Modelling the architecture of hazelnut (Corylus avellana) Tonda di Giffoni over two successive years

F. Grisafi, S. Tombesi, D. Farinelli, Evelyne Costes, Jean-Baptiste Durand, Frédéric Boudon

Modelling the architecture of hazelnut (*Corylus avellana*) Tonda di Giffoni over two successive years.

F. Grisafi[1*], S. Tombesi[1], D. Farinelli[2], E. Costes[3], J.-B. Durand[5,6], F. Boudon[3,4]

[1]Department of Sustainable Crop Production, Università Cattolica del Sacro Cuore, Via Emilia Parmense 84, 29122 Piacenza, Italy

[2]Department of Agricultural, Food and Environmental Sciences, Università degli studi di Perugia, Via Borgo XX giugno 74 , 06123 Perugia, Italy

[3]UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, F-34398 Montpellier, France

[4] CIRAD, UMR AGAP Institut, F-34398 Montpellier, France

[5] CIRAD, UMR AMAP, F- 34398 Montpellier, France

[6] AMAP, Univ Montpellier, CIRAD, CNRS, INRAE, IRD, Montpellier, France

F. Boudon and J.-B. Durand have been supported by the MaCS4Plants CIRAD network, initiated from the AGAP Institute and AMAP joint research units.

[*]For correspondence. Email francesca.grisafi1@unicatt.it

F. Boudon and J.-B. Durand should be considered as joint senior author

Hazelnut (*Corylus avellana*) cultivation is increasing worldwide. A 3D model of its structure could improve the managerial techniques such as pruning. This study aims to analyse, over two successive years, hazelnut architectural development to implement a functional structural plant model. 104 one-year-old shoots of own-rooted hazelnut trees were selected and analyzed in winter 2020 and 2021. Exploratory analyses, generalized linear models, and multinomial regression models were used to describe the architectural processes. The existence of sylleptic shoots on hazelnut one-year-old shoots, characterized by the presence of the male inflorescence on apical position, was detected. Along proleptic shoots the branching pattern was described by (1) blind nodes located in the proximal part (2) sylleptic shoots and mixed buds in the median part (3) vegetative buds in the distal part. Apical bud died during the growing season, suggesting that Tonda di Giffoni has a sympodial branching. The models revealed dependencies among buds located at the same node, in the case of proleptic shoots. Especially, the probability of a bud to burst depended on both its type (i.e., mixed or vegetative) and the presence of other buds, either mixed or vegetative. Based on these local models and on a flow diagram, which defines the steps that lead to the construction of hazelnut tree architecture, a first functional-structural plant model of hazelnut tree architecture was built. Further experiments will be needed and should be repeated over following years to extend this study toward the juvenility phase and tree architecture over time.

Keywords: Hazelnut tree, *Corylus avellana*, architecture, growth, branching, buds, modelling.

1. INTRODUCTION

In recent decades, functional-structural plant models (FSPMs) have been developed to describe the growth and development of trees and crops (DeJong 2019; Louarn and Song 2020). FSPMs have the peculiarity of coupling two different sub-models: one reproducing the architectural part of the tree or crop and the other simulating its functions, in interaction with environmental conditions (Room *et al.* 1996; Sievänen *et al.* 2014). The construction of an FSPM, usually, starts from the architectural part to which physiological models are added (for reviews see Grisafi et al., (2021) and Vos et al., (2010)). Architecture has a crucial role in the growth and development of a fruit tree crop. Indeed, it influences how the light reaches the leaves in the canopy and, consequently, the photosynthetic process. It determines where the different organs are within the plant and, therefore, the carbon partitioning. Thus, knowing the position of the flowers on the branches and, consequently the location of the fruits, enables defining where the production is and how fruits can receive water and carbohydrates during their development.

Hazelnut is an emerging fruit crop and its cultivation is increasing worldwide (FAOSTAT 2020). The world leader producer is Turkey, followed by Italy, United States and Azerbaijan (FAOSTAT 2020). In Italy, "Tonda di Giffoni" is one of the most appreciated cultivar thanks to its nut quality (Petriccione *et al.* 2010). Despite its importance, few attempts have been made to model its development so far. A process-based model that simulates the yield was recently developed (Bregaglio *et al.* 2020), but, nowadays, a structural model is still missing. Architectural analyses of several fruit tree crops, such as almond (Negrón *et al.* 2013), kiwifruit (Cieslak *et al.* 2007), and apple (Costes *et al.* 2008) have already been performed. However, few architectural studies have been achieved on bushy trees, except on coffee tree (Motisi *et al.* 2019) and on the ornamental rose plant

(Crespel *et al.* 2013). Since hazelnut (*Corylus avellana*) has a bushy shape, this study could be particularly interesting for this species.

Hazelnut is a monoecious species with male flowers, grouped into inflorescences called catkins, and female flowers, grouped into inflorescences, called glomerules, that are located in mixed buds. Male and female flowers bloom in full winter while the fecundation process occurs in late spring. During this period of time, bud breaking occurs followed by stem elongation. Along one-year-old shoots, the successive nodes can be latent, bear a sylleptic shoot with male flowers (catkin), or have more than one axillary buds whose fate can be mixed, or vegetative. Vegetative buds will give birth, the following year, to a new vegetative shoot called proleptic shoot. Mixed buds behave as vegetative buds except for the presence of the glomerules at their apex that, if the fecundation succeeded, will become a cluster of nuts (Germain and Sarraquigne 2004). Several studies describe the biology and physiology of reproductive organs of hazelnut (Mehlenbacher 1991; Germain 1994). However, they paid no attention to the dynamic of axillary shoot emergence and on the position of catkins in sylleptic shoots i.e., shoot developing immediately without a resting period (Germain and Sarraquigne 2004).

Plant structure is the result of two fundamental processes: **apical growth**, from the apical meristem and **branching processes** from axillary meristems (Gifford and Foster 1987). In perennial plants growing in temperate climates, those processes are usually investigated, over different years, during winter because the plant structure is more accessible without leaves. To investigate branching, it has been proposed to focus on one-year-old shoots and record the type of bud at each node (Godin *et al.* 1997; Costes and Guédon 2002). The node scale is appropriate when there is a strong effect of its position along the shoot on the type of bud and the development of new shoots (Caraglio and Barthélémy 1997). Moreover, the position of the longest lateral shoots along a parent shoot allows distinguishing

acrotonic (i.e., apple (Lauri 2007)) or basitonic (i.e., olive (Bongi and Palliotti 1994)) branching. Those **qualitative descriptions** can be complemented by **quantitative analyses** that permit gaining a deeper understanding of the topological relationship between the different organs that compose the plant (Guédon *et al.* 2001; Durand *et al.* 2005). The most widely used analyses for branching patterns have been **Markovian models** (Taylor and Karlin 1998) that permit the identification of homogeneous zones (i.e. parts of the shoot where buds of consecutive metamers have the same fate) within the shoot and the estimation of their characteristics, e.g., zone composition and length (Guédon *et al.* 2001; Durand *et al.* 2005). Recent studies have proposed **generalized-linear models** (GLMs) to analyse the relationship between variables related to new shoots (e.g., burst, length) with one or more predictors linked to the bearer (e.g., length of the bearer shoot, node rank) (Boudon *et al.* 2020). Those models are easy to estimate and can be used *per se* or as complementary analysis to be incorporated into more complex ones (Dambreville *et al.* 2013).

In this study, a detailed characterization of "Tonda di Giffoni" architectural development including observation of shoot emergence positions and dynamics was performed. For this, specific observations, over two successive years, were conducted to understand how plant elements are connected to each other and how they contribute to tree architecture. Those architectural information were then used to code the first hazelnut FSPM.

In this work, a focus is made on the analysis, by successive steps, of hazelnut "Tonda di Giffoni" architectural development and the quantification of bud fates depending on the developmental time and location along their parent shoot, which constitute elementary knowledge required before the implementation of a model of hazelnut (*C. avellana)* architecture.

2. MATERIAL AND METHODS

2.1. DATASET

The experiment was carried out in Deruta, Perugia (Italy) in 2020 and 2021. The experimental orchard contained 140 seven-years-old own-rooted hazelnut trees (*C. avellana* Tonda di Giffoni), planted in 2014 at 4x4m distances. Standard horticultural practices were applied. In January 2020, 104 one-year-old shoots (26 shoots per tree on four plants) of hazelnut trees were selected according to four length categories: short (Sh) when shorter than five cm, medium (Me) when between five and 20 cm, long (Lo) when between 20 and 40 cm, and very long (VLo) when longer than 40 cm. On those shoots, biometrical measurements (i.e., diameter, length, and number of nodes) were performed during winter. From the base to the shoot tip, at each node, the number and fates of buds were recorded. Four types of fates were known in *C. avellana*: latent bud (i.e., when no bud was present, B), vegetative bud (V), mixed bud (M), and male flower (catkins = C). Each of them can be easily recognized, on the shoot, thanks to their particular shape (Figure 1A)

In January 2021, the same biometrical measurements and node buds' inventory were conducted on lateral one-year-old-shoots born from vegetative and mixed buds of shoots previously selected in 2020. In the following, one-year-old shoots of 2020 will be considered as **parent shoots** or **bearers**, while the shoots born from their buds in 2021 will be named **children shoots.**

2.2. STRUCTURE OF THE ANALYSIS

A logigram was designed to formalise developmental processes involved in the generation of the growth and branching patterns in hazelnut (Figure 2). Each box in the diagram represents a question that needs to be addressed to permit the development of a hazelnut architectural model. Developmental processes occurred at different scales: shoot

scale (represented by circled boxes), node scale (square boxes), and bud scale (rhomboid boxes). The logigram started with the assumption that the first shoot was proleptic. The first question step was to estimate the number of nodes of this shoot. In a second step, the probability, for each node, to be blind had to be estimated. If the node was blind, the model considers that node as latent and the process aborts. In case the node is not blind, the probability of carrying or not a sylleptic lateral shoot had to be computed at node scale. According to the value of this probability, different developmental options were considered (YES and NO arrows). In the case of a sylleptic lateral shoot, the number of buds composing this shoot and, subsequently, the proportion of V and M were determined. Then the probability of bursting was considered and the length of new shoots was computed. In the opposite case (i.e., NO arrow), the number of lateral buds and their fate (i.e. V or M) were determined at each node along proleptic shoots. For the shoots developed from either mixed or vegetative buds, the length of the new lateral shoots was computed. At the very end, the logigram updated the order of the shoot considered from n (i.e., order of the bearer shoot) to n+1 (i.e., order of the new shoot), to start again to the different steps for a new year of development (Figure 2).

2.3. STATISTICAL ANALYSES

FSPM parametrization. All statistical analyses were performed using RStudio (R Core Team 2022). For each box in the logigram, exploratory analyses were performed to identify the variables that could be involved in the current phenomenon. When a bud fate was observed in the median part of proleptic one-year-old shoots, the distance of each node from the median one was considered as a predictor. This distance was normalized by dividing it by the total number of nodes in the shoot. The variables selected from the exploratory analyses were, then, tested through more complex statistical models (i.e. GLMs). The relationship between the length of a shoot in cm and the number of nodes carried by that shoot was estimated using

a polynomial linear model (S1, **[Supplementary material]**). The probability, of a node, to be blind was computed using a binomial GLM with node rank as predictor (S2, **[Supplementary material]**).

For estimating the probability of a node to develop a sylleptic shoot, depending on the absolute value of the normalized distance from the median node rank, a binomial GLM was used (MOD1, S3 **[Supplementary material]**). The number of buds, either in proleptic or in sylleptic shoots were estimated using a Poisson GLM (S4 and S5, **[Supplementary material]**). In sylleptic shoots, the number of V was estimated as the proportion of V buds, computed with a binomial GLM (S6, **[Supplementary material]**) multiplied by the number of buds in sylleptic shoots (S4, **[Supplementary material]**). Then the number of M was computed subtracting the number of V buds from the total number of buds of that shoot. This strategy was chosen because there was no significant correlation between the number of V or M and any variable. In proleptic shoots, the proportion of M, V buds was estimated using a multinomial regression model (MRM) (MOD2, S7 **[Supplementary material]**). This model was chosen because those proportions can be considered as an ordinal multinomial distribution. The probability of a bud to burst (or not) into a new shoot was predicted using a binomial GLM with interactions with other buds in proleptic and sylleptic shoots (MOD3, MOD4, S8, and S9 **[Supplementary material]**). The interactions were examined between the probability of bursting and the fate of the bud itself (i.e., mixed or vegetative) and the presence, in the same sylleptic shoot or in the same node, of other buds (mixed and/or vegetative). The next step was to predict the length of the new shoot. In the case of new shoots born from sylleptic shoots, "fitdistrplus" package (Delignette-Muller and Dutang 2015) was used to find the best length distribution (probability density function) (MOD5, S10 **[Supplementary material]**). The length of the new shoots from proleptic shoots was

modelled using a gaussian GLM with, as predictors, bud fate (either V or M), shoot length (cm) and normalized distance (MOD6, S11 **[Supplementary material]**).

All the GLMs were estimated using "stats" package (R Core Team 2022), while "nnet" package (Ripley and Venables 2022) was used to run MRM.

Each model was first run including in the equation all the possible predictors selected through exploratory analyses. The least significant component of the equation (i.e., the one with the highest p-value) was selected and tested into a permutational model (with p-value set as 0.001) that could either exclude or maintain it in the equation, based on AIC (Burnham and Anderson 1998). AIC was used to compare several, possibly non-nested, regression models: the model with lowest AIC was selected. However, AIC was subjected to some random noise present in data, so that AIC differences were also noisy. The logic of permutation models consisted in removing the effect to be tested by shuffling the data set and comparing the AIC in the true data set (potentially containing the effect) with the AIC in the model obtained from shuffled data set (not containing the effect) by construction and referred to as "shuffled model". This was expected to prevent AIC differences from being explained by chance. The shuffled model had the same equation as the original model and the same dataset, except for the selected predictor, whose data were shuffled. The shuffled model was run 10000 times and, each time the difference between its AIC and the AIC of the null model was computed. If the AIC differences between the shuffled model and the null one was lower than the AIC differences between the original model and the null model, at least in one of the permutations, the selected parameter was discarded. Then, a new model was run again excluding, from the original equation, the selected predictor. This process was reiterated until all the predictors in the equation showed significance. Table 1 summarizes all the variables (either discarded, selected and not tested) for each model shown in the paper.

FSPM implementation. The statistical outputs of the former models (i.e. estimates, standard errors) were used to develop a first model of hazelnut architectural development. The implementation was done using L-Py (Boudon *et al.* 2012) in the OpenAlea platform (Pradal *et al.* 2008). L-Py provides a programming environment that combines L-system formalisms (Prusinkiewicz and Lindenmayer 1990) in Python language and has been already used in other fruit tree crop models (Costes *et al.* 2008; Boudon *et al.* 2020). The simulation starts in 2018 with an axiom composed of a trunk with 10 proleptic buds. Then, the juvenile phase is simulated for one year. During juvenility, the plant, is mainly focused on vegetative production (Borchert 1976). Thus, in the model, during this phase, all the buds are proleptic and vegetative. Due to the lack of data regarding the length of new shoots during the juvenility phase, this equation, in hazelnut FSPM, was obtained changing MOD6 estimates and (S11 **[Supplementary material]**). In particular, it was increased the effect of the normalized distance from median node rank. After the juvenile phase, in 2020, the production starts and lasts for two years as described in the previous paragraph.

FSPM evaluation. Model validation was performed using Python and RStudio (R Core Team 2022) by comparing the number and proportions of axillary productions, i.e. blind nodes, sylleptic shoots and types of buds, between a subset of one year old simulated proleptic shoots and the observed datasets collected in Deruta in 2020. To perform this evaluation 1000 simulations were run in L-Py. To make the two dataset (i.e. observed and simulated) comparable, for each L-Py simulation, 104 one-year-old shoots were randomly sampled according to the same length category distribution of the observed dataset (i.e. 26 Sh shoots, 25 Me shoots, 28 Lo shoots, 25 VLo shoots). Just 576 simulations out of 1000 had enough one year old shoots to enable the random sampling. The number of nodes, the proportion of blind nodes, the proportion of sylleptic shoots and the proportion of M and V buds in proleptic shoots was compared between the two subsets of shoots. Prop.test of the “stats”

package in R (R Core Team 2022) was used to evaluate if there were significant differences in the proportion of sylleptic shoot, blind nodes, and M and V buds in proleptic shoots.

3. RESULTS

The branching behaviour was observed at both the tree and shoot scales. At the tree scale, long shoots are present at the base and are responsible for its bushy global shape (Figure 3). On the contrary, at the shoot scale, the longest lateral shoots were present in the distal position of one-year-old shoots and this was characteristic of an acrotonic gradient. The direct observation of proleptic shoots also revealed that the apical bud died during the growing season (Figure 1C). This suggested Tonda di Giffoni has a sympodial branching mode.

In the following, results related to the growth and branching of one-year-old shoots will be presented through two sub-paragraphs: the first one for sylleptic shoots, with the analysis of the production within the same year and the second one for proleptic shoots, with the analysis of the production within the following year.

3.1. **Growth and branching of sylleptic shoots.**

The direct observation of developmental time of laterals revealed that the shoots bearing catkins were sylleptic shoots (Figure 1D). This diagnostic was based on the presence of those axillary shoots prior to the other (proleptic) ones and on the absence of visible scars at their base. Moreover, those shoots were located in the median zone of the one-year-old bearer shoot. We thus used the absolute value of normalized distance, in MOD1, to estimate the probability of having a sylleptic shoot at a particular node. The maximal probability was 0.43 in the median part of the shoot and decreased toward the shoot extremities (Figure 3 and S3 **[Supplementary material]**). The number of buds in sylleptic shoots was estimated based on the length of the bearer shoot (cm) and the absolute value of normalized distance. The

total number of buds (vegetative and mixed) increased with the length of the bearer shoot and with the proximity of the median zone from a minimum of one to a maximum of seven. The proportion of vegetative was estimated using a binomial GLM (S6 **[Supplementary material]**). Then, the probability of each bud (either M or V) to burst was computed using MOD4. For mixed buds, the probability of bursting decreased from 0.72 to 0.06 depending on the number of other M present within the sylleptic shoot and from 0.53 to 0.18 depending on the number of other V present within the sylleptic shoot (Figure 4 and S9 **[Supplementary material]**). For vegetative buds, the probability of bursting was not influenced by the presence of other M within the sylleptic shoot while it decreased from 0.73 to 0.13 with the number of other V present within the sylleptic shoot (Figure 4 and S9 **[Supplementary material]**). The length of new shoots was not related to any of the tested predictors. Thus, the new shoot length distribution, born either from M or V buds, was approximated to a Gamma distribution with α=2.37 and β=1.20 (Figure 5, S10 **[Supplementary material]**).

3.2. **Growth and branching of proleptic shoots.**

The nodes that did not bear a sylleptic shoot were having one or more axillary buds at each node. The mean number of buds (i.e., M, V) per node rank was equal to 1.13 ± 0.02 (± standard error). Then, the proportion of each bud type at a given node was estimated by MOD2, i.e., a multinomial regression (Figure 6 and S7 **[Supplementary material]**). Mixed buds (M) were mostly located (0.68%) at the median part of the shoot (rank 4-11) and less frequent (0.34%) in the proximal and distal parts (ranks <4 and >11, respectively)). Vegetative buds had their maximum frequency (0.60%, 0.54%) in the proximal and distal parts of the shoot (rank < 4 and rank >13, respectively) (Figure 6 and S7 **[Supplementary material]**).

Regarding the probability of bud burst, if the fate of the axillary bud was blind (B), there was no shoot to be developed and the probability was zero. If the fate was either V or M, the probability of bursting was computed by MOD3**.** The probability of V buds bursting depends on their position on the shoot. It decreased from 0.97 to 0.62 from the distal to the proximal part, independently from the presence of other buds at the same node. However, the presence of other buds (either M or V) in the same node decreased the bursting probability. For example, in the distal part, the probability decreased from 0.97 to 0.02 when the number of other buds in the same node (either M or V) increased from zero to eight (Figure 7, S8 **[Supplementary material]**). The probability of M buds bursting depends on their position on the shoot. It decreased from 0.96 to 0.68 from the distal to the proximal part, independently from the presence of other buds at the same node. However, the presence of other buds (either M or V) in the same node decreased the bursting probability. For example, in the distal part, the probability decreased from 0.96 to 0.03 when the number of other buds in the same node (either M or V) increased from zero to eight (Figure 7, S8 **[Supplementary material]**).

When all the proleptic buds, whatever they will sprout or not, were considered, the mean number of lateral shoots per node increased with node rank from 0.78 ± 0.05 (± standard error) to 1.30 ± 0.15 (± standard error) whatever they were bearing more than one bud or not (S12 **[Supplementary material]**). When nodes with at least one lateral shoot were considered only, the mean number of lateral shoots per node was no longer affected by the node rank and was 1.03 ± 0.01 (± standard error) (S13 **[Supplementary material]**). However, in both cases, the mean number of laterals per node presented a large variation depending on the total number of buds (either V or M) at the node. Indeed, this mean number of laterals increased with the number of buds per node, and when all buds were considered, it varied from 0.89 ± 0.04 (± standard error) to 4.65 ± 1.75 (± standard error) from nodes with

one bud to nodes with nine buds (S12 **[Supplementary material]**). This mean number slightly varied when we considered nodes with one lateral at least (from 1.01 ± 0.04 (± standard error) to 4.14 ± 1.58 (± standard error) from nodes with one to nine buds (S13 **[Supplementary material]**).

The length in cm of the new lateral shoots (Figure 8) was estimated by MOD6. It varied depending on the type of the bud from which the shoot was born (V or M), on the length of the bearer shoot, and the normalized distance from the median node rank. The length of new shoots developed from V buds was 4.14 ± 0.29 cm (± standard error) on average but largely depended on the node rank. It decreased from 6.84 ± 0.28 cm when the bud was located in distal position to 0.58 ± 0.35 cm in proximal position (Figure 8 and S11 **[Supplementary material]**). The length of new shoots developed from V buds also depended on the bearer shoot length. It increased from 2.13 ± 0.25 cm when the bearer shoot was short (5cm) to 6.52 ± 0.33 cm for long bearer shoots (70cm) (Figure 8 and S11 **[Supplementary material]**). The length of new shoots developed from M buds was 2.89 ± 0.08 cm on average and was shorter than those developed from V buds. As for laterals developed from V buds, it depended on the node rank and the bearer shoot length. It decreased from 3.42 ± 0.30 cm to 1.96 ± 0.40 cm when the bud was located in distal or proximal position, respectively (Figure 8 and S11 **[Supplementary material]**). The length of new shoots developed from M buds increased from 1.57 ± 0.24 (± standard error) to 4.15 ± 0.32 cm when the bearer shoot was short (5cm) or long (70cm), respectively (Figure 8 and S11 **[Supplementary material]**).

3.3. Model evaluation

The overall correlation between the length of one year old shoots and number of nodes is similar in the observed and simulated datasets. All the black points (i.e. observed data) appear to be in the range of the simulated data (i.e. red dots) (Figure 9). However, when shoot length is between 10cm and 25cm, the simulated number of nodes is overestimated (p-value < 0.05, t-test). In fact, the simulated number of nodes is 8.71 ± 0.02 while the observed number of nodes is 8.00 ± 0.31. The proportion of blind nodes was slightly but statistically different (p-value < 0.05) between observed and simulated data (0.13 and 0.11, respectively) (Figure 10). The proportion of sylleptic shoots was comparable and not statistically different between observed and simulated data (0.22 and 0.23, respectively) (Figure 11). The proportion of mixed and vegetative buds, along one year old proleptic shoots, was in a similar range between observed and simulated data: 0.46 and 0.48 , for mixed buds (Figure 12); 0.40 and 0.41, for vegetative buds (Figure 12).

4. DISCUSSION

Few attempts have been made so far to investigate the activity of meristems in hazelnut trees to understand the tree and shoot development. Although it was already described that catkins were located laterally on proleptic shoots (Germain and Sarraquigne 2004), we clarified that they are located terminally only on shoots that had grown in the same year as the parent shoot and therefore are sylleptic shoots (Figure 1D). Indeed, throughout direct observations and measurements, we noticed that catkins are present in the apical position of quite short shoots, with five-six axillary buds. Furthermore, the branching of hazelnut has been previously described, at the tree level, as basitonic (Botta and Valentini 2018). Here, we proposed to make a distinction between the branching mode at tree scale and at shoot scale (Figure 1B and Figure 1C) as suggested by Germain and Sarraquigne (2004).

Our observation clearly highlighted that a basitonic branching at the tree scale is compatible with an acrotonic at the shoot scale. While acrotony is usually considered as being driven by apical dominance, the basitony and the activation of buds at the tree base have generated few investigations. Even though Champagnat (1978) has described basitony as an autumnal gradient, resulting from a higher axillary bud growth potential of basal buds prior to dormancy, which is inverted during winter in acrotonic species contrary to basitonic species (Champagnat 1978), the physiological conditions in which such a gradient is established and maintained after dormancy are still to be investigated. In addition, few studies have been carried out for defining if hazelnut branching mode was monopodial (i.e., when the apical meristem remains alive and dominant) or sympodial (i.e., when the apical meristem dies) (for a review see Costes et al. (2006)). Solar and Štampar (2005) have observed a sympodial branching on two Slovenian cultivars, 'Istrska dolgoplodna leska' and 'Pauetet'. In the current study, we consistently observed a sympodial branching mode in *C.avellana* Tonda di Giffoni where the death of the apical bud was observable in proleptic shoots during the growing season (Figure 1C).

We used a logigram to organise the successive steps that need to be considered for modelling shoot growth and branching depending on the shoot and bud types (Figure 2). This approach, previously used to implement the FSPM V-Mango, was helpful in the hazelnut case as the elementary processes are shared among species. In hazelnut, the branching pattern was found to be organised by zones along the bearer shoot, as previously found in several fruit species (apple tree (Guédon *et al.* 2001), peach tree (Fournier 1994), almond (Negrón *et al.* 2013), apricot (Costes and Guédon 1996)). Sylleptic shoots were found more frequently in the middle part of the shoot. This is in accordance with previous literature on the above mentioned fruit species, but also on the hazelnut where in another cultivar, 'Tonda Gentile delle Langhe', the male flowers have been described as being mainly in the median part of

the shoots (Tombesi and Farinelli 2014). On the proleptic shoots, mixed buds (M) were also present mainly in the median part of the shoot. Such a location has been previously described along a neoformed parts in several prunus trees, such as peach and almond (Costes *et al.* 2014). However, this result appears contradictory with previous literature on another cultivar of hazelnut, 'Tonda Gentile delle Langhe', where female flowers have been found mainly in the distal part of the shoot (Tombesi and Farinelli 2014). This suggests that the difference among hazelnut cultivars should be further investigated. The hazelnut branching pattern also included the presence of the longest new shoots in the distal part of the parent shoot. This was consistent with the visual observation of the tree shape (Figure 1B). This acrotonic behaviour at the shoot level is in accordance with what is described in literature (Tombesi 1985; Germain and Sarraquigne 2004).

The new shoot lengths were found to depend on the type of buds they were coming from, with longer shoots developing from V buds than those developing from M buds. This may result from the determinate and preformed nature of the mixed shoots developing from M buds, with the female flower located in the terminal position (Germain and Sarraquigne 2004). On the contrary, the shoots developing from V buds are constituted of a preformed part but are likely to include neoformed part, even though the number of preformed organs remains to be determined (Gordon *et al.* 2006). We have sampled one-year-old shoots at the tree periphery, on seven years old trees. Therefore, additional observations and measurements on the basal and central part of the trees would be required to complement the current dataset and analyses with longer bearer and children shoots.

In nature there are many species that have more buds at the same node (Bell and Bryan 1991). In *Prunus* spp. one vegetative structure (i.e., vegetative bud or sylleptic shoot) could be associated with one or more reproductive structures (i.e., female flowers) (Costes *et al.* 2014). In *Vitis vinifera*, the same node can bear a prompt mixed bud and a compound

latent bud (Vasconcelos *et al.* 2009). The current study assessed that in hazelnut, several vegetative buds, or mixed buds or a combination of the two could be present at a given node. Most usually, when there are multiple buds per node, one of them is more outstanding than the others and it will be the only one to burst, as described in *Eucalyptus* spp. (Bell and Bryan 1991) and at the compound bud axil of *V. vinifera* (Vasconcelos *et al.* 2009). This could result from competition between buds, for resources such as carbohydrates and water (Rameau *et al.* 2015) as it was suggested for *Juglans regia* (Bonhomme *et al.* 2010) and *Prunus armeniaca* (Costes and Guédon 1996). However, there are cases in which two buds or more, present at the same node, could develop into different structures. In *Leucaena* spp. all axillary floral buds develop into inflorescences (Bell and Bryan 1991). In hazelnut, when there are more than one bud at the same node, most frequently just one of them will develop into a shoot (Figure 7). However, when just vegetative buds are present (i.e., no mixed buds associated), two or more of them could burst in the successive year (data not shown for proleptic shoots). This phenomenon could result from a greater competition between reproductive (flowering and fruiting) and vegetative growth than between two growth units as it has been suggested for *P. armeniaca* (Costes and Guédon 1996).

4.1. Model evaluation

The relationship between the length of one year old shoots and number of nodes could be simulated with an acceptable accuracy, when comparing the overall correlation between these two variables in observed and simulated data. Despite the simulated data included observations, when the shoot length was between 10 and 25cm, the simulation had an overestimated number of nodes (8.71 ± 0.02 and 8.00 ± 0.31, simulated and observed data respectively) (Figure 9). This problem could be due to two reasons: the first is the reduced sample size (i.e. 104 one year old shoots, approximately 25 shoots per length category). A larger sample size could lead to a more precise estimation of the correlation between length

and number of nodes. Secondly, the lack of information of the architectural rules during the juvenile phases of the hazelnut tree, has forced to obtain such rules changing MOD6 equation (S11 **[Supplementary material]**). This inference could not be correct because tree architectural traits and rules could change as the plant grows (de Reffye *et al.* 1988, Barthélémy *et al.* 1989). In addition, the model underestimates the proportion of blind nodes on proleptic shoot (Figure 10) while the proportion of sylleptic shoot and mixed and vegetative buds is correctly simulated (Figure 11 and Figure 12). Again, enhancing the sample size and repeating the same experiment for a larger number of years could reduce the model errors and improve the evaluation process.

5. MODEL AND DATA AVAIABILITY

Data are available in the open source GitHub repository: https://github.com/Corilana/L-HAZELNUT.

6. CONCLUSIONS AND PERSPECTIVES

A first developmental FSPM of hazelnut growth and branching was built (Figure 13). The development of hazelnut buds and shoots was monitored over two successive years. The logical diagram proposed, and the equations presented in this study were used to draw the visual representation of hazelnut development, in L-Py through L-systems formalism (Prusinkiewicz and Lindenmayer 1990). However, the evaluation analysis suggested that such a structural model will need additional field experiments to collect and estimate the architectural behaviour of the tree during the juvenile phase (Borchert 1976) and over a larger number of years. For this, observations should be collected over multiple successive years or on trees of different ages, to have a better view of how architectural traits change over time. Further studies will also be required to create a functional model that describes carbon partitioning within the tree in order to merge the architectural model with the functional one to obtain a more complex FSPM of hazelnut.

ACKNOWLEDGMENTS

F. Boudon and J.-B. Durand have been supported by the MaCS4Plants CIRAD network, initiated from the AGAP Institute and AMAP joint research units.

REFERENCES

Barthélémy D, Edelin C, Hallé F. **1989**. Some architectural aspects of tree ageing. *Annales des sciences forestières.* **46**: 194–198.

Bell AD, Bryan A. **1991**. *Plant form: an illustrated guide to flowering plant morphology*. Oxford: Oxford University Press.

Bongi G, Palliotti A. **1994**. Olive In: Palliotti A, ed. *Handbook of Environmental Physiology of Fruit Crops:Temperate Crops*. Cleveland: CRC Press. Inc., 165–187.

Bonhomme M, Peuch M, Ameglio T, *et al.* 2010. Carbohydrate uptake from xylem vessels and its distribution among stem tissues and buds in walnut (Juglans regia L.). *Tree Physiology* **30**: 89–102.

Borchert R. **1976**. The concept of juvenility in woody plants. *Acta Horticulturae* **5**: 21–36.

Botta R, Valentini N. **2018**. Il nocciolo. Progettazione e coltivazione del corileto. Edagricole–Edizioni Agricole New Business Media srl.

Boudon F, Persello S, Jestin A, *et al.* 2020. V-Mango: A functional-structural model of mango tree growth, development and fruit production. *Annals of Botany* **126**: 745–763.

Boudon F, Pradal C, Cokelaer T, Prusinkiewicz P, Godin C. **2012**. L-Py: An L-system simulation framework for modeling plant architecture development based on a dynamic language. *Frontiers in Plant Science* **3**: 76.

Bregaglio S, Giustarini L, Suarez E, Mongiano G, de Gregorio T. **2020**. Analysing the behaviour of a hazelnut simulation model across growing environments via sensitivity analysis and automatic calibration. *Agricultural Systems* **181**: 102794.

Burnham KP, Anderson DR. **1998**. *Model Selection and Multimodel Inference*. New-York, USA: Springer-Verlag.

Caraglio Y, Barthélémy D. **1997**. Revue critique des termes relatifs à la croissance et à la ramification des tiges des végétaux vasculaires In: Bouchon J, de Reffye P, Barthélémy D, eds. *Modélisation et simulation de l'Architecture des végétaux*. Science Update INRA, 11–88.

Champagnat P. **1978**. Formation of the trunk in woody plants In: *Tropical trees as living systems*. New York: Cambridge University Press, 401–422.

Cieslak M, Seleznyova AN, Hanan J. **2007**. Virtual kiwifruit: modelling annual growth cycle and light distribution In: *Proceedings of the 5th international workshop on functional-structural plant models*. Napier, New Zealand: Print Solutions Hawke's Bay Limited, 41–46.

Costes E, Crespel L, Denoyes B, *et al.* 2014. Bud structure, position and fate generate various branching patterns along shoots of closely related Rosaceae species: A review. *Frontiers in Plant Science* **5**: 666.

Costes E, Guédon Y. **1996**. Modelling the annual shoot structure of the apricot tree (CV lambertin) in terms of axillary flowering and vegetative growth In: *Acta Horticulturae*.21–28.

Costes E, Guédon Y. **2002**. Modelling branching patterns on 1-year-old trunks of six apple cultivars. *Annals of Botany* **89**: 513–524.

Costes E, Lauri PE, Regnard JL. **2006**. Analyzing fruit tree architecture: implications for tree management and fruit production. In: Janick J, ed. *Horticultural reviews*. Montpellier, France.: John Wiley & Sons, Inc., .

Costes E, Smith C, Renton M, Guédon Y, Prusinkiewicz P, Godin C. **2008**. MAppleT: Simulation of apple tree development using mixed stochastic and biomechanical models. *Functional Plant Biology* **35**: 936–950.

Crespel L, Sigogne M, Donè N, Relion D, Morel P. **2013**. Identification of relevant morphological, topological and geometrical variables to characterize the architecture of rose bushes in relation to plant shape. *Euphytica* **191**: 129–140.

Dambreville A, Lauri P-É, Trottier C, Guédon Y, Normand F. **2013**. Deciphering structural and temporal interplays during the architectural development of mango trees. *Journal of Experimental Botany* **64**: 2467–2480.

DeJong TM. **2019**. Opportunities and challenges in fruit tree and orchard modelling. *European Journal of Horticultural Science* **84**: 117–123.

Delignette-Muller ML, Dutang C. **2015**. fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software* **64**: 1–34.

Durand JB, Guédon Y, Caraglio Y, Costes E. **2005**. Analysis of the plant architecture via tree- structured statistical models: the hidden Markov tree models. *New Phytologist* **166**: 813–825.

FAOSTAT. **2020**. Food and Agriculture Organization of the United Nations. (1997). FAOSTAT statistical database. [Rome] :FAO. .

Fournier D. **1994**. Analyse et modélisation des processus de croissance et développement qui contribuent aux performances agronomiques du pêcher Prunus persica (L.) Batsch. Thesis École nationale supérieure agronomique (Montpellier)

Germain E. **1994**. The reproduction of hazelnut(Corylus avellana L.): a Review. *Acta Horticulturae*: 195–209.

Germain E, Sarraquigne J-P. **2004**. *Le noisetier* (Ctifl, Ed.). Paris.

Gifford EM, Foster CE. **1987**. *Morphology and Evolution of Vascular Plants.* New York, USA: W.H Freeman and Company.

Godin C, Costes E, Caraglio Y. **1997**. Exploring plant topological structure with the AMAPmod software: An outline. *Silva Fennica* **31**: 357–368.

Gordon D, Damiano C, DeJong T. 2006. Preformation in vegetative buds of Prunus persica: factors influencingnumber of leaf primordia in overwintering buds. *Tree Phisiology* **26**: 537–544.

Grisafi F, DeJong TM, Tombesi S. **2021**. Fruit tree crop models: an update. *Three Phisiology*.

Guédon Y, Barthélémy D, Caraglio Y, Costes E. **2001**. Pattern analysis in branching and axillary flowering sequences. *Journal of theoretical biology* **212**: 481–520.

Lauri PE. **2007**. Differentiation and growth traits associated with acrotony in the apple tree (Malus Xdomestica, Rosaceae). *American Journal of Botany* **94**: 1273–1281.

Louarn G, Song Y. **2020**. Two decades of functional-structural plant modelling: Now addressing fundamental questions in systems biology and predictive ecology. *Annals of Botany* **126**: 501–509.

Mehlenbacher SA. **1991**. Hazelnuts (Corylus). *Acta Hort.* **290**: 791–836.

Motisi N, Ribeyre F, Poggi S. **2019**. Coffee tree architecture and its interactions with microclimates drive the dynamics of coffee berry disease in coffee trees. *Scientific Reports*: 9–2544.

Negrón C, Contador L, Lampinen BD, *et al.* 2013. Systematic analysis of branching patterns of three almond cultivars with different tree architectures. *Journal of the American Society for Horticultural Science* **138**: 407–415.

Petriccione M, Ciarmiello LF, Boccacci P, ***et al.* 2010.** Evaluation of "Tonda di Giffoni" hazelnut (Corylus avellana L.) clones. *Scientia Horticulturae*, **124**(2): 153–158.

Pradal C, Dufour-Kowalski S, Boudon F, Fournier C, Godin C. **2008**. OpenAlea: A visual programming and component-based software platform for plant modelling. *Functional plant biology* **35**: 751–760.

Prusinkiewicz P, Lindenmayer A. **1990**. *The algorithmic beauty of plants.* (P Prusinkiewicz and A Lindenmayer, Eds.). New York: Springer-Verlag New York.

R Core Team. **2022**. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rameau C, Bertheloot J, Leduc N, *et al.* 2015. Multiple pathways regulate shoot branching. *Frontiers in Plant Science* **5**.

de Reffye P, Edelin C, Françon J, Jaeger M, Puech C. **1988**. Plant models faithful to botanical structure and development In: *Proceedings of the 15th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1988*.151–158.

Ripley B, Venables W. **2002**. Modern Applied Statistics with S, Fourth edition. Springer, New York. ISBN 0-387-95457-0,

Room P, Hanan J, Prusinkiewicz P. **1996**. Virtual plants: new perspectives for ecologists, pathologists and agricultural scientists. *Trends in Plant Science* **1**: 33–38.

Sievänen R, Godin C, de Jong TM, Nikinmaa E. **2014**. Functional-structural plant models: A growing paradigm for plant studies. *Annals of Botany* **114**: 599–603.

Solar A, Štampar F. **2005**. The architectural analysis of a fruiting branch in two hazelnut cultivars. *Acta Horticulturae* **686**: 179–186.

Taylor HM, Karlin S. **1998**. *An introduction to stochastic modeling*. Orlando, FL: Academic Press.

Tombesi A. **1985**. Il nocciolo: manuale pratico. Reda.

Tombesi S, Farinelli D. **2014**. Relationships between flower density and shoot length in hazelnut (Corylus avellana L.) In: *Acta Horticulturae*. Thompson, 137–142.

Vasconcelos MC, Greven M, Winefield CS, Trought MCT, Raw V. **2009**. The Flowering Process of Vitis vinifera: A Review. *Am. J. Enol. Vitic.* **60**: 411–434.

Vos J, Evers JB, Buck-Sorlin GH, Andrieu B, Chelle M, de Visser PHBB. **2010**. Functional-structural plant modelling: A new versatile tool in crop science. *Journal of experimental Botany* **61**: 2101–2115.

TABLES

Legend of all the tested predictors

d= distance from median rank node;

|d|=absolute value of d;

D= normalized distance from median rank node;

|D|= absolute value of D;

L= shoot length(cm)/bearer shoot length (in case of sylleptic);

R= rank of node/bearer rank of node (in case of sylleptic);

F= fate is *k* (either V or M);

N= shoot length (node)/bearer shoot length(in case of sylleptic)

S= other V and M buds in the same node;

M= number of M buds;

V= number of V buds;

Model ID	Model type	d	|d|	D	|D|	L	R	F	N	S	M	V	Equation of linear predictor in *Y*
MOD1	Binomial GLM	NO	NO	NO	**YES**	X	X	X	X	X	X	X	$b_0+b_1\|D\|$
MOD2	Multinomial regression (categories: M and V)	NO	X	X	NO	NO	**YES**	X	NO	X	X	X	M: $b_0+b_1R+b_2R^{0.5}+b_3R^2+b_4R^3+b_5R^4$
MOD3	Binomial GLM	X	X	**YES**	X	NO	NO	**YES**	X	**YES**	X	X	$b_0+b_1F_VS+b_2F_MS+b_3F_VD+b_4F_MD$
MOD4	Binomial GLM	X	NO	X	NO	NO	NO	**YES**	X	X	**YES**	**YES**	$b_0+b_1F_VM+b_2F_MM+b_3F_VV+b_4F_MV$

Legend of all the tested predictors

d= distance from median rank node;

|d|=absolute value of d;

D= normalized distance from median rank node;

|D|= absolute value of D;

L= shoot length(cm)/bearer shoot length (in case of sylleptic);

R= rank of node/bearer rank of node (in case of sylleptic);

F= fate is *k* (either V or M);

N= shoot length (node)/bearer shoot length(in case of sylleptic)

S= other V and M buds in the same node;

M= number of M buds;

V= number of V buds;

Model ID	Model type	d	|d|	D	|D|	L	R	F	N	S	M	V	Equation of linear predictor in *Y*
MOD6	Gaussian GLM	X	NO	**YES**	NO	**YES**	NO	**YES**	X	NO	X	X	$b_0+b_1F_VL+b_2F_ML+b_3F_VD+b_4F_MD$

Table 1: List and description of all models estimated. For each one, the name, all the tested predictors, and the final equation are mentioned. Predictors marked as X were not considered in the GLMs or MRM. Variables with NO are the ones excluded through the permutation steps (see. Material and Methods)

FIGURE LEGENDS

Figure 1: A) Real picture (on the right) and simplified draw (on the left) with 4 types of hazelnut structures (catkins in yellow, mixed buds in pink and vegetative buds in green). B) Schema and photo illustrating the branching behaviour in hazelnut Tonda di Giffoni, at both tree and shoot scale. C) Schema of the development of a Hazelnut shoot Tonda di Giffoni over two successive years. D) Schema and photos illustrating the branching pattern of a proleptic shoot and the sylleptic nature of shoots bearing catkins in hazelnut Tonda di Giffoni.

Figure 2: Logical diagram built to formalise the developmental processes involved in the generation of growth and branching in hazelnut. Each box represents a step that needs to be followed to draw the architecture of hazelnut. Different scales are possible: shoot scale (circled boxes), node scale (squared boxes) and bud scale (rhomboid boxes). Different shoots are considered: proleptic shoots (in blue) and sylleptic shoots (in orange). The red shapes represent different distributions: binomial (two arrows); Poisson (star); Gaussian, multinomial or gamma (triangle).

Figure 3: MOD1. Proportion of sylleptic shoots developed per node rank (computed as number of sylleptic shoots / total number of nodes), along a one-year old shoot of hazelnut Tonda di Giffoni, depending on the absolute value of the normalized distance of the current node from the median node of the parent shoot (computed as distance from median rank node / total number of nodes of the parent shoot). Bars represent observed data while the line corresponds to the MOD1 GLM equation.

Figure 4: MOD4. Proportion of new shoots developed from sylleptic shoots of hazelnut Tonda di Giffoni, depending on the bud fate (V or M) and the presence of other M or V buds in the same sylleptic. GLMs output shows significance in the interaction between fate M and

the presence of other buds (either M or V) and in the interaction between fate V and other M buds. Points represent the observed data while lines represent the predicted values with confidence interval, using MOD4 GLM equation.

Figure 5: MOD5. Density distribution of length of new shoots, born in sylleptic shoots of Hazelnut Tonda di Giffoni, whatever the type of bud (M or V). The distribution of length follows a gamma distribution with α=2.38 and β=1.20.

Figure 6: MOD2. Proportion of bud types (V, M) present at a node along a one-year old proleptic shoot of hazelnut Tonda di Giffoni, depending on the parent rank node. Blue, and yellow boxes represent the proportion of V and M respectively while the blue and yellow lines represent their predicted values, respectively, using MOD2 GLM equation.

Figure 7: MOD3. Proportion of new shoots developed at a node along a one-year old shoot of hazelnut Tonda di Giffoni, depending on the presence of other buds at the same node (either M or V), and the normalized distance of the node from the median node of the parent shoot. On the left: proportion of new shoots from V buds; on the right: proportion of new shoots from M buds. Different colours represent different normalized distances: red= distal position, blue = median position, green = proximal. Squares, dots and triangles represent observed data while lines represent the predicted values with their confidence interval, using MOD3 GLM equation.

Figure 8: MOD6. Length of new lateral shoots developed at a node along one year old proleptic shoots of hazelnut Tonda di Giffoni, depending on bud fate, parent shoot length and the normalized distance of the node from the median rank node of the parent shoot. On the left: length of new shoots from V buds; on the right: length of new shoots from M buds. Different colours represent different normalized distances: red= distal position, blue =

median position, green = proximal. Squares, dots and triangles represent real data while lines are the predicted values and their confidence interval, using MOD6 GLM equation.

Figure 9: Number of nodes of one year old proleptic shoots of hazelnut Tonda di Giffoni, depending on their length. Different colours represent different datasets: black = data observed in Deruta fields in January 2020; red = data obtained from 576 simulations in L-Py using S1 [Supplementary material] linear equation.

Figure 10: A) Proportion of blind nodes in one year old proleptic shoots of Hazelnut Tonda di Giffoni. Different colours represent different datasets: black = data observed in Deruta fields in January 2020; red = data obtained from 576 simulations in L-Py using S2 [Supplementary material] linear equation. B) Proportion of nodes with sylleptic shoots in one year old proleptic shoots of Hazelnut Tonda di Giffoni. C) Proportion of M and V buds in one year old proleptic shoots of Hazelnut Tonda di Giffoni.

Figure 11: Visual result of Tonda di Giffoni FSPM. The architecture is simulated over two successive years in L-Py. The simulation was run without leaves (A) and with leaves (B).

Figure 1

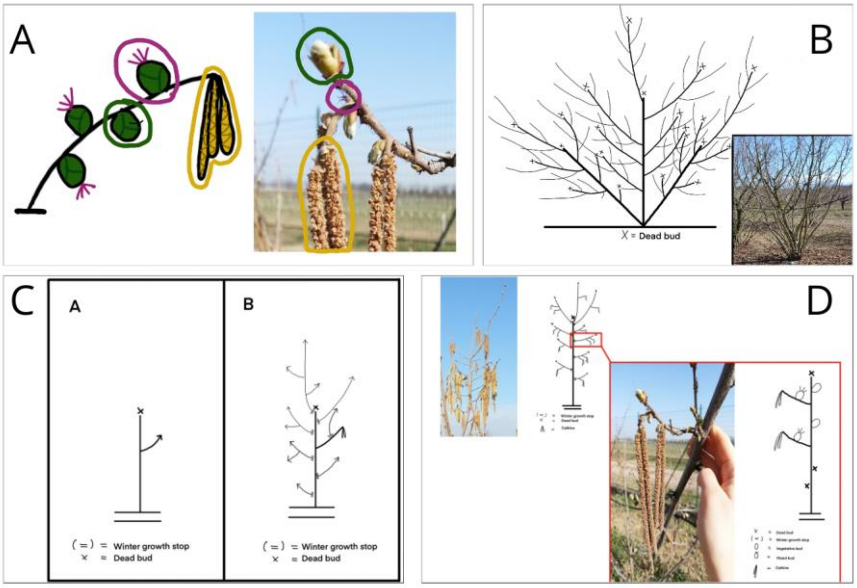

Figure 2

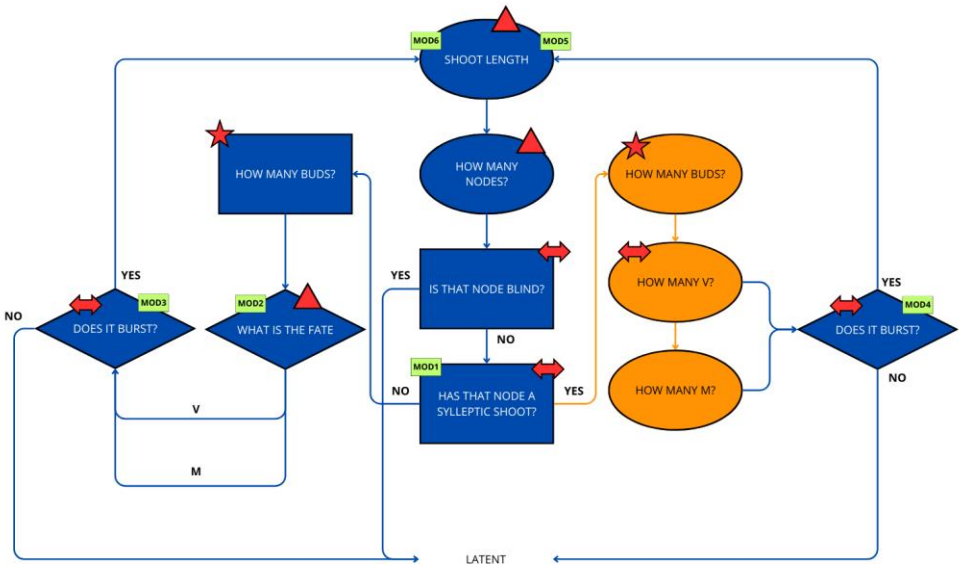

Figure 3

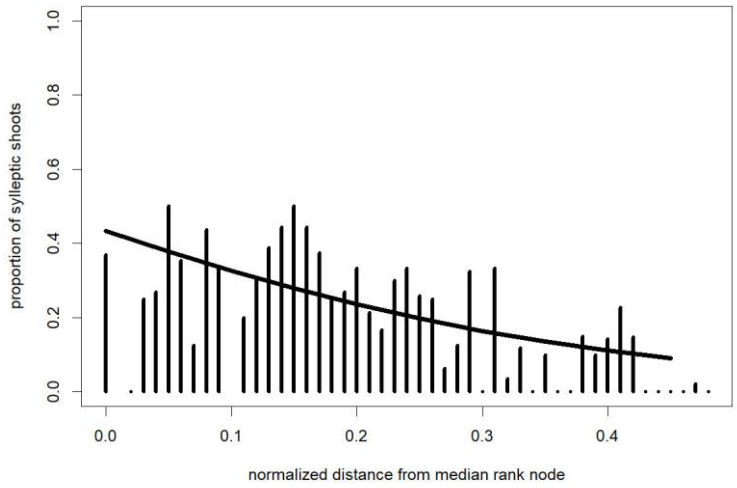

Figure 4

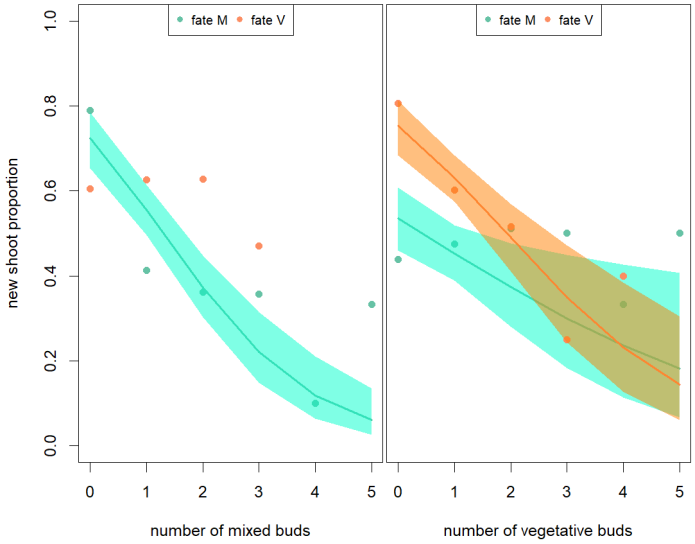

Figure 5

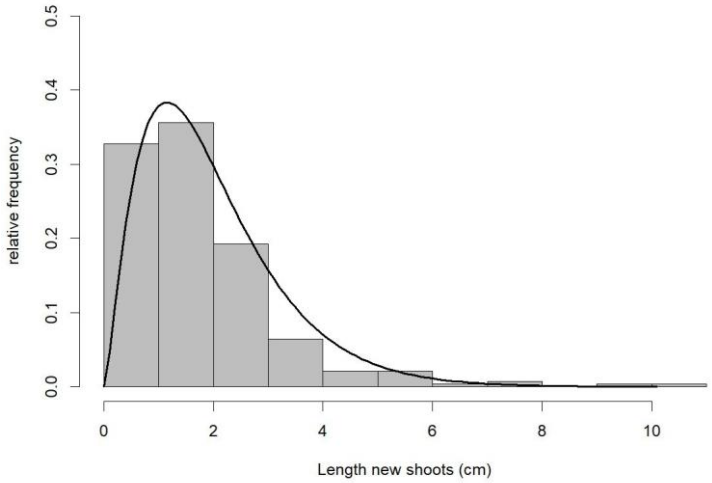

Figure 6

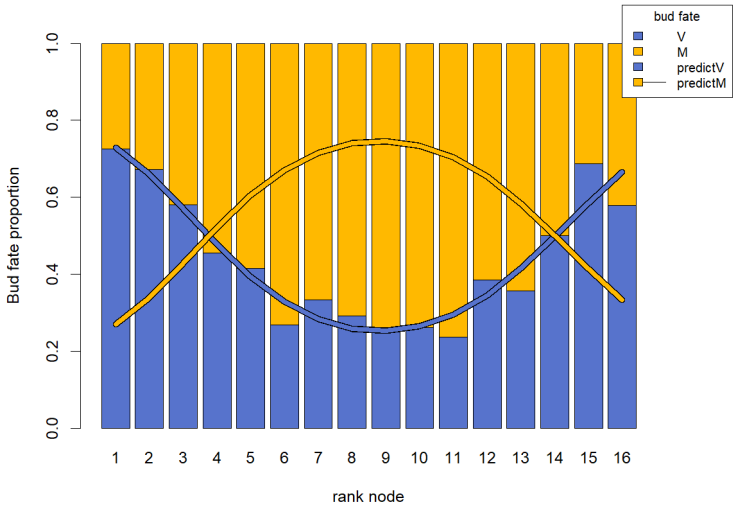

Figure 7

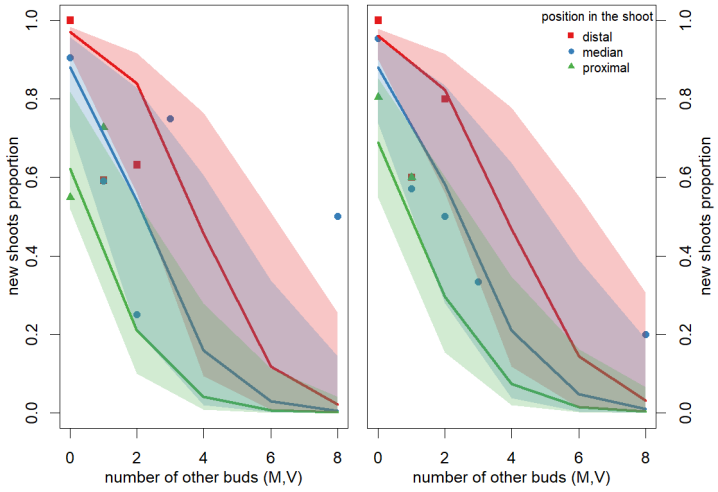

Figure 8

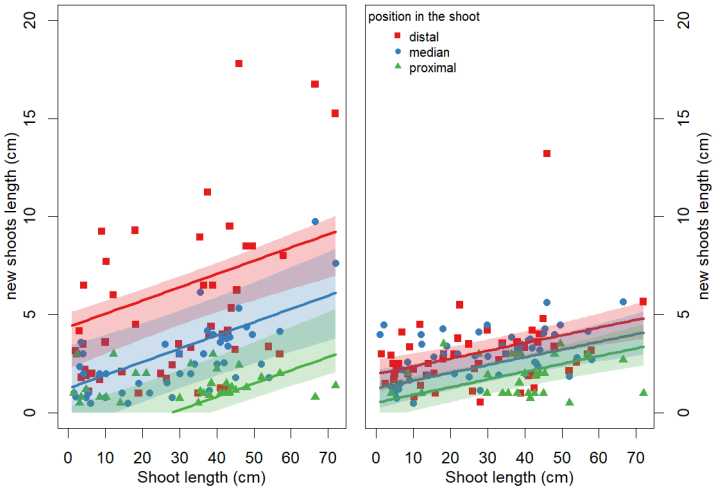

Figure 9

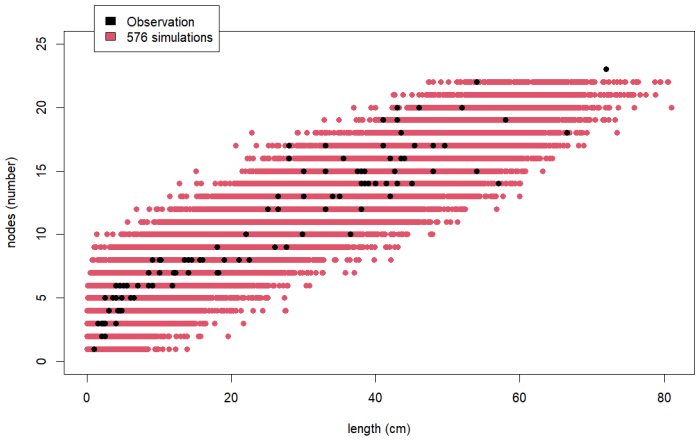

Figure 10

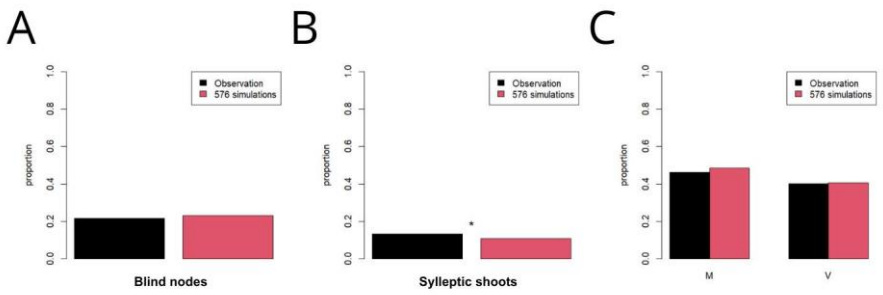

Figure 11