well as to Gaston Courrier and Didier Garrigou at INRA-Bordeaux who conducted the field surveys in the Nezer Forest. This work was funded by an INRA scientific package awarded to Barry Gardiner and by grant ANR-12-AGRO-0007-04 (ANR, Agrobiosphere, France, project “FOR-WIND”) and developed out of work undertaken in grant EP/J021628/1 to Emma Hart (EPSRC, UK)

References

Albrecht, A., Hanewinkel, M., Bauhus, J., Kohnle, U., 2010. How does silviculture affect storm damage in forests of south-western Germany? Results from empirical modeling based on long-term observations. Eur. J. For. Res. 131, 229–247. doi:10.1007/s10342-010-0432-x

Albrecht, A., Kohnle, U., Hanewinkel, M., Bauhus, J., 2012. Storm damage of Douglas-fir unexpectedly high compared to Norway spruce. Ann. For. Sci. 70, 195–207. doi:10.1007/s13595-012-0244-x

Alpaydin, E., 2014. Introduction to Machine Learning, 3rd ed. MIT Press, Cambridge.

Biging, G.S., Dobbertin, M., 1995. Evaluation of competition indices in individual tree growth models. For. Sci. 41, 360–377.

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

Bruno, E., Bartoli, M., 2001. Premiers enseignements de l’utilisation de logiciel ecoflore pour traiter les relevés botaniques du l’IFN. Rev. For. Fr. 53, 391–396. doi:10.4267/2042/5254

Chehata, N., Orny, C., Boukir, S., Guyon, D., Wigneron, J.P., 2014. Object-based change detection in wind storm-damaged forest using high-resolution multispectral images. Int. J. Remote Sens. 35, 4758–4777. doi:10.1080/01431161.2014.930199

Chen, S.H., Jakeman, A.J., Norton, J.P., 2008. Artificial Intelligence techniques: An introduction to their use for modelling environmental systems. Math. Comput. Simul. 78, 379–400. doi:10.1016/j.matcom.2008.01.028

Colin, F., I., V., Rou-Nivert, P., Renaud, J.-P., Hervé, J.-C., Bock, J., Piton, B., 2009. Facteurs de risques de chablis dans les peuplements forestiers : les leçons tirées des tempêtes de 1999, in: Birot, Y., Landmann, G., Bonhême, I. (Eds.), La Forêt Face Aux Tempêtes. Editions Quae, pp. 177–228.

Csilléry, K., Kunstler, G., Courbaud, B., Allard, D., Lassègues, P., Haslinger, K., Gardiner, B., 2017. Coupled effects of wind-storms and drought on tree mortality across 115 forest stands from the Western Alps and the Jura mountains. Glob. Chang. Biol. doi:10.1111/gcb.13773

Cucchi, V., Meredieu, C., Stokes, A., Berthier, S., Bert, D., Najar, M., Denis, A., Lastennet, R., Lamberts, L., 2004. Root anchorage of inner and edge trees in stands of Maritime pine (Pinus pinasterAit .) growing in different podzolic soil conditions 460–466. doi:10.1007/s00468-004-0330-2

Cucchi, V., Merediu, C., Stokes, A., De Coligny, F., Suarez, J., Gardiner, B. a. B.A., Meredieu, C., Stokes, A., De Coligny, F., Suarez, J., Gardiner, B. a. B.A., 2005. Modelling the windthrow risk for simulated forest stands of Maritime pine (Pinus pinaster Ait.). For. Ecol. Manage. 213, 184–196. doi:10.1016/j.foreco.2005.03.019

Dobbertin, M., 2002. Influence of stand structure and site factors on wind damage comparing the storms Vivian and Lothar. For. Snow Landsc. Res. 77, 187–205.

Dupont, S., Pivato, D., Brunet, Y., 2015. Wind damage propagation in forests. Agric. For. Meteorol. 214–215, 243–251. doi:10.1016/j.agrformet.2015.07.010

Frank, E., Hall, M.A., Witten, I.H., 2016. The WEKA Workbench. Online Appendix for “Data Mining: Practical Machine Learning Tools and Techniques,” 4th ed. Morgan Kaufmann.

Gardiner, B., Byrne, K., Hale, S., Kamimura, K., Mitchell, S.J., Peltola, H., Ruel, J.C., 2008. A review of mechanistic modelling of wind damage risk to forests. Forestry.

Gardiner, B., Peltola, H., Kellomaki, S., 2000. Comparison of two models for predicting the critical wind speeds required to damage coniferous trees. Ecol. Modell. 129, 1–23.

Gardiner, B., Welten, P., 2013. Mitigation of forest damage, in: Gardiner, B., Schuck, A., Schelhaas, M.-J., Orazio, C., Blennow, K., Nicoll, B. (Eds.), Living with Storm Damage to Forests: What Science Can Tell Us. European Forest

Institute, Joensuu, pp. 81–88.

Gardiner, B.A., 1992. Mathematical modelling of the static and dynamic characteristics of plantation trees, in: Franke, J., Roeder, A. (Eds.), Mathematical Modelling of Forest Ecosystems. Sauerländers Verlag, Frankfurt am Main, p. 40–61.

GISsol, 2011. L'état des sols de France. Nancy.

Guan, B.T., Gertner, G., 1995. Modeling individual tree survival probability with a random optimization procedure: an artificial neural network approach. AI-Applications 9, 39–52.

Guan, B.T., Gertner, G., Parysow, P., 1997. A framework for uncertainty assessment of mechanistic forest growth models: a neural network example. Ecol. Modell. 98, 47–58.

Haarsma, R.R.J., Hazeleger, W., Severijns, C., de Vries, H., Sterl, A., Bintanja, R., van Oldenborgh, G.J., van den Brink, H.W., 2013. More hurricanes to hit western Europe due to global warming. Geophys. Res. Lett. 40, 1783–1788. doi:10.1002/grl.50360

Hale, S., Gardiner, B., Peace, A., Nicoll, B., Taylor, P., Pizzirani, S., 2015. Comparison and validation of three versions of a forest wind risk model. Environ. Model. Softw. 68, 27–41. doi:10.1016/j.envsoft.2015.01.016

Hanewinkel, M., 2005. Neural networks for assessing the risk of windthrow on the forest division level : a case study in southwest Germany. Eur. J. For. Res. 124, 243–249. doi:10.1007/s10342-005-0064-8

Hanewinkel, M., Peltola, H., Soares, P., 2010. Recent approaches to model the risk of storm and fire to European forests and their integration into simulation and decision support tools. For. Syst. 19, 30–47.

Hanewinkel, M., Zhou, W., Schill, C., 2004. A neural network approach to identify forest stands susceptible to wind damage. For. Ecol. Manage. 196, 227–243. doi:10.1016/j.foreco.2004.02.056

Hasenauer, H., Merkl, D., Weingartner, M., 2001. Estimating tree mortality of Norway spruce stands with neural networks. Adv. Environ. Res. 5, 405–414.

Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression, 2nd ed. John Wiley & Sons, Inc., New York.

Kamimura, K., Gardiner, B., Dupont, S., Guyon, D., Meredieu, C., 2016. Mechanistic and statistical approaches to predicting wind damage to individual maritime pine (*Pinus pinaster*) trees in forests. Can. J. For. Res. 100, 88–100.

Kourtz, P., 1990. Artificial intelligence: a new tool for forest management. Can. Jounal For. Res. 20, 428–437. doi:https://doi.org/10.1139/x90-060

Kunkel, K.E., Karl, T.R., Brooks, H., Kossin, J., Lawrimore, J.H., Arndt, D., Bosart, L., Changnon, D., Cutter, S.L., Doesken, N., Emanuel, K., Groisman, P.Y., Katz, R.W., Knutson, T., O'brien, J., Paciorek, C.J., Peterson, T.C., Redmond, K., Robinson, D., Trapp, J., Vose, R., Weaver, S., Wehner, M., Wolter, K., Wuebbles, D., 2013. Monitoring and understanding trends in extreme storms: State of knowledge. Bull. Am. Meteorol. Soc. 94, 499–514. doi:10.1175/BAMS-D-11-00262.1

Lagerquist, R., Flannigan, M.D., Wang, X., Marshall, G.A., 2017. Automated prediction of extreme fire weather from synoptic. Can. J. For. Res. 1183, 1175–1183. doi:10.1139/cjfr-2017-0063

Le Cessie, S., Van Houwelingen, J.C., 1992. Ridge Estimators in Logistic Regression. J. R. Stat. Soc. Ser. C (Applied Stat. 41, 191–201. doi:10.2307/2347628

Lindner, M., Maroschek, M., Netherer, S., Kremer, A., Barbati, A., Garcia-Gonzalo, J., Seidl, R., Delzon, S., Corona, P., Kolström, M., Lexer, M.J., Marchetti, M., 2010. Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. For. Ecol. Manage. 259, 698–709. doi:10.1016/j.foreco.2009.09.023

Mortensen, N.G., Landberg, L., Troen, I., Petersen, E.L., 1993. Wind Atlas Analysis and Application Program (WAsP)., 1st ed. Risø National Laboratory, Roskilde, Denmark.

Neild, S.A., Wood, C.J., 1999. Estimating stem and root-anchorage flexibility in trees. Tree Physiol. 19, 141–151.

Nicoll, B.C., Gardiner, B.A., Rayner, B., Peace, A.J., 2006. Anchorage of coniferous trees in relation to species, soil type, and rooting depth. Can. J. For. Res. 36, 1871–1883. doi:10.1139/x06-072

Patterson, D.W., 1996. Artificial neural networks: theory and applications. Prentice Hall, Englewood Cliffs.

Robert, N., Vidal, C., Colin, A., Hervé, J.C., Hamza, N., Cluzeau, C., 2009. 12.1 Development of France's National Forest Inventory, in: National Forest Inventories. p. 207.

Schelhaas, M.J., Nabuurs, G.J., Schuck, A., 2003. Natural disturbances in the European forests in the 19th and 20th centuries. Glob. Chang. Biol. 9, 1620–1633. doi:10.1046/j.1365-2486.2003.00684.x

Seidl, R., Schelhaas, M.-J., Lexer, M.J., 2011. Unraveling the drivers of intensifying forest disturbance regimes in Europe. Glob. Chang. Biol. 17, 2842–2852. doi:10.1111/j.1365-2486.2011.02452.x

Valinger, E., Fridman, J., 2011. Factors affecting the probability of windthrow at stand level as a result of Gudrun winter storm in southern Sweden. For. Ecol. Manage. 262, 398–403. doi:10.1016/j.foreco.2011.04.004

Zeng, H., Pukkala, T., Peltola, H., 2007. The use of heuristic optimization in risk management of wind damage in forest planning. For. Ecol. Manage. 241, 189–199. doi:10.1016/j.foreco.2007.01.016

Appendix A: Supplementary Material for “Use of Machine Learning Techniques to Model Wind Damage to Forests”

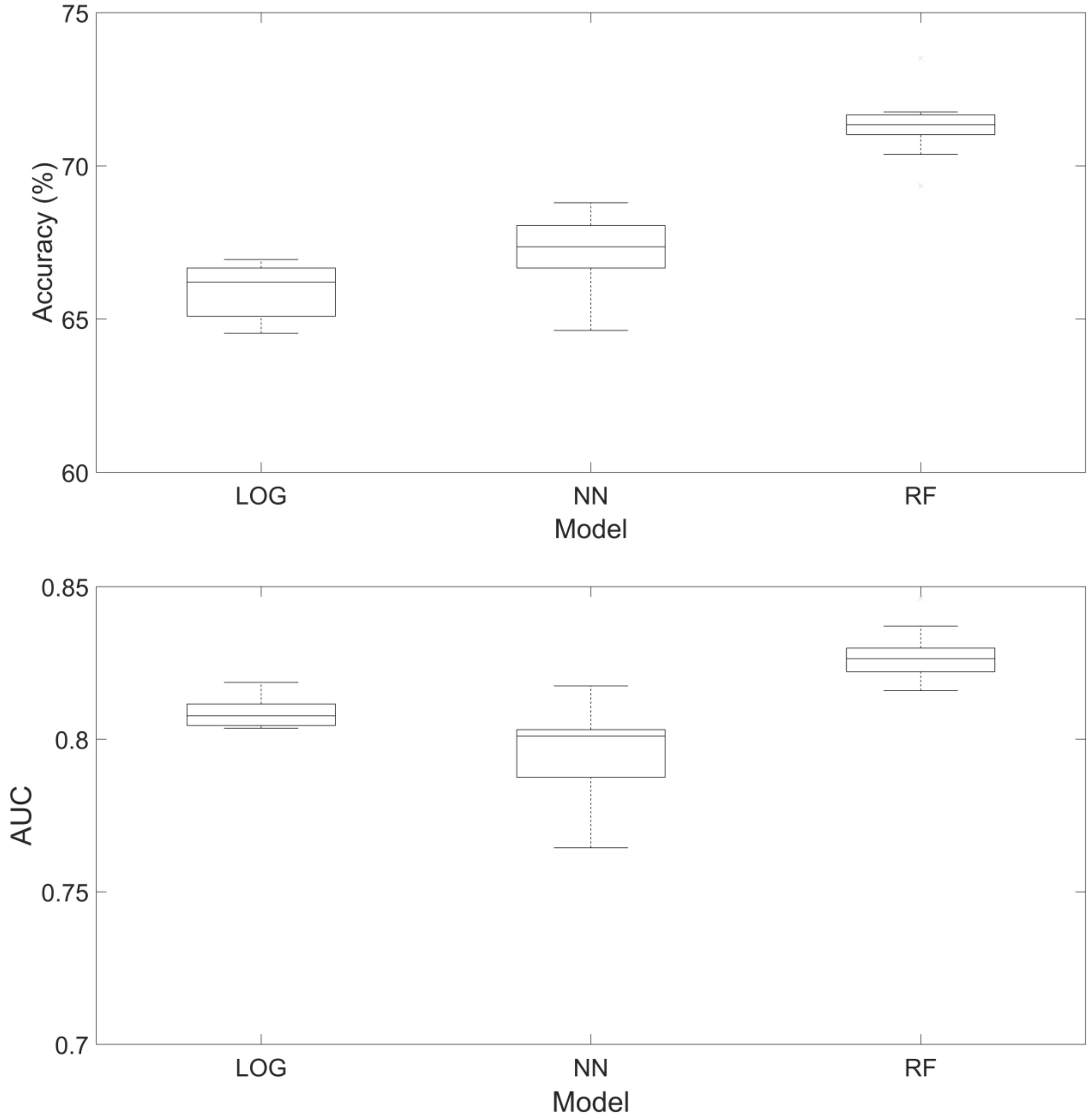


Fig. A 1: Accuracy and AUC for the LOG, NN and RF model predictions using the GALES predicted CWSs at *d*+10 m against the Nezer damage data. All variables in Table 3 were used except the WAsP derived wind speeds because these are only calculated at a single height above the ground and the *d*+10 m results are for variable heights above the ground depending on the calculated value of *d*.

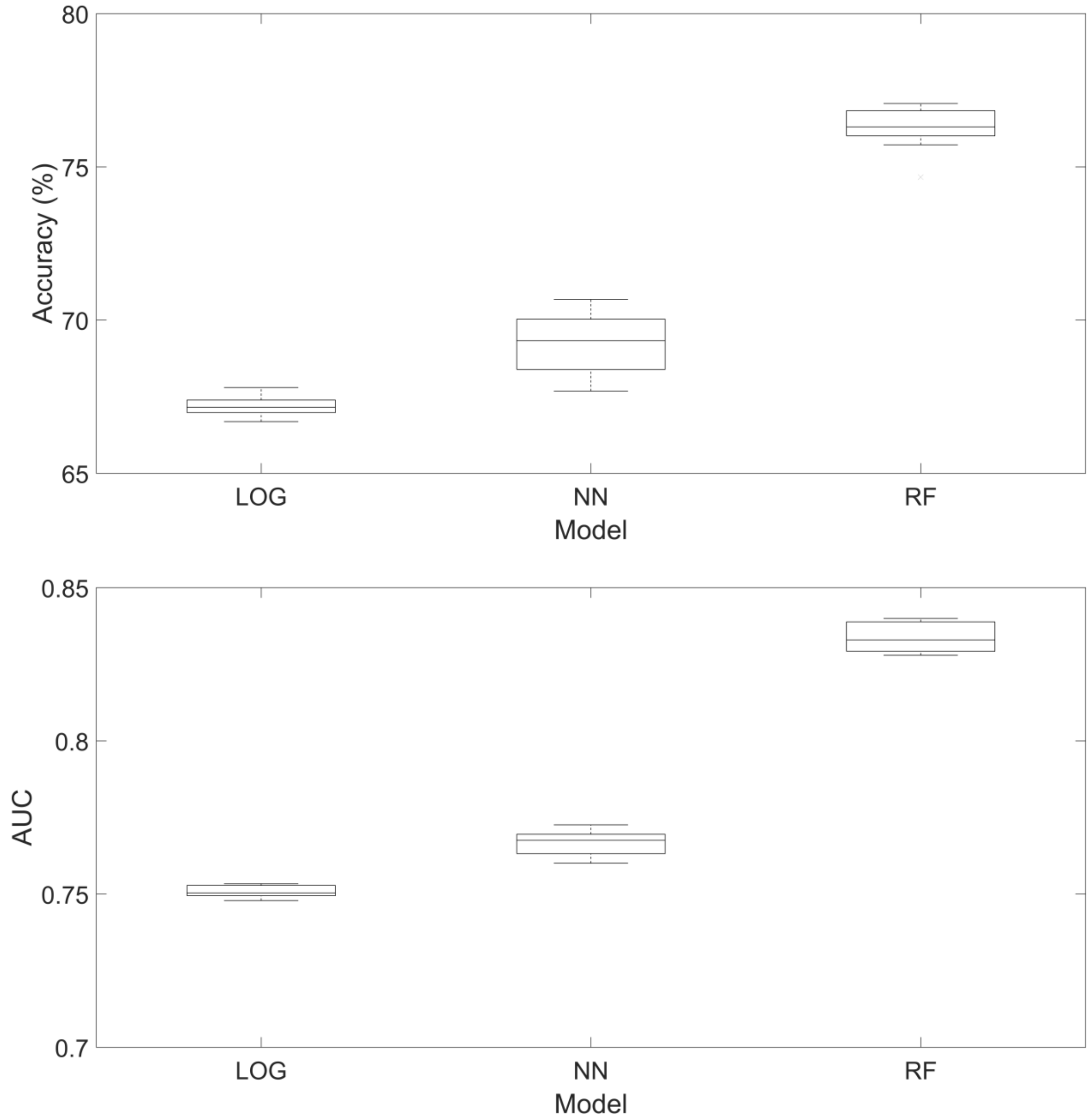


Fig. A 2: Accuracy and AUC for the LOG, NN and RF model predictions using the GALES predicted CWSs at *d*+10 m against the NFI damage data. All variables in Table 3 were used except the WAsP derived wind speeds because these are only calculated at a single height above the ground and the *d*+10 m results are for variable heights above the ground depending on the calculated value of *d*.

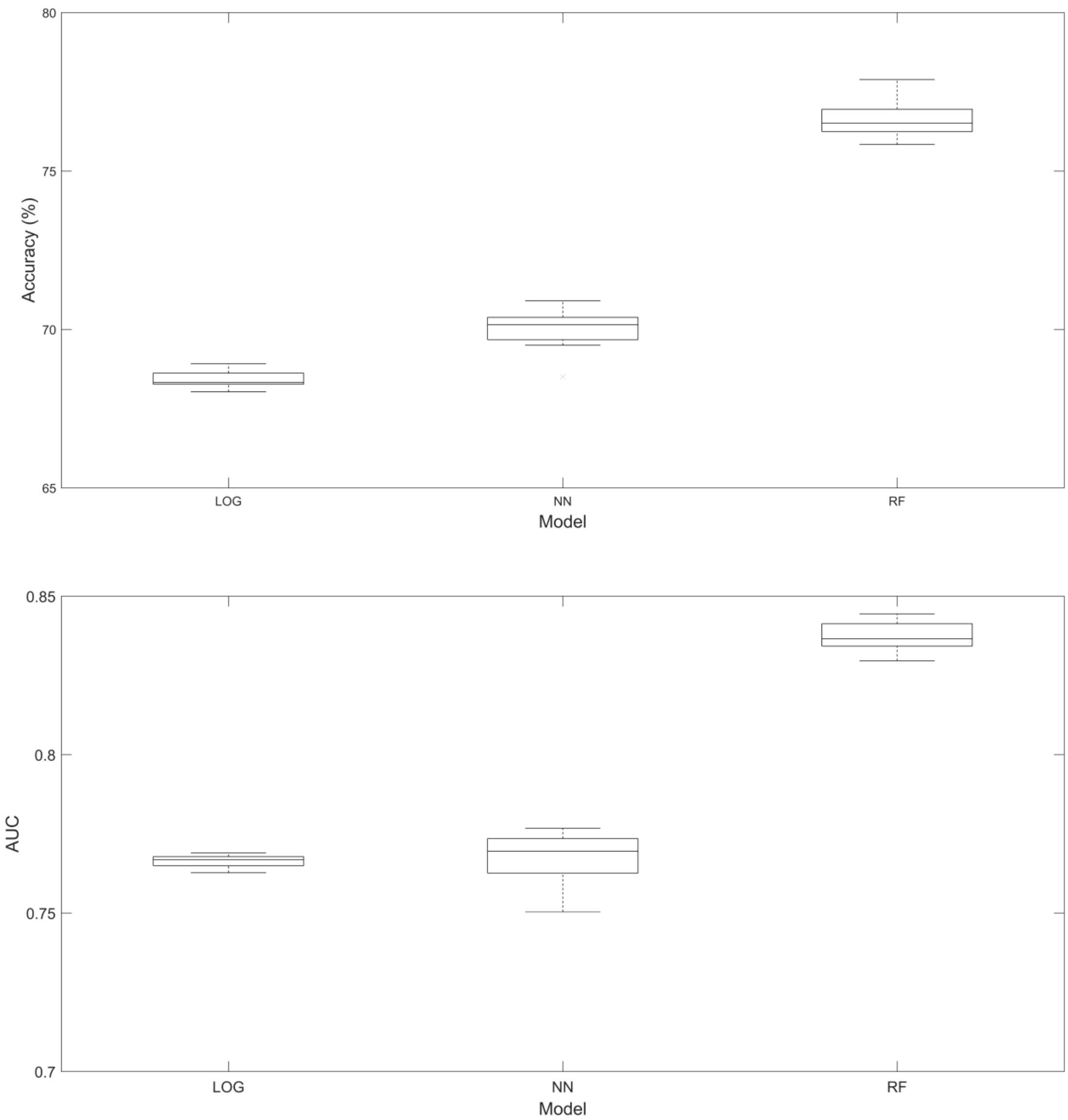


Fig. A 3: Accuracy and AUC for the LOG, NN and RF model predictions using the GALES predicted CWSs at 29 m against the NFI damage data

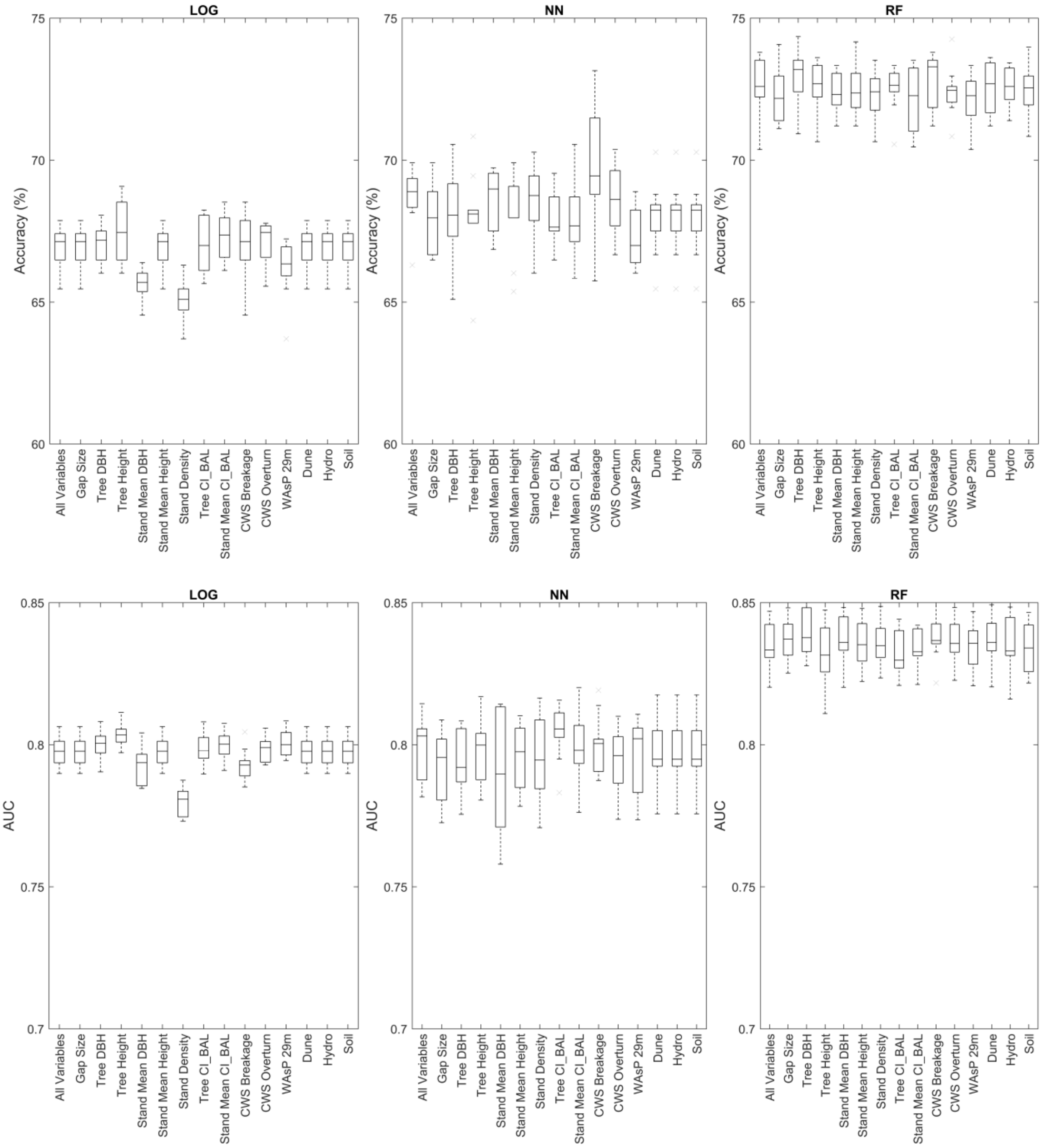


Fig. A 4: Test of impact of leaving out each parameter in the models on the overall model accuracy and ability to discriminate between damage and no damage (AUC) for the Nezer Forest using CWS and WAsP wind speed at 29 m.

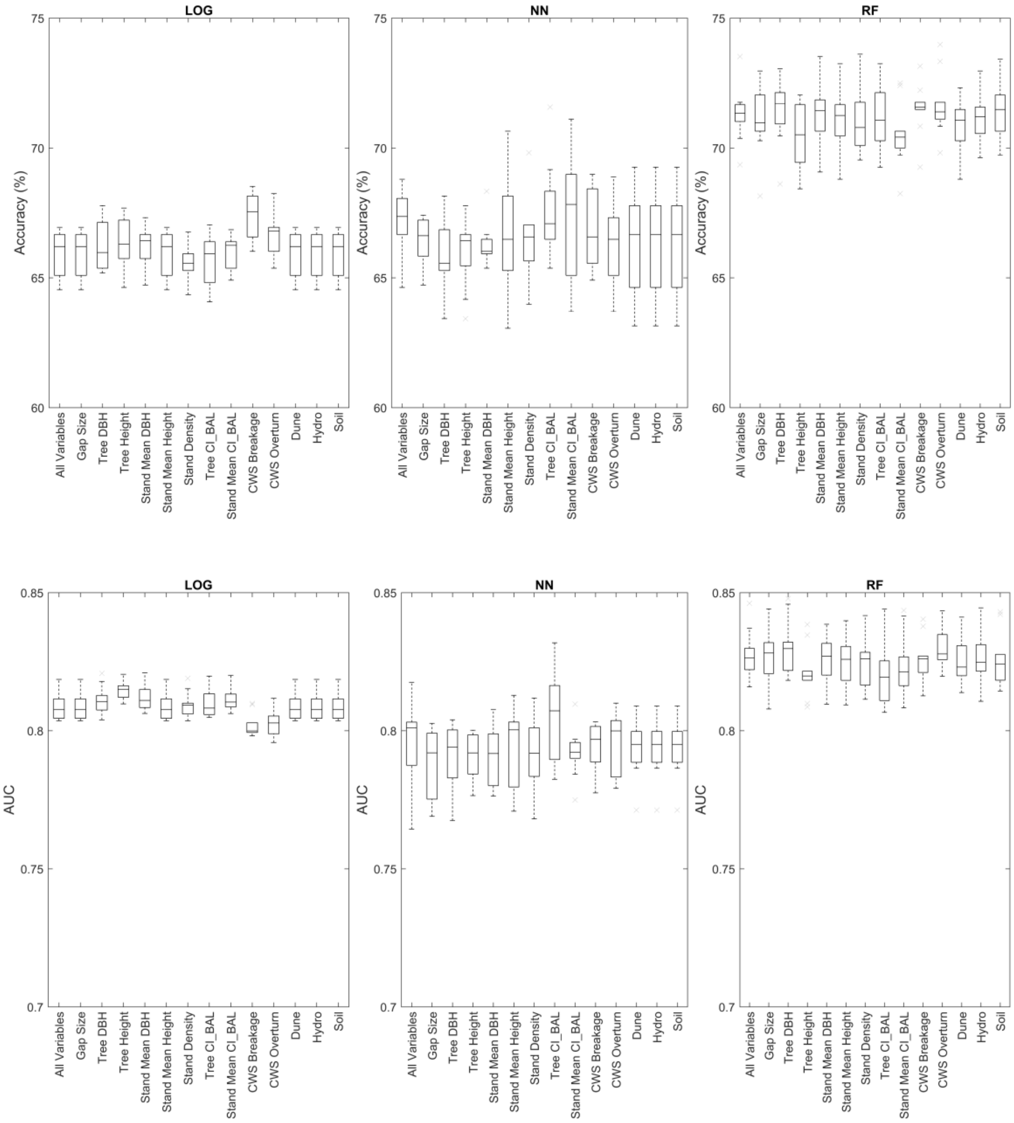


Fig. A 5: Test of impact of leaving out each parameter in the models on the overall model accuracy and ability to discriminate between damage and no damage (AUC) for the Nezer Forest using CWS at *d*+10 m.

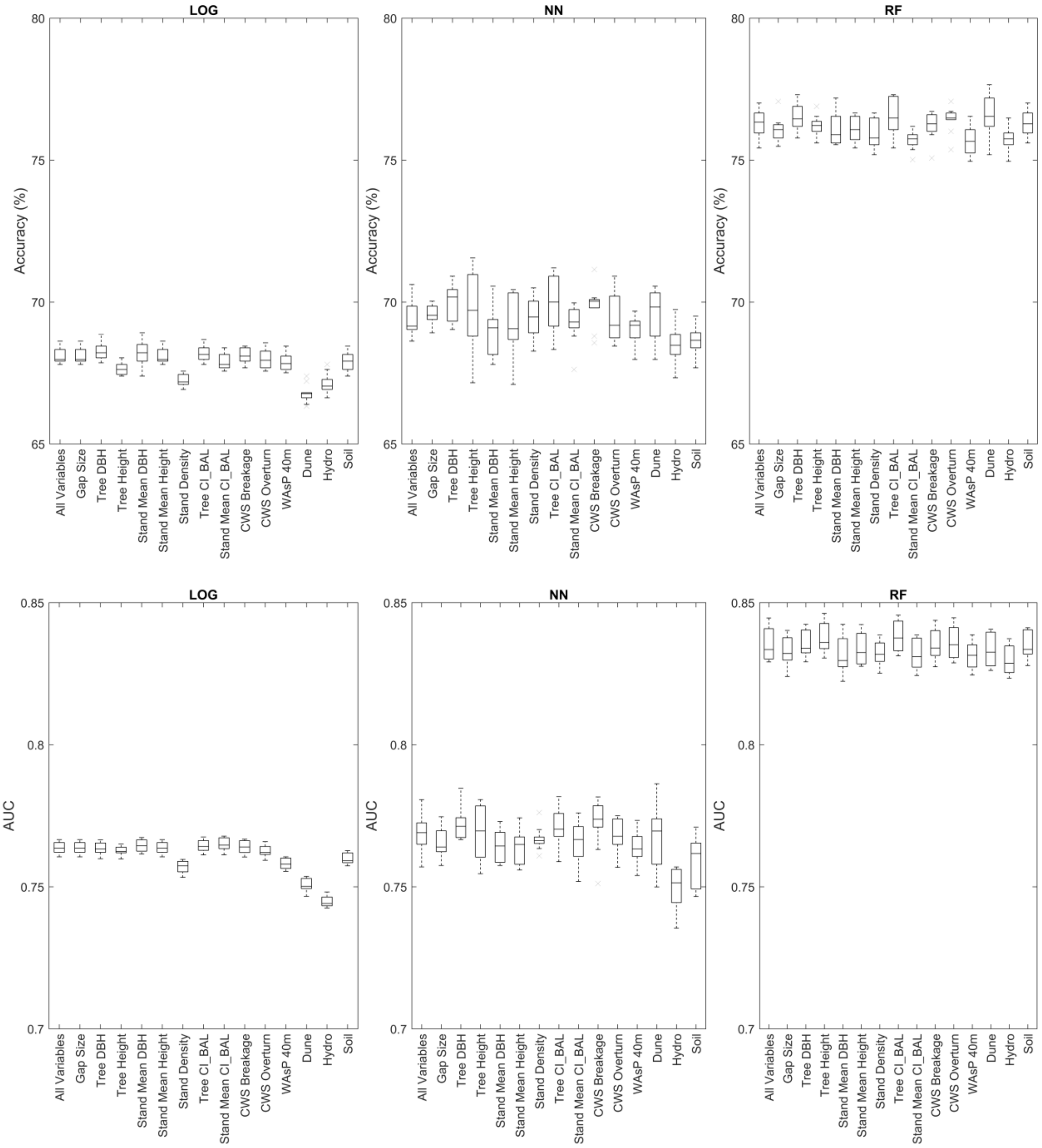


Fig. A 6: Test of impact of leaving out each parameter in the models on the overall model accuracy and ability to discriminate between damage and no damage (AUC) for the NFI data using CWS and WAsP wind speed at 40 m. Note change of scale from Fig. A 4 and Fig. A 5 for Accuracy.

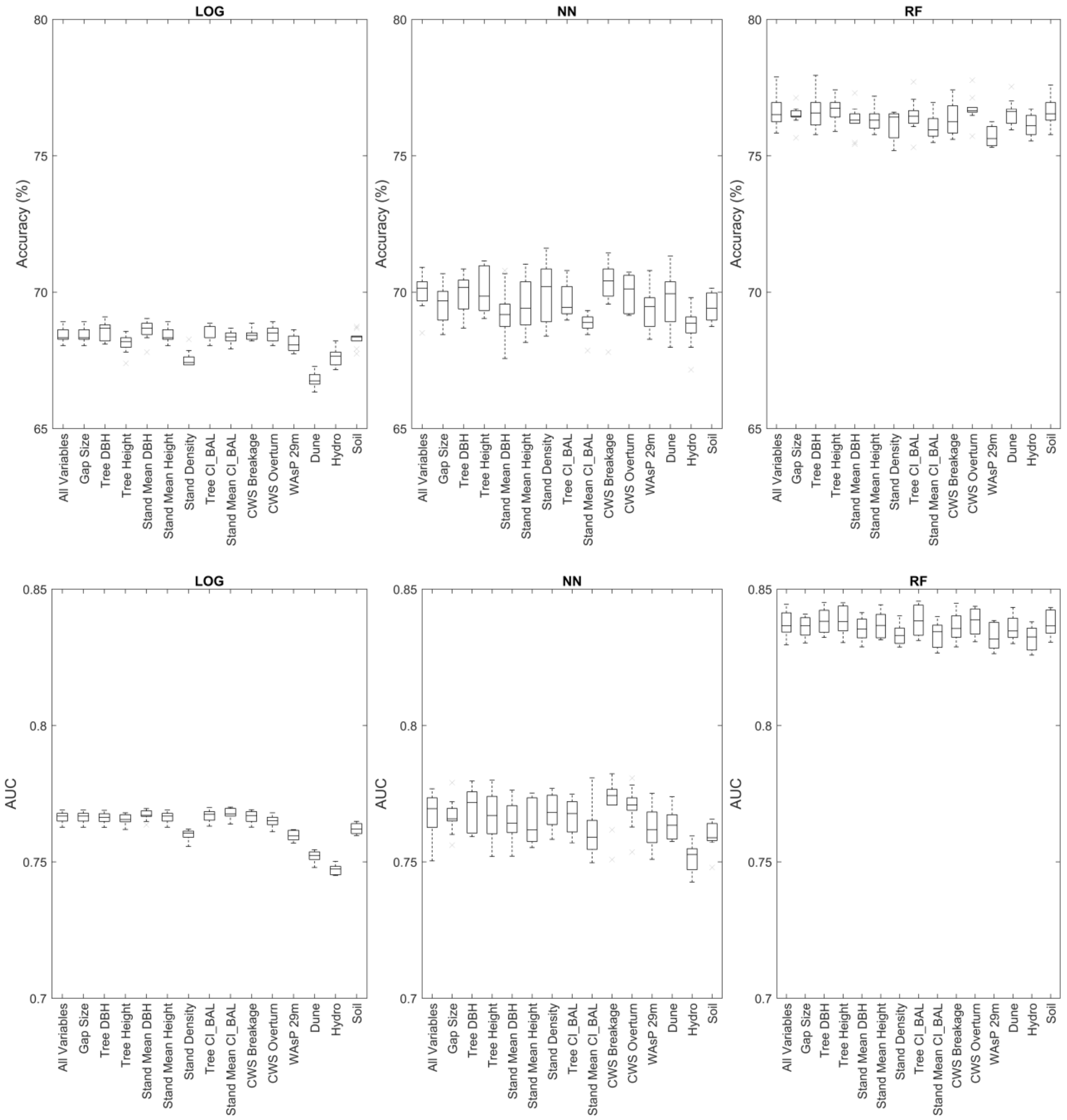
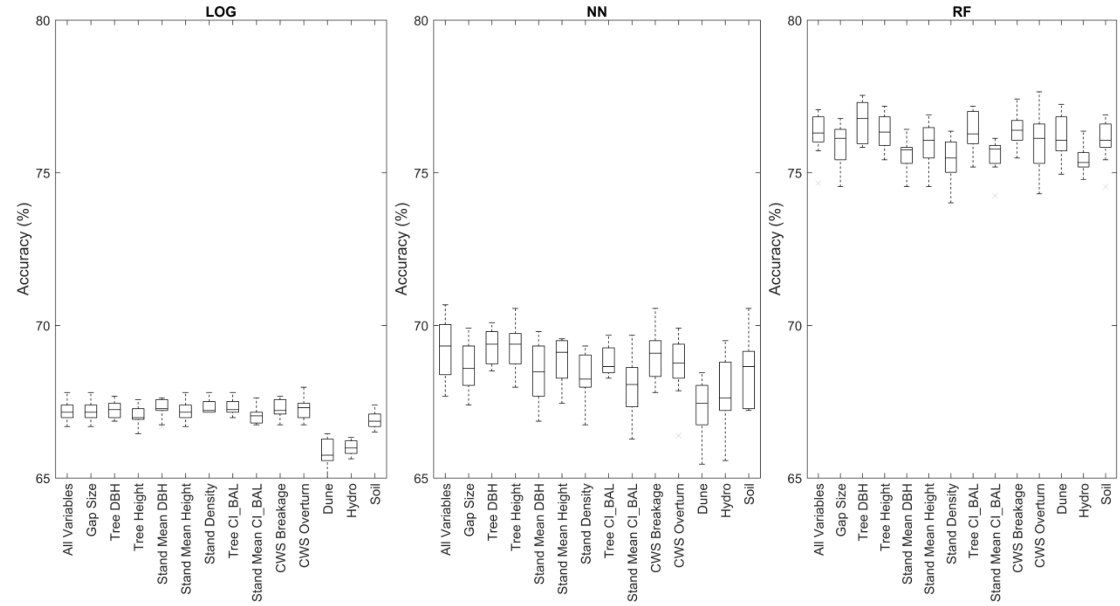


Fig. A.7: Test of impact of leaving out each parameter in the models on the overall model accuracy and ability to discriminate between damage and no damage (AUC) for the NFI data using CWS and WAsP wind speed at 29 m.



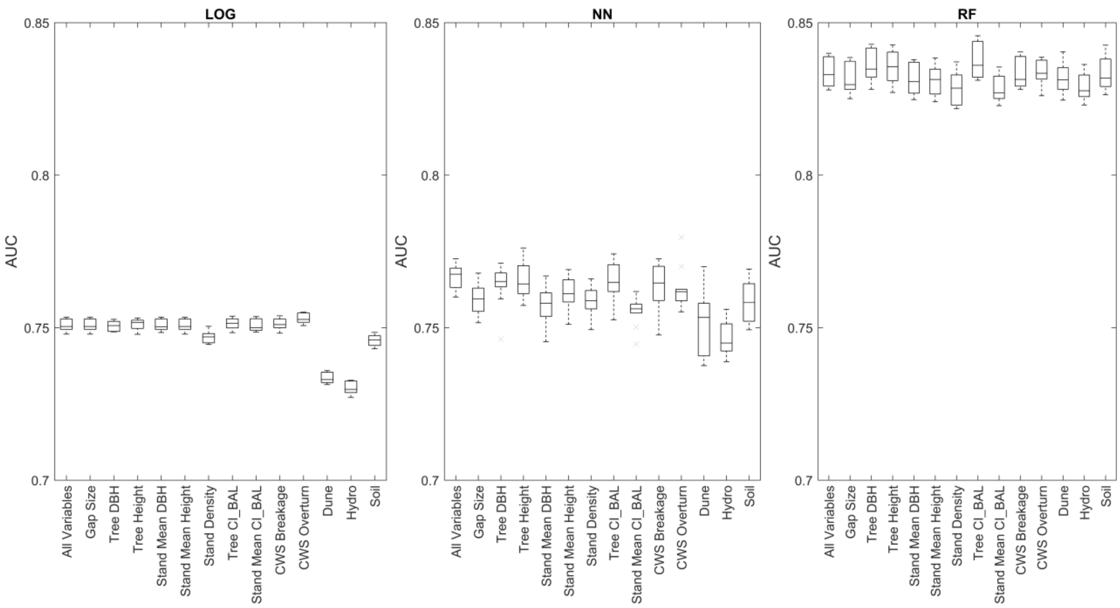


Fig. A.8: Test of impact of leaving out each parameter in the models on the overall model accuracy and ability to discriminate between damage and no damage (AUC) for the NFI data using CWS at *d*+10 m.

Table A 1: Results of comparison of predictions from trained LIN/NN/RF models and GALES for Nezer and NFI data at *d*+10 m. Numbers are correlation coefficient between trained model results and GALES predictions and root-mean square (RMS) error is given in brackets in ms^{-1}.

Training Set	*Test Set*	*Output*	*LIN*	*NN*	*RF*
GALES *d*+10m predictions from artificial data	Nezer	CWS for breakage	0.9668 (3.5072)	0.9738 (6.8321)	0.9616 (5.0649)
GALES *d*+10m predictions from artificial data	Nezer	CWS for overturning	0.9760 (3.2098)	0.9877 (2.8620)	0.9773 (3.1491)
GALES *d*+10m predictions from artificial data	NFI	CWS for breakage	0.9362 (3.1918)	0.9495 (4.2107)	0.9493 (3.0381)
GALES *d*+10m predictions from artificial data	NFI	CWS for overturning	0.9562 (2.0722)	0.9777 (1.7025)	0.9744 (1.7226)

Table A 2: Mean accuracy of different models with each model variable removed in turn. Standard deviation is given in brackets. * indicates value significantly different ($p<0.05$) from the value with *All Variables*. The superscript letters against the values in the All column (a, b, or c) indicate whether there are significant differences between the models for that particular height of CWS calculation at the $p=0.5$ level.

Data Set	*Model*	*CWS Height*	*All Variables*	*Average CI_BAL*	*CI_BAL*	*Tree DBH*	*Stand Density*	*Dune*	*Gap Size*	*Hydro*	*Stand DBH*	*Soil*	*Stand Height*	*Tree Height*	*CWS Break*	*CWS Overturn*	*WAsP Wind Speed*
Nezer	LOG	*d*+10 m	65.972[a] (0.839)	65.944 (0.667)	65.713 (0.984)	66.25 (0.915)	65.565 (0.681)	65.972 (0.839)	65.972 (0.839)	65.972 (0.839)	66.148 (0.838)	65.972 (0.839)	65.972 (0.839)	66.352 (0.97)	67.361* (0.863)	66.648 (0.913)	
	NN		67.176[a] (1.346)	67.259 (2.433)	67.509 (1.854)	65.824 (1.339)	66.528 (1.504)	66.509 92.003)	66.407 (0.98)	66.509 (2.003)	66.269 (0.822)	66.509 (2.003)	66.741 (2.223)	66.019 (1.346)	66.778 (1.464)	66.417 (1.708)	
	RF		71.306[b] (1.066)	70.519 (1.238)	71.167 (1.225)	71.426 (1.243)	71.000 (1.178)	70.917 (0.991)	71.046 (1.313)	71.093 (0.939)	71.231 (1.201)	71.407 (1.075)	71.167 (1.273)	70.454 (1.195)	71.509 (0.99)	71.657 (1.197)	
NFI	LOG	*d*+10 m	67.202[a] (0.309)	67.056 (0.270)	67.308 (0.248)	67.226 (0.287)	67.349 (0.246)	65.801* (0.518)	67.202 (0.309)	65.982* (0.233)	67.261 (0.327)	66.897 (0.270)	67.202 (0.309)	67.050 (0.303)	67.267 (0.307)	67.284 (0.382)	
	NN		69.267[b] (0.996)	67.971 (1.067)	68.868 (0.496)	69.261 (0.562)	68.305 (0.779)	67.290* (0.995)	68.657 (0.864)	67.713* (1.172)	68.481 (0.985)	68.540 (1.064)	68.880 (0.716)	69.273 (0.814)	69.021 (0.834)	68.639 (0.999)	
	RF		76.240[c] (0.693)	75.572 (0.559)	76.364 (0.664)	76.663 (0.680)	75.367 (0.763)	76.117 (0.773)	75.900 (0.664)	75.384 (0.452)	75.613 (0.510)	76.059 (0.700)	75.930 (0.683)	76.293 (0.622)	76.375 (0.567)	76.006 (0.950)	
	LOG	29 m	68.405[a] (0.270)	68.364 (0.232)	68.428 (0.277)	68.569 (0.330)	67.560* (0.303)	66.798* (0.285)	68.405 (0.270)	67.613* (0.327)	68.604 (0.368)	68.311 (0.303)	68.405 (0.270)	68.117 (0.340)	68.434 (0.193)	68.469 (0.275)	68.129 (0.306)
	NN		69.988[b] (0.673)	68.815 (0.435)	69.672 (0.623)	69.947 (0.726)	70.041 (1.064)	69.760 (0.999)	69.601 (0.654)	68.698* (o.722)	69.273 (0.962)	69.455 (0.521)	69.537 (0.965)	70.065 (0.809)	70.205 (1.015)	69.994 (0.643)	69.372 (0.729)
	RF		76.604[c] (0.619)	76.065 (0.507)	76.475 (0.636)	76.669 (0.660)	76.158 (0.505)	76.587 (0.462)	76.481 (0.370)	76.123 (0.426)	76.299 (0.550)	76.622 (0.528)	76.364 (0.456)	76.663 (0.496)	76.352 (0.610)	76.710 (0.512)	75.695* (0.354)

Table A 3: Mean AUC of different models with each model parameter removed in turn. Standard deviation is given in brackets. * indicates value significantly different ($p<0.05$) from the value with *All Variables*. The superscript letters against the values in the All column (a, b, or c) indicate whether there are significant differences between the models for that particular height of CWS calculation at the $p=0.5$ level.

Data Set	*Model*	*CWS Height*	*All Variables*	*Average CI_BAL*	*CI_BAL*	*Tree DBH*	*Stand Density*	*Dune*	*Gap Size*	*Hydro*	*Stand DBH*	*Soil*	*Stand Height*	*Tree Height*	*CWS Break*	*CWS Overturn*	*WAsP Wind Speed*
Nezer	LOG	*d*+10 m	0.809[a] (0.005)	0.812 (0.004)	0.81 (0.005)	0.811 (0.005)	0.809 (0.005)	0.809 (0.005)	0.809 (0.005)	0.809 (0.005)	0.812 (0.005)	0.809 (0.005)	0.809 (0.005)	0.815 (0.003)	0.802 (0.004)	0.803 (0.005)	
	NN		0.796[b] (0.015)	0.792 (0.009)	0.805 (0.016)	0.791 (0.012)	0.791 (0.014)	0.794 (0.01)	0.789 (0.012)	0.794 (0.01)	0.791 (0.011)	0.794 (0.01)	0.795 (0.014)	0.791 (0.008)	0.793 (0.009)	0.795 (0.011)	
	RF		0.827[c] (0.009)	0.823 (0.011)	0.82 (0.012)	0.83 (0.01)	0.825 (0.01)	0.825 (0.009)	0.827 (0.01)	0.826 (0.01)	0.826 (0.009)	0.826 (0.01)	0.826 (0.01)	0.821 (0.009)	0.826 (0.009)	0.83 (0.008)	
NFI	LOG	*d*+10 m	0.751[a] (0.002)	0.751 (0.002)	0.751 (0.002)	0.751 (0.002)	0.747* (0.002)	0.733* (0.002)	0.751 (0.002)	0.730* (0.002)	0.751 (0.002)	0.746* (0.002)	0.751 (0.002)	0.751 (0.002)	0.751 (0.002)	0.753 (0.002)	
	NN		0.767[b] (0.004)	0.755* (0.005)	0.765 (0.006)	0.764 (0.007)	0.759 (0.005)	0.752* (0.011)	0.760 (0.005)	0.747* (0.006)	0.757 (0.007)	0.759 (0.006)	0.761 (0.006)	0.765 (0.006)	0.764 (0.008)	0.763 (0.007)	
	RF		0.834[c] (0.005)	0.828 (0.004)	0.838 (0.006)	0.836 (0.005)	0.828 (0.005)	0.832 (0.005)	0.831 (0.005)	0.829 (0.004)	0.831 (0.005)	0.833 (0.006)	0.831 (0.005)	0.836 (0.006)	0.834 (0.005)	0.833 (0.005)	
	LOG	29 m	0.766[a] (0.002)	0.767 (0.002)	0.767 (0.002)	0.766 (0.002)	0.760* (0.002)	0.752* (0.002)	0.766 (0.002)	0.747* (0.002)	0.767 (0.002)	0.762* (0.002)	0.766 (0.002)	0.766 (0.002)	0.767 (0.002)	0.765 (0.002)	0.760* (0.002)
	NN		0.768[a] (0.008)	0.761 (0.009)	0.766 (0.007)	0.769 (0.008)	0.768 (0.007)	0.764 (0.005)	0.767 (0.006)	0.752* (0.005)	0.765 (0.007)	0.760 (0.005)	0.764 (0.008)	0.766 (0.009)	0.772 (0.009)	0.770 (0.008)	0.763 (0.008)
	RF		0.837[b] (0.005)	0.833 (0.005)	0.839 (0.006)	0.838 (0.005)	0.834 (0.004)	0.836 (0.004)	0.836 (0.004)	0.832 (0.004)	0.836 (0.004)	0.837 (0.005)	0.837 (0.005)	0.839 (0.005)	0.836 (0.005)	0.838 (0.005)	0.832 (0.005)

Table A 4: Accuracy and AUC values for the models when tested on other data set (Nezer models tested on NFI data and NFI models tested on Nezer data). Standard deviations are given in parentheses. Mean value when models were tested on the data sets from which they were developed on second line in square brackets (from Table 5 and 6).

			LOG		*NN*		*RF*	
CWS Height	*Model Source Area*	*Model Test Area*	*Accuracy (%)*	*AUC*	*Accuracy (%)*	*AUC*	*Accuracy (%)*	*AUC*
d+10m	Nezer	NFI	53.443 (1.208) [65.972]	0.563 (0.013) [0.809]	50.358 (3.857) [67.176]	0.556 (0.040) [0.796]	54.411 (1.717) [71.306]	0.521 (0.019) [0.827]
29m	Nezer	NFI	52.657 (1.098) [66.954]	0.549 (0.011) [0.798	53.836 (2.828) [68.741]	0.531 (0.531) [0.799]	52.440 (1.735) [72.528]	0.578 (0.026) [0.834]
d+10m	NFI	Nezer	69.981 (1.673) [67.202]	0.766 (0.005) [0.751]	59.676 (7.145) [69.267]	0.741 (0.054) [0.767]	73.778 (1.794) [76.240]	0.735 (0.022) [0.834]
29m	NFI	Nezer	73.102 (1.770) [68.405]	0.756 (0.008) [0.766]	60.787 (10.062) [69.988]	0.724 (0.032) [0.768]	75.537 (1.223) [76.604]	0.713 (0.026) [0.837]