THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Variance-based sensitivity analysis of a wind risk model - Model behaviour and lessons for forest modelling

OPEN ACCESS

1 **Title:** Variance-based sensitivity analysis of a wind risk model - model behaviour and lessons for
2 forest modelling

3 **Highlights:**

4 - The Sobol' method for correlated variables is applied to a complex wind-risk model.
5 - The results are interpreted from the viewpoints of model users and modellers.
6 - The variance-based approach is sensitive to the variables correlation structure.
7 - Rooting depth and soil type provide minor contribution to the outputs variance.
8 - ForestGALES models the dynamics of wind damage to forest stands very effectively.

9 **Abstract:**

10 We submitted the semi-empirical, process-based wind-risk model ForestGALES to a variance-based
11 sensitivity analysis using the method of Sobol' for correlated variables proposed by Kucherenko et al.
12 (2012). Our results show that ForestGALES is able to simulate very effectively the dynamics of wind
13 damage to forest stands, as the model architecture reflects the significant influence of tree height,
14 stocking density, dbh, and size of an upwind gap, on the calculations of the critical wind speeds of
15 damage. These results highlight the importance of accurate knowledge of the values of these
16 variables when calculating the risk of wind damage with ForestGALES. Conversely, rooting depth and
17 soil type, i.e. the model input variables on which the empirical component of ForestGALES that
18 describes the resistance to overturning is based, contribute only marginally to the variation in the
19 outputs. We show that these two variables can confidently be fixed at a nominal value without
20 significantly affecting the model's predictions. The variance-based method used in this study is
21 equally sensitive to the accurate description of the probability distribution functions of the
22 scrutinised variables, as it is to their correlation structure.

23 **Keywords:**

24 Method of Sobol'; Assessment of model performance; Copula method; Correlated variables

25 **Software availability:**

26 Name of software: ForestGALES Developers: Forest Research, INRA, and the University of Edinburgh.
27 Contact address: Forest Research, Northern Research Station, Roslin, Midlothian EH25 9SY, United
28 Kingdom Email: forestgales.support@forestry.gsi.gov.uk Availability and Online Documentation: The
29 software along with supporting material is freely available. Go to
30 http://www.forestresearch.gov.uk/forestgales to find out how to obtain the software or email
31 forestgales.support@forestry.gsi.gov.uk Year first available: 2000 Hardware required: IBM
32 compatible PC Software required: MS Windows Programming language: Borland Delphi 5.0®.
33 Versions have also been written in Python, Fortran, R and Java. Contact Prof. Barry Gardiner
34 (barry.gardiner@bordeaux.inra.fr) for further details. Contact the corresponding author
35 (tom.locatelli@ed.ac.uk) for information on the R version. Program size: 10 MB. With all additional
36 support files and manuals ¼ 25 MB. For free professional tools for sensitivity analysis please visit the
37 European Commission Joint Research Centre sensitivity analysis page at
38 https://ec.europa.eu/jrc/en/samo/simlab Please contact Dr. Stefano Tarantola
39 (stefano.tarantola@jrc.ec.europa.eu) for information on the Matlab scripts of the Sobol' method for
40 the case of correlated variables.

1. Introduction

Environmental modelling has become a crucial part of the study of environmental phenomena. Significant advances in the fields of hardware and computing now allow for the creation of complex, computationally-demanding, process-based models, aimed at the investigation of natural systems (e.g. Nossent el al., 2011). These complex models are extensively adopted in support of decision-making and for environmental policy settings (e.g. Rahmstorf et al. (2007) on IPCC projections). While a large amount of time and resources are spent to formalise nature in mathematical terms, considerably less effort is often made to investigate the behaviour of mathematical models, which is often done as an "afterthought" (Saltelli and Funtowicz, 2014). As elegantly discussed by Oreskes et al. (1994), the same practices of model validation, evaluation, and confirmation, are philosophical and practical minefields. Modellers are confronted with these issues for a number of reasons: natural systems, which are inherently open in nature, are forced into closed systems to obtain mathematical solutions; scaling issues can arise when the scales at which some elements of a model are calculated differ from the scale of application of the model; nonuniqueness of modelling approaches might result in a faulty model providing "reasonable" outputs (Oreskes et al., 1994). Ultimately, however, the main issue with environmental modelling is the same reason why models are built: we can never exactly know all the data, and those that we do know, we do so with a degree of uncertainty. With regards to the modelling process, in our paper we refer to uncertainty as incomplete knowledge of parameter values (Gaber et al., 2009). Deterministic approaches to modelling require elimination of these uncertainties, thus effectively further removing a model from its intended representation of reality. The inadequacy of the attempts to eliminate at all costs the uncertainties of the parameters and variables of a model, in order to produce completely deterministic results, is nowadays generally accepted (e.g. Penman et al., 2003). The transparency of model predictions is an important requirement especially when models are applied for decision-making, and in policy frameworks (e.g. the US Environmental Protection Agency, see Gaber et al. (2009)). To this end, uncertainty analysis is normally applied to quantify the uncertainties of the input variables, parameters, and outputs of a model, thus providing some insight on the reliability and the applicability range of the model.

On the other hand, the issue of sensitivity of model predictions to variation in model parameters and variables is still relatively underestimated. Quoting Saltelli et al. (2004), a sensitivity analysis is *"The study of how uncertainty in the output of a model (…) can be apportioned to different sources of uncertainty in the model input"*. However, when performed appropriately (Saltelli and Annoni, 2010), sensitivity analysis (SA) of mathematical models is a tool that can help with fundamental issues about the robustness and the behaviour of a model (Tarantola et al., 2002; Norton, 2015). A number of techniques exist to perform sensitivity analysis (see https://ec.europa.eu/jrc/en/samo/methods). These can be broadly divided in two groups, typically referred to as "local" and "global", on the basis of the region of the input space that is scrutinised in the analysis. Local SA are normally based on derivatives of the output Y with respect to one factor X_i (e.g. $\delta Y / \delta X_i$), where by factor here we denote either a model parameter or an input variable. These derivatives are often normalised by the input-output standard deviations (they are said to be sigma-normalised) to produce more robust sensitivity indices, as recommended by the Intergovernmental Panel on Climate Change in their guidelines on the inventories of greenhouse gases (IPCC, 1999; IPCC, 2000). However, with this approach only the base point where the derivatives are computed is investigated, which is an issue

84 when the model is of unknown linearity (Saltelli et al., 2008). Local derivatives-based methods are
85 mostly adopted within the context of one-at-a-time (OAT) approaches, where only one factor is
86 perturbed while all the others are fixed at a nominal value (usually the mean). Therefore, the effects
87 of factors interactions on the output variance are neglected with OAT methods, which are therefore
88 only applicable for strictly additive models (Campolongo and Saltelli, 1997). Global SA (GSA)
89 methods, on the other hand, allow for the exploration of the entire range of the factors, and for
90 simultaneous perturbation of all the factors. The most powerful GSA methods are variance-based
91 techniques that decompose the total variance of the output into conditional variances for single
92 factors and for sets of factors. These techniques include the importance measures of Iman and Hora
93 (1990) and of Sacks et al. (1989), the FAST (Fourier Amplitude Sensitivity Test) method (Cukier et al.,
94 1973; Cukier et al., 1978) and the extended FAST (Saltelli et al., 1999), and the method of Sobol'
95 (Sobol', 2001). The last two approaches can be solved numerically with Monte Carlo methods.
96 Derivatives-based methods have been developed for global sensitivity measures (DGSM, e.g.
97 Kucherenko et al., 2009; Sobol' and Kucherenko, 2009). The values of DGSM is exactly equal to that
98 of total sensitivity indices calculated with the Sobol' method (see section 2.2.1) in a number of cases,
99 e.g. for linear models, while in a general case they correspond to the upper bound of the total Sobol'
100 indices, with the advantage of a much shorter computational time. Variance-based GSA methods
101 have a number of advantages: they are model-independent; they can capture the influence of the
102 full range of variation of each input variable; they allow for the investigation of interaction effects
103 amongst variables; and they provide the possibility of grouping factors (Saltelli et al., 2008). Their
104 drawback is the high computational cost required for performing such techniques, due to the large
105 number of model executions required for the convergence of the values of the sensitivity indices
106 (Kucherenko et al., 2012). For this reason, a large body of research has been devoted to devise
107 efficient algorithms for their computation (e.g. Kucherenko et al., 2012; Mara and Tarantola, 2012;
108 Most, 2012; Saltelli, 2002).

109 Of the aforementioned variance-based GSA techniques, the method of Sobol' has found favour with
110 modellers in the environmental sciences, because of the relatively straightforward interpretation of
111 the sensitivity indices calculated with this method, and because it very efficiently samples the factors
112 space (Sobol', 1990; Yang, 2011; Kucherenko et al., 2015). The Sobol' method is often used as a
113 benchmark against which to compare the results of other SA techniques (Confalonieri et al., 2010).
114 In a previous issue of this journal, Nossent et al. (2011) successfully applied the Sobol' method to the
115 identification of the most, and the least, important factors in a SWAT model (Soil and Water
116 Assessment Tool). The authors also provided an exhaustive description of the Monte Carlo
117 procedures required for the calculation of the Sobol' sensitivity indices. Song et al. (2012) used the
118 method of Sobol' for the SA of the 3-PG2 forest growth model, aimed at model calibration. A known
119 issue with variance-based GSA techniques is how to account for correlation between factors when
120 calculating the conditional variances. Indeed, correlation amongst factors in environmental models is
121 typical. A number of studies propose methods to obviate the issue of dependent factors in GSA (e.g.
122 Mara and Tarantola, 2012; Most, 2012).

123 In this paper, we submit ForestGALES, a forest wind-risk model, to a variance-based GSA using the
124 method of Kucherenko et al. (2012), a generalisation of the method of Sobol' for correlated factors.
125 The rationale of ForestGALES, together with the most important model calculations for the context
126 of our GSA, is discussed in the Methods section. For a thorough description of the model, the
127 interested reader is referred to Hale et al. (2015), published in a previous issue of this journal.

128   Variance-based GSA are normally applied to complex models composed of a large number of factors,
129   sometimes in excess of one hundred, mostly for the direct benefit of the modelling community. In
130   this paper, we limit our GSA to the inputs of ForestGALES that are controllable by the end-users.
131   Focussing on those input variables that are user-modifiable extends the benefits of a GSA to the end-
132   user base of an environmental model, and facilitates the interpretation of the results of the SA in a
133   practical setting. To extend the results of our GSA to a large user-base community, we perform our
134   GSA on three species *(Picea sitchensis* (Bong.) Carr., *Pinus pinaster* (Ait.), and *Eucalyptus globulus*
135   (Labill.))*, representative of three of the most extensively planted and highly productive tree genera
136   worldwide: spruces, pines, and eucalypts. We also investigate the differences in the ranking of the
137   influential variables between the three species, to evaluate whether the sensitivity of the model to
138   its input parameters is the same across the species used in the simulations. In this paper we focus
139   our attention on two questions that SA can help with: (1) What model inputs should a user of
140   ForestGALES focus on knowing more accurately to maximally reduce the uncertainty in the model
141   predictions? (2) What model inputs contribute the least to the variation in the output? The first
142   question can be answered under the Factor Prioritisation setting of GSA, while the second pertains
143   to the Factor Fixing setting (Saltelli et al., 2008), discussed in the Methods section.

144

145   **2. Materials and methods**

146   *2.1 The ForestGALES model*

147   ForestGALES is a semi-mechanistic, process-based model that uses tree and stand characteristics for
148   the calculation of the critical wind speeds (CWS) that result in tree uprooting and stem breakage,
149   and combines them with information on the local wind climate to estimate the associated risks of
150   wind damage to forest stands (Hale et al., 2015). The model is based on engineering principles
151   (Gardiner et al., 2008) and requires only a small number of user inputs. These include: tree species,
152   which is used as a trigger to call species-specific sub-modules; tree-level variables, such as dominant
153   tree height (i.e. the average height of the hundred largest trees per hectare) and diameter at breast
154   height (*dbh*, measured at 1.3m height on the stem); and stand-level variables, such as stocking
155   density (*Sph*, stems per hectare), soil type, rooting depth, and size of any upwind gaps. Estimation of
156   the risk of damage requires a characterisation of the local wind climate. In Britain, where the model
157   was initially developed, this is done with DAMS scores (Detailed Aspect Method of Scoring), which
158   relate topographic characteristics of a site to the rate of tatter of tatter flags to estimate the
159   windiness of the site (Quine and White, 1993). For instance, a DAMS score of 10 represents very
160   sheltered conditions, a score of 17 is considered as quite exposed (Mason, 2003), and a score of 20
161   often corresponds to the limit for commercial forestry (Quine, 2000). DAMS scores are converted
162   internally to the model to scale and shape parameters of a Weibull distribution. Weibull distributions
163   have been recently reviewed as the most reliable and effective methods for describing distributions
164   of wind speed and directions (Seguro and Lambert, 2000). Because of the extreme sensitivity of the
165   Weibull distribution, small changes in CWS can result in large differences in the probability of
166   extreme events (Gardiner et al., 2008). The rationale of ForestGALES and the interplay between the
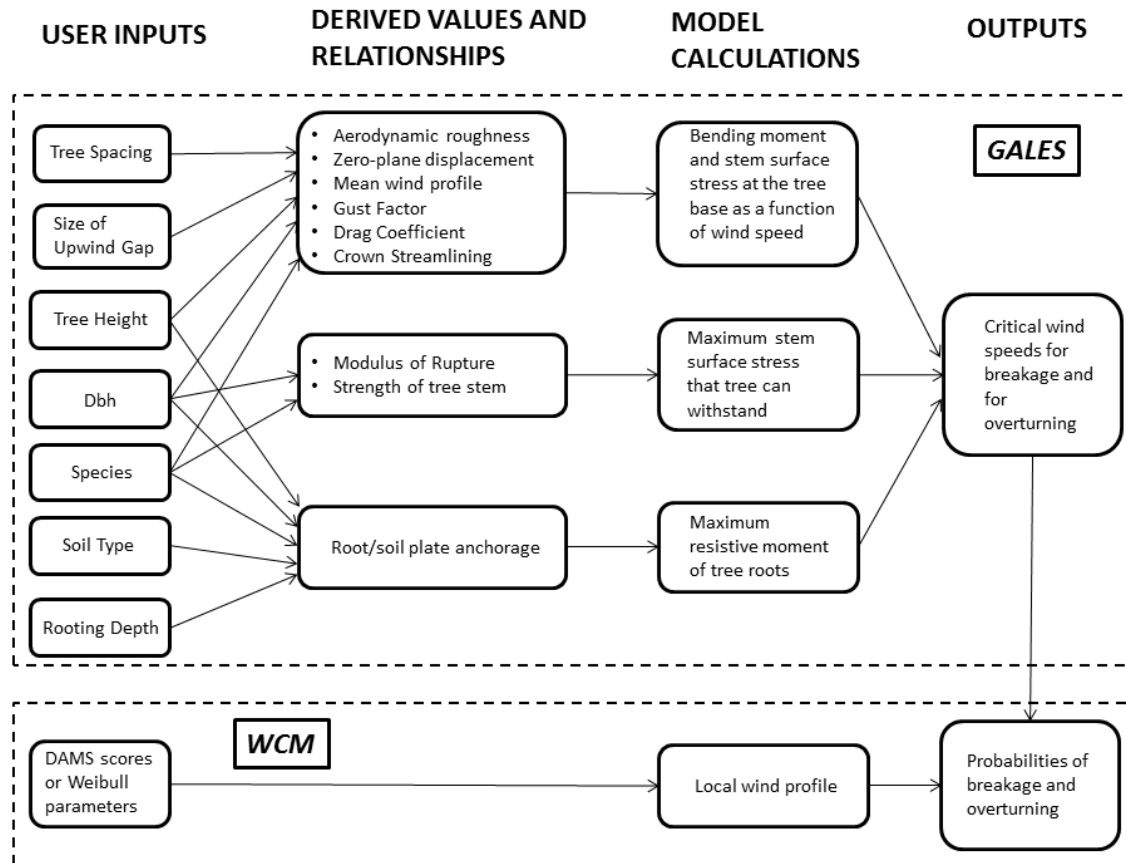167   inputs are shown in Figure 1.

168

169

170

171

Figure 1: Basic schematic representation of the rationale of ForestGALES from the point of view of user-defined input variables (adapted from Gardiner et al., 2000). The dashed boxes delimit the two main modules. GALES calculates the critical wind speeds for breakage and overturning. WCM: Wind Climate Module, where the probabilities of breakage and overturning are calculated.

177

The diagram in Figure 1 shows that ForestGALES is composed of two main modules: the first one, often referred to simply as GALES (Gardiner et al., 2000), makes use of tree and stand variables to calculate the CWS for breakage and overturning (henceforth: $CWS_{(B, O)}$). In GALES, tree height and *dbh* are featured in the calculations of canopy dimensions and aerodynamic properties, as well as being involved in the calculations of the mean wind profile, together with *Sph* and size of an upwind gap. Soil type and rooting depth are used in species-specific sub-modules to retrieve the values of species-specific coefficients (C_{reg}) of linear regressions of total overturning moment, as measured empirically in the field, against stem weight under different soil types and rooting depths. These relationships are derived from tree-pulling fieldwork data used in species parameterisations of ForestGALES (see Nicoll et al. (2006) for *P. sitchensis*; Cucchi et al. (2005) for *P. pinaster*; and Locatelli et al. (2016) for *E. globulus*). For *E. globulus*, only one value is available because tree-pulling was

189 performed only in one location with homogeneous soil type (162.32 N m kg^{-1}, from Locatelli et al.,
190 2016). The values of C_{reg} for *P. sitchensis* and *P. pinaster* are shown in Table 1.

191

192

193 Table 1: Values of the linear regressions of total overturning moment vs stem weight, for combinations of rooting depth
194 (shallow, medium, and deep) and soil type (freely draining, gleys, mineral peats, and deep peats). *P. sitchensis* and *P.*
195 *pinaster* are shown here. Only one value available for *E. globulus* (162.3, from Locatelli et al., 2016). Units are N m kg^{-1}. s.d.
196 denotes the standard deviation.

Picea sitchensis

		Rooting depth			Soil Average	Soil s.d.
		Shallow	Medium	Deep		
Soil Type	Freely draining	153.2	156.2	178.1	162.5	13.6
	Gleys	135.4	138.5	157.9	143.9	12.2
	Mineral peats	147.8	151.2	172.5	157.2	13.4
	Deep peats	168.1	172.1	196.2	178.8	15.2

	Shallow	Medium	Deep
Depth average	*151.1*	*154.5*	*176.2*
Depth s.d.	*13.5*	*13.9*	*15.8*

Pinus pinaster

		Rooting depth			Soil Average	Soil s.d.
		Shallow	Medium	Deep		
Soil Type	Freely draining	125.8	168.8	144.5	146.4	21.6
	Gleys*	124.2	144.4	126.7	131.8	11.0
	Mineral peats*	135.6	157.7	138.3	143.9	12.1
	Deep peats*	154.3	179.4	157.3	163.7	13.7

	Shallow	Medium	Deep
Depth average	*135.0*	*162.6*	*141.7*
Depth s.d.	*13.8*	*15.0*	*12.8*

*C_{reg} values of soil types: gleys, mineral peats, deep peats for *P. pinaster* are taken from those of *P .sylvestris* (L.)

197

198 Soil type and rooting depth do not contribute to the calculation of CWS_B, and therefore for this
199 output they are expected not to be flagged as important in the SA. Hale et al. (2015) exhaustively
200 describe the structure of ForestGALES. In this paper we limit ourselves to showing the final formulas
201 for the calculations of $CWS_{(B, O)}$, for the discussion of the results of the SA

202 $$CWS_B = \frac{1}{kD}\left[\frac{\pi * MOR * dbh^3}{32\rho G(d-1.3)}\right]^{\frac{1}{2}}\left[\frac{f_{knot}}{f_{CW}}\right]^{\frac{1}{2}}\ln\left(\frac{h-d}{z_0}\right) \qquad (1)$$

203 $$CWS_O = \frac{1}{kD}\left[\frac{C_{reg} * SW}{\rho G d}\right]^{\frac{1}{2}}\left[\frac{1}{f_{CW}}\right]^{\frac{1}{2}}\ln\left(\frac{h-d}{z_0}\right) \qquad (2)$$

204

205 Where k is von Karman's constant (0.4, dimensionless); D (m) is the mean tree spacing, calculated as
206 $1/\sqrt{sph}$; MOR is the Modulus of Rupture (Pa) of green wood; ρ is the air density (kg m^{-3}); G is a "gust
207 factor" (dimensionless), used to convert the calculated mean bending moments to maximum
208 bending moments; d is the zero-plane displacement (m), i.e. the height up the tree at which the
209 wind can be regarded as acting on a single point (Raupach, 1994); f_{knot} (dimensionless) is a factor to
210 account for the presence of wood knots, the values of which vary with species and typically range
211 between 0.8 and 1 (Ruel et al., 2010), and can decrease wood strength (Lavers, 1969); f_{CW} is a tree
212 mass factor (dimensionless), which accounts for the additional turning moment provided by the tree
213 canopy as a tree sways from its vertical axis under the action of the wind; h is the average tree
214 height (m); z_0 is the canopy roughness (m); C_{reg} (dimensionless) is described above; and SW is stem
215 weight (kg). Table 2 shows the input variables involved in the calculations of the non-constant terms
216 in Eq. (1) and (2).

217

218 Table 2: Mapping between ForestGALES input variables and the non-constant terms of the formulas for the calculations of
219 the critical wind speeds for breakage (CWS$_B$) and overturning (CWS$_O$). Species is used as a trigger to select species-specific
220 sub-modules; [a]Rooting depth and Soil type are involved in the calculations of d, f_{CW}, and z_0 only for CWS$_O$; [b]C_{reg} values are
221 empirically derived (see main text for a brief explanation, or Nicoll et al. (2006) for a complete description of tree-pulling
222 fieldwork techniques). The symbol (✓) denotes which formula variables are present in the corresponding calculations of
223 the critical wind speeds for breakage and overturning.

Formulas Variables		CWS$_B$	CWS$_O$	Input Variables
D	Mean tree spacing	✓	✓	Sph
G	Gust Factor	✓	✓	Tree height; Sph; Gap Size
d	Zero-plane displacement	✓	✓	Tree height; Dbh; Sph; Gap Size; Rooting depth[a]; Soil type[a]
f_{CW}	Tree mass factor	✓	✓	Tree height; Dbh; Sph; Gap Size; Rooting depth[a]; Soil type[a]
h	Mean tree height	✓	✓	Tree height
z_0	Canopy roughness	✓	✓	Tree height; Dbh; Sph; Gap Size; Rooting depth[a]; Soil type[a]
C_{reg}[b]	Overturning moment multiplier		✓	Rooting depth; Soil type
SW	Stem weight		✓	Tree height; Dbh

224

225 In the second main module of ForestGALES (see Figure 1), the local wind climate is used to calculate
226 the return period of a storm with maximum mean wind speeds that would match, or exceed, the
227 calculated CWS$_{(B, O)}$. The probabilities of breakage and overturning (henceforth: Prob$_{(B, O)}$) are
228 calculated as the inverse of the return period. We refer to the second main module as "wind climate
229 module" (WCM).

230

231 *2.2 Procedure for the GSA of ForestGALES*

232 The variance-based GSA method for correlated variables described in Kucherenko et al. (2012) is a
233 generalisation of the traditional Sobol' method for the calculation of the first-order and total
234 sensitivity indices (Sobol', 2001). First-order indices convey information on the direct contribution of
235 an input to the output's variance, and are therefore sometimes referred to as "importance

236 measures" (e.g. in Homma and Saltelli, 1996). Total indices account for the total contribution of an
237 input to the output's variance, by including all interaction effects with other variables (Saltelli, 2002).
238 Nossent et al. (2011) provide a thorough and accessible explanation of the theory and the
239 mathematics adopted in the method of Sobol'. Here, we limit ourselves to providing a general
240 description of this approach, before briefly discussing the modifications contributed by Kucherenko
241 et al. (2012) for the case of correlated inputs.

242

243 *2.2.1 The original method of Sobol'*

244 The original method of Sobol' is based on the propagation of the uncertainties in the inputs to the
245 outputs. The uncertainty in the inputs is expressed via independent marginal distribution functions.
246 This method is similar to ANOVA techniques, in that the model output variance V is decomposed into
247 summands (partial variances) of increasing dimensionality. These partial variances are calculated for
248 each X_i, (representing the importance of main effects) for pairs X_i and X_j, (representing the
249 importance of interactions between pairs) and finally for groups of inputs, such as $X_i, X_j, ... X_m$,
250 (representing the importance of higher order interactions)(Sobol', 2001). First order sensitivity
251 indices of the output Y to the inputs X_i can then be introduced, simply by normalizing the first order
252 variances V_i by the total variance V (Sobol', 2001). This is commonly done using the variance of the
253 expectation of Y conditional on a fixed value of X_i, by averaging for all values of X_i (Saltelli et al.,
254 1999), as shown in Eq. (3).

$$S_i = \frac{V(E(Y|X_i))}{V_Y}$$

(3)

255

256 Where the S_i is the first order sensitivity index for X_i, the numerator is the conditional variance, and
257 V_Y is the total variance of Y. With the method of Sobol', total sensitivity indices (S^T_i) can be computed
258 by grouping together all the variables but the one for which the S^T_i are calculated. Total indices can
259 be calculated with Eq. (4).

$$S_i^T = 1 - V_{X_{-i}}/V_Y$$

(4)

260

261 where $V_{X_{-i}}$ is the variance of all the variables except X_i, for which the indices are calculated. By
262 repeating for all X_i, it is possible to calculate the total effects of all the input variables. First order and
263 total sensitivity indices are computed with Monte Carlo techniques, at the cost of $N(m + 2)$ model
264 runs, where N is the sample size and m the number of model inputs (Archer et al., 1997; Chan et al.,
265 1997).

266

267 *2.2.2 Generalisation of the Sobol' method for the case of correlated variables*

268 In the case of correlated inputs, describing the inputs uncertainty solely with independent marginal
269 distribution functions is inadequate. Kucherenko et al. (2012) suggest sampling from both the joint
270 and conditional distribution functions of the inputs, which requires providing the correlation matrix

271 of the inputs. When at least one of the inputs is non-normally distributed, the authors suggest the
272 use of a Gaussian copula to generate the correlated sample. This technique is widely used in studies
273 of financial risk (e.g. Cherubini et al., 2004). In this setting, correlated inputs are described by their
274 marginal distribution and a measure of their correlation with the other inputs:

275 $$C(G_1(X_1), \ldots, G_n(X_n); \Sigma_X) = F_n(F^{-1}(G_1(X_1)), \ldots, F^{-1}(G_n(X_n)); \Sigma) \qquad (5)$$

276 Where C denotes the copula; $G_{1\ldots n}$ are the marginal univariate cumulative distribution functions; $X_{1\ldots n}$
277 are the original inputs; Σ_X is the original correlation matrix; F_n is the multivariate cumulative *normal*
278 distribution function; F^{-1} is the inverse normal cumulative distribution function; and Σ is the
279 correlation matrix of the corresponding joint *normal* distribution function. With this approach, the
280 original correlation matrix Σ_X of an original input vector $\boldsymbol{X}$ is mapped to the correlation matrix Σ. The
281 adoption of copulas is convenient for SA because all the information on the dependencies between
282 inputs is contained in the copula, while the information contained in the marginal distributions of
283 the inputs is provided by the marginal univariate *cumulative* distribution functions of the inputs
284 (Sklar, 1973). In our study we used eq. 5.4 and 5.6 in Kucherenko et al. (2012) for the calculations of
285 S_i and S^T_i, respectively, at a cost of $N(2m + 2)$ (35,200 in our case) model runs, using the Quasi-Monte
286 Carlo method of Sobol' applied to the case of correlated variables (Sobol', 1990; Kucherenko et al.,
287 2012). Quasi-Monte Carlo methods are commonly used in GSA because they provide enhanced
288 convergence properties in comparison to traditional Monte Carlo methods, and hence require
289 considerably less model executions to achieve a given precision in the estimates (Sobol', 1998;
290 Kucherenko et al., 2012).

291

292 *2.2.4 Data used in the GSA of ForestGALES*

293 Both the original method of Sobol' and that of Kucherenko et al. (2012) require knowledge of the
294 probability distribution functions of the inputs for the generation of the samples. Describing the
295 inputs with appropriate PDFs is a requirement of GSA methods, in order to ensure that the pseudo-
296 random numbers generated with Monte Carlo methods are representative of the variables
297 distributions. We described some variables (*Sph*, Rooting depth, Soil type, Gap size, and DAMS) with
298 uniform or discrete uniform distributions, using the same distribution parameters for all the species,
299 because we wanted to explore as large an input space as possible with regards to these variables.
300 The gap used in this version of ForestGALES is what is normally referred to as "green edge gap"; that
301 is, a gap that has been in place for some time, rather than a newly exposed one. For tree height and
302 dbh, we found that our data source was well represented with Gaussian distributions. We have
303 therefore fitted normal distributions to tree-pulling data for the three species used in this paper: for
304 *P. sitchensis*, we used data from the UK Forestry Commission tree-pulling database (see Nicoll et al.,
305 2006); for *P. pinaster*, we combined tree-pulling data from Cucchi et al. (2005) with data from a
306 similar species (Scots pine, *Pinus sylvestris* (L.)), from the UK database, because complete data for *P.*
307 *pinaster* was not available. In fact, tree-characteristics such as canopy structure and the architecture
308 of the rooting system are very similar between *P. pinaster* and *P. sylvestris*. For *E. globulus*, we used
309 tree-pulling data used in Locatelli et al. (2016) for the parameterisation of ForestGALES for this
310 species. The lack of variability in Soil type and Rooting depth for the parameterisation of this
311 particular species means that our study cannot evaluate the sensitivity of ForestGALES to these two
312 variables, and of their contributions to the interactions with each other and the other input

313 variables, for *E. globulus*. We imposed lower bounds to the distributions of tree height (4m) and dbh
314 (3cm), to ensure that the functional limits of ForestGALES were not exceeded. The parameters of the
315 inputs distributions are shown in Table 3. Our proposed approach is data-driven, i.e. it is an
316 exploration of the behaviour and sensitivity of the ForestGALES model from the point of view of the
317 data used for its parameterisations. This ensures that the operational limits of the model are
318 respected as much as possible, while allowing for large amount of data within the probability density
319 functions of the inputs to explore the input space as thoroughly as possible.

320

321 Table 3: Parameters of the probability distribution functions used for the generation of quasi-random samples used for the
322 calculation of sensitivity indices. Parameters calculated from tree-pulling data. Dbh: diameter at breast height (1.3m); Sph:
323 stems per hectare; DAMS: Detailed aspect method of scoring, a measure of the windiness of a site. [a]Lower bound for tree
324 height: 4m. [b]Lower bound for dbh: 3cm.

Variable	Species	Parameters of the Normal Distribution	
		Mean	sd
Tree height[a] (m)	*Picea sitchensis*	13.64	2.64
Dbh[b] (cm)		19.89	4.52
Tree height[a] (m)	*Pinus pinaster*	13.70	2.75
Dbh[b] (cm)		18.98	3.51
Tree height[a] (m)	*Eucalyptus globulus*	23.17	4.59
Dbh[b] (cm)		21.78	7.51
	Type of distribution	Min	Max
Sph	Discrete Uniform	300	3300
Rooting depth	Discrete Uniform	1	3
Soil type	Discrete Uniform	1	4
Gap size (m)	Uniform	0	1000
DAMS	Discrete Uniform	7	22

325

326 Based on the data available, we calculated the correlation matrix (Σ_X) of the input variables for the
327 three species for the Gaussian copula required for the GSA, shown in Table 4. As expected, for all the
328 species the most relevant correlation was between tree height and *dbh* (*P. sitchensis*: 0.54; *P.*
329 *pinaster*: 0.72; *E. globulus*: 0.91. All *p*-values < 0.001). We calculated the correlation between these
330 two variables from species-specific tree-pulling data. Due to the fact that correlation data involving
331 the other variables were only available for *P. sitchensis*, we have applied this to all the species. Gap
332 size and DAMS are not correlated to any other variable. The large values of the Pearson correlation
333 coefficients between tree height and dbh justify the use of the GSA method for correlated variables.

334

335

336

337

338

339

340

341

342

343

344

Table 4: Correlation matrices for the ForestGALES input variables for *P. sitchensis*. Dbh: diameter at breast height; Sph:
stems per hectare; Gap size and DAMS are not shown because uncorrelated to the other variables. Correlations involving
the variables: Sph, Rooting depth, and Soil type, are calculated from *P. sitchensis* data from the UK Forestry Commission's
tree-pulling database, and applied to the other two species. [a] The significance of the correlation between pairs of variables
with Pearson's correlation coefficient > 0.3 (thus retained in the calculation of the copula) is large (all *p*-values < 0.001)

Picea sitchensis	Height	Dbh	Sph	Rooting depth	Soil type
Height	1	0.54[a]	-0.07	0.22	-0.07
Dbh		1	0.04	0.13	-0.02
Sph			1	0	-0.28
Rooting depth				1	-0.35[a]
Soil type					1

350

In Figure 2 we show the effect of the correlation between Rooting depth and Soil type (Pearson
correlation coefficient: -0.35; *p*-value < 0.001) on the sampling matrix calculated with the copula
method. The occurrence of the different Soil types modelled in Figure 2 is representative of the *P.
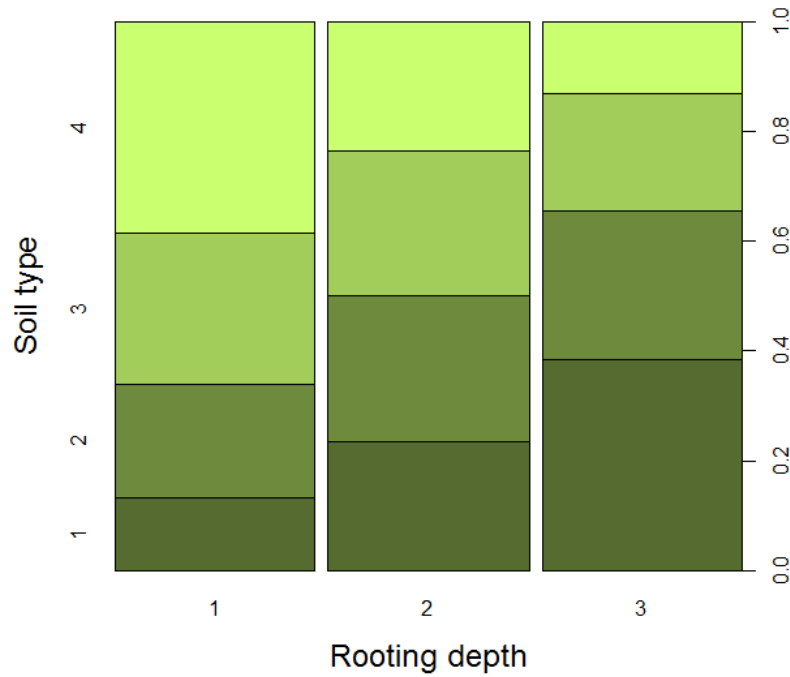sitchensis* tree-pulling database used for the calculation of the correlation coefficients.

355

356

357 Figure 2: Distribution of the levels of Rooting depth and Soil type in the generation of the sample for the sensitivity analysis
358 using the copula method. The Pearson correlation coefficient between the two variables is -0.351. Rooting depth levels:
359 1=shallow; 2=medium; 3=deep. Soil type levels: 1=Freely draining; 2=Gleys; 3=Mineral peats; 4=Deep peats. This Figure
360 shows that the Quasi-random values of Rooting depth and Soil type generated with the method of Sobol' are influenced by
361 the correlation structure of the copula. E.g.: for shallow Rooting depth (level: 1), the likelihood of a Soil type "Deep peats"
362 (level: 4) is higher than that the other types of soil. For deep Rooting depth (level: 3), the likelihood of a Soil type "Freely
363 draining" (level: 1) is higher than that of the other types of soil. For medium Rooting depth (level: 2) all Soil types are
364 almost equally likely to be selected with the Quasi-random number generator under the correlation structure of the
365 copula.

366

367 *2.2.5 ForestGALES simulations*

368 From the samples generated for each of the three species, we used ForestGALES to calculate two
369 sets of outputs, $CWS_{(B, O)}$ and $Prob_{(B, O)}$. Because of the structure of the model, and the nature of the
370 outputs, we performed different analyses on the two sets of outputs, as described below.

371

372 *2.2.5.1 Sensitivity analysis of the output: Critical wind speeds*

373 In order to investigate what variables the model users should focus on knowing more accurately to
374 maximally reduce uncertainty in the model's predictions of CWS for each species, we applied the SA
375 Factor Prioritisation (FP) setting. Following this setting, data collection of the variables with the
376 largest S_i should be prioritised for reducing the uncertainty of the inputs and optimising predictions
377 calculated with ForestGALES. Conversely, to identify those variables which negligibly contribute to
378 the variance of the CWS, we adopted the SA setting Factor Fixing (FF), which is based on S^T_i. When
379 the S^T_i of a variable was found to be close to zero, we fixed that variable to different values within its
380 range, and re-ran the simulations with ForestGALES, while maintaining the variation in all the other
381 variables. For discrete variables, we explored all the possible values. For continuous variables, we

382   focussed on the minimum, maximum, and mean values. We then compared the original CWS$_{(B, O)}$
383   with those calculated after fixing a non-influential variable using scatterplots. We used the method
384   of Sobol' et al. (2007) for the estimation of the approximation error when fixing non-influential
385   variables. The authors have shown that for orthogonal (i.e. non-correlated) variables, the
386   approximation error when fixing uninfluential variables is equal to twice the value of the S^T_i of the
387   uninfluential variables, and that it also applies to groups of variables. In the case of multiple non-
388   influential variables, we repeated these procedures for all the combinations of the values of the
389   relevant variables.

390

391   *2.2.5.2 Sensitivity analysis of the output: Probabilities of damage*

392   For Prob$_{(B, O)}$ we applied the Factor Mapping (FM) setting of SA, which is based on Regionalised
393   Sensitivity Analysis (e.g. Hornberger and Spear, 1981), a procedure belonging to the family of Monte
394   Carlo Filtering methods. We divided the output space in a behavioural region and a non-behavioural
395   region, setting the threshold at Prob$_{(B, O)}$ = 0.1 (i.e. a 10% probability of damage) to differentiate
396   between endemic and catastrophic wind damage (Mitchell, 1998). That is, when the calculated
397   Prob$_{(B, O)}$ were smaller than 0.1, we assigned the model run to the corresponding behavioural
398   regions, and to the non-behavioural ones otherwise. We mapped the outputs to the values of the
399   input variables, and investigated the sensitivity of ForestGALES to DAMS, and to the other inputs
400   that were flagged as influential in the calculations of the respective CWS. In fact, DAMS is the only
401   input variable directly involved in the wind climate model, and logically a non-influential variable for
402   the calculation of CWS$_{(B,O)}$ cannot drive variation in the corresponding Prob$_{(B, O)}$. To investigate the
403   sensitivity of the wind climate module to these variables, we plotted their marginal cumulative
404   distribution functions (CDF) conditional on the behavioural and non-behavioural realisations of the
405   model. We estimated the sensitivity of the model to these variables with Smirnov two-sample (two
406   sided) tests, which determine the significance of the differences between the behavioural and non-
407   behavioural CDFs (Saltelli et al., 2008). In order to investigate the second order interactions between
408   these variables, we used two-dimensional density plots to identify the regions in the bivariate input
409   space that are more likely to result in behavioural or non-behavioural realisations of ForestGALES.

410

411   *2.2.6 Differences between species*

412   Species is used in ForestGALES as a trigger to activate the corresponding sub-modules where a
413   number of tree characteristics such as canopy dimensions are calculated. Similarly, the values of C_{reg}
414   for different combinations of soil type and rooting depth, the species-specific values of the density
415   of green wood, and the Modulus of Rupture and the Modulus of Elasticity of green wood, are stored
416   in these sub-modules. To explore whether differences between species exist in the ranking of the
417   variables that drive most of the variance in CWS$_{(B,O)}$, we used a ranking method based on Savage
418   scores (Savage, 1956). Coefficients of concordance were used as described by Iman and Conover
419   (1987) and Helton et al. (2005) to compare the importance ranking of input variables between
420   species. For each species, the *m* input variables are ranked in order of their importance (expressed
421   as their S^T_i), reversely in comparison to the procedure used in standard rank regressions; that is, the
422   variable with the highest S^T_i is given a rank $r(S^T_i)$ of 1, the variable with the second highest S^T_i a rank

423 of 2, and so on. This procedure is often used to calculate Kendall's coefficients of concordance (KCC),
424 which assign equal weight to each rank. However, because the purpose of a sensitivity analysis is to
425 identify the most important variables (i.e. those with the top ranks), KCC are inappropriate (Helton
426 et al., 2005). Therefore, the ranks are replaced by the corresponding Savage scores to emphasise the
427 importance of the top ranks, as described in Iman and Conover (1987). For this, we have adopted Eq.
428 (6), adapted from Helton et al. (2005).

$$ss(S_i^T) = \sum_{j=r(S_i^T)}^{m} 1/j \tag{6}$$

429

430 where $ss(.)$ indicates the Savage scores of variable X_i, and the m variables in the model are ranked in
431 descending order of importance according to their S_i^T. In the present study $m = 6$, as species is the
432 variable under scrutiny here. Averages are calculated in the event of ties (Iman and Conover, 1987).
433 This substitution allows the calculation of the top-down coefficient of concordance amongst all
434 species (C_T), with Eq. (7), adapted from Iman and Conover (1987) and Helton et al. (2005):

$$C_T = \left\{ \sum_{j=1}^{m} \left[\sum_{q=1}^{nS} ss_q(S_i^T) \right]^2 - nS^2 m \right\} \bigg/ \left\{ nS^2 \left(m - \sum_{j=1}^{m} 1/j \right) \right\} \tag{7}$$

435 where nS is the number of species (here, $nS = 3$). To calculate a value of correlation amongst all the
436 species we used an index developed by Iman and Conover (1987). The authors showed that

$$T_{all} = nS(m - 1)C_T \tag{8}$$

437

438 follows a chi-square distribution with $(m - 1)$ degrees of freedom, with the subscript "*all*" indicating
439 a value of correlation amongst all the species. Hence, a comparison between T_{all} and a chi-square
440 distribution with d.f. = 5 provides a p-value to test the null hypothesis of independence of the
441 importance rankings of the input variables between different species. That is, a small p-value would
442 prompt us to reject the null hypothesis, and would indicate that the rankings of the most important
443 input variables are similar between species.

444 In order to delve further into the differences between pairs of species, a similar test was carried out
445 with Pearson correlation coefficients adapted for Savage scores when $nS = 2$, as shown by Iman and
446 Conover (1987). For each pairing of species α and β, Savage scores were applied to the rankings of
447 their S_i^T with Equation (6). The top-down coefficient of concordance between each pair was then
448 calculated with Eq. (9), adapted from Iman and Conover (1987) and Helton et al. (2005).

$$r_T = \left(\sum_{j=1}^{m} ss_{\alpha_j} ss_{\beta_j} - m \right) \bigg/ \left(m - \sum_{j=1}^{m} 1/j \right) \tag{9}$$

449

450 where $ss_{\alpha j}$ and $ss_{\beta j}$, ($j = 1,...,m$) are the Savage scores for the rankings of the m S_i^T for species α and β,
451 respectively. The significance of the differences between each pair of species is then evaluated
452 against a chi-square distribution with $(m - 1)$ degrees of freedom, using an adaptation of Eq. (8).

$$T_{\alpha,\beta} = 2(m-1)r_T \qquad\qquad\qquad (10)$$

453

454   **3. Results**

455   The Results section is divided in three parts: the first one shows the results of the GSA for the
456    $CWS_{(B,O)}$ calculated in the GALES module. The second part shows the similarities in ranking of
457   sensitivity indices between species. The third part describes the GSA results for the $Prob_{(B,O)}$
458   calculated with the wind climate module. In the first and third parts the three species are presented
459   one at a time.

460

461   *3.1 Critical Wind Speeds – Sensitivity in the GALES module*

462   Throughout this section, we first show the results for *P. sitchensis*, followed by *P. pinaster*, and lastly
463   *E. globulus*. We then introduce the results of the Factor Fixing setting.

464

465   *3.1.1 Picea sitchensis*

466   For *P. sitchensis*, the sensitivity of the GALES module to the input variables summarised in Table 2
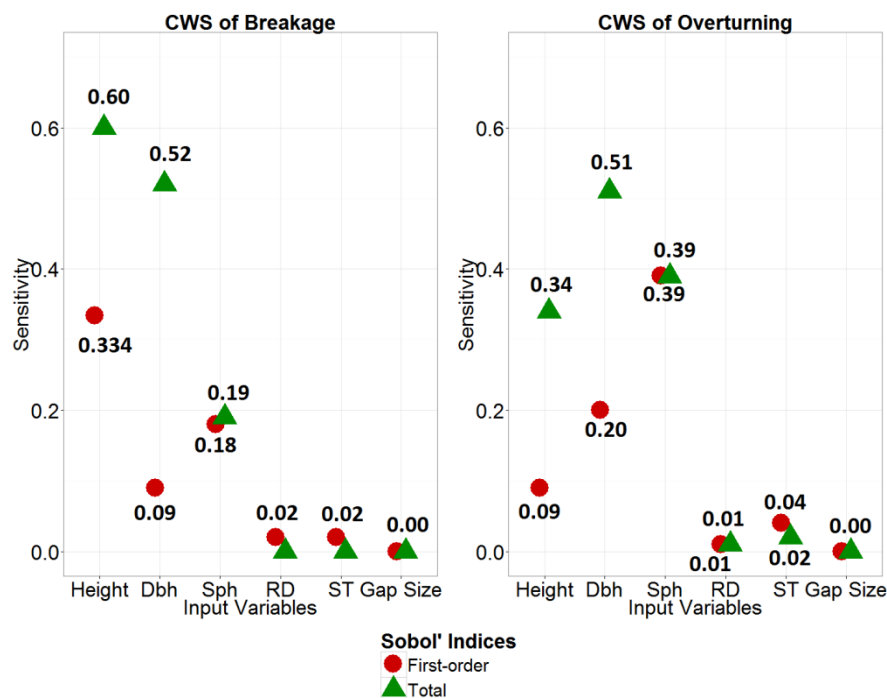467   (DAMS aside) is shown in Figure 3. The Sobol' sensitivity indices are shown for $CWS_{(B,O)}$.

468

469

470   Figure 3: Sobol' first-order and total sensitivity indices for the critical wind speeds for breakage and overturning for *P.*
471   *sitchensis*. Dbh: diameter at breast height; Sph: stems per hectare. RD: rooting depth. ST: soil type. The total indices of RD,
472   ST, and Gap size for CWS_B are 0.00. The total index of Gap size for $CWS_{(B,O)}$ is 0.00.

473

474   As seen in Figure 3, Tree height, *dbh*, and *Sph* are the variables to which the $CWS_{(B,O)}$ are most

475   sensitive, although their importance ranking differs between the two CWS, as shown by their S^T_i. The

476   large differences between the S^T_i (green triangles) and S_i (red circles) of Tree height and *dbh* indicate

477   that these variables are involved in a large number of interactions with other variables, for both

478   breakage and overturning. Seen that the S^T_i of the other variables are never significantly larger than

479   their S_i, these interactions are mainly between Tree height and *dbh* themselves. As expected,

480   Rooting depth and Soil type do not contribute to the variation of CWS_B. These two variables are only

481   marginally influential with regards to CWS_O, with Rooting depth being more important than Soil

482   type. The size of an upwind green edge Gap is uninfluential to the calculations of $CWS_{(B,O}$

483

484   *3.1.2 Pinus pinaster*

485   Figure 4 shows the indices of Sobol' for *P. pinaster*.
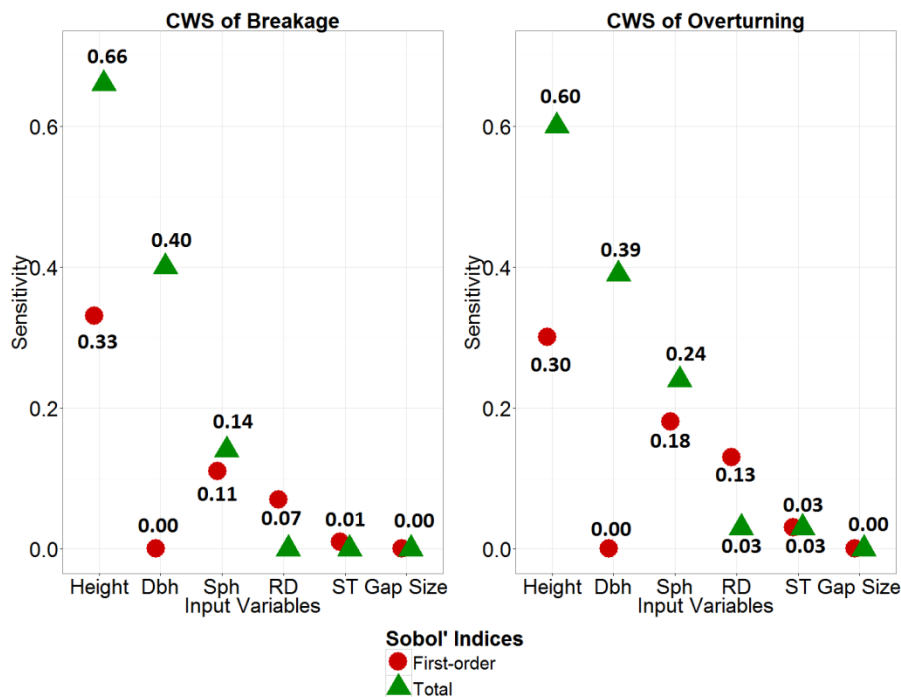

487   Figure 4: Sobol' first-order and total sensitivity indices for the critical wind speeds for breakage and overturning for *P.*
488   *pinaster*. Dbh: diameter at breast height; Sph: stems per hectare. RD: rooting depth. ST: soil type. The total indices of RD,
489   ST, and Gap size for CWS_B are 0.00. The total index of Gap size for CWS_O is 0.00.

490

491   Figure 4 shows that, for *P. pinaster*, the most influential variables are, in decreasing order: Tree

492   height, *dbh*, and *Sph*, for both CWS. The large differences between the S^T_i and S_i seen for *P. sitchensis*

493   are found also for *P. pinaster*, indicating large interactions. *Sph* is likely to participate more

494   prominently than for *P. sitchensis*, as its S^T_i is slightly larger than its S_i, especially for CWS_O. Rooting

495   depth's S_i for CWS_B is marginal (0.07), suggesting that the moderate correlation with the highly

496   influential Tree height (0.22, see Table 4) is responsible for this non-zero value. For CWS_O, Rooting
497   depth is more influential than Soil type. Gap size is not influential for either CWS.

498

499   *3.1.3 Eucalyptus globulus*

500   Figure 5 shows the indices of Sobol' for *E. globulus*.
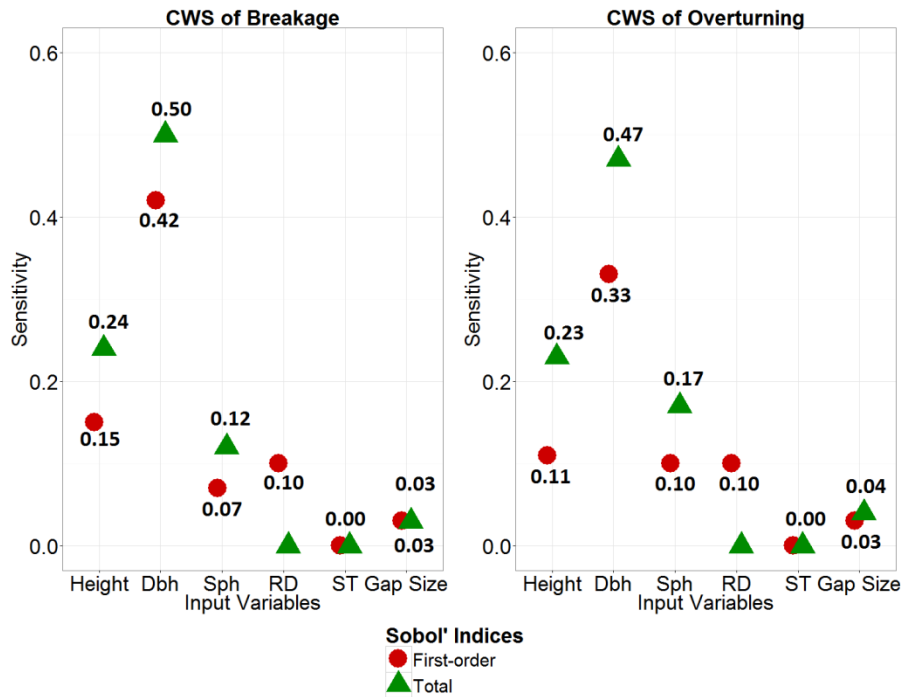

501

502   Figure 5: Sobol' first-order and total sensitivity indices for the critical wind speeds for breakage and overturning for *E.*
503   *globulus*. Dbh: diameter at breast height; Sph: stems per hectare. RD: rooting depth. ST: soil type. The total indices of RD,
504   and ST for CWS_B and for CWS_O are 0.00.

505

506   As seen in Figure 5, the most influential variables for *E. globulus* are, in decreasing order: *dbh*, Tree
507   height, and *Sph* for both CWS. The differences between the S^T_i and S_i seen for the other two species
508   are less pronounced for *E. globulus*, indicating somewhat smaller interactions. As for *P. pinaster*, *Sph*
509   is likely to participate in the interactions for *E. globulus*, as its S^T_i is slightly larger than its S_i,
510   especially for CWS_O. The S_i for Rooting depth for both $CWS_{(B,O)}$ are small but significant (0.10). For
511   CWS_B, this is larger than for *Sph*. For CWS_O, Rooting depth has the same S_i value of Tree height and
512   just smaller than that of *Sph* for CWS_O. This is partially attributable to the correlations between
513   Rooting depth and the influential variables Tree height and *dbh*, and perhaps to a numerical
514   imprecision in the estimation of Rooting depth's S_i. In fact, the C_{reg} values in the Rooting depth vs Soil
515   type matrix for *E. globulus* are all equal (i.e. Rooting depth and Soil type function as constants for *E.*
516   *globulus*). This is because tree-pulling for this species was only performed on one site with a
517   homogeneous soil, and no significant differences were found between C_{reg} values for different
518   rooting depths (Locatelli et al., 2016). This is reflected in the S_i values of Soil type for $CWS_{(B,O)}$ (0 in
519   both cases). Gap size is more influential for both CWS than for the previous two species.

520

3.1.4 Fixing uninfluential variables

Despite the moderate correlation between Rooting depth and Soil type (-0.35, see Table 4), we calculated the average approximation errors when the three uninfluential variables are fixed on a nominal value, using the method of Sobol' et al. (2007). This method has only been tested for non-correlated variables; therefore the error estimate might not be entirely accurate. The errors are shown in Table 5.

Table 5. Average approximation errors in the calculations of the critical wind speeds when fixing Rooting depth, Soil type, and Gap size

	P. sitchensis		*P. pinaster*		*E. globulus*	
	CWS_B	CWS_O	CWS_B	CWS_O	CWS_B	CWS_O
Rooting depth	0%	2%	0%	6%	0%	0%
Soil type	0%	4%	0%	6%	0%	0%
Gap size	0%	0%	0%	0%	6%	8%

For *P. sitchensis*, fixing Rooting depth or Soil type to a value within their range would result in an average approximation error of 2% and 4%, respectively, of CWS_O. If Rooting depth and Soil type were completely non-correlated, fixing both variables would therefore cause an approximation error of 6%. For *P. pinaster*, fixing Rooting depth and Soil type would result in a 6% average approximation error for CWS_O, and potentially a 12% cumulative error if both were fixed. Lastly, for *E. globulus*, as expected fixing Rooting depth and Soil type would have no average approximation error for $CWS_{(B,O)}$, while fixing Gap would result in a 6% and 8% errors for CWS_B and CWS_O, respectively. The lack of correlation between Gap size and the other input variables ensures the reliability of this estimate.

Because the S^T_i of Rooting depth, Soil type, and Gap size are practically zero for both CWS, we re-ran the simulations with ForestGALES using the same dataset apart from the values of these variables, which were fixed one at a time to different values within their range. For Rooting depth and Soil type, these are the discrete values in Table 3. For Gap size, we chose 0m, 2 times Tree height, 1000m, and 10 times Tree height. The latter was chosen because ForestGALES contains a trap in its code by which an upwind gap cannot be larger than that. We chose to fix Gap size to 2 times Tree height because Gardiner et al. (1997) have shown that it corresponds to the limit of the increase in bending moment coefficient with increasing Gap size. The scatterplots for Rooting depth and Soil type investigated one at a time are shown, for the three species and both CWS, in the Appendix. We anticipate here that the effect of these two variables on CWS_B was confirmed as null by the scatterplots shown in the Appendix. For this reason, in this section we limit ourselves to showing the scatterplots for Gap size alone, and for combinations of values of Rooting depth and Soil type, for Gap size fixed at 10 times Tree height. For Gap size, we show the scatterplots for both CWS because, despite not being numerically identified as influential by our GSA, it shows some influence on $CWS_{(B,O)}$. For Rooting depth vs Soil type, we show the effect of fixing them for CWS_O alone. Figures 6 and 7 show the Gap size scatterplots for $CWS_{(B,O)}$ for the three species*.
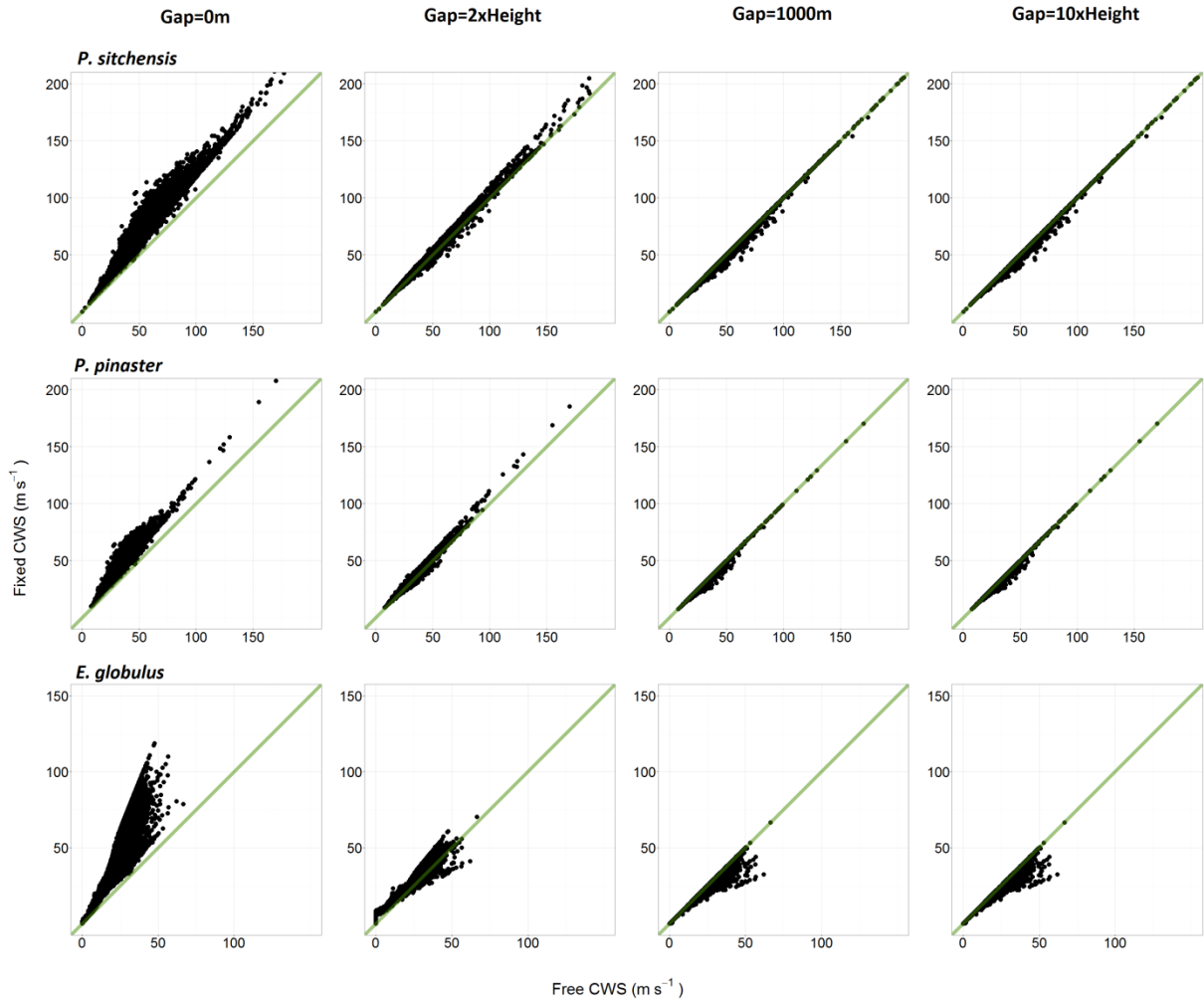
555

556

557

558

Figure 6: Scatterplots of critical wind speed for breakage for Gap size allowed to vary within its range vs Gap size fixed at four different values. Free: all variables allowed to vary within their ranges; Fixed: Gap size fixed at one of the following values: Gap=0m, Gap size fixed at 0m; Gap=2xHeight, Gap size fixed at 2 times Tree height; Gap=1000m, Gap size fixed at 1000m; Gap=10xHeight, Gap size fixed at 10 times Tree height. The green diagonal line represents the slope through origin [0; 0], i.e.a 1:1 relationship between the results.

565

Despite the S^{T}_{i} of Gap size being null for *P. sitchensis* and *P. pinaster*, the effect of fixing Gap is evident for all the species, especially when Gap size is fixed at 0m. Fixing Gap size to 0m results in a marked overestimation of CWS_B. Fixing Gap size to 2 times Tree height results in minor overestimation of CWS_B for *P. sitchensis* and *P. pinaster*, while the effect on *E. globulus* is more complex, with equal probabilities of over and underestimation of CWS_B. Fixing Gap size at 1000m or 10 times Tree height has the same result of underestimating CWS_B. The latter is particularly evident

572   for low to medium CWS_B, which correspond to the area of higher wind risk for a forest. The effect of
573   Gap size on CWS_B is more pronounced for *E. globulus*, which is consistent with the approximation
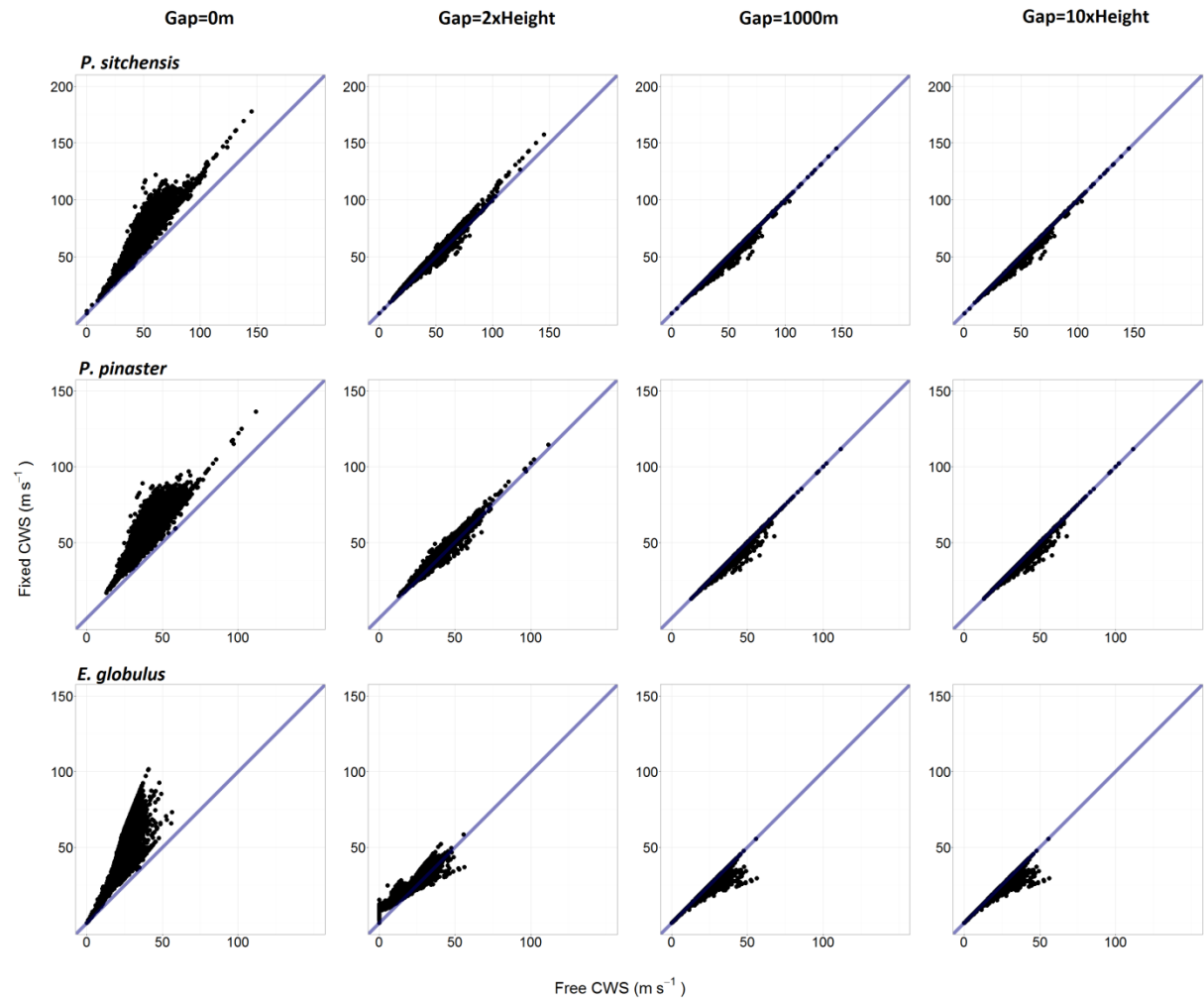574   error shown in Table 5.

575

576

577

578   Figure 7: Scatterplots of critical wind speed for overturning for Gap size allowed to vary within its range vs Gap size fixed at
579   four different values. Free: all variables allowed to vary within their ranges; Fixed: Gap size fixed at one of the following
580   values: Gap=0m, Gap size fixed at 0m; Gap=2xHeight, Gap size fixed at 2 times Tree height; Gap=1000m, Gap size fixed at
581   1000m, Gap=10xHeight: Gap size fixed at 10 times Tree height. The green diagonal line represents the slope through origin
582   [0; 0], i.e.a 1:1 relationship between the results.

583

584   The scatterplots in Figure 7 show the effect of fixing Gap size on CWS_O. The plots mirror those for
585   CWS_B shown in Figure 6, with the difference that the range of damaging wind speeds is smaller for
586   overturning than for breakage. For *E. globulus*, fixing Gap size at 2 times Tree height results in a
587   pattern even more complex for CWS_O than for CWS_B, as also very low values are affected. Next, we
588   present the results of fixing Rooting Depth and Soil type, one species at a time.

589

590   *P. sitchensis:*

591   Figure 8 shows the *P. sitchensis* CWS$_O$ scatterplots for the different levels of Rooting depth and Soil
592   type, with Gap size fixed at 10 times Tree height.
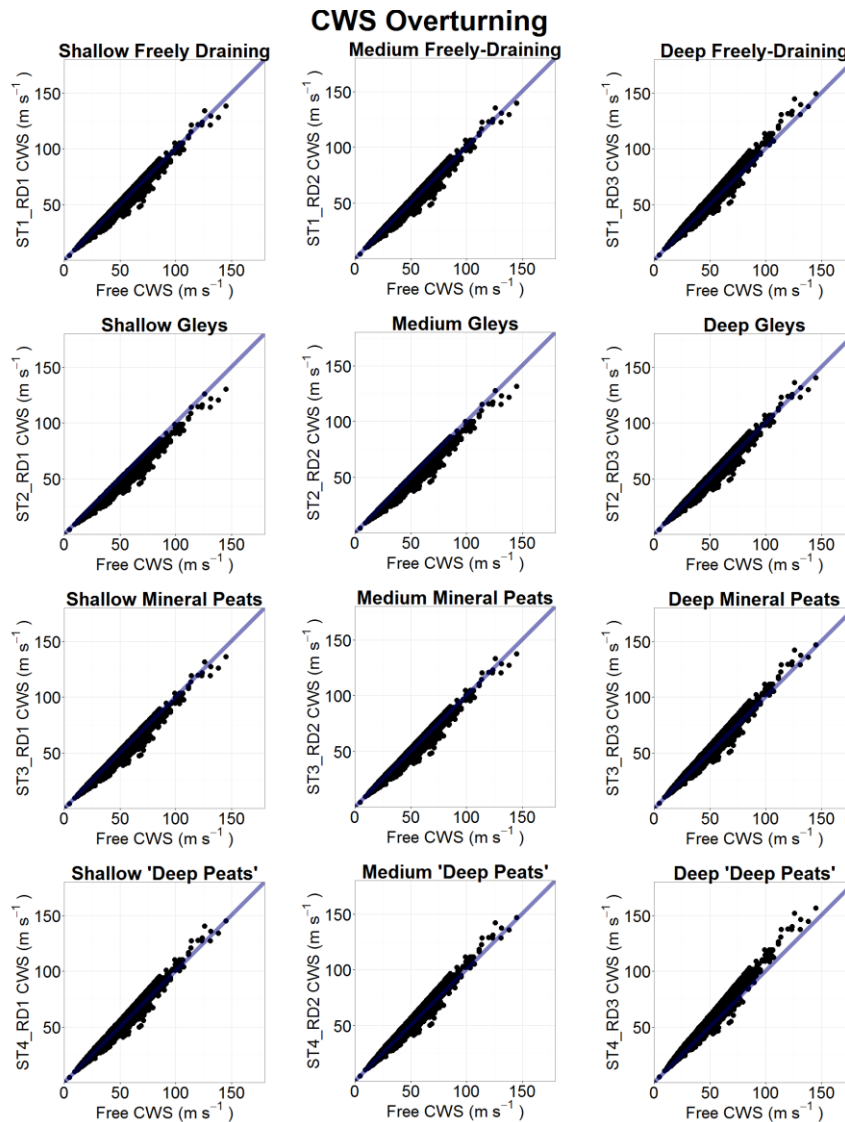
593

594

595

596   Figure 8: Scatterplots of critical wind speed for overturning for *P. sitchensis*. Simulations with variation in all the inputs vs
597   simulations where Rooting depth and Soil type were fixed at a value within their range. Gap size fixed at 10 times tree
598   height. The blue diagonal lines represent the slope through origin [0; 0], i.e. a 1:1 relationship between the results. "Free"
599   on the abscissa indicates simulations for which all variables were allowed to vary within their range. The coding for the
600   ordinate axis indicates the values at which Soil type (ST) and Rooting depth (RD) were fixed. E.g. "ST1_RD1" indicates Soil
601   type 1 (Freely draining) and Rooting depth 1 (Shallow).

602

603   Fixing Rooting depth at "deeply rooted" results in slight overestimations of the CWS$_O$ (i.e. the cloud
604   of points shifts upwards), regardless of Soil type. The same effect is obtained when fixing Soil type at
605   "Freely Draining" and "Deep Peats", regardless of Rooting depth. Fixing Soil type at "Gleys" and

606   "Mineral Peats" results in a slight underestimation of the CWS$_O$, apart from the already mentioned
607   case of deep Rooting depth. The shape of the clouds of points is partially affected by fixing Gap size
608   (see Figure 7).

609

610   *P. pinaster*

611   Figure 9 shows the scatterplots for CWS$_O$ for *P. pinaster*.
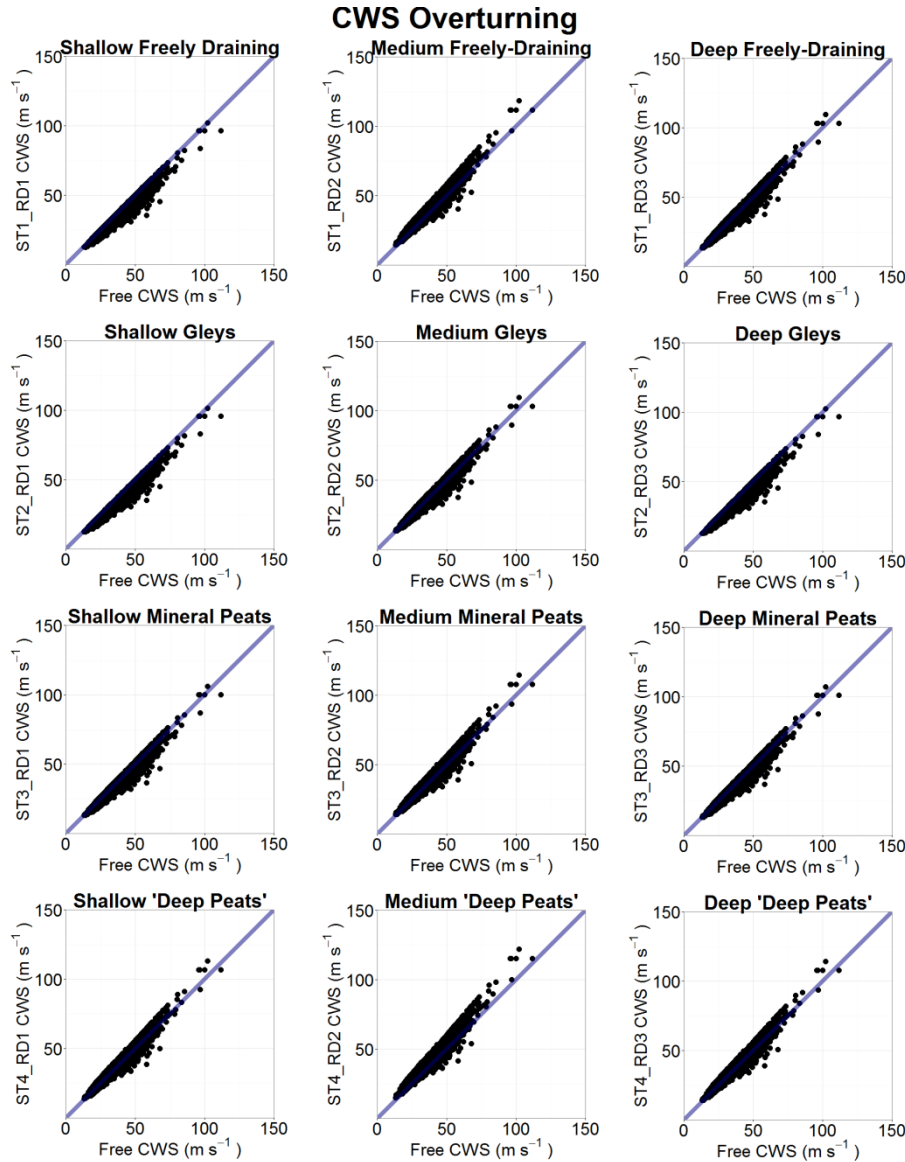
612

613

614   Figure 9: Scatterplots of critical wind speed for overturning for *P. pinaster*. Simulations with variation in all the inputs *vs*
615   simulations where Rooting depth and Soil type were fixed at a value within their range. Gap size fixed at 10 times tree
616   height. The blue diagonal lines represent the slope through origin [0; 0], i.e.a 1:1 relationship between the results. "Free"
617   on the abscissa indicates simulations for which all variables were allowed to vary within their range. The coding for the
618   ordinate axis indicates the values at which Soil type (ST) and Rooting depth (RD) were fixed. E.g. "ST1_RD1" indicates Soil
619   type 1 (Freely draining) and Rooting depth 1 (Shallow).

620

621 The scatterplots in Figure 9 show the effect of fixing Gap size, Rooting depth and Soil type on CWS_O

622 calculated for *P. pinaster*. As for *P. sitchensis*, the effect is mainly visible at medium to high CWS_O. In

623 fact, for both species the S^T_i and the average approximation error of Gap size (Table 5) are similar. As

624 shown in Figure 7, fixing Gap size to values other than 0m has the same effect as for *P. sitchensis*,

625 causing an underestimation of CWS_O. Fixing Rooting depth at "medium rooted" results in slight

626 overestimations of the CWS_O, regardless of Soil type. Simulations where Soil type was fixed at "Deep

627 Peats" show a slight overestimation of the CWS_O regardless of Rooting depth, while the simulations

628 where Soil type was fixed at "Gleys" show a slight underestimation of the CWS_O (excluding the case

629 of medium Rooting depth).

630

631 *E. globulus*

632 Figure 10 shows the scatterplots for CWS_O for *E. globulus*, comparing "Free" simulations with "Fixed"

633 simulations. Gap sized was fixed at 10 times Tree height, as for the previous species.

634
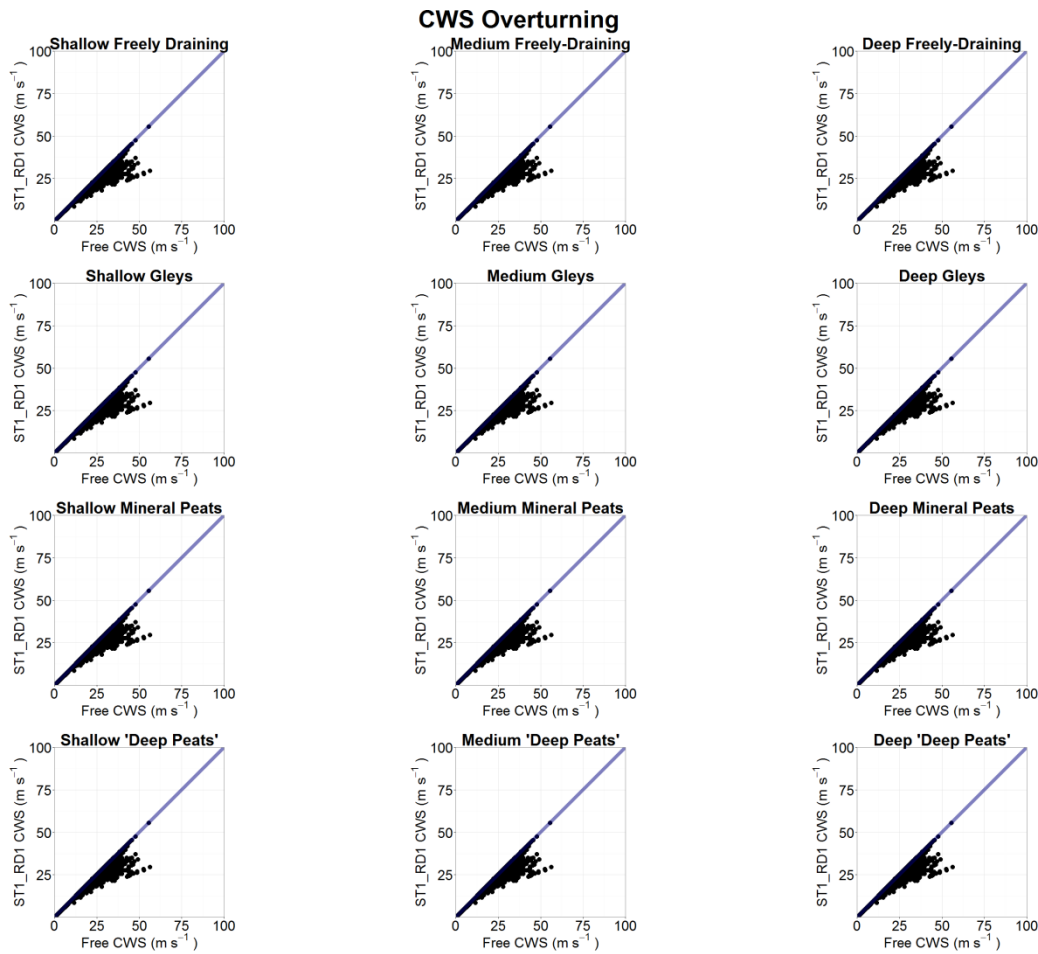
635

636 Figure 10: Scatterplots of critical wind speed for overturning for *E. globulus*. Simulations with variation in all the inputs *vs*

637 simulations where Rooting depth and Soil type were fixed at a value within their range. Gap size fixed at 10 times tree

638 height. The blue diagonal lines represent the slope through origin [0; 0], i.e. a 1:1 relationship between the results. "Free"

639 on the abscissa indicates simulations for which all variables were allowed to vary within their range. The coding for the
640 ordinate axis indicates the values at which Soil type (ST) and Rooting depth (RD) were fixed. E.g. "ST1_RD1" indicates Soil
641 type 1 (Freely draining) and Rooting depth 1 (Shallow).

642

643 The scatterplots in Figure 10 confirm that fixing Rooting depth and Soil type has no effect on CWS_O
644 calculated for *E. globulus*, as expected from the lack of variability in the soil parameters. In fact,
645 these plots are identical to those for overturning for different values of Gap size, as shown in Figure
646 5, as they are entirely driven by the error in fixing Gap size.

647

648 *3.2 Differences between species*

649 Table 6 shows the *p*-values of the comparisons between the coefficients of concordance for all the
650 species (T_{all}), and for pairs of species ($T_{\alpha,\beta}$), and a chi-square distribution with d.f. = 5. Small *p*-values
651 indicate that the null hypothesis of independence of the importance rankings of the input variables
652 between different species can be rejected, i.e. that the rankings of the most important variables are
653 similar across the three species. Conversely, large *p*-values indicate that the rankings are significantly
654 different between species. Rankings are based on S^T_i values.

655

656 Table 6: Significance of correlation between rankings of S^T_i of different species. The H_0 is of independence of rankings
657 between species. Low *p*-values suggest the H_0 should be rejected (i.e. high *p* -values suggest independence of S^T_i rankings
658 between species).

	CWS$_B$	CWS$_O$
All species	0.029	0.044
P. sitchensis / P. pinaster	0.087	0.448
P. sitchensis / E. globulus	0.247	0.128
P. pinaster / E. globulus	0.247	0.273

659

660 As shown in the first row of Table 6, the *p*-values for the comparison between all the species suggest
661 that the rankings of the influential variables for $CWS_{(B,O)}$ are similar for *P. sitchensis*, *P. pinaster*, and
662 *E. globulus*. In fact, as seen in Figures 3, 4, and 5, for the three species the top three ranks of the
663 total Sobol' indices are shared between Tree height, *dbh*, and *Sph*, although not in the same order
664 for all species and for both types of damage (breakage and overturning). The species-specific
665 differences in the ranking of these three variables are mostly responsible for the large *p*-values for
666 the pairwise comparisons. As shown by Iman and Conover (1987), the calculation of the top-down
667 coefficient of concordance amongst all species (C_T) is strongly driven by agreement between the top
668 ranks. This can be seen in the calculation of the Savage scores with Eq. (6). The importance of at
669 least a degree of accordance between the top three ranks in driving the calculation of the T_{all} index is
670 further corroborated by the numerator of Eq. (7): the more similar the top ranks are amongst
671 species, and the larger the number of species, the larger C_T (and therefore T_{all}) becomes. As a result,
672 the probability of the T_{all} index to be larger than the upper-tail critical value of a chi-square
673 distribution (in our case, with 5 degrees of freedom) diminishes as C_T and T_{all} increase.

674 With regards to the pairwise comparisons between our species, certain differences are evident. For
675 CWS$_B$, the ranking of the S^T_i of *E. globulus* is markedly different from the other two (Figures 3, 4, and
676 5). This is reflected by the *p*-values in Table 6 for the comparisons that include *E. globulus* being the
677 largest (0.247), while the *p*-value of the comparison between *P. sitchensis* and *P. pinaster* (0.087)
678 indicates that differences between these two species are almost non-significant at the 95%
679 confidence interval. It should be noted that the fact that the S^T_i rankings of these two species are
680 identical (Figures 3 and 4) suggests that the accuracy of our statistical test is not perfect. For CWS$_B$,
681 the large *p*-values (0.237 and 0.246) for the comparisons between *P. sitchensis* / *E. globulus*, and *P.*
682 *pinaster* / *E. globulus*, are driven by the fact that the ranking of the top two variables, Tree height
683 and *dbh*, are inverted between the pair (*P. sitchensis*, *P. pinaster*), and *E. globulus*. In fact, the
684 difference between the Savage scores calculated for the top rank, and the second or the third rank,
685 are quite significant (top rank: 2.45; second rank: 1.45; third rank: 0.95), while further ranks,
686 especially in the case of ties (e.g. for Rooting Depth and Soil type) have similar low values, below 0.5.
687 Therefore, in pairwise comparisons a small disagreement at the top three ranks can result in the S^T_i
688 rankings of the 2 species being flagged as substantially different. This is unlike in the calculations of
689 C_T and T_{all}, for which partial agreements between the top S^T_i rankings of the three species
690 contributes to the degree of similarity between all three species. Similarly, the small, but non-zero S^T_i
691 of Gap for *E. globulus* contributes marginally to these pairwise differences that involve *E. globulus*,
692 as its rank is higher than for the other two species. For CWS$_O$, the difference between *P. sitchensis*
693 and *P. pinaster* (*p*-value 0.448) is due to the different ranks of the top three variables (Tree height,
694 *dbh*, and *Sph*). For the *P. pinaster* / *E. globulus* pair, the large *p*-value (0.273) is due to the rankings of
695 Tree height, *dbh*, and Gap size. Similarly, the difference between *P. sitchensis* and *E. globulus* (*p*-
696 value 0.128) is attributable to the rankings of Tree height, *Sph*, and Gap size.

697

698 *3.3 Probabilities of damage – Sensitivity in the wind climate module*

699 As stated in the Methods section, we have set the threshold for the probabilities of damage at 10%
700 (i.e. a probability of 0.1), to differentiate between behavioural (Prob$_{(B,O)}$ < 0.1) and non-behavioural
701 (Prob$_{(B,O)}$ > 0.1) regions. The CDF plot and the 2D density plots throughout this section are based on
702 this categorisation. In the CDF plot, the values of the D-statistic represent the maximum distance
703 between the CDF curves: the larger this statistic is, the further apart the CDF curves are, suggesting
704 that the variable's importance in differentiating between behavioural and non-behavioural
705 realisations of ForestGALES is large We show only the first CDF plot, for Prob$_B$ for *P. sitchensis*, to
706 illustrate the interpretation of the Smirnov test and its D-statistic. The other CDF plots are not
707 shown, while the values of the D-statistic are summarised in Table 7, which will be referenced
708 throughout this section.

709

710 *3.3.1 Picea sitchensis*

711 Figure 11 shows the CDF plots for *P. sitchensis* for Prob$_B$, for all the ForestGALES input variables.
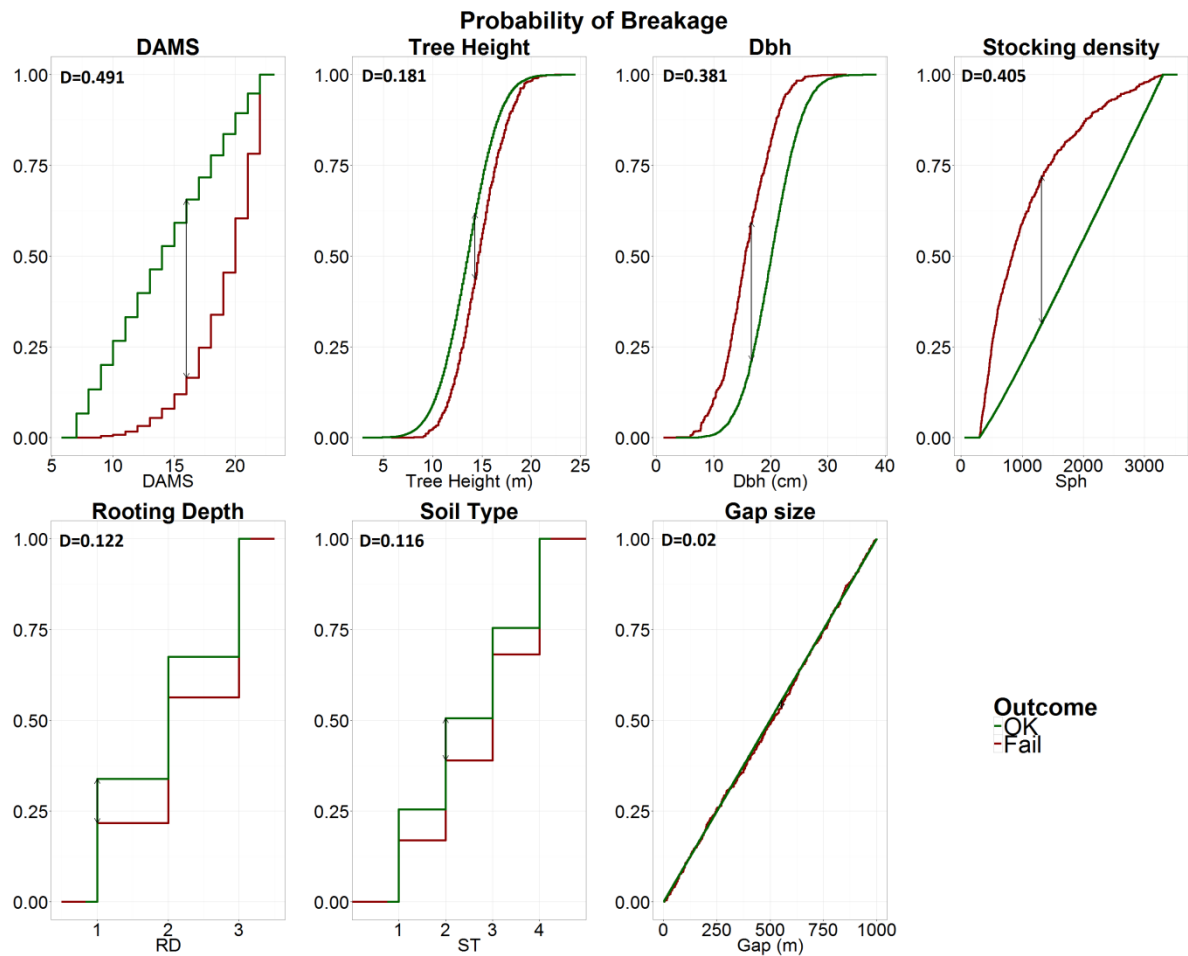
Figure 11: Cumulative density function plots of the probability of breakage for *P. sitchensis*. DAMS: Detailed Aspect Methods of Scoring, a measure of the windiness of the site. Dbh: diameter at breast height; Sph: stems per hectare. The green lines represent model realisation in the behavioural region (probability < 0.1). The red lines represent model realisations in the non-behavioural region (probability > 0.1). The arrows indicate the maximum distance between the cumulative distributions. Large distances indicate that the variable is influential, as described by the values of the D-statistic.

Figure 11 shows that, for *P. sitchensis*, the main drivers of the realisations of ForestGALES in the behavioural and non-behavioural regions of $Prob_B$ are *dbh*, DAMS, and *Sph*, followed by Tree Height, while the other variables contribute only marginally. Rooting depth and Soil type show some importance, which is attributable to their correlation with *dbh* and *Sph*, respectively (see Table 4). The results for $Prob_O$ are shown in Table 7.

Table 7: Values of the D-statistic of the Smirnov two-sample (two sided) tests for probabilities of damage (breakage, $Prob_B$; overturning, $Prob_O$). The larger the value of the D-statistic, the more influential a variable is on differentiating between the behavioural (probability of damage < 10%) and non-behavioural (> 10%) realisations of our ForestGALES simulations.

	P. sitchensis		*P. pinaster*		*E. globulus*	
	$Prob_B$	$Prob_O$	$Prob_B$	$Prob_O$	$Prob_B$	$Prob_O$
DAMS	0.491	0.563	0.584	0.642	0.572	0.648
Tree height	0.181	0.107	0.256	0.321	0.157	0.089
Dbh	0.381	0.392	0.081	0.099	0.294	0.2
Sph	0.405	0.493	0.27	0.373	0.105	0.114
Rooting Depth	0.122	0.087	0.126	0.232	0.135	0.122
Soil Type	0.116	0.113	0.07	0.064	0.017	0.019
Gap size	0.02	0.032	0.014	0.022	0.033	0.041

As shown in Table 7, for *P. sitchensis*, the variables driving the variation in $Prob_O$ are similar to those of $Prob_B$, with the only difference being that Tree Height is less important than Soil Type. The influence of Rooting depth is smaller than for $Prob_B$, and that of Soil type is very similar for the two probabilities of damage. These results mirror the relative differences between the respective S^T_i for CWS_B and CWS_O, shown in Figure 3.

In order to investigate the two-way interactions between influential variables for $Prob_{(B,O)}$, we used 2D-density plots. We also include Tree height, despite it being flagged as marginally influential for $Prob_{(B,O)}$, because of its high values of S^T_i and S_i. Figures 12 and 13 show these two-way interactions for *P. sitchensis* for $Prob_B$ and $Prob_O$, respectively. The first three plots (*a* to *c*) show the interactions between DAMS and one of Tree height, *dbh*, and *Sph*. The last three plots (*d* to *f*) show the interactions between these three variables. Irregular shapes are due to the pattern of pseudo-random numbers generated with the Quasi-Monte Carlo method of Sobol' (1990, 1998).
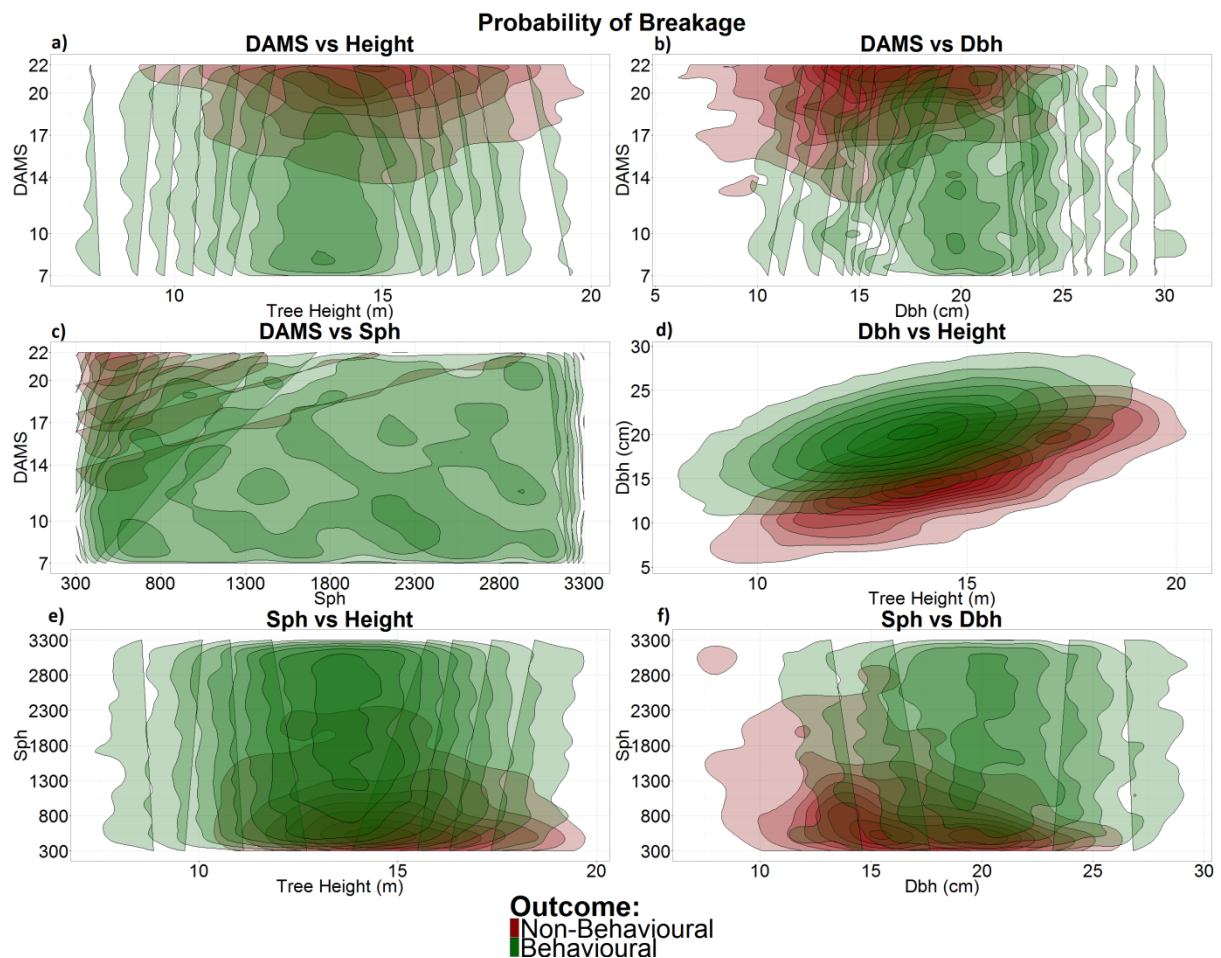
Figure 12: 2D – density plots for probability of breakage for *P. sitchensis*. DAMS: Detailed Aspect Method of Scoring, a measure of the windiness of a site; Dbh: diameter at breast height; Sph: stems per hectare. The green areas represent the "Behavioural" realisations of ForestGALES that correspond to a probability of damage < 0.1. The red areas correspond to "non-behavioural" realisations, i.e. probabilities > 0.1. The intensity of the colour corresponds to the density of the outputs: darker areas have higher densities.
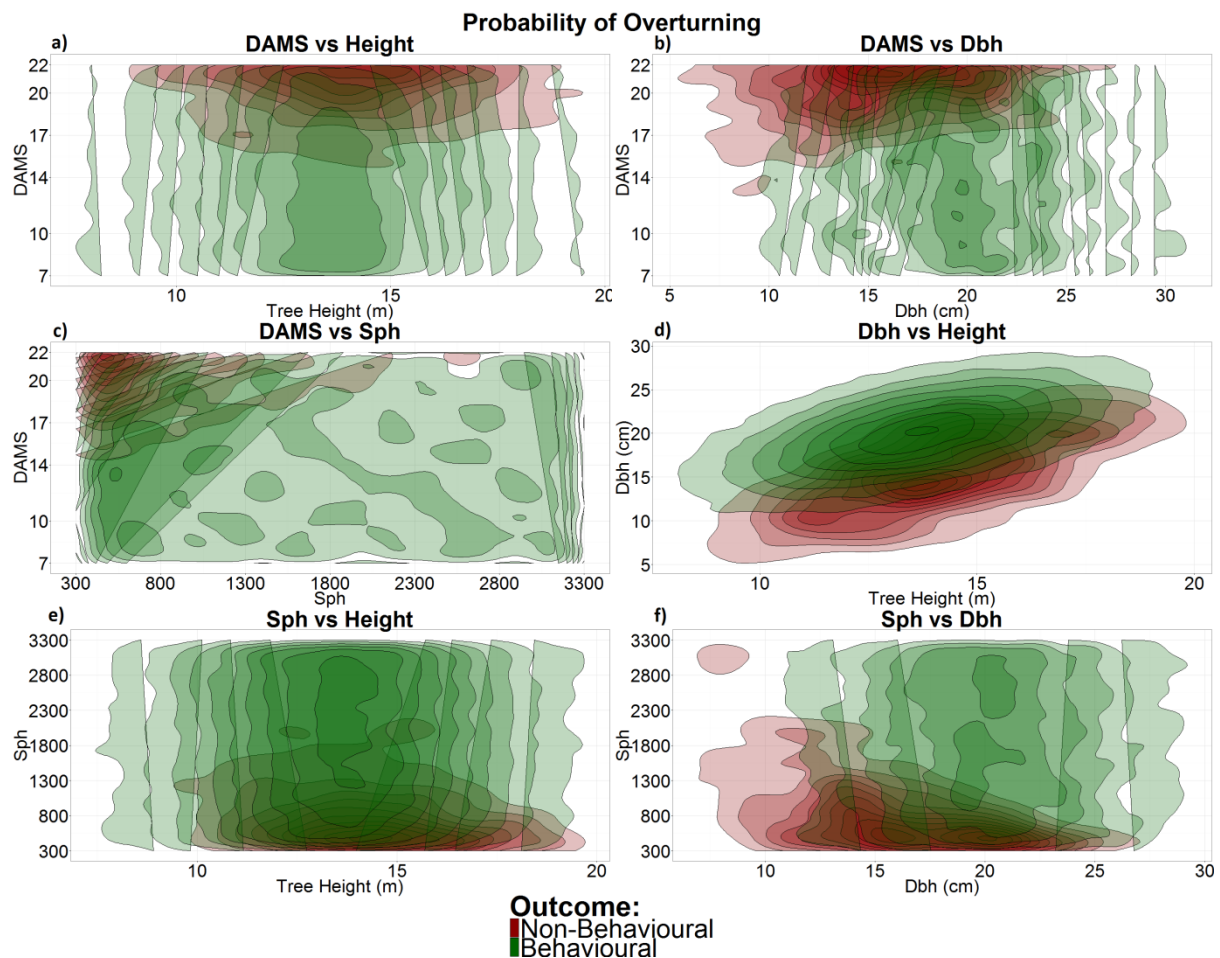
Probability of Overturning

Figure 13: 2D – density plots for probability of overturning for *P. sitchensis*. DAMS: Detailed Aspect Method of Scoring, a measure of the windiness of a site; Dbh: diameter at breast height; Sph: stems per hectare. The green areas represent the "Behavioural" realisations of ForestGALES that correspond to a probability of damage < 0.1. The red areas correspond to "Non-behavioural" realisations, i.e. probabilities > 0.1. The intensity of the colour corresponds to the density of the outputs: darker areas have higher densities.

The trends in Figures 12 and 13 are very similar. Most of the plots show large areas where either damage or no damage can result for the same combinations of the variables in the plots (i.e. red and green areas overlap), suggesting that bivariate interactions are not sufficient to discriminate between behavioural (Prob$_{(B,O)}$ < 0.1) and non-behavioural (Prob$_{(B,O)}$ > 0.1) realisations of ForestGALES. The patterns of green and red areas in both Figures show that for *P. sitchensis* ForestGALES predicts damage for medium-sized trees for exposed sites (DAMS over 12, graphs *a* and *b* in Figures 12 and 13), while shorter trees and trees with a large *dbh* are at less risk of damage. The model predicts that trees with *dbh* above ~25cm will be safe from damage, regardless of the severity of the wind, while trees of small diameters will be prone to damage even at low DAMS (Figures 12*b* and 13*b*). The DAMS vs *Sph* plots show that ForestGALES predicts that the stands most at risk are those of low stocking densities, although areas of higher probabilities of damage are present for other stocking densities (Figure 13*c*). The *dbh* vs Height plots show that the model predicts higher probabilities of damage to trees with small *dbh*, especially for short trees. The *Sph* vs Height plots indicate that, regardless of the height of the trees, ForestGALES predicts more damage to stands with low stocking densities, and higher probabilities of damage for trees taller than 10m. As for the

784   interaction between *Sph* and *dbh*, ForestGALES predicts more damage to trees of small diameter,
785   especially for stands of low to medium stocking densities (Figures 12*f* and 13*f*).

786

787   *3.3.2 Pinus pinaster*

788   As shown in Table 7,for *P. pinaster* the most influential variables with regards to $Prob_B$ are DAMS,
789   *Sph*, and Tree height, while the other variables contribute little to the variation in the output. As for
790   *P. sitchensis*, Rooting depth and Soil type show minimal importance, likely attributable to their
791   correlation with Tree height and *Sph*, respectively, as shown in Table 4.

792   The main drivers of variation of $Prob_O$ for *P. pinaster* are the same as those of $Prob_B$ (in decreasing
793   order: DAMS, *Sph*, Tree height). The value of the D-statistic of Rooting depth is higher for $Prob_O$
794   (0.232) than for $Prob_B$ (0.126), which reflects the variable's higher S_i for CWS_O than for CWS_B, as seen
795   in Figure 4. In Figures 14 and 15 we show the two-way interactions for *P. pinaster* for the
796   probabilities of breakage and overturning, respectively.
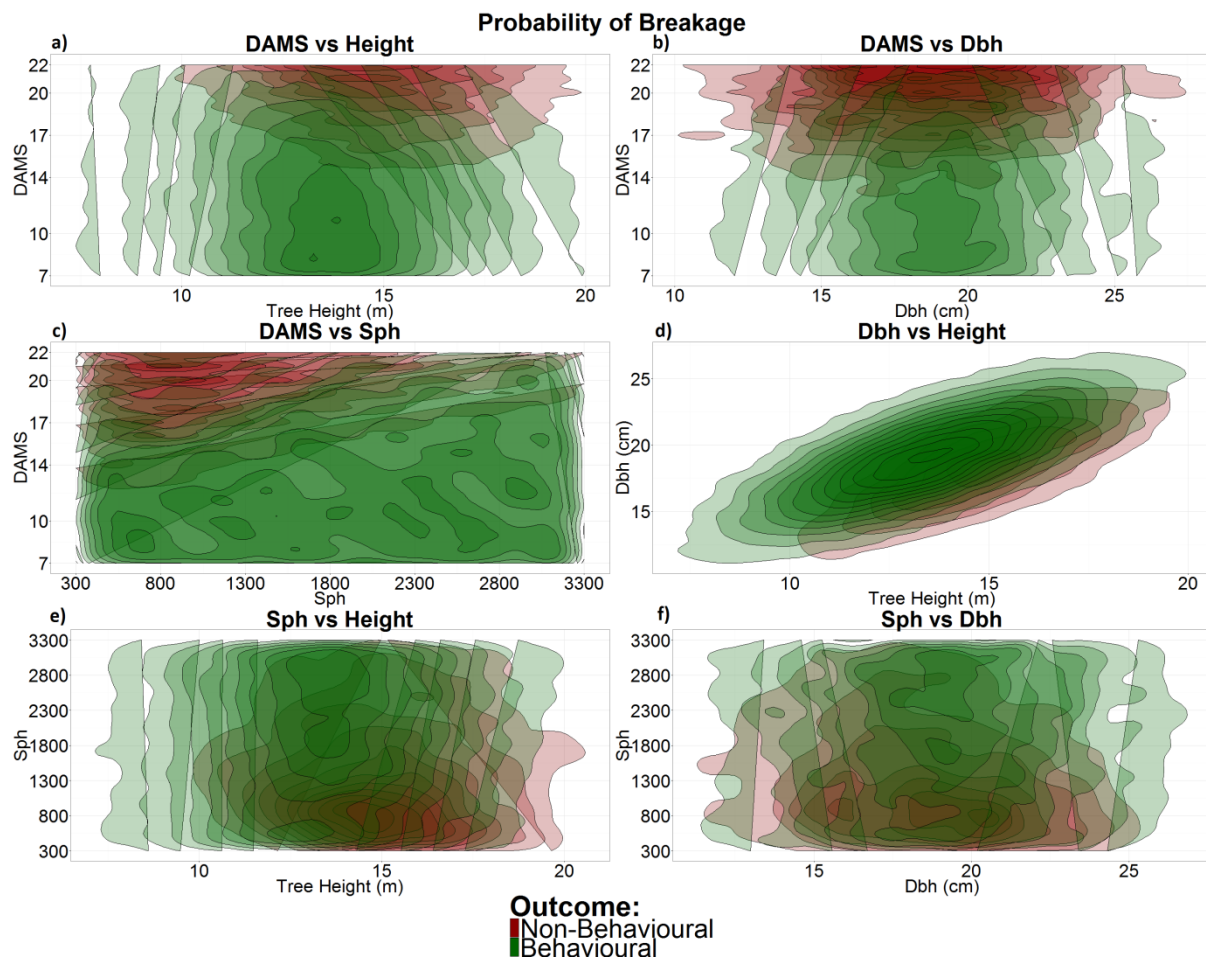
797

798

799

800

801

Figure 14: 2D – density plots for probability of breakage for *P. pinaster*. DAMS: Detailed Aspect Method of Scoring, a measure of the windiness of a site; Dbh: diameter at breast height; Sph: stems per hectare. The green areas represent the "Behavioural" realisations of ForestGALES that correspond to a probability of damage < 0.1. The red areas correspond to "Non-behavioural" realisations, i.e. probabilities > 0.1. The intensity of the colour corresponds to the density of the outputs: darker areas have higher densities.

The patterns of green and red areas in Figure 14 show that, with regards to the interaction between DAMS and Tree height, ForestGALES predicts that short (<10m) *P. pinaster* trees are safe from breakage, and that DAMS lower than 14 are generally safe (Figure 14*a*). In fact, below this DAMS values, non-behavioural realisations of ForestGALES ($Prob_{(B,O)}$ > 0.1, shown in red), are absent. This threshold is slightly lower (DAMS =13) for *dbh*, whereby larger trees (*dbh* > 24cm) are at low risk of breakage for DAMS up to 20 (Figure 14*b*). As for *P. sitchensis*, low stocking densities are associated with higher probabilities of damage, although ForestGALES predicts damage to *P. pinaster* stands of high densities (up to 3,300 *Sph*) for DAMS as low as 17 (Figure 14*c*). With regards to the interaction between *dbh* and Tree height, ForestGALES predicts that tall trees (height > 11m) are as likely to survive as they are to break, regardless of their *dbh*, while short trees with small to medium *dbh*, and tall trees with large *dbh* are less likely to break (Figure 14*d*). The *Sph* vs Tree height plot shows that short *P. pinaster* trees are at lower risk of breakage regardless of the stocking density of the stand, while tall trees are exposed to a higher risk for low *Sph* (Figure 14*e*). With regards to the interaction between *Sph* and *dbh*, areas of high probability of breakage are widespread in the plot (Figure 14*f*).

823 However, ForestGALES predicts that large trees at high stocking densities have a lower probability of
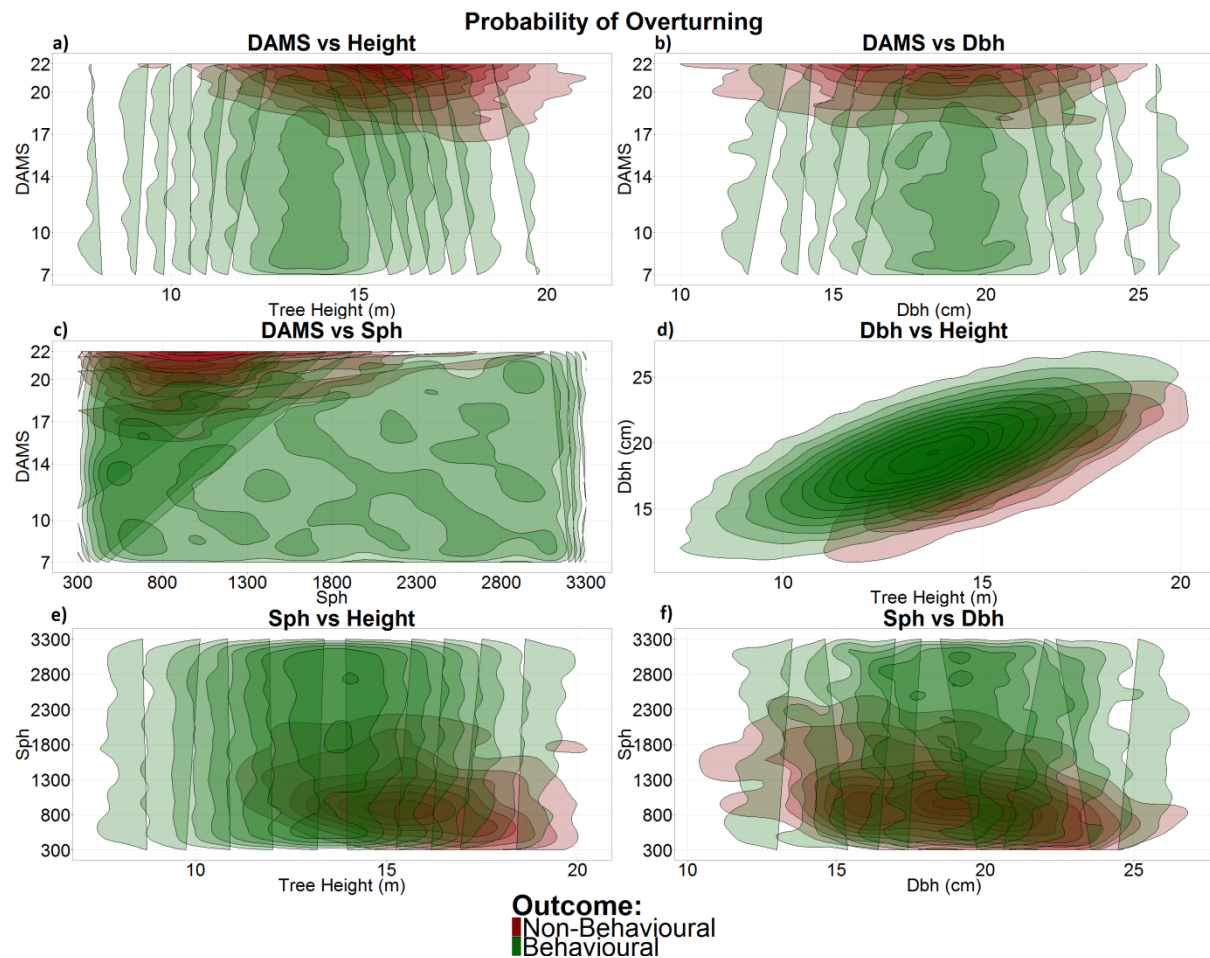824 damage.

Figure 15: 2D – density plots for probability of overturning for *P. pinaster*. DAMS: Detailed Aspect Method of Scoring, a measure of the windiness of a site; Dbh: diameter at breast height; Sph: stems per hectare. The green areas represent the "Behavioural" realisations of ForestGALES that correspond to a probability of damage < 0.1. The red areas correspond to "Non-behavioural" realisations, i.e. probabilities > 0.1. The intensity of the colour corresponds to the density of the outputs: darker areas have higher densities.

832 The DAMS plots in Figure 15 are considerably clearer than those for Prob_B shown in Figure 14, as the
833 areas that describe high and low probabilities of damage (red and green areas, respectively) are
834 more distinct than those in Figure 14. The plots that show the interactions between Tree height,
835 *dbh*, and *Sph*, are similar to those for Prob_B. With regards to overturning, ForestGALES predicts that
836 *P. pinaster* trees above a height of ~10m are likely to fail for DAMS > 17 (Figure 15*a*). Similarly,
837 ForestGALES predicts that *P. pinaster* trees will overturn when exposed to a wind climate
838 corresponding to DAMS > 17, regardless of their *dbh* (Figure 15*b*). The DAMS vs *Sph* plot shows that
839 low stocking densities are at higher risk of overturning for DAMS as low as 11, while DAMS of 20 and
840 above are required to overturn trees in very dense stands. The d*bh* vs Tree height plot (figure 15*d*) is
841 almost identical to that in Figure 14*d*, with trees of height larger than 11m being as likely to survive
842 as to uproot, while short trees are less likely to uproot, regardless of their *dbh*. The interactions
843 between *Sph* and Tree height, and *Sph* and *dbh*, show that higher stocking densities are associated

844   with lower probabilities of overturning, regardless of the height or diameter of the trees. However,

845   short trees (height < 11m) show higher probabilities of survival also at low stocking densities.

846

847   *3.3.3 Eucalyptus globulus*

848   As shown in Table 7, for *E. globulus* DAMS is the most important variable for the behavioural and

849   non-behavioural realisations of ForestGALES for $Prob_B$, followed by *dbh* and Tree height. For $Prob_O$,

850   DAMS and *dbh*, and marginally *Sph*, are the most important variables to differentiate between the

851   behavioural and non-behavioural regions of the output space. The correlations shown in Table 4

852   between Rooting depth and these two variables are responsible for the relatively high value of the

853   D-statistic for Rooting depth, and for the non-zero value of that for Soil type, for both $Prob_{(B,O)}$.

854   Figures 16 and 17 show the two-way interactions for *E. globulus* for the probabilities of breakage

855   and overturning, respectively.
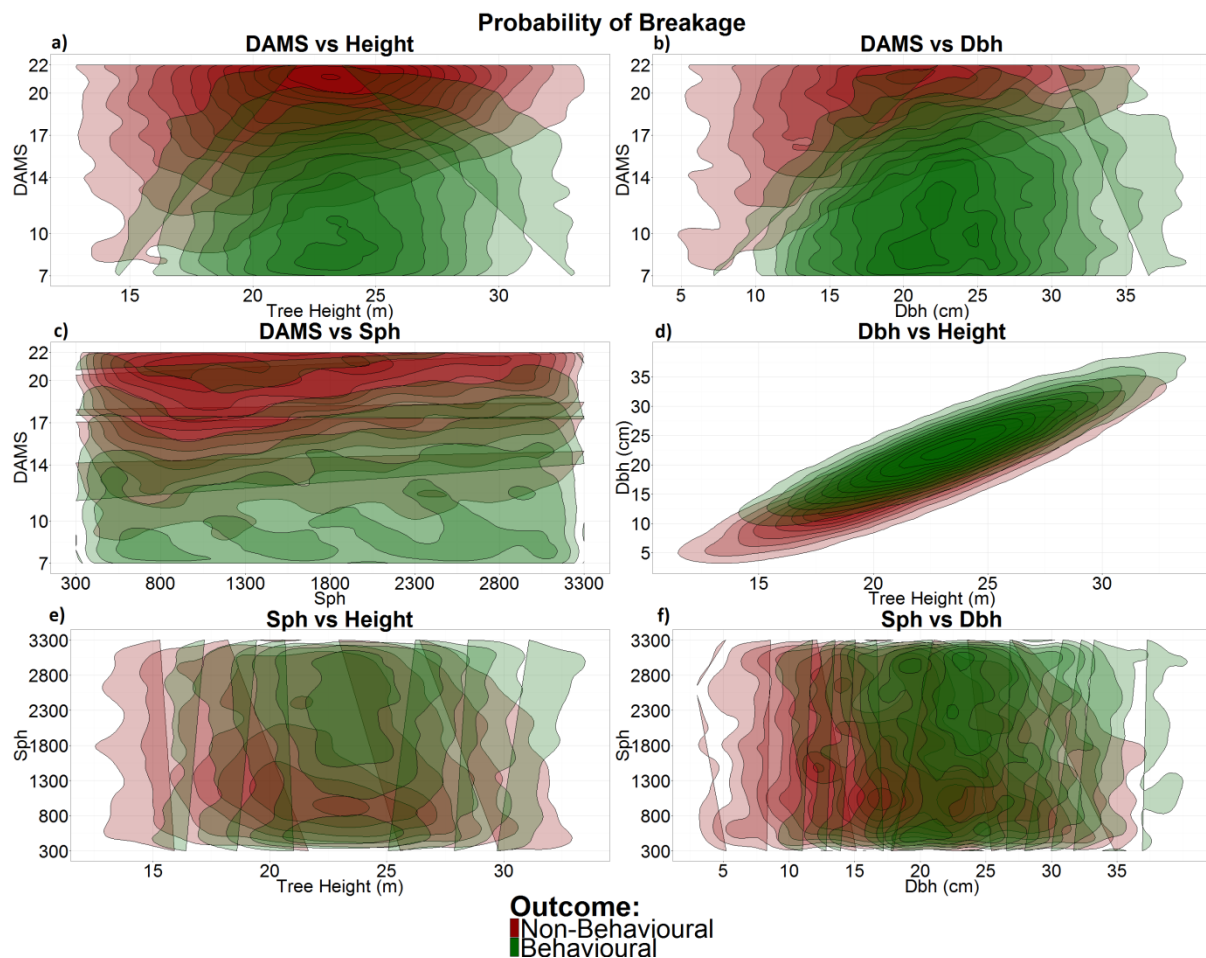
856

857

858

859

860

861

862

Figure 16: 2D – density plots for probability of breakage for *E. globulus*. DAMS: Detailed Aspect Method of Scoring, a measure of the windiness of a site; Dbh: diameter at breast height; Sph: stems per hectare. The green areas represent the "Behavioural" realisations of ForestGALES that correspond to a probability of damage < 0.1. The red areas correspond to "Non-behavioural" realisations, i.e. probabilities > 0.1. The intensity of the colour corresponds to the density of the outputs: darker areas have higher densities.
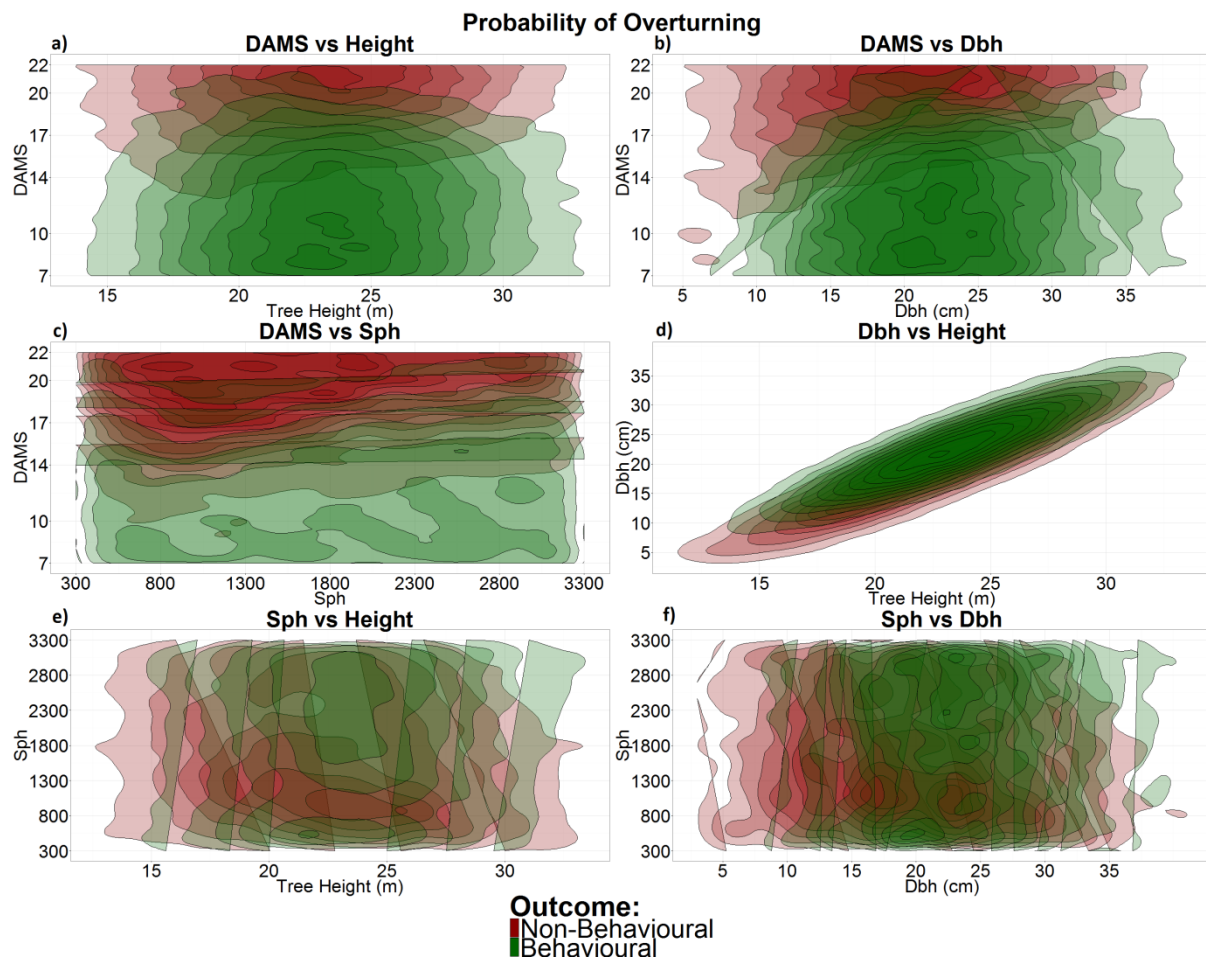
Figure 17: 2D – density plots for probability of overturning for *E. globulus*. DAMS: Detailed Aspect Method of Scoring, a measure of the windiness of a site; Dbh: diameter at breast height; Sph: stems per hectare. The green areas represent the "Behavioural" realisations of ForestGALES that correspond to a probability of damage < 0.1. The red areas correspond to "Non-behavioural" realisations, i.e. probabilities > 0.1. The intensity of the colour corresponds to the density of the outputs: darker areas have higher densities.

For *E. globulus*, the two-way interactions between DAMS, Tree height, *dbh*, and *Sph*, are almost identical for the probabilities of breakage and overturning. For breakage, short trees are at risk of failing for DAMS as low as 8 (Figure 16*a*), while for overturning *E. globulus* trees are at low risk regardless of their height, for DAMS lower than 14 (Figure 17*a*). Similarly, trees with a small *dbh* (< 10cm) are at risk of both breakage and overturning, while as *dbh* increases the windiness required to damage a tree increases linearly, with trees of *dbh* > 30cm being at risk of damage only for DAMS > 15 (figures 16*b* and 17*b*). For the interaction between DAMS and *Sph*, stands of low stocking densities are more prone to damage. However, especially for breakage (Figure 16*c*), ForestGALES predicts that low values of DAMS (between 10 and 14) can result in damage to a stand regardless of its stocking density. The *dbh* vs Tree height plots in Figures 16*d* and 17*d* show that for *E. globulus* ForestGALES cannot discriminate between high and low risk solely on the basis of these two variables, as the density areas overlap almost entirely. However, short trees with small *dbh* are predicted to be at higher risk of damage. The last two plots (*e* and *f*) in Figures 16 and 17 show that short trees, and trees with small *dbh*, are at higher risk of damage regardless of the stocking density.

901 Tall trees (>30m) and trees with large *dbh* (>30cm) are at lower risk when associated with high
902 stocking densities.

903

904 **4. Discussion**

905 In this study we have performed a variance-based sensitivity analysis (SA) on the forest wind-risk
906 model ForestGALES (Hale et al., 2015). We have used the method of Kucherenko et al. (2012), a
907 generalisation of the Sobol' method (Sobol', 2001) for the case of correlated variables. To provide
908 wide silvicultural and geographical applicability of our results, we have performed our analysis on
909 the performance of ForestGALES for three tree species, representative of three of the most
910 extensively planted and highly productive tree genera in the world: spruces (*P. sitchensis*), pines (*P.*
911 *pinaster*), and eucalypts (*E. globulus*). We have focussed our sensitivity analysis only on the model
912 input variables that are modifiable by the end-users in order to contextualise our results for practical
913 applications of the model, as well as for the forest wind-risk modelling community. A number of
914 settings are available when performing variance-based SA. These settings make use of different
915 results of the SA, and provide information on different processes within the architecture of a model.
916 In this study we have focussed on three SA settings: Factor Prioritisation (FP), Factor Fixing (FF), and
917 Factor Mapping (FM). ForestGALES provides two pairs of outputs: the critical wind speeds for
918 breakage and overturning ($CWS_{(B,O)}$), and the associated probabilities of damage ($Prob_{(B,O)}$). In this
919 section we first discuss the FP and FF settings that were applied to the $CWS_{(B,O)}$, which we
920 complement with a discussion of the similarities between species in the ranking of the most
921 influential input variables, as identified with the FF setting. We then follow with a discussion of the
922 results of the FM setting which was applied to the $Prob_{(B,O)}$. We conclude this section with an
923 evaluation of the performance of the Sobol' method for correlated variables applied to our study.

924

925 *4.1 Critical wind speeds – Factor Prioritisation setting*

926 The aim of the FP setting is to identify the variables with the highest first-order sensitivity indices
927 (S_i), not taking into account any interactions in the model between the variables. The identification
928 of the variables with the highest S_i values allows optimising the resources required for the
929 acquisition of accurate data for model execution. In fact, high S_i values highlight the variables that, if
930 the uncertainty associated with their measurement or collection is reduced the most, will cause the
931 largest reduction in the uncertainty of the outputs. The two most important variables identified by
932 our SA are Tree height and *Sph*, regardless of tree species. For *P. sitchensis* differences exist in the
933 ranking of these variables between CWS_B and CWS_O, with *Sph* being largely more important for the
934 latter than Tree height. Gardiner et al. (1997) have shown that low *Sph* results in increased wind
935 loading on a tree, promoting overturning over breakage as type of damage. For *P. sitchensis* and
936 particularly for *E. globulus*, *dbh* is also influential. For *P. pinaster*, *dbh* is completely uninfluential,
937 while variation in Rooting depth contributes marginally to the variation in CWS_O.

938 Recent advancements in the field of remote sensing can help with fast and cost-effective forest
939 mensuration (McInerney et al., 2011; Rosette et al., 2011). For large applications of the model (i.e. in
940 the Capsis software platform used in France, Dufour-Kowalski et al., 2012), knowledge of species

941 geographical distributions within the area of interest is recommended. However, in the case of
942 mixed-species stands the similarities between the species most influential variables provide some
943 confidence that an average level of measurement accuracy across the variables would ensure that
944 estimates of $CWS_{(B,O)}$ for large-scale investigations of vulnerability are reliable.

945 The species differences in the FP setting results pose some questions on the inner workings of
946 ForestGALES for different species. For *P. sitchensis*, the influence of Tree height is larger than that of
947 *dbh* for CWS_B, while the opposite is true for CWS_O. This is surprising since, as seen in Eq. (1) and (2),
948 dbh^3 is involved in the calculation of CWS_B, while CWS_O is calculated with stem weight, of which Tree
949 height * dbh^2 is a good approximation, as shown by Gardiner et al. (1997). However, as shown in
950 Table 2 Tree height is involved in a large number of components of CWS_B, and actually one more
951 than for CWS_O. The FP setting result of *dbh* being uninfluential for *P. pinaster* is also surprising. We
952 ascribe this to the fact that the variance of our sample data for *dbh* was the smallest of the three
953 species (see Table 3). Conversely, the variance of *dbh* was largest for *E. globulus*, and its S_i for both
954 CWS is the largest for this species. With regards to *E. globulus*, it is interesting to note that our SA
955 identified some contribution of Gap size to the variance of $CWS_{(B,O)}$, while for the other two species
956 this variable was completely uninfluential. We ascribe this to the fact that in our *E. globulus* sample
957 there is no variation of Rooting depth and Soil type, which therefore cannot outweigh the influence
958 of Gap size, as it is likely to be the case for the other two species. While it is true that Rooting
959 depth's S_i for *E. globulus* is quite high (0.10, Figure 5), its S^T_i is zero. Given that there is no variation in
960 Rooting depth and Soil type for *E. globulus*, the non-zero S_i value of Rooting depth is necessarily a
961 mistake in the numerical estimation of the sensitivity indices. Rooting depth is more influential for *P.*
962 *pinaster* than for *P. sitchensis* (S_i values of 0.13 and 0.01 for CWS_O, Figures 4 and 3, respectively). The
963 likely reason for this is the larger variation in the C_{reg} values for *P. pinaster* than for *P. sitchensis*
964 (Table 1). While *P. pinaster*'s C_{reg} values for three soil types (Gleys, Mineral peats, and Deep peats)
965 are actually taken from tree-pulling on *P. sylvestris* (in Nicoll et al., 2006), the largest variation can be
966 seen in the *P. pinaster* bespoke tree-pulling experiments (Cucchi et al., 2004).

967

968 *4.2 Critical wind speeds – Factor Fixing setting*

969 The FF setting is based on the total sensitivity indices (S^T_i) and provides information on the
970 interactions within the model between variables, and on which variables can be confidently fixed at
971 any value within their range without significantly affecting the predictive potential of the model. We
972 first discuss the interactions in ForestGALES, before discussing the issue with fixing variables.

973 A large difference between a variable's S_i and S^T_i indicate that the variable is involved in a large
974 number of interactions. As expected from Eq. (1) and (2) and Table 2, the largest interactions are
975 expected to be found between Tree height and *dbh*, with *Sph* contributing in a much smaller
976 measure. In fact, our results show that the other variables are not involved in significant
977 interactions, as their S^T_i are basically zero. The largest interactions between Tree height and *dbh* are
978 found for *P. pinaster*, with *Sph* marginally involved (Figure 4). Based on the Sobol' indices in Figure 3,
979 for *P. sitchensis* the interactions between Tree height and *dbh* are of similar magnitude to those of *P.*
980 *pinaster*, while *Sph* is not involved in interactions, with CWS_B showing larger interactions than CWS_O.
981 For *E. globulus*, the interactions between Tree height and *dbh* are much smaller than for the other
982 species, while those with *Sph* are of similar magnitude to *P. pinaster*.

983 Besides the estimation of interactions between variables, the most significant outcome of applying
984 the FF setting is to identify the variables that contribute negligibly to the variance of the output. This
985 has practical advantages, in that when data collection is costly or impractical, resources can be
986 displaced from sampling variables with low S^T_i to those that show a large influence on the output. In
987 the case of ForestGALES, our results calculated negligible S^T_i for Rooting depth, Soil type, and Gap
988 size. The first two are related in ForestGALES as they are used to retrieve the C_{reg} values used in the
989 calculation of CWS$_O$ (Eq. 2), and as such are discussed together. Our results suggest that accurate
990 knowledge of Rooting depth and Soil type is not necessary, and that only minor approximation
991 errors would follow from fixing these variables to any value within their ranges. These findings
992 suggest that, when Rooting depth and Soil type are not known and would be expensive/impractical
993 to investigate (as for owners and managers of small forested stands), or when they are very
994 heterogeneous (as in the case of large-scale studies of wind damage), these variables can be quite
995 confidently ignored. In fact, the approximation errors shown in Table 5 for fixing Rooting depth or
996 Soil type are never larger than 6%. For *P. sitchensis* and *P. pinaster*, when we fixed these variables to
997 the values within their ranges, we did notice some small differences with the "Free" simulations
998 where all the variables were allowed to vary at the same time. There is no variation in the C_{reg} values
999 for *E. globulus* (Locatelli et al., 2016), which is therefore not discussed here. We ascribe the relatively
1000 high S_i value of Rooting depth for *E. globulus* (0.10) to a numerical imprecision in the calculations.
1001 Indeed, the same value is found for CWS$_B$, and the associated S^T_i were zero. For *P. sitchensis*, fixing
1002 Rooting depth to deep rooting resulted in an overestimation of CWS$_O$, showing more resistance to
1003 overturning (Figure 8). Similarly, overestimation of CWS$_O$ was found when we fixed Soil type to
1004 "Freely draining" and "Deep peats", as the average C_{reg} values shown in Table 1 for these soil types
1005 are the largest for *P. sitchensis.* Conversely, underestimation of CWS$_O$ was found when Soil type was
1006 fixed at "Gleys" and "Mineral Peats". For *P. pinaster*, fixing Rooting depth to medium resulted in
1007 overestimating CWS$_O$ (Figure 9), as did fixing Soil type to "Deep peats". Underestimation of CWS$_O$
1008 was found when Soil type was fixed at "Gleys". These results are consistent with the C_{reg} values
1009 shown in Table 1 for combinations of Rooting depth and Soil type. However, as seen in Figures 8 and
1010 9, these over- and underestimations are relatively minor and mostly affect simulated trees
1011 associated with very large CWS$_O$, suggesting that the contribution of Rooting depth and Soil type to
1012 the calculations of ForestGALES is mostly relevant in cases of catastrophic wind speeds, i.e. when
1013 forest management practices are of minor importance (Kohnle et al., 2003). This reinforces our
1014 conclusion that Rooting depth and Soil type are of minor importance to the use of ForestGALES.

1015 With regards to the influence of Gap size on CWS$_{(B,O)}$, the Sobol' indices in Figures 3, 4, and 5,
1016 suggest that while for *P. sitchensis* and *P. pinaster* Gap size has no influence, it does have an effect
1017 on *E. globulus*. However, the scatterplots in Figures 6 and 7 show that fixing Gap size had an effect
1018 on the calculations of the critical wind speeds for all three species. This is especially evident for Gap
1019 size = 0m, which resulted in large overestimations, especially for *E. globulus*. Fixing Gap size to the
1020 other values (1000m and 10 times Tree height) resulted in underestimation of CWS$_{(B,O)}$, especially for
1021 wind speeds over 25 m s^{-1}, but never for very high CWS$_{(B,O)}$. The effect is more evident for *E.
1022 globulus*. Gap size fixed at 2 times Tree height caused minor overestimation of CWS$_{(B,O)}$ for *P.
1023 sitchensis* and *P. pinaster*, also for low wind speeds. For *E. globulus*, the effect is complex, with over
1024 and underestimation of the CWS. This is particularly marked for overturning, where very low CWS
1025 are affected by the change. Fixing Gap size at 0m means that the stand is part of a continuous forest.
1026 As shown by Gardiner et al. (1997), this has the effect of modifying the wind profile, resulting in less

1027 loading on the trees (i.e. higher CWS are required to damage a tree within a forest continuum). Our
1028 results confirm this, showing that ForestGALES is able to simulate the effect of Gap size even in the
1029 case of green edges. In their study on the risk of wind damage to three conifer species, Mitchell et al.
1030 (2001) have shown that the likelihood of damage increased with creation of new edges, and
1031 persisted for a period of time insufficient for the newly exposed trees to acclimate to the new wind
1032 loading at the edges. Their findings on the effect of newly created edges were confirmed by Scott
1033 and Mitchell (2005) in their study on the effect of large upwind gaps. In their simulation study on the
1034 susceptibility of stands composed of two conifer species to wind damage in Finland, Zeng et al.
1035 (2010) concluded that the presence of upwind gaps increases the risk of wind damage to a stand
1036 more than its species composition. A number of simulation studies centred on the use of complex
1037 airflow models (either based on Large – Eddy Simulation techniques, or high – resolution Reynolds –
1038 averaged type models) have shed some light on some of the possible reasons for the effect of gap
1039 creation on the risk of wind damage. Dupont et al. (2015a, 2015b) have shown that as gaps are
1040 formed, more gusts penetrate the forest canopy, increasing the wind loading on trees. As shown in
1041 Figures 6 and 7, the effect of fixing Gap size to 0m is evident also at very high CWS, while fixing it to
1042 the other values only results in underestimation for medium to moderately high CWS. As the results
1043 of the Savage scores indicate (Table 6), Tree height, *dbh*, and *Sph* are the most influential variables in
1044 driving the CWS, regardless of tree species. Therefore, high CWS must be associated with short
1045 trees, large *dbh*, and mid-to-high values of *Sph* (the latter is shown in the discussion of the
1046 probabilities of damage). This is confirmed with data shown in the Appendix: when we isolated the
1047 simulated trees associated with extreme CWS (over 75 m s^{-1} for *P. sitchensis*; over 50 m s^{-1} for *P.*
1048 *pinaster*; over 35 m s^{-1} for *E. globulus*), we noticed that these trees were mostly short and their *dbh*
1049 was large (trees had a large taper), and stocking densities were medium to high. Therefore, for Gap
1050 size to be able to have such a large effect when fixed at 0m, it must be more important than
1051 estimated with the indices of Sobol'. The Sobol' indices were however able to identify Gap size as
1052 more important for *E. globulus* than for the other two species, which is confirmed by our
1053 investigation of tree characteristics for extreme CWS shown in the Appendix. In fact, there is much
1054 more variation in Tree height, *dbh*, and *Sph* for *E. globulus* trees than for the other species. That is,
1055 these trees are taller, their *dbh* is not necessarily large, and stocking densities are as low as ~300
1056 *Sph*. Considering that our simulated *E. globulus* trees are much slender than those of the other two
1057 species, and are therefore at higher risk of wind damage, the larger effect of fixing Gap size to 0m is
1058 to be expected. Locatelli et al. (2016) performed a GSA with the method of Kucherenko et al. (2012)
1059 on their parameterisation of ForestGALES for *E. globulus*. Their results show that, when Gap was
1060 allowed to vary within the same range as in our study, its influence on the CWS was significant, and
1061 the interactions between Tree height, *dbh*, *Sph*, and Gap size were very large. The authors used a
1062 version of ForestGALES where upwind gaps were brown edges, i.e. gaps recently formed. This
1063 suggests that variation of Gap size for newly created gaps has a larger influence on the dynamics of
1064 ForestGALES than when the gaps are of the green edge type.

1065 The FF setting can also highlight inadequacies in the modelling of processes that include variables
1066 that are found to be important in the natural world. In our study, this is likely to be the case of the
1067 influence of Rooting depth and Soil type on overturning. A number of surveys of wind damaged
1068 stands have provided information on the factors associated with wind damage to forests. Although
1069 these studies often do not discriminate between breakage and uprooting, they suggest that
1070 variation in soil type and rooting depth are important drivers of wind damage (e.g. Hanewinkel et al.,

1071   2008; Mayer et al., 2005; Nilsson et al., 2004; Schindler et al., 2009). The authors report that shallow

1072   rooting, waterlogged soils, and acidic soils, increase the risk of wind damage. These stand

1073   characteristics are often time-consuming, expensive, and difficult to measure both during tree-

1074   pulling experiments and in forest inventories. These constraints, together with the current scarcity

1075   of easily implementable, fully mechanistic methods of modelling tree anchorage, have forced wind

1076   damage modellers to adopt an empirical approach to model trees' resistance to uprooting that is

1077   limited to easily obtainable characteristics of the root and soil system, such as coarse classifications

1078   of soil type and rooting depth. Examples of the complexities of the effect of soil conditions on the

1079   risk of uprooting can be found, e.g. in Ray and Nicoll (1998), and more recently in Kamimura et al.

1080   (2012). In their investigation of the effect of waterlogging soils prior to tree-pulling experiments,

1081   Kamimura et al. (2012) identified that soil water content influenced tree anchorage in a complex

1082   fashion, highly dependent on whether water was concentrated within the soil-root plate, or below it.

1083   Heavier, water-saturated root plates provided the trees with initial higher stability, while large water

1084   content below the plates can lead to hydraulic fracturing in the soil.

1085

1086   *4.3 Difference between species*

1087   As shown by Iman and Conover (1987), the calculation of the top-down coefficient of concordance

1088   amongst all species (C_T) is strongly driven by agreement between the top ranks. This can be seen in

1089   the calculation of the Savage scores with Eq. (6). The importance of at least a degree of accordance

1090   between the top three ranks in driving the calculation of the T_{all} index is further corroborated by the

1091   numerator of Eq. (7): the more similar the top ranks are amongst species, and the larger the number

1092   of species, the larger C_T (and therefore T_{all}) becomes. As a result, the probability of the T_{all} index to

1093   be larger than the upper-tail critical value of a chi-square distribution (in our case, with 5 degrees of

1094   freedom) diminishes as C_T and T_{all} increase.

1095   With regards to the pairwise comparisons between our species, certain differences are evident. For

1096   CWS$_B$, the ranking of the S^T_i of *E. globulus* is markedly different from the other two (Figures 3, 4, and

1097   5). This is reflected by the *p*-values in Table 6 for the comparisons that include *E. globulus* being the

1098   largest (0.247), while the *p*-value of the comparison between *P. sitchensis* and *P. pinaster* (0.087)

1099   indicates that differences between these two species are almost non-significant at the 95%

1100   confidence interval. It should be noted that the fact that the S^T_i rankings of these two species are

1101   identical (Figures 3 and 4) suggests that the accuracy of our statistical test is not perfect. For CWS$_B$,

1102   the large *p*-values (0.237 and 0.246) for the comparisons between *P. sitchensis* / *E. globulus*, and *P.*

1103   *pinaster* / *E. globulus*, are driven by the fact that the ranking of the top two variables, Tree height

1104   and *dbh*, are inverted between the pair (*P. sitchensis*, *P. pinaster*), and *E. globulus*. In fact, the

1105   difference between the Savage scores calculated for the top rank, and the second or the third rank,

1106   are quite significant (top rank: 2.45; second rank: 1.45; third rank: 0.95), while further ranks,

1107   especially in the case of ties (e.g. for Rooting Depth and Soil type) have similar low values, below 0.5.

1108   Therefore, in pairwise comparisons a small disagreement at the top three ranks can result in the S^T_i

1109   rankings of the 2 species being flagged as substantially different. This is unlike in the calculations of

1110 C_T and T_{all}, for which partial agreements between the top S^T_i rankings of the three species

1111   contributes to the degree of similarity between all three species. Similarly, the small, but non-zero S^T_i

1112   of Gap for *E. globulus* contributes marginally to these pairwise differences that involve *E. globulus*,

1113 as its rank is higher than for the other two species. For CWS_O, the difference between *P. sitchensis*
1114 and *P. pinaster* (*p*-value 0.448) is due to the different ranks of the top three variables (Tree height,
1115 *dbh*, and *Sph*). For the *P. pinaster* / *E. globulus* pair, the large *p*-value (0.273) is due to the rankings of
1116 Tree height, *dbh*, and Gap size. Similarly, the difference between *P. sitchensis* and *E. globulus* (*p*-
1117 value 0.128) is attributable to the rankings of Tree height, *Sph*, and Gap size.

1118

1119 *4.4 Probabilities of damage – Factor Mapping setting*

1120 The FM setting is a form of Monte Carlo filtering, in that it divides the output space in acceptable
1121 and unacceptable regions, and maps the realisations of the model in these regions back to the input
1122 space. We have applied this setting to the probabilities of breakage and overturning, and we have
1123 chosen the probability threshold to differentiate between the two regions in the output space as 0.1
1124 (10% probability of damage). Our results show that DAMS, the variable that describes the intensity
1125 of the wind, is the most responsible for realisation of $Prob_{(B,O)}$ in the two regions of the output space,
1126 regardless of tree species. DAMS values range from 7 to 22, with 20 often regarded as the limit for
1127 commercial forestry (Quine, 2000). Our SA shows that DAMS values over ~15 are much more likely
1128 to result in damage to a stand, regardless of the mode of failure (breakage or overturning) and tree
1129 species. This finding suggests that ForestGALES describes well the effect of the wind climate on the
1130 probabilities of damage. In fact, it is well known that in the case of extreme wind storm, the effect of
1131 silvicultural practices on tree survival is minor (Kohnle et al., 2003), while for sheltered stands the
1132 effect of stand and tree characteristics is more prominent (Albrecht et al., 2012). This raises the issue
1133 of the availability of accurate wind speed data, which can be obtained only from localised climate
1134 stations. While methods to extrapolate spatially the wind speeds exist (e.g. the Wind Atlas Analysis
1135 and Application Program (WAsP) used for estimates of wind energy productivity), their
1136 extrapolations are not devoid of error, especially in the case of complex terrain or large distances
1137 from the data source (Venäläinen et al., 2004). Mayer et al. (2005) have shown that when
1138 knowledge of wind speeds during a storm is accurate, wind speeds become important in statistical
1139 models used to discriminate between areas with and without wind damage.

1140 The results of the FM setting show some similarities between *P. sitchensis* and *P. pinaster*, and
1141 therefore these two species are discussed together. *E. globulus* behaved quite differently, probably
1142 due to the tall trees with low taper used in our simulations, and is discussed separately. The results
1143 of the Smirnov tests generally agree with the Sobol' S_i calculated for $CWS_{(B,O)}$ for all the species. For
1144 *P. pinaster*, this means that *dbh* was found to be uninfluential with the Smirnov tests. The role of
1145 *dbh* in driving the outputs is important for *P. sitchensis*. Our results (Figures 11 and Table 7) show
1146 that the risk of breakage and overturning decreases for trees with a *dbh* larger than 10cm. Both
1147 conifer species respond similarly with regards to *Sph*, as intermediate values (~1500 sph) correspond
1148 to lower risk. No threshold could be identified for Tree height. For $Prob_O$, DAMS values lower than 10
1149 did not cause any damage to the simulated stands of the two species (Table 7). The results of our
1150 bivariate investigations of the relationship between significant variables with regards to $Prob_{(B,O)}$
1151 show some differences between the two species. For *P. sitchensis*, the two modes of damage show
1152 similarities (Figures 12 and 13). The DAMS vs Tree height density plots show that short trees (Tree
1153 height < 10m) are at lower risk of damage regardless of the wind speed, and that no damage was
1154 estimated below DAMS ~ 13. We ascribe this behaviour to the form of the *P. sitchensis* trees in our

1155 simulations, which are characterised by high taper (Table 3). For breakage, *P. pinaster* behaves very
1156 similarly, while for overturning higher wind speeds are required for damage (DAMS > 16, Figure 15).
1157 For both species and both Prob$_{(B,O)}$, the DAMS vs *dbh* plots show a diminished importance of wind
1158 speed when trees have large *dbh*. The role of taper in driving the probabilities of damage is evident
1159 from the *dbh* vs Tree height density plots. The separation between the clouds of points is quite clear,
1160 showing that trees of the same height are at lower risk of damage when their *dbh* is large. This
1161 finding confirms that ForestGALES simulates effectively the well-known effect of taper on the risk of
1162 wind damage (e.g. Peltola and Kellomaki, 1993; Quine et al., 1995). For *P. pinaster*, only trees taller
1163 than ~10.5m were considered at risk of damage. The role of *Sph* is the same for both species and
1164 both modes of damage. The DAMS vs *Sph* plots show that at low stocking densities the probabilities
1165 of damage are higher than 10% for DAMS as low as 12, while at high stocking densities damage is
1166 predicted only at high DAMS. The relationship between *Sph* and Tree height in ForestGALES is such
1167 that short trees in sparse stands are at lower risk of damage, while trees above 11m are associated
1168 with damage regardless of stocking density. Conversely, trees of medium to large *dbh* are mostly at
1169 risk for low *Sph*, while low *dbh* is associated with damage regardless of the stocking density, but
1170 mostly at low *Sph*. The effect of dense stands is quite constant for the two species, with large
1171 stocking densities being associated with lower risk of damage. With regards to *Sph*, ForestGALES
1172 behaves in accordance with the findings of Coutts (1986), who reported less damage in dense *P.
1173 sitchensis* stands. Gardiner et al. (1997) showed that low stocking densities result in higher wind
1174 loading on the trees, despite the gustiness of the wind decreases with decreasing *Sph*. The authors
1175 showed that the maximum bending moment increases faster than the Gust Factor decreases, as
1176 stocking densities decrease.

1177 The FM results for *E. globulus* differ partially from those of the other two species. For both types of
1178 damage, the *Sph* thresholds are lower (~ 1000 sph), as are the associated D-statistic values (Table 7),
1179 suggesting that for *E. globulus Sph* is less important in discriminating between high and low
1180 probabilities of damage. The density plots (Figures 16 and 17) show that our simulated *E. globulus*
1181 trees are at higher risk of damage than the other species, and that DAMS is much more important
1182 than *Sph*, which is in turn much more important than Tree height and *dbh* in discriminating between
1183 damage and no damage. The DAMS vs Tree height and *dbh* plots are very similar, showing that small
1184 trees are vulnerable to damage even at low wind speeds (DAMS ~8) while larger trees require DAMS
1185 to exceed 14 for the probabilities of damage to exceed 10%. This trend is confirmed by the Tree
1186 height vs *dbh* plot. We ascribe the similarities between the response of Tree height and *dbh*, and the
1187 relatively low importance of *Sph*, to the high taper of our *E. globulus* trees, which probably largely
1188 influenced the calculations of Prob$_{(B,O)}$, as it did for CWS$_{(B,O)}$.

1189

1190 *4.5 Evaluation of the performance of our GSA*

1191 The most important step of variance-based methods of sensitivity analysis is the characterisation of
1192 the variables with reliable probability distribution functions. In order to investigate the behaviour
1193 and sensitivity of the ForestGALES model, in our study we have adopted a data-driven approach, as
1194 we fitted PDFs to Tree height and *dbh* from available tree-pulling data, to ensure that the model is
1195 investigated within the limits of its parameterisations for different species. Large trees are typically
1196 under-represented in these field experiments for safety reasons and technical limitations (Nicoll et

1197   al., 2006), even more so for data that was gathered in the past (Fraser and Gardiner, 1967). As a
1198   consequence, tall trees are under-represented in the *P. sitchensis* and *P. pinaster* tree-pulling
1199   datasets that were used for the PDFs in our study, while short trees with a high taper are common.
1200   The high correlation coefficient between Tree height and *dbh* (*P. sitchensis*: 0.54; *P. pinaster*: 0.73; *p*-
1201   values < 0.001) resulted in our simulated trees for the two conifer species to be quite short and with
1202   a high taper. The stability of high tapering trees with regards to wind damage is well known,
1203   especially for breakage (e.g. Slodicak and Novak, 2006; Valinger and Lundqvist, 1992). As shown by
1204   Zubizarreta-Gerendiain et al. (2012), regardless of species trees with a high taper are associated with
1205   lower risk of uprooting, a tenet which is often interpreted as trees with larger diameters in the lower
1206   stem having likely allocated more resources in the formation of extensive root systems (Nicoll and
1207   Ray, 1996). This is reflected in our ForestGALES simulations, with low-tapered *P. sitchensis* and *P.*
1208   *pinaster* trees being at lower risk of uprooting (i.e. higher CWS_O) than the more slender *E. globulus*
1209   (Figures 8, 9, and 10).

1210   In addition to this, the taper of our simulated trees was independent of the stocking density, which
1211   is not representative of reality. As discussed in the previous paragraphs, these factors have a
1212   significant impact on the ForestGALES calculations of the critical wind speeds and their associated
1213   probabilities of damage. Consequently, the influence of *Sph* on the outputs of ForestGALES was
1214   probably underestimated in our GSA, as shown by the small differences between total and first-
1215   order sensitivity indices for *Sph*. This is likely to be the case also for *E. globulus*, for a similar but
1216   opposite reason. In fact, our simulated eucalypt trees are fairly tall and slender, with a very high
1217   taper. Because of the importance of Tree height, *dbh*, and taper in ForestGALES, the outputs are
1218   likely to have been mostly driven by Tree height and *dbh*, while the role of *Sph* was probably
1219   outweighed. The effect of this can be seen in the smaller range of $CWS_{(B,O)}$ for *E. globulus* in
1220   comparison to the other species, as seen in Figures 6 and 7. Although Tree height and *dbh*, and their
1221   high correlation (0.91, *p*-value < 0.001) are representative of the fieldwork data used to calculate
1222   the parameters of the corresponding PDFs, the source of our *E. globulus* data has some limitations
1223   and might not be representative of all eucalypt stands (Locatelli et al., 2016). Nevertheless, our GSA
1224   allowed us to highlight the limitations of the parameterisation of ForestGALES for this species. Our
1225   study shows that the GSA method of Kucherenko et al. (2012) is very sensitive to the correlations
1226   between variables in the correlation matrix of the copula. Therefore, it requires an accurate
1227   characterisation not only of the PDFs of the inputs, but also of their initial correlation matrix. This is
1228   important, to ensure that the generated dataset used for the SA is representative of the original
1229   sample's structure.

1230   We described Rooting depth and Soil type with uniform discrete distributions in order to explore the
1231   input space more thoroughly, but we did impose a correlation based on *P. sitchensis* data. This might
1232   not be representative of *P. pinaster*, and might have marginally influenced our results for this
1233   species. Gap size is the variable that is most likely to have been poorly characterised using the range
1234   of values in our simulations. As shown in Figures 6 and 7, fixing Gap size to 0m has a large effect on
1235   the calculations of the CWS; at just two tree heights, however, the effect is largely diminished, while
1236   fixing Gap size at values as low as 10 times Tree height has an almost negligible effect. However, the
1237   model runs where Gap had values close to 0m were not sufficient to influence the calculations of the
1238   Sobol' indices. A Gap of the size of twice the mean Tree heights shown in Table 3 belongs to the first
1239   quartile of the range of Gap size. Therefore, most of our simulations had values that exceed this

1240   average value, and this likely influenced the calculations of the Sobol' indices. We expect that a
1241   narrower range of Gap size would have resulted in larger Sobol' indices for this variable.

1242

1243   **5. Conclusion**

1244   In this study we have performed a variance-based sensitivity analysis on the forest wind-risk model
1245   ForestGALES, for three species (*P. sitchensis*, *P. pinaster*, and *E. globulus*) representative of three of
1246   the most extensively planted and highly productive tree genera worldwide: spruces, pines, and
1247   eucalypts. The application of the variance-based sensitivity analysis method for correlated variables
1248   shows great sensitivity not only to the characterisation of the variables with appropriate probability
1249   density functions, but also to the correlation matrix of the variables. Therefore, particular care must
1250   be exercised when describing the input data for this method of sensitivity analysis. Our results show
1251   that Tree height, *dbh*, and stocking density are the tree and stand variables mostly responsible for
1252   the variation in the critical wind speeds for breakage and overturning, regardless of tree species,
1253   although minor intraspecific differences exist in the ranking of these variables. These variables,
1254   together with the wind climate local to a stand, as expressed by the DAMS variable, are the major
1255   drivers of variation in the associated probabilities of damage. Therefore, for practical applications of
1256   ForestGALES, users should focus their resources on sampling accurately these three tree and stand
1257   variables to maximally reduce the uncertainty in the predictions of the model. Our study shows that
1258   Rooting depth and Soil type are only marginally important for the calculations of the critical wind
1259   speed of overturning. This finding suggests that these variables can be fixed at any value within their
1260   ranges without significantly affecting the output of the model, thus suggesting that ForestGALES can
1261   be successfully applied to large-scale studies of wind damage when information on these variables is
1262   coarse at best. While there is not a general consensus on the role of rooting depth and soil type on
1263   the vulnerability of a stand to wind damage, our study suggests that the empirical component of
1264   ForestGALES that calculates the resistance to overturning requires further development. The effect
1265   of an upwind gap on the critical wind speeds is confirmed also for pre-existing forest edges,
1266   suggesting that recommendations made with ForestGALES on the careful management of forest
1267   edges to reduce the risk of damage are to be considered seriously, namely avoiding Gap sizes larger
1268   than twice the mean height of the stand.

1269

1277

1278   **Reference list**

1279 Albrecht, A., Hanewinkel, M., Bauhus, J. & Kohnle, U. 2012. How does silviculture affect storm
1280 damage in forests of south-western Germany? Results from empirical modeling based on long-term
1281 observations. European Journal of Forest Research, 131, 229-247.

1282 Archer, G.E.B., Saltelli, A. & Sobol', I.M. 1997. Sensitivity measures,anova-like techniques and the use
1283 of bootstrap. Journal of Statistical Computation and Simulation, 58, 99-120.

1284 Campolongo, F. & Saltelli, A. 1997. Sensitivity analysis of an environmental model: an application of
1285 different analysis methods. Reliability Engineering & System Safety, 57, 49-69.

1286 Chan, K., Saltelli, A. & Tarantola, S. 1997. Sensitivity analysis of model output: variance-based
1287 methods make the difference. Proceedings of the 29th conference on Winter simulation. Atlanta,
1288 Georgia, USA: IEEE Computer Society.

1289 Cherubini, U., Luciano, E. & Vecchiato, W. 2004. Copula methods in finance, John Wiley & Sons.

1290 Confalonieri, R., Bellocchi, G., Bregaglio, S., Donatelli, M. & Acutis, M. 2010. Comparison of
1291 sensitivity analysis techniques: a case study with the rice model WARM. Ecological Modelling, 221,
1292 1897-1906.

1293 Coutts, M.P. 1986. Components of tree stability in Sitka Spruce on peaty gley soil. Forestry, 59, 173-
1294 197.

1295 Cucchi, V. & Bert, D. 2003. Wind-firmness in Pinus pinaster (Ait.) stands in Southwest France:
1296 influence of stand density, fertilisation and breeding in two experimental stands damaged during the
1297 1999 storm. Annals of Forest Science, 60, 209-226.

1298 Cucchi, V., Meredieu, C., Stokes, A., Berthier, S., Bert, D., Najar, M., Denis, A. & Lastennet, R. 2004.
1299 Root anchorage of inner and edge trees in stands of Maritime pine (Pinus pinaster Ait.) growing in
1300 different podzolic soil conditions. Trees-Structure and Function, 18, 460-466.

1301 Cucchi, V., Meredieu, C., Stokes, A., De Coligny, F., Suarez, J. & Gardiner, B.A. 2005. Modelling the
1302 windthrow risk for simulated forest stands of Maritime pine (Pinus pinaster Ait.). Forest Ecology and
1303 Management, 213, 184-196.

1304 Cukier, R.I., Fortuin, C.M., Shuler, K.E., Petschek, A.G. & Schaibly, J.H. 1973. Study of the sensitivity of
1305 coupled reaction systems to uncertainties in rate coefficients. I Theory. The Journal of Chemical
1306 Physics, 59, 3873-3878.

1307 Cukier, R.I., Levine, H.B. & Shuler, K.E. 1978. Nonlinear sensitivity analysis of multiparameter model
1308 systems. Journal of computational physics, 26, 1-42.

1309 Dufour-Kowalski, S., Courbaud, B., Dreyfus, P., Meredieu, C. & De Coligny, F. 2012. Capsis: an open
1310 software framework and community for forest growth modelling. Annals of Forest Science, 69, 221-
1311 233.

1312 Dupont, S., Ikonen, V.P., Väisänen, H. and Peltola, H. 2015. Predicting tree damage in fragmented
1313 landscapes using a wind risk model coupled with an airflow model. Canadian Journal of Forest
1314 Research, 45(8), 1065-1076.

1315   Dupont, S., Pivato, D. and Brunet, Y. 2015. Wind damage propagation in forests. Agricultural and
1316   Forest Meteorology, 214, 243-251.

1317   Fraser, A.I. & Gardiner, J.B.H. 1967. Rooting and stability in Sitka spruce. Forestry Commission
1318   Bulletin, 331-4.

1319   Gaber, N., Foley, G., Pascual, P., Stiber, N. & Sunderland, E. 2009. Guidance on the development,
1320   evaluation, and application of environmental models. In: Agency, E.P. (ed.). Washington, D.C.: EPA.

1321   Gardiner, B., Byrne, K., Hale, S., Kamimura, K., Mitchell, S.J., Peltola, H. & Ruel, J.-C. 2008. A review of
1322   mechanistic modelling of wind damage risk to forests. Forestry, 81, 447-463.

1323   Gardiner, B., Peltola, H. & Kellomaki, S. 2000. Comparison of two models for predicting the critical
1324   wind speeds required to damage coniferous trees. Ecological Modelling, 129, 1-23.

1325   Gardiner, B.A., Stacey, G.R., Belcher, R.E. & Wood, C.J. 1997. Field and wind tunnel assessments of
1326   the implications of respacing and thinning for tree stability. Forestry, 70, 233-252.

1327   Hale, S.A., Gardiner, B., Peace, A., Nicoll, B., Taylor, P. & Pizzirani, S. 2015. Comparison and validation
1328   of three versions of a forest wind risk model. Environmental Modelling & Software, 68, 27-41.

1329   Hanewinkel, M., Breidenbach, J., Neeff, T. & Kublin, E. 2008. Seventy-seven years of natural
1330   disturbances in a mountain forest area - the influence of storm, snow, and insect damage analysed
1331   with a long-term time series. Canadian Journal of Forest Research-Revue Canadienne De Recherche
1332   Forestiere, 38, 2249-2261.

1333   Helton, J.C., Davis, F.J. & Johnson, J.D. 2005. A comparison of uncertainty and sensitivity analysis
1334   results obtained with random and Latin hypercube sampling. Reliability Engineering & System Safety,
1335   89, 305-330.

1336   Homma, T. & Saltelli, A. 1996. Importance measures in global sensitivity analysis of nonlinear
1337   models. Reliability Engineering & System Safety, 52, 1-17.

1338   Hornberger, G.M. & Spear, R.C. 1981. Approach to the preliminary analysis of environmental
1339   systems. Journal Name: J. Environ. Manage.; (United States); Journal Volume: 12:1, Medium: X; Size:
1340   Pages: 7-18.

1341   Iman, R.L. & Conover, W.J. 1987. A measure of top-down correlation. Technometrics, 29, 351-357.

1342   Iman, R.L. & Hora, S.C. 1990. A robust measure of uncertainty importance for use in fault tree
1343   system analysis. Risk Analysis, 10, 401-406.

1344   Ipcc. 1999. IPCC expert meetings on good practice guidance and uncertainty management in
1345   national greenhouse gas inventories. Background papers. Available: http://www.ipcc-
1346   nggip.iges.or.jp/public/gp/gpg-bgp.htm.

1347   Ipcc. 2000. Good practice guidance and uncertainty management in national greenhouse gas
1348   inventories. Available: http://www.ipcc-nggip.iges.or.jp/public/gp/gpgaum.htm.

1349 Kamimura, K. Kitagawa, K. Saito, S. and Mizunaga, H. 2012. Root anchorage of hinoki (*Chamaecyparis*
1350 *obtuse* (Sieb. Et Zucc.) Endl.) under the combined loading of wind and rapidly supplied water on soil:
1351 analyses based on tree-pulling experiments. European Journal of Forest Research, 131(1), 219-227.

1352 Kohnle, U., Gauckler, S., Risse, F. & Stahl, S. 2003. Der Orkan Lothar im Spiegel von Betriebsinventur
1353 und Einschlagsbuchführung: Auswirkungen auf einen baden-württembergischen Forstbezirk im
1354 Randbereich des Sturms. AFZ-Der Wald, 58, 1203-1207.

1355 Kucherenko, S., Albrecht, D. & Saltelli, A. 2015. Exploring multi-dimensional spaces: a Comparison of
1356 Latin Hypercube and Quasi Monte Carlo Sampling Techniques. arXiv preprint arXiv:1505.02350.

1357 Kucherenko, S., Rodriguez-Fernandez, M., Pantelides, C. and Shah, N. 2009. Monte Carlo evaluation
1358 of derivative-based global sensitivity measures. Reliability Engineering & System Safety, 94(7), 1135-
1359 1148.

1360 Kucherenko, S., Tarantola, S. & Annoni, P. 2012. Estimation of global sensitivity indices for models
1361 with dependent variables. Computer Physics Communications, 183, 937-946.

1362 Lavers, G. 1969. The strength properties of timbers. Forest Products Research Laboratory Bulletin.
1363 London: HMSO.

1364 Locatelli, T., Gardiner, B., Tarantola, S., Nicoll, B., Bonnefond, J.-M., Garrigou, D., Kamimura, K. &
1365 Patenaude, G. 2016. Modelling wind risk to Eucalyptus globulus (Labill.) stands. Forest Ecology and
1366 Management, 365, 159-173.

1367 Mara, T.A. & Tarantola, S. 2012. Variance-based sensitivity indices for models with dependent
1368 inputs. Reliability Engineering & System Safety, 107, 115-121.

1369 Mason, W.L. 2003. Continuous Cover Forestry: Developing close to nature forest management in
1370 conifer plantations in upland Britain. Scottish Forestry, 57, 141-149.

1371 Mayer, P., Brang, P., Dobbertin, M., Hallenbarter, D., Renaud, J.P., Walthert, L. & Zimmermann, S.
1372 2005. Forest storm damage is more frequent on acidic soils. Annals of Forest Science, 62, 303-311.

1373 Mcinerney, D., Suarez, J. & Nieuwenhuis, M. 2011. Extending forest inventories and monitoring
1374 programmes using remote sensing: A review. Irish Forestry, 68, 6-22.

1375 Mitchell, S.J. 1998. A diagnostic framework for windthrow risk estimation. The Forestry Chronicle,
1376 74(1), 100-105.

1377 Most, T. 2012. Variance-based sensitivity analysis in the presence of correlated input variables. In:
1378 Proc. 5th Int. Conf. Reliable Engineering Computing (REC), Brno, Czech Republic, 2012.

1379 Nicoll, B.C., Gardiner, B.A., Rayner, B. & Peace, A.J. 2006. Anchorage of coniferous trees in relation to
1380 species, soil type, and rooting depth. Canadian Journal of Forest Research-Revue Canadienne De
1381 Recherche Forestiere, 36, 1871-1883.

1382 Nicoll, B.C. and Ray, D. 1996. Adaptive growth of tree root systems in response to wind action and
1383 site conditions. Tree physiology, 16(11-12), 891-898.

1384   Nilsson, C., Stjernquist, I., Bärring, L., Schlyter, P., Jönsson, A.M. & Samuelsson, H. 2004. Recorded
1385   storm damage in Swedish forests 1901–2000. Forest Ecology and Management, 199, 165-173.

1386   Norton, J. 2015. An introduction to sensitivity assessment of simulation models. Environmental
1387   Modelling & Software, 69, 166-174.

1388   Nossent, J., Elsen, P. & Bauwens, W. 2011. Sobol' sensitivity analysis of a complex environmental
1389   model. Environmental Modelling & Software, 26, 1515-1525.

1390   Oreskes, N., Shrader-Frechette, K. & Belitz, K. 1994. Verification, validation, and confirmation of
1391   numerical models in the earth sciences. Science, 263, 641-646.

1392   Peltola, H. & Kellomaki, S. 1993. A mechanistic model for calculating windthrow and stem breakage
1393   of Scots pines at stand edge. Silva Fennica, 27, 99-111.

1394   Penman, J., Gytarsky, M., Hiraishi, T., Krug, T., Kruger, D., Pipatti, R., Buendia, L., Miwa, K., Ngara, T.
1395   & Tanabe, K. 2003. Good practice guidance for land use, land-use change and forestry, Institute for
1396   Global Environmental Strategies.

1397   Quine, C., Coutts, M.P., Gardiner, B. & Pyatt, G. 1995. Forests and wind: management to minimize
1398   damage, FC.

1399   Quine, C.P. 2000. Estimation of mean wind climate and probability of strong winds for wind risk
1400   assessment. Forestry, 73, 247-258.

1401   Quine, C.P. 2000. Estimation of mean wind climate and probability of strong winds for wind risk
1402   assessment. Forestry, 73, 247-258.

1403   Quine, C.P. & White, I.M.S. 1993. Revised windiness scores for the windthrow hazard classification:
1404   The revised scoring method, Edinburgh, Forestry Commission Publications.

1405   Rahmstorf, S., Cazenave, A., Church, J.A., Hansen, J.E., Keeling, R.F., Parker, D.E. & Somerville, R.C.J.
1406   2007. Recent climate observations compared to projections. Science, 316, 709-709.

1407   Raupach, M.R. 1994. Simplified expressions for vegetation roughness length and zero-plane
1408   displacement as functions of canopy height and area index. Boundary-Layer Meteorology, 71, 211-
1409   216.

1410   Ray, D. and Nicoll, B.C. 1998. The effect of soil water-table depth on root-plate development and
1411   stability of Sitka spruce. Forestry, 71(2), 169-182.

1412   Rosette, J., Suarez, J., North, P. & Los, S. 2011. Forestry Applications for Satellite Lidar Remote
1413   Sensing. Photogrammetric Engineering and Remote Sensing, 77, 271-279.

1414   Ruel, J.-C., Achim, A., Herrera, R.E., Cloutier, A. & Brossier, B. 2010. Wood degradation after
1415   windthrow in a northern environment. Forest Products Journal, 60, 200-206.

1416   Sacks, J., Welch, W.J., Mitchell, T.J. & Wynn, H.P. 1989. Design and analysis of computer
1417   experiments. Statistical science, 409-423.

1418 Saltelli, A. 2002. Making best use of model evaluations to compute sensitivity indices. Computer
1419 Physics Communications, 145, 280-297.

1420 Saltelli, A. & Annoni, P. 2010. How to avoid a perfunctory sensitivity analysis. Environmental
1421 Modelling & Software, 25, 1508-1517.

1422 Saltelli, A. & Funtowicz, S. 2014. When all models are wrong. Issues in Science and Technology, 30,
1423 79-85.

1424 Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. & Tarantola, S.
1425 2008. Global sensitivity analysis: the primer, John Wiley & Sons.

1426 Saltelli, A. Tarantola, S. Campolongo, F, and Ratto, M. (2004). Sensitivity analysis in practice: A guide
1427 to assessing scientific models. John Wiley & Sons.

1428 Saltelli, A., Tarantola, S. & Chan, K.P.S. 1999. A quantitative model-independent method for global
1429 sensitivity analysis of model output. Technometrics, 41, 39-56.

1430 Savage, I.R. 1956. Contributions to the theory of rank order statistics-the two-sample case. The
1431 Annals of Mathematical Statistics, 590-615.

1432 Schindler, D., Grebhan, K., Albrecht, A. & Schoenborn, J. 2009. Modelling the wind damage
1433 probability in forests in Southwestern Germany for the 1999 winter storm 'Lothar'. International
1434 Journal of Biometeorology, 53, 543-554.

1435 Scott, R.E. and Mitchell, S.J. 2005. Empirical modelling of windthrow risk in partially harvested stands
1436 using tree, neighbourhood, and stand attributes. Forest Ecology and Management, 218(1), 193-209.

1437 Seguro, J.V. & Lambert, T.W. 2000. Modern estimation of the parameters of the Weibull wind speed
1438 distribution for wind energy analysis. Journal of Wind Engineering and Industrial Aerodynamics, 85,
1439 75-84.

1440 Sklar, A. 1973. Random variables, joint distribution functions, and copulas. Kybernetika, 9, (449)-460.

1441 Slodicak, M. and Novak, J. 2006. Silvicultural measures to increase the mechanical stability of pure
1442 secondary Norway spruce stands before conversion. Forest Ecology and Management, 224(3), 252-
1443 257.

1444 Sobol', I.M. 1998. On quasi-Monte Carlo integrations. Mathematics and Computers in Simulation, 47,
1445 103-112.

1446 Sobol̇, I.M. 1990. Quasi-Monte Carlo methods. Progress in Nuclear Energy, 24, 55-61.

1447 Sobol', I.M. & Kucherenko, S. 2009. Derivative based global sensitivity measures and their link with
1448 global sensitivity indices. Mathematics and Computers in Simulation, 79, 3009-3017.

1449 Sobol', I.M., Tarantola, S., Gatelli, D., Kucherenko, S.S. & Mauntz, W. 2007. Estimating the
1450 approximation error when fixing unessential factors in global sensitivity analysis. Reliability
1451 Engineering & System Safety, 92, 957-960.

1452 Sobol', I.M. 2001. Global sensitivity indices for nonlinear mathematical models and their Monte
1453 Carlo estimates. Mathematics and Computers in Simulation, 55, 271-280.

1454 Song, X., Bryan, B.A., Paul, K.I. & Zhao, G. 2012. Variance-based sensitivity analysis of a forest growth
1455 model. Ecological Modelling, 247, 135-143.

1456 Tarantola, S., Giglioli, N., Jesinghaus, J. & Saltelli, A. 2002. Can global sensitivity analysis steer the
1457 implementation of models for environmental assessments and decision-making? Stochastic
1458 Environmental Research and Risk Assessment, 16, 63-76.

1459 Valinger, E. & Fridman, J. 2011. Factors affecting the probability of windthrow at stand level as a
1460 result of Gudrun winter storm in southern Sweden. Forest Ecology and Management, 262, 398-403.

1461 Valinger, E. and Lundqvist, L. 1992. The influence of thinning and nitrogen fertilisation on the
1462 frequency of snow and wind induced stand damage in forests. Scottish Forestry, 46(4), 311-320.

1463 Venäläinen, A., Zeng, H.C., Peltola, H., Talkkari, A., Strandman, H., Wang, K.Y. & Kellomaki, S. 2004.
1464 Simulations of the influence of forest management on wind climate on a regional scale. Agricultural
1465 and Forest Meteorology, 123, 149-158.

1466 Yang, J. 2011. Convergence and uncertainty analyses in Monte-Carlo based sensitivity analysis.
1467 Environmental Modelling & Software, 26, 444-457.

1468 Zeng, H., Garcia-Gonzalo, J., Peltola, H. and Kellomäki, S. 2010. The effects of forest structure on the
1469 risk of wind damage at a landscape level in a boreal forest ecosystem. Annals of Forest Science,
1470 67(1), 111.

1471 Zubizarreta-Gerendiain, A., Pellikka, P., Garcia-Gonzalo, J., Ikonen, V.P. and Peltola, H. 2012. Factors
1472 affecting wind and snow damage of individual trees in a small management unit in Finland:
1473 assessment based on inventoried damage and mechanistic modelling. Silva Fennica, 46(2), 181-196.

1474