Statistical characterization of the root system architecture model CRootBox

A. Schnepf[1,*], K. Huber[1], M. Landl[1], F. Meunier[2], L. Petrich[3], V. Schmidt[3]

[1] Institut für Bio- und Geowissenschaften: Agrosphäre (IBG-3), Forschungszentrum Jülich GmbH, Wilhelm-Johnen-Str., D-52425 Jülich, Germany.

[2] Earth and Life Institute, Université catholique de Louvain, 15-de Serres, Croix du Sud 2, 1348 Louvain-la-Neuve, Belgium

[3] Ulm University, Institute of Stochastics, Helmholtzstr. 18, D-89069 Ulm, Germany.

* corresponding author, Email: a.schnepf@fz-juelich.de

Keywords: root architecture models, CRootBox, statistical analysis, meta-models, stochasticity

Core ideas

- The stochasticity of 3D root architecture models needs to be recognized by statistical analysis.
- Probability density functions, regression and correlation analyses elucidate the impact of model input parameters on different CRootBox measures.
- Distributions of ratios of root system measures are highly asymmetric.
- Multivariate approaches (copulas) are envisioned for future root architecture model analysis.

Abstract

Background and aims The connection between the parametrization of 3D root architecture models and characteristic measures of the simulated root systems is often not obvious. We used statistical methods to analyze the simulation outcome of the root architecture model CRootBox and build meta-models that determine the dependency of root system measures on model input parameters.

Methods Starting with a reference parameter set, we varied selected input parameters one at a time and used CRootBox to compute 1000 root system realizations as well as their root system measures. The obtained data sets were then statistically analyzed with regard to dependencies between input parameters, as well as distributions and correlations between different root system measures.

Results While absolute root system measures (e.g. total root length) were approximately normal distributed, distributions of ratios of root system measures (e.g. root tip density) were highly asymmetric and could be approximated with inverse gamma distributions. We derived regression models (meta-models) that link significant model parameters to 18 widely used root system measures, and determined correlations between different root system measures.

Conclusions Statistical analyses of 3D root architecture models improve understanding of the impacts of input parametrization on specific root architectural measures. Our developed meta-models can be used to determine the effect of parameter variations on the distribution of root system measures without running a full simulation. Model intercomparison and benchmarking of root architecture models is still missing. Our approach provides a means to compare different models with each other and with experimental data.

Introduction

Root architecture and phenotypic plasticity determine a plant's success to acquire below-ground resources such as water and nutrients (*Lynch* 2007). Experimental investigation of root system development, however, is a laborious task due to the opaque nature of the soil that makes direct measurements difficult. Researchers therefore usually resort to simplified experimental designs and derive root system measures from plants grown in artificial systems such as hydroponics or rhizotron boxes (*Nagel et al*. 2015, *Atkinson* et al. 2015) or from observations through rhizotubes in the field, where only small sections of the root system are visible (*Garré et al.* 2012). Several studies, however, showed that lab and field derived phenotypic root properties are poorly related and extrapolation from single root observations to the entire root system is delicate (*Wojciechowski et al*. 2009, *Poorter* et al. 2016).

A possibility to overcome these limitations are 3D root architecture models, which allow to generate large numbers of different root systems from a range of plant-specific and environmentally influenced input parameter sets (*Schnepf et al*. 2018) and use them to deduce characteristic root traits. Models of root architecture and function have become readily available (*Dunbabin* et al. 2013, *Schnepf et al.* 2018, *Postma et al.* 2017, Pagès *et al.* 2014) and provide a means to efficiently analyze different plant species and their performance in different environments (*Meunier et al.* 2016). Most models use comparable input parameter sets that include parameters influencing the total size of a root system (e.g. growth speed, branching density) as well as the shape of a root system (e.g. tropism and tortuosity parameters) (*Bingham and Wu* 2011). Furthermore, most models have stochastic components such that the same parameter set may produce many different realizations. However, to our

knowledge, none of these models has so far been subject to thorough statistical analysis regarding the dependence of root system measures on model parameterization. We chose a terminology which explicitly distinguishes between the “model input parameters”, from which we compute the 3D root architecture using CRootBox, and statistical “measures”, which we compute from the resulting 3D root architecture. Both could be ecological “traits”. For example both the model input parameter “branching angle” and the “root length density” could be considered traits.

Measures to characterize root architectures

Classical measures to characterize root system architecture include total root length, root surface area and root volume. While total root length is related to the soil volume explored by the root system, root surface area is important for uptake and exudation mechanisms that occur across the root-soil interface, and root volume can be seen as a measure of carbon investment into a specific root structure. The number of branches (or number of root tips) gives information about the degree of branching within a root system. Maximum rooting depth and maximum horizontal spread of the root system are negatively correlated and determine whether the root system is of steep and deep (*Lynch* 2013) or of shallow appearance which has direct implications on root foraging. While deep rooting plants can take up water from deeper soil layers and are thus advantageous in dry climates and during drought periods, shallow rooting plants enhance the exploration of topsoil layers where nutrient availability is greatest in many soils (Lynch and Brown 2001). Irrespective of the specific location, the convex hull determines the smallest convex set that encloses the whole root system, while the rhizosphere volume is a measure of the soil volume that is actually influenced by the root system. The size of the rhizosphere volume depends on the effective soil diffusion coefficient, which varies for different nutrients, as well as on

root age, root length and radius and overlap between rhizospheres of individual rootaxes, i.e. the foraging performance (Landl et al. 2018).

To compare root systems of different plant species with each other, ratios of root system measures are used. Normalizing the number of root tips, total root length, root surface area and root volume by the volume of the convex hull results in root tip density, root length density, root surface area density and root volume density. We chose the convex hull since our simulations refer to the root system of a single plant. In field conditions, the volume of convex hull would be replaced by the volume of the sampled soil and potentially, root systems from several plants would contribute to the root system measures. While root length density is one of the most widely measured traits in many lab or field experiments (*Zuo et al.* 2004, *Van Noordwijk et al.* 1985), root surface area density is the most relevant parameter in water flow and solute transport models (*Couvreur et al*. 2014).

Root water and nutrient uptake as well as transport towards the shoot is determined by root hydraulic properties, which are thus - next to root architecture parameters - key components for root system functioning (*Vadez* 2014). To describe the hydraulic architecture of an entire root system, root hydraulic properties such as radial and axial conductivities are related to root system measures. Root hydraulic architecture measures include root system equivalent conductance (K_{rs}) and standard root water uptake fraction (SUF), which represent respectively the ability of the root system to take up a certain water volume under a given water potential difference between the root collar and an homogeneous soil and the water uptake by a root segment relative to the total water uptake of the root system. These variables were calculated by solving the water flow in the generated root system architectures under homogeneous soil conditions, using the algorithm of Meunier et al. (2017). A water potential was

imposed at the root collar and the resulting stem sap flow was used to calculate the root system conductance. The water uptake by each single segment served then to derive the standard uptake fraction of each individual segment. The root hydraulic conductivities (function of root age and order) were taken from Doussan et al. (2006) and considered as identical for all root systems. The mean depth of standard root water uptake is then the product of SUF and the depth of the respective root segment summed up over all segments. To allow comparisons of the hydraulic architecture of differently sized root systems, the root system equivalent conductance is normalized by root length respectively root surface area (*Couvreur et al.* 2012). A further important measure for the characterization of root water uptake is the root half mean distance (*HMD*) which affects water depletion in the soil and is

Table 1: Characteristic root system measures

Variable	Units	Name and description
RL	cm	total root length
RSA	cm^2	total root surface area
RV	cm^3	total root volume
z_{max}	cm	maximum rooting depth
r_{max}	cm	maximum horizontal spread (radius of the confining cylinder)
Conv	cm^3	volume of the convex hull
NR	-	number of root tips / branches
V_{rhizo}	cm^3	rhizosphere volume for phosphate
RND	cm^3	*NR/conv* (root tip density)
RLD	cm cm^{-3}	*RL/conv* (root length normalized with the volume of the convex hull)
RSAD	cm^2 cm^{-3}	*RSA/conv* (root surface area normalized with the volume of the convex hull)
RVD	cm^3 cm^{-3}	*RV/conv* (root volume normalized with the volume of the convex hull)
VD_{rhizo}	cm^3 cm^{-3}	V_{rhizo}*/conv* (rhizosphere volume for phosphate normalized with the volume of the convex hull)
K_{rs}	cm^2 d^{-1}†	equivalent conductance of the root system
$K_{rs\ A}$	d^{-1}	K_{rs}*/RSA* (equivalent conductance of the root system per unit root area)
$K_{rs\ L}$	cm d^{-1}	K_{rs}*/RL* equivalent conductance of the root system per unit root length
z_{SUF}	cm	mean depth of standard root water uptake
HMD	cm	half mean distance between roots

†$(cm^2\ d^{-1}) \cong (cm^3\ hPa^{-1}\ d^{-1})$

approximated with the classical approach by *Newman* (1969) as $HMD = (\pi\ RLD)^{-0.5}$ where *RLD* is the root length density. An overview of the different characteristic root system measures analyzed in this work is given in Table 1.

Statistical tools to analyze characteristic root system measures from root architecture model outputs

Statistical methods have so far been mainly used to group different root systems into plant functional types (*Bodner et al*. 2013) or similar genotypes (*Chen et al*. 2017) based on principal component analysis. Furthermore, nonlinear least-square fitting has been used to fit model parameters based on modelled and measured root lengths in homogeneous root groups (*Zhang et al*. 2003). *Bingham and Wu* (2011) analyzed the effect of varying model input parameters on two root system measures, total root length and root distribution in the soil profile in a sensitivity analysis. Pagès (2011) investigated the impact of different model input parameters as well as interactions between these parameters on the foraging performance of a root system.

The connection between the complex parametrization of 3D root architecture models and characteristic measures of the simulated root systems is often not obvious. Metamodels, which can determine the effect of parameter variations on any of the different measures without running a full simulation, have recently been developed for root hydraulic measures (*Meunier et al*. 2017); for root system measures, however, they are completely missing. To our knowledge, no stochastic root architecture model has so far been thoroughly analyzed with statistical methods.

Objectives

The objective of this work is to use statistical analysis methods to investigate the dependency of key root system measures on model input parameters using the example root architecture model CRootBox, which was chosen for its speed, efficiency and flexibility as well as its acceptance within the root modeling community. For future work, it would, however, be beneficial to also apply the presented analysis methods on other root architecture models and even experimental data sets to allow comparisons of different simulators and validate dependencies between root system measures and model input parameters with experimental data.

In this study, we

- Analyze the distributions of characteristic measures of root architecture, e.g. maximum rooting depth with respect to model parameters of interest, e.g. initial growth speed.
- Derive statistical meta-models to link changes in individual model input parameters to the distributions of characteristic measures of root architecture. These meta-models will be helpful for CRootBox users to estimate the impact of a parameter change on the model's outcome.
- Elucidate correlations between different measures of root architecture
- Compute the correlation changes between different architecture measures due to variations in individual model parameters

Material and methods

The root architecture model CRootBox

CRootBox is a recent root architecture model (*Schnepf et al*. 2018). It is written in C++, but also has a Python binding that allows scripting in Python for most applications. Furthermore, a web application that is based on CRootBox enables the user to quickly create, modify and export root architectures from a database of currently 22 plant species.

CRootBox is fully described in *Schnepf et al*. (2018). Briefly, it is a generic model that is not focused on a specific plant species but is able to simulate the root architectures of any monocotyledonous and dicotyledonous plant. It distinguishes different types of roots, i.e., tap root, basal roots, shoot borne roots and lateral roots, and each root type is defined by a set of different model parameters. Basal and apical root zone define the length of the unbranched root before the first and behind the last branch, respectively. Branch spacing describes the distance between two successive branches and thereby determines branching density respectively the number of branches for a specified maximum root length. Root elongation is defined by a negative exponential function whose initial slope is determined by the initial growth speed (following the approach by *Pagès et al*. 2004) and whose asymptote is specified by the maximal root length. The insertion angle defines the angle from the vertical under which primary roots emerge (a larger angle thus leads to a more shallow root system), while the branching angle describes the initial angle between a branch and its parent root. The model is stochastic because of two aspects. Firstly, the reorientation of a newly emerged root segment of defined length is determined by a random optimization algorithm that selects from N randomly computed values of the deflection angle σ the value with the closest proximity to the desired growth direction (tropism).

We refer the reader to Appendix A, section “Changes in root tip heading” of Schnepf et al. (2018) for details. Secondly, all parameters are assumed to be normal distributed with user-defined mean and standard deviation. Thus, each realization of the same parameter set results in a different root system with variability depending on the standard deviations of the model input parameters, the type of tropism, and the random deflection angle σ. For our statistical analysis, we used the model parameters of the sample plant *Zea mays* as a reference data set. The parameter set was derived from the CRootBox model parameter database (*Schnepf et al.* 2018, *Leitner et al.* 2014) and is shown in Table 2. If the model structure remains the same, the dependency of root system measures on model input parameters is expected to remain qualitatively similar for different parametrizations, and the use of one single reference parameter set is thus justified. As specified in the analysis section below, we performed altogether 53 000 simulations of 3D root architectures based on the reference parameter set and its variations. A 3D visualization of a particular realization of this root system is shown in Fig 1.

Table 2: Model parameter set of *Zea Mays* derived from the CRootBox model parameter database (Zea_maize_4_Leitner_2014) (*Schnepf et al.*, 2018, *Leitner et al.*, 2014)

Variable	Units	Description	Axial roots	1st order laterals
			Value (std)	Value (std)
l_b	cm	basal root zone	0.5 (0.5)	0 (-)
l_a	cm	apical root zone	18.1 (1.81)	8.8 (-)
l_n	cm	branch spacing	0.5 (0.5)	1.2 (0.12)
l_{max}	cm	maximum root length	180.1 (18.01)	8.8 (-)
r	cm d^{-1}	initial elongation rate	2.94 (0.294)	0.75 (0.075)
a	cm	root radius	0.13 (0.013)	0.05 (0.005)
Θ	°	insertion/branching angle	80 (8.0)	85 (8.5)
tropism N	-	number of additional trials	0.5 (-)	0.5 (-)
tropism σ	° cm^{-1}	range of the random deflection angle	5.7 (-)	5.7 (-)
tropism type	-	-	gravitropism	exotropism
max_B	-	maximal number of basal roots	5 (-)	
simtime	d	simulation time	60	

Analysis methods

We selected six model input parameters from the reference parameter set in Table 2 that we expected to have diverse effects on root architecture development. For each of these six parameters we chose a minimum and maximum value based on lower and upper parameter bounds of *Zea Mays* found in literature (*Leitner et al*. 2010, *Leitner et al*. 2014, *Pagès et al*. 2014, *Postma and Lynch* 2011). Standard deviations of the parameters were kept constant (Table 3). We then varied the means of each of the six parameters one at a time in 4 increments in the case of *N*, 9 increments in the case of σ and in 10 increments otherwise between the identified minimum and maximum value resulting in 53 different parameter sets altogether. Subsequently, we used CRootBox to simulate $n = 1000$ root system realizations for each of the 53 parameter sets and computed their characteristic root system measures (Table 1). Fig. 1 shows a 3D visualization of a particular realization of a root system generated with the standard input parameter set given in Table 1. Fig. 2 shows 3D visualizations of the root systems generated with the minimum and maximum

Table 3: Ranges of analyzed model input parameters of basal roots

Model parameter	Unit	Reference (std)	Minimum (std)	Maximum (std)
l_n, axial root	(cm)	0.5 (0.5)	0.5 (0.5)	2 (0.5)
max_B	(-)	5 (-)	4 (-)	40 (-)
r	(cm d^{-1})	2.94 (0.294)	1 (0.294)	3 (0.294)
σ	(° cm^{-1})	5.7 (-)	0 (-)	5.7 (-)
θ	(°)	80 (8.0)	40 (8.0)	85 (8.0)
N	(-)	0.5 (-)	0 (-)	2 (-)

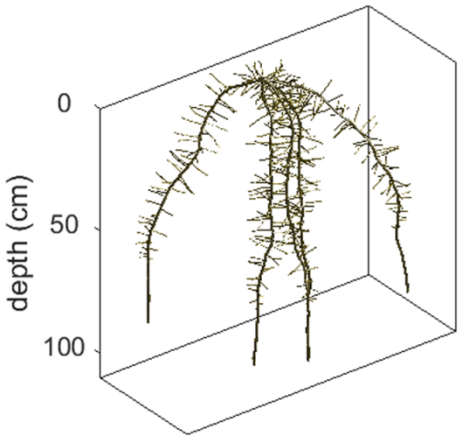


Fig. 1 3D visualization of root architectures simulated by CRootBox using the reference parameter set of Table 1.

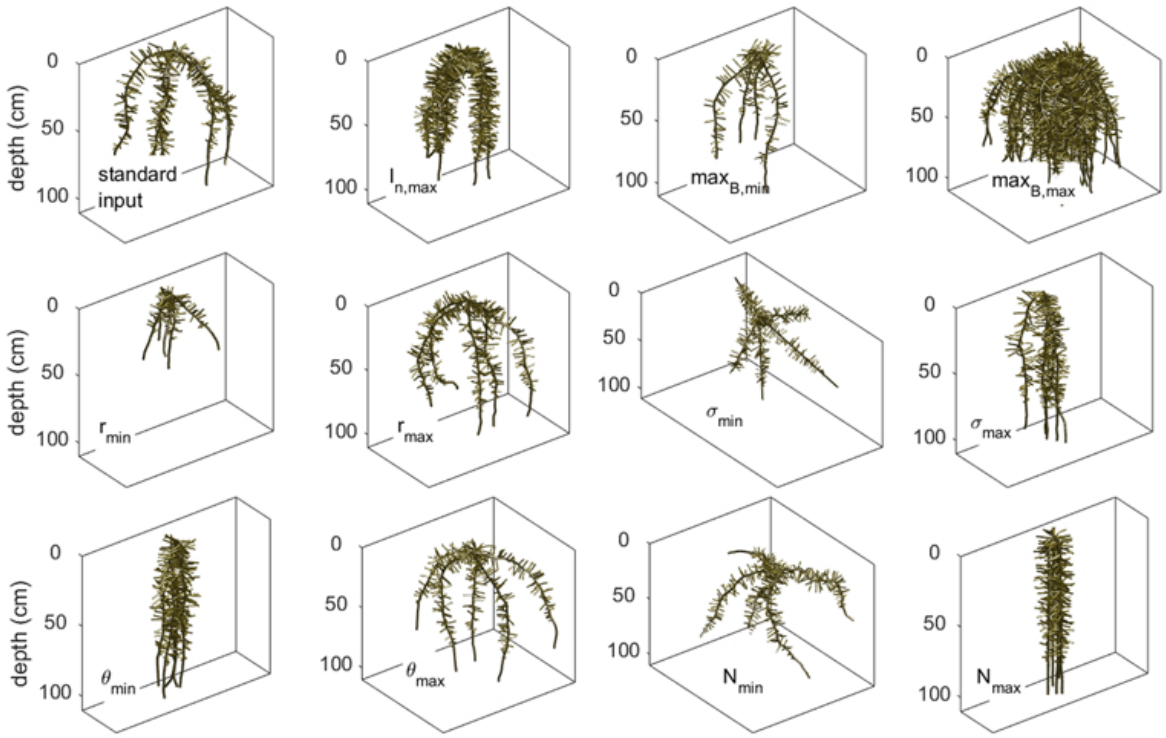


Fig. 2 3D visualization of root architectures simulated by CRootBox using the reference parameter set of Table 1 and setting the six input parameters selected for variation one at a time to minimum and maximum values of the respective parameter range (Table 2).

parameter values specified in Table 2. The obtained data sets were then analyzed statistically as described in the following paragraphs.

The probability distribution of a random root system measure X is described by *n* sample values $x_1, \ldots, x_n$ that can be visualized e.g. by a histogram. As a first step, we estimated the probability density function f_X of X using a nonparametric approach, namely a kernel density estimator (KDE). A KDE of f_X is defined as $\hat{f}_{X,KDE}(s) = \frac{1}{\lambda n}\sum_{i=1}^{n} K_\lambda(s, x_i)$, where the Gaussian kernel $K_\lambda(s,x) = \Phi(|x-s|/\lambda)$ is used with $\Phi(s) = (2\pi)^{-\frac{1}{2}} exp(-\frac{1}{2}s^2)$. The bandwidth $\lambda > 0$ is selected using Scott's rule (*Scott* 1979). Intuitively, each point s is assigned a value $K_\lambda(s, x_i)$ corresponding to the distance between s and the i-th realization x_i of the random root system measure X. The superposition of all these values of every sample forms a KDE. One advantage of a KDE compared to directly investigating the histogram is that it is much easier to get an understanding of the underlying probability distribution because a KDE is not based on discrete bins and produces smooth estimates of the density function. However, being a nonparametric estimator it does not give the concise representation of a parametric approach.

We account for this in a further step where we fit and compare several types of parametric distributions. An important type of probability distributions is the normal distribution given by the density $f(s) = (2\pi\sigma^2)^{-\frac{1}{2}} exp(-\frac{1}{2\sigma^2}(s-\mu)^2)$ with mean μ and variance $\sigma^2 > 0$. Furthermore, for skewed data, several parametric distributions are available. The log-normal distribution is given by $X = exp(Y)$ where Y denotes a normal distributed random variable. Another important example is the gamma distribution given by the density $f(s) = exp(-(s-\alpha)/\theta)\,(s-\alpha)^{k-1}/(\Gamma(k)\,\theta^k)$ for $s > 0$ with shape $k > 0$, scale $\theta > 0$ and location parameter α, where Γ denotes the gamma

function. The inverse-gamma distribution is defined as follows: Let Y be gamma-distributed, then $X = 1/Y$ is inverse-gamma-distributed and the density of X is given by $f(s) = \beta^{\alpha}/\Gamma(\alpha)\ (x-\theta)^{-\alpha-1} exp(-\beta/(x-\theta))$ with shape $\alpha > 0$, scale $\beta > 0$ and location parameter θ. The parameters of the above (parametric) probability distributions were estimated using the maximum-likelihood method; the type of the parametric distribution was chosen manually by comparing the visual fit of its density to the KDE.

Additionally, so-called Q-Q (quantile-quantile) plots were used to evaluate the goodness of fit. More precisely, a Q-Q plot is a method to visually compare two distributions. Let F_X and F_Y be cumulative distribution functions. The Q-Q plot is then given by $s \to ({F_X}^{-1}(s), {F_Y}^{-1}(s))$ for $s \in (0,1)$ where F^{-1} denotes the inverse function of F, the so-called quantile function. In our case F_Y is the empirical distribution function of the sample $x = (x_1, \ldots, x_n)$ and it is thus possible to rewrite the formula above as $i \to ({F_X}^{-1}(\frac{i}{n}), x_{(i)})$ for $i = 1 \ldots n$ where $x_{(i)}$ is the (standardized) order statistics of the sample x. It is clear that if the two distributions fit perfectly, we get a straight line with a 45° angle. If the Q-Q plot is steeper than this line, then F_Y is more dispersed than F_X and vice versa. This allows us in particular to compare the skewness ("S-shape") and the tails of the two distributions graphically. For further information on Q-Q plots, see *Gibbons and Chakraborti* (2010).

To analyze the dependency of a descriptor z of a root system measure (dependent variable) on an input parameter p (regressor variable) we use polynomial regression models. That is, our model is $z = \beta_0 + \beta_1 p + \beta_2 p^2$. The coefficients β_i are determined by the ordinary least squares method. The decision if a polynomial of degree 1 or 2 is used, i.e. whether β_2 is fixed to be 0, is done manually by maximizing the coefficient

of determination R^2 which is computed using leave-one-out cross-validation (*Hastie* et al. 2009). In our case, for a selected model input parameter we are mapping its value p, for example, to the mean value $z = \frac{1}{n}\sum_{i=1}^{n} x_i$ of all realizations $x_1, \ldots, x_n$ of a root system measure given p.

In the above paragraph we discussed how descriptors of single root system measures can depend on a given input parameter, but it is also of great interest if and how root system measures depend on each other. For this reason we consider the sample correlation coefficient $\hat{r}(x,y) = \sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})/((n-1)\sqrt{{s_x}^2 {s_y}^2})$ where $\overline{x}, \overline{y}$ are the sample means and ${s_x}^2, {s_y}^2$ are the sample variances of $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. In our case $\boldsymbol{x}$ and $\boldsymbol{y}$ correspond to samples of two characteristic root system measures X and Y of a given input parameter configuration.

Results and discussion

All the analysis can be reproduced by using the data in HDF5 format and the Jupyter Notebook which can be downloaded from our publication archive on the github repository of CRootBox: https://github.com/Plant-Root-Soil-Interactions-Modelling/CRootBox/tree/master/publication%20archive/VZJ%202018. The resulting pdf of the Analysis can also be downloaded there.

In the first part of this section, we present probability density functions of different root system measures and their variation with different input parametrization. In the second part, dependencies between input parameters and root system measures are described in detail. The third part is dedicated to correlations between individual architecture measures and variations in these correlations due to parametrization. Due

to the large amount of data, we only focus on selected results; the complete analysis can be found in the supplementary material S1: Complete Analysis.pdf (available at "https://github.com/Plant-Root-Soil-Interactions-Modelling/CRootBox/tree/master/publication%20archive/VZJ%202018").

Probability density functions

All absolute root system measures (i.e. z_{max}, r_{max}, *conv, RL, RSA, RV, NR,* V_{rhizo}) were approximately normal distributed. The parameters of the fitted normal distributions, however, varied with changing model input parameters. Root measures based on ratios (i.e. *RND, RLD, RSAD, RVD, and* VD_{rhizo}) were distributed according to the complex distribution function for ratios of correlated normal distributed variables derived by *Hinkley* (1969). In most cases they showed skewed probability distributions, which could be well approximated with inverse gamma distributions, whose parameters, again, depended on the model input parametrization. This can be illustrated by the example of root tip density (*RND*): For low values of max_B, the probability distribution of *RND* is strongly positively skewed, while the skewness becomes less for larger values of max_B (Fig 3). In most cases the Q-Q plots showed good agreement between sample data and theoretical fit suggesting that the fitted normal respectively inverse gamma distributions are valid approximations. The probability distributions for all root system measures are provided in the supplementary material S1. It follows from the differences in skewness that it is required to correctly sample from these distributions to appropriately represent root system/plant diversity. Otherwise, we could miss some 'extreme' root systems that might be the most suited in extreme conditions such as water stress, drought stress, nutrient limitations etc.

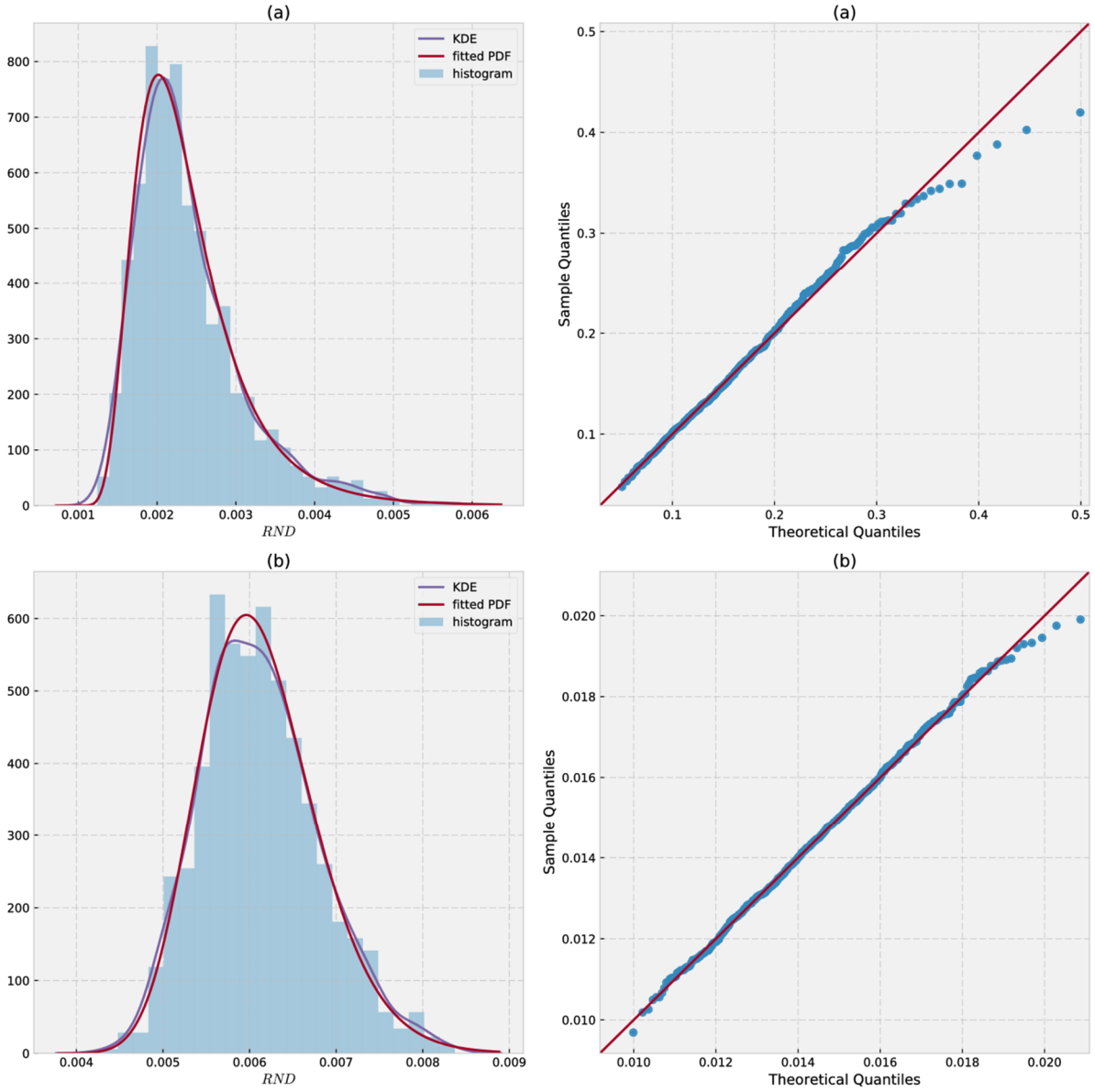


Fig. 3 Contrasting probability density functions, shown as histogram (blue bars), kernel density estimation (blue line) and fitted parametric model, and Q-Q plots for root tip density at A) a low (max_B = 4.0) and B) a high value (max_B = 40.0) of max_B (=number of axial roots). The parametric model in this case is an inverse-gamma distribution with parameters A) $\alpha = 8.068$, $\beta = 0.01237$, $\theta = 0.0006525$ and B) $\alpha = 71.16$, $\beta = 0.3982$, $\theta = 0.0004433$.

Regressions

In this section, we describe and interpret the impact of selected model input parameters on root system measures using the fitted regression models. An overview of the relationships is given in Table 4. Fig 4 visualizes the dependency of root system measures on model input parameters by means of three examples.

Variations in branch spacing l_n have no influence on root system measures defining the shape of the root system (z_{max}, r_{max}, *conv*), but significantly impact total root system size, V_{rhizo} and root hydraulic architecture measures (K_{rs}, $K_{rs\ L}$, $K_{rs\ A}$). Absolute measures defining the total size of a root system (*RL, RSA, RV, NR*), V_{rhizo} as well as ratios including these measures (*RND, RLD, RSAD, RVD,* VD_{rhizo}) decrease nonlinearly with larger branch spacing. The half mean distance between roots (*HMD*) increases linearly with greater values of l_n. As expected, K_{rs}, which predominantly depends on the surface of the root system, decreases nonlinearly with greater values of l_n. Measures of unit root system conductance ($K_{rs\ L}$, $K_{rs\ A}$), however, increase with greater l_n, which is caused by a greater decline rate in K_{rs} than in *RL* or *RSA*. Due to the length difference between apical and basal root zone (apical root zone is generally longer), which becomes important when laterals are scarce, Z_{SUF} increases nonlinearly with greater values of l_n.

In contrast to branch spacing, variations in the insertion angle θ have no impact on measures defining the total size of a root system (*RL, RSA, RV, NR*), V_{rhizo} or root hydraulic architecture measures (K_{rs}, $K_{rs\ L}$, $K_{rs\ A}$), but strongly influence the shape of a root system. A greater insertion angle leads to a shallower root system and thus to lower z_{max} and z_{SUF} and larger r_{max}. The volume of the convex hull increases linearly with greater θ. Ratios including a measure describing the total size of the root system as well as conv (*RND, RLD, RSAD, RVD,* VD_{rhizo}) decrease nonlinearly with greater

θ. A larger insertion angle leads to a more widespread root system and thus also to a larger *HMD*.

A greater number of axial roots max_B leads to a linear increase in measures describing the total size of the root system (*RL, RSA, RV, NR*) as well as V_{rhizo}. It also increases the likelihood of single roots to grow deeper and to spread wider and thus results in greater Z_{max}, r_{max} and *conv*. An increase in z_{SUF} is perceptible, however, not statistically significant. Ratios including *conv* (*RND, RLD, RSAD, RVD,* VD_{rhizo}) increase linearly with increasing max_B. A larger number of axial roots leads to a denser root system and thus to a decrease in *HMD*. While K_{rs} increases with greater values of max_B, measures of unit root system conductance ($K_{rs\ L}$, $K_{rs\ A}$) decrease due to a lower growth rate of K_{rs} than of *RL* or *RSA*.

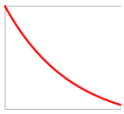
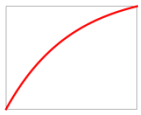
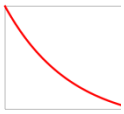
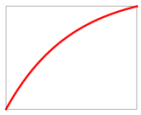
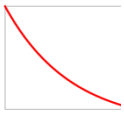
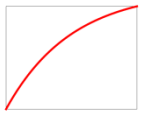
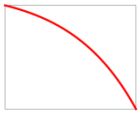
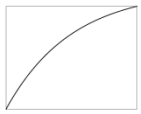
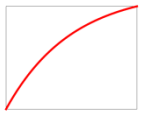
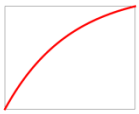
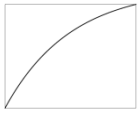
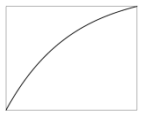
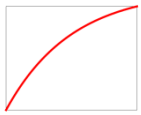
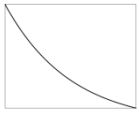
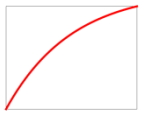
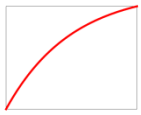
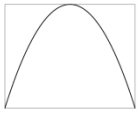
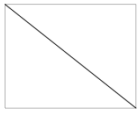
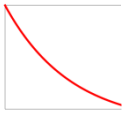
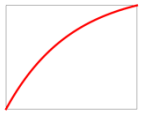
Higher initial growth speed *r* leads to increases in root system measures defining the total size of a root system (*RL, RSA, RV, NR*) as well as V_{rhizo} and K_{rs}. This increase is nonlinear, because root elongation follows a negative exponential function whose initial slope is determined by the initial growth speed and whose asymptote is specified by the maximal root length. Higher values of *r* also lead to greater z_{max}, r_{max}, *conv* and z_{SUF}. Ratios that include both a measure describing the total size of the root system and *conv* (*RND, RLD, RSAD, RVD,* VD_{rhizo}) decrease nonlinearly with increasing values of *r.* This is caused by a larger growth rate of *conv* than of root measures describing the size of the root system with increasing *r*. Greater values of *r* lead to a larger and thus wider spread root system, which results in a nonlinear increase of *HMD*. Again, the growth rate of K_{rs} with greater values of *r* is lower than that of *RL* or *RSA*, which leads to a nonlinear decrease of $K_{rs\ L}$ and $K_{rs\ A}$.

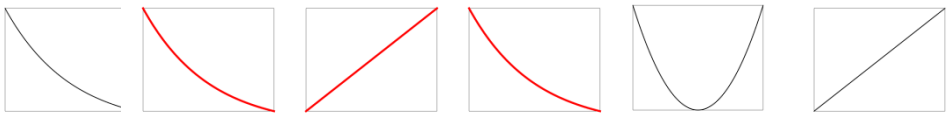
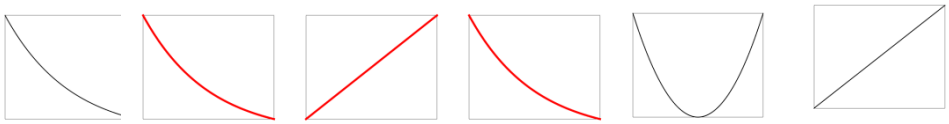
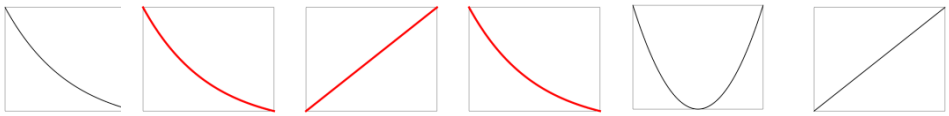
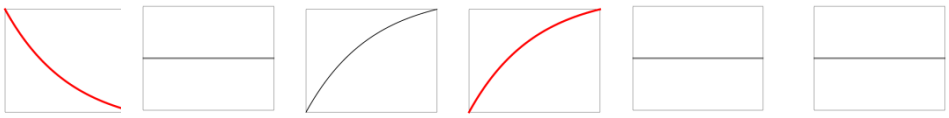
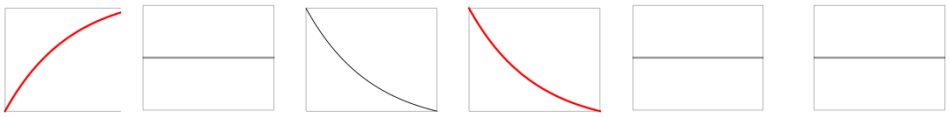
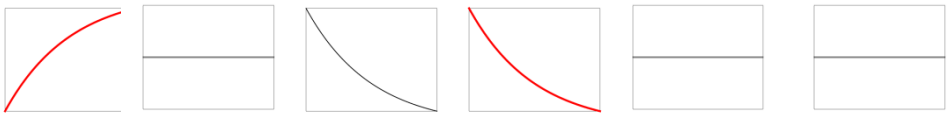
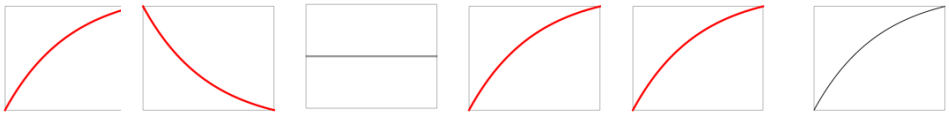
Similar to the insertion angle θ, variations in the range of the random angle deflection σ have no impact on measures defining the total size of a root system (*RL, RSA, RV,*

NR), on V_{rhizo} or on root hydraulic architecture parameters (K_{rs}, $K_{rs\,L}$, $K_{rs\,A}$), but greatly influence the shape of a root system. This influence, however, is complex: Greater values of σ lead to both greater root tortuosity and a stronger impact of gravitropism. The way in which variations of σ influence root system measures thus depends on the parametrization of the insertion angle θ as well as on the number of trials N. For $N > 0$ (tropism type 'gravitropism') and $\theta > 1$, greater values of σ increase the probability of vertical reorientation of a root segment and thus lead to higher values of z_{max} and z_{SUF} and lower values of r_{max}. The measures *conv* and *HMD*, in contrast, first increase up to a certain threshold value with greater values of σ and then decrease again. This is explained with the predominant influence of tortuosity for smaller values of σ, which leads to a less dense root system. When σ becomes larger, the influence of gravitropism outweighs tortuosity and the root system becomes denser. Ratios including *conv* (*RND, RLD, RSAD, RVD,* VD_{rhizo}) as a denominator first decrease down to a certain threshold value for increasing values of σ and then increase again.

Variations in the number of trials N also do not affect measures defining the total size of a root system (*RL, RSA, RV, NR*), V_{rhizo} or root hydraulic architecture parameters (K_{rs}, $K_{rs\,L}$, $K_{rs\,A}$), but greatly influence the shape of a root system. A larger number of N leads to a stronger gravitropic response and thus to higher z_{max} and z_{SUF}, respectively lower r_{max}. The volume of the convex hull decreases linearly with increasing N. Ratios including a measure describing the total size of the root system as well as the volume of the convex hull (*RND, RLD, RSAD, RVD,* VD_{rhizo}) increase linearly with increasing N. A larger number of trials leads to a denser root system and thus to a smaller *HMD*.

Table 4: Dependences between root architecture measures (y-axis) and model input parameters (x-axis): direct or indirect and linear or nonlinear relationships respectively no correlation, regressions with R^2>0.9 are shown in red

	Root input parameters					
Root system measures	Branch spacing (ln)	Insertion angle (θ)	Maximal number of basal roots (max_B)	Initial elongation rate (r)	Range of the random deflection angle (σ)	Number of trials (N)
Total root length (RL)						
Total root surface area (RSA)						
Total root volume (RV)						
Maximum rooting depth (z_{max})						
Horizontal spread (r_{max})						
Volume of the convex hull (conv)						
Number of root tips (NR)						
Rhizosphere volume (V_{rhizo})						
Half mean distance between roots (HMD)						
Root tip density (RND)						
Root length density (RLD)						

Root surface area density (RSAD)	
Root volume density (RVD)	
Rhizosphere volume per volume of the convex hull (VD_{rhizo})	
Equivalent conductance of the root system (K_{rs})	
Equivalent conductance of the root system per unit of root length ($K_{rs\ L}$)	
Equivalent conductance of the root system per unit of root area ($K_{rs\ A}$)	
Mean depth of standard root water uptake (Z_{SUF})	

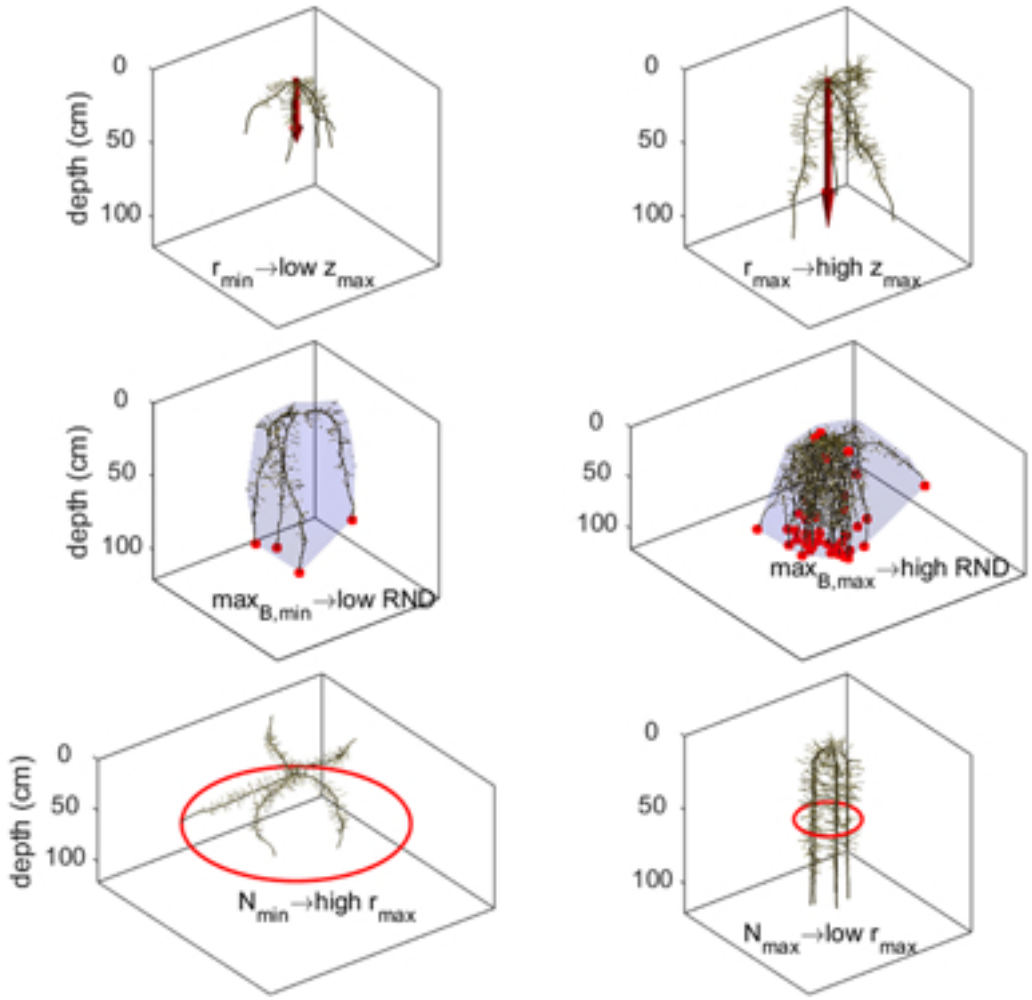


Fig. 4 Three examples visualizing the dependency of root system measures on model input parameters: maximum rooting depth z_{max} as influenced by elongation rate *r,* maximum horizontal spread r_{max} as influenced by the range of the random angle deflection σ and root tip density *RND* (number of root tips per convex hull) as influenced by the number of primary roots max_B.

Researchers from both plant and soil communities agree that it is important to understand the interactions between root and soil to better understand plant water and nutrient acquisition (Vetterlein et al. 2018) and soil science (Gregory 2006). Plant breeding increasingly focused on roots (Bodner 2015, Lynch 2007). Wasson et al. (2012) for example discuss root traits that help increase deep water uptake. They are directly related to the root hydraulic architecture and are included in the set of our

model input parameters: axial and radial resistance, maximal root length, branch spacing. With our approach, it is now possible to quantify the effect of changing both the mean and the variance of those parameters on different root system measures. For example, a deeper root system is postulated to be desirable for increased deep water uptake, and two strategies are discussing for achieving this goal: (a) an increased elongation rate of basal roots or (b) a steeper insertion angle. Our Table 4 shows that the mean depth of root water uptake Z_{SUF} increases with increasing insertion angle as well as with increasing elongation rate r. However, increasing *r* results in an increased root volume RV while this parameter is not influenced by changing the insertion angle. Taking RV as a proxy for carbon costs, our approach thus is a tool to quantify increased costs associated with strategy (a).

Correlations between different root system measures

We quantified correlations between all root system measures.

An interesting finding was that some of the correlation coefficients varied across the parameter space. Each entry in the correlation matrix shown in Fig. 5a is again a matrix in which each line corresponds to one of the selected model input parameters and the different values that were chosen for each parameter. Each color in the small matrix thus corresponds to the basic model setup in which one parameter was changed according to the value outlined in the small matrix in Fig. 5b. We observe that all of the density measures (RND, RLD, RSAD, RVD) have constant strong correlations between each other, irrespective of the chosen parameters. Root surface area density RSAD is for example always strongly correlated with the root length

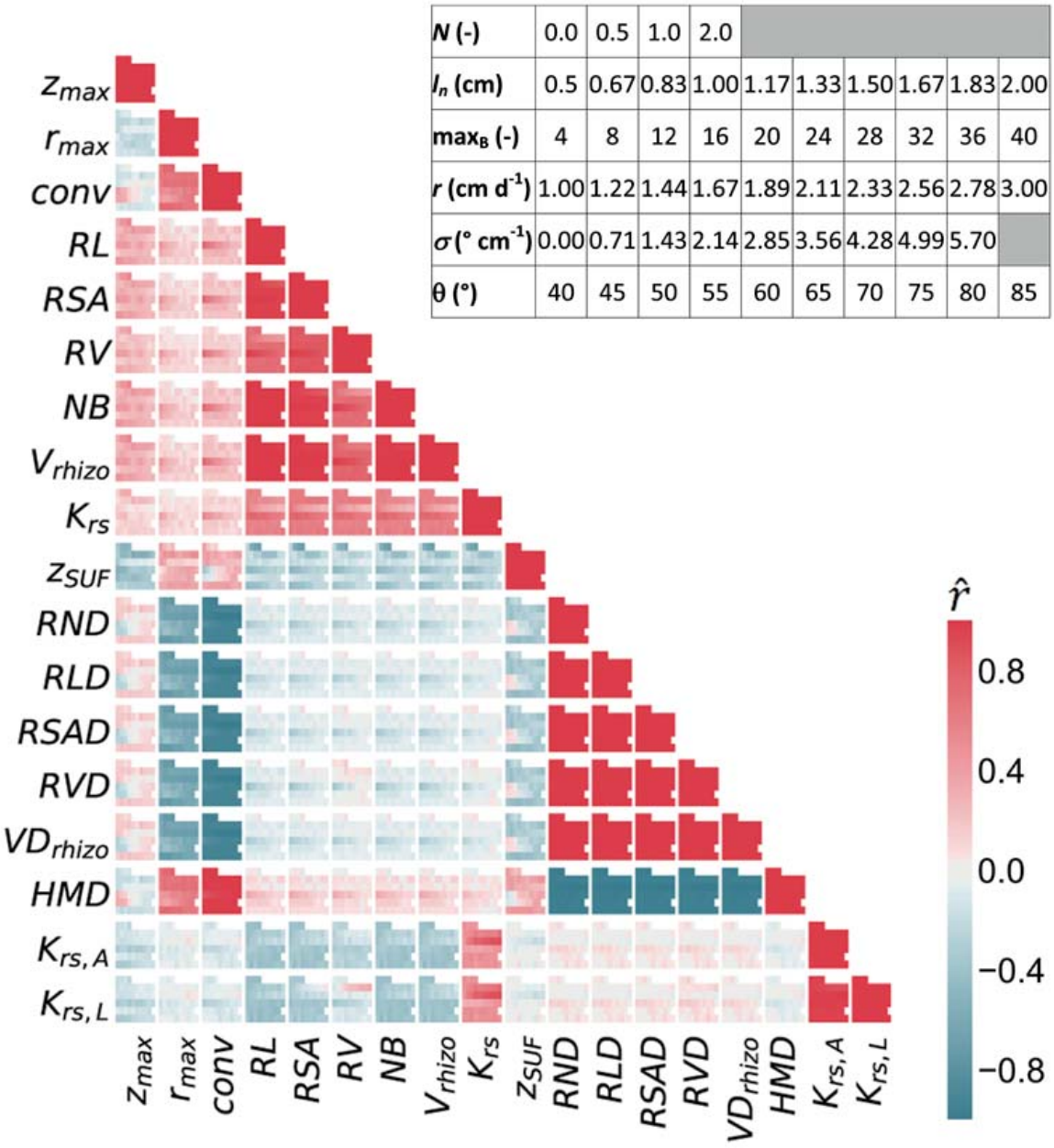


Fig. 5 Correlations between the different root system measures and their variations across the parameter space. Each entry in the correlation matrix (a) is again a matrix in which each line corresponds to one of the selected model input parameters. Each color in the small matrix corresponds to the basic model setup in which one parameter was changed according to the value outlined in the small matrix (b).

density RLD. In a similar way, the rhizosphere volume is always strongly positive correlated with the root length, whereas there is a constant strongly negative correlation

between the half mean distance between roots, HMD, and the densities such as the root length density. As anticipated, maximum horizontal spread and maximum rooting depth are negatively correlated, although the strength of the correlation varies across the parameter space. Such changes in strength of correlation can also be observed for the correlation between equivalent conductance of the root system K, and the maximum horizontal spread r_{max}.

Other root system measures show correlations that even change from slightly negative to slightly positive, as in the case of the correlation between the equivalent conductance of the root system per unit root length, $K_{rs\ L}$, and the densities such RLD, although the correlations are only weak. This makes sense as per definition $K_{rs\ L}$ is dependent of the root length itself (RL) and not on the soil volume explored by this root system. Consequently, its (negative) correlation with the root length RL itself is stronger.

In the correlation between the volume of the convex hull, conv, and the maximal rooting depth, z_{max}, we also observe that there are a few parameterizations in which the correlation is positive. In the basic setup, the root system is parameterized such that it shows gravitropism. Hence, as long as the parameter σ is large enough, the root system becomes steeper over time due to gravitropism. This is at the cost of volume of convex hull - the steeper the root system the smaller the convex hull. Hence there is a negative correlation between the maximum rooting depth and the volume of the convex hull. However, if σ is small, then the root system does not show gravitropism, such that an increase in maximum rooting depth will also mean an increase in volume of convex hull and thus a positive correlation between conv and z_{max}.

Fig. 6a shows the nonlinear change in correlation between maximum rooting depth and horizontal spread with a varying range of the random deflection angle σ, which could be approximated with a quadratic polynomial with a minimum at $\sigma = 0.7$. For increasing values of σ below 0.7, the decrease rate in maximum rooting depth is thus larger than the growth rate in horizontal spread. For increasing values of σ above 0.7, the decrease rate in maximum rooting depth is then smaller than the growth rate in horizontal spread. Fig. 6b shows the nonlinear change in correlation between maximum rooting depth and horizontal spread with a varying range of the insertion angle θ: While the correlation is approximately constant for smaller values of θ, it becomes more negative as θ increases, meaning that the decrease rate in maximum rooting depth is larger than the growth rate in horizontal spread with increasing values of θ.

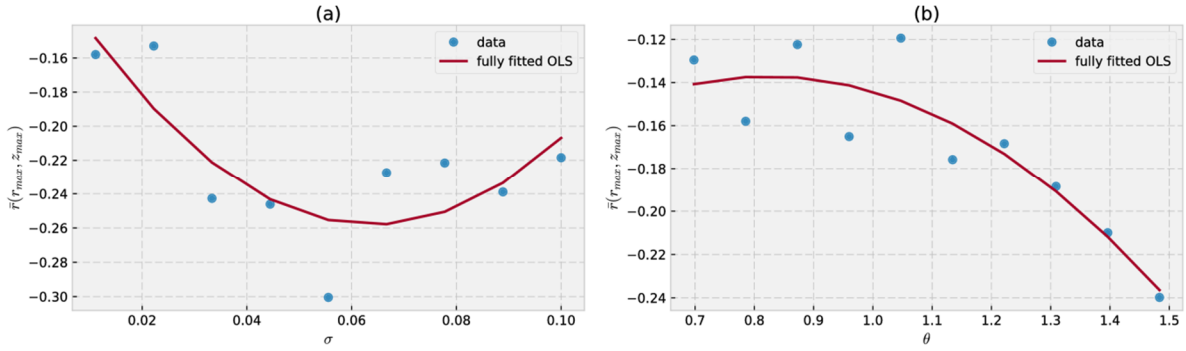


Fig. 6 Correlation between rooting depth and horizontal spread as a function of the model input parameter σ (a) and θ (b).

From the point of view of ecology, our modelling and analysis approach provides a tool to predict which model input parameters ("root traits") will/need to be tuned in order to achieve a given measure in a given environment ("root system trait") and

how the link between the two also depends on the parameter variance (see also Landl et al. 2018).

Conclusions

CRootBox is a mechanistic root architecture model with stochastic components. Using statistical methods, we analyzed characteristic root system measures of root systems simulated with CRootBox. We found that absolute root system measures have a normal probability distribution while ratios of root system measures have an inverse gamma distribution. The general shapes of the regression curves that determine the effect of parameter variations on root system measures as illustrated in Table 4 are expected to remain qualitatively similar for other parameterizations. Furthermore, we found that correlations between different root system measures are also variable across the parameter space.

In conclusion, we presented statistical analysis of the simulation outcome of the 3D root architecture model CRootBox that helped to understand the impact of model structure and model parameterization on characteristic root system measures. Other root architecture models use different approaches to describe processes of root system development. The effect of these different approaches on simulated root system measures, however, has never been analyzed systematically. Our presented method can be applied to select root traits (model input parameters) in order to achieve a certain outcome (root system measure). It can also be used to intercompare different root architecture models including the effects of their stochasticity. It can also be used to compare probability distributions of root system measures obtained from simulations and experimental data (e.g. from 3D root images). Compared to experi-

ments, model simulation can usually be repeated much more often; in this work, the number of replications for each parameter set was 1000. Thus, experimental data could be used to inform and parameterize root architecture models, which can then in turn be used to create realistic, data-driven scenarios for further investigations. The meta-models derived here are simple polynomial regression models. In the future, this work shall be extended to more complex statistical methods including multivariate approaches such as copulas, which provide mathematical tools to build meta-models for vectors of (correlated) root system measures, instead of doing this for each individual measure separately.

Supplementary material

The complete statistical analysis is available as supplemental data S1. It is structured according to the different root system measures considered in this paper. For each individual root system measure and input parametrization, a probability density function as well as kernel density estimation (purple line) are plotted. Fitted normal respectively inverse-gamma distributions are represented by a red line. Q-Q plots show the goodness of fit of the estimated functions. Thereafter, regressions showing the dependency between variations of input parameters and root system measures are presented. Finally, correlations are shown between root system measures and their dependency on variations of input parameters.

Acknowledgments

M.L. acknowledges funding from the German Research Foundation, DFG, within the Research Unit DFG PAK 888.

References

Atkinson, J.A., L.U. Wingen, M. Griffiths, M.P. Pound, O. Gaju, M.J. Foulkes, J. Le Gouis, S. Griffiths, M.J. Bennett, and J. King. 2015. Phenotyping pipeline reveals major seedling root growth QTL in hexaploid wheat. Journal of Experimental Botany 66(8): 2283–2292.

Bingham, I.J., and L. Wu. 2011. Simulation of wheat growth using the 3D root architecture model SPACSYS: validation and sensitivity analysis. European Journal of Agronomy 34: 181–189.

Bodner, G., D. Leitner, A. Nakhforoosh, M. Sobotik, K. Moder, and H.-P. Kaul. 2013. A statistical approach to root system classification. Frontiers in Plant Science 4: 292.

Chen, Y., M.E. Ghanem, and K.H. Siddique. 2017. Characterising root trait variability in chickpea (Cicer arietinum L.) germplasm. Journal of Experimental Botany 68(8): 1987–1999.

Couvreur, V., J. Vanderborght, X. Draye, and M. Javaux. 2014. Dynamic aspects of soil water availability for isohydric plants: Focus on root hydraulic resistances. Water Resources Research 50(11): 8891–8906.

Couvreur, V., J. Vanderborght, and M. Javaux. 2012. A simple three-dimensional macroscopic root water uptake model based on the hydraulic architecture approach. Hydrology and Earth System Science 16(8): 2957–2971.

Dunbabin, V.M., A.J. Diggle, Z. Rengel, and R. van Hugten. 2002. Modelling the interactions between water and nutrient uptake and root growth. Plant and Soil 239(1): 19–38.

Garré, S., E. Laloy, M. Javaux, and H. Vereecken. 2012. Parameterizing a dynamic architectural model of the root system of spring barley from minirhizotron data. Vadose Zone Journal 11(4). vzj2011.0179.

Ge, Z., G. Rubio, and J.P. Lynch. 2000. The importance of root gravitropism for inter-root competition and phosphorus acquisition efficiency: results from a geometric simulation model. Plant and Soil 218(1): 159–171.

Gibbons, J.D., and Chakraborti, S. 2010. Nonparametric Statistical Inference. 5th ed. Boca Raton, FL: Chapman and Hall, CRC.

Hinkley, D.V. 1969. On the ratio of two correlated normal random variables. Biometrika 56(3): 635–639.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. The Elements of Statistical Learning. Springer, New York, NY.

Leitner, D., S. Klepsch, A. Knieß, and A. Schnepf. 2010. The algorithmic beauty of plant roots–an L-system model for dynamic root growth simulation. Mathematical and Computer Modelling of Dynamical Systems 16(6): 575–587.

Leitner, D., F. Meunier, G. Bodner, M. Javaux, and A. Schnepf. 2014. Impact of contrasted maize root traits at flowering on water stress tolerance–A simulation study. Field Crops Research 165: 125–137.

Lynch, J.P. 2007. Roots of the second green revolution. Australian Journal of Botany 55(5): 493–512.

Lynch, J.P. 2013. Steep, cheap and deep: an ideotype to optimize water and N acquisition by maize root systems. Annals of Botany 112(2): 347–357.

Lynch, J.P., and K.M. Brown. 2001. Topsoil foraging–an architectural adaptation of plants to low phosphorus availability. Plant and Soil 237(2): 225–237.

Meunier, F., V. Couvreur, X. Draye, J. Vanderborght, and M. Javaux. 2017. Towards quantitative root hydraulic phenotyping: novel mathematical functions to calculate plant-scale hydraulic parameters from root system functional and structural traits. Journal of Mathematical Biology 75(5): 1133–1170.

Meunier, F., M. Javaux, V. Couvreur, X. Draye, and J. Vanderborght. 2016. A new model for optimizing the water acquisition of root hydraulic architectures over full crop cycles. IEEE International Conference on Functional-Structural Plant Growth Modeling, Simulation, Visualization and Applications (FSPMA), Qingdao, 2016, pp. 140-149.

Nagel, K.A., D. Bonnett, R. Furbank, A. Walter, U. Schurr, and M. Watt. 2015. Simultaneous effects of leaf irradiance and soil moisture on growth and root system architecture of novel wheat genotypes: implications for phenotyping. Journal of Experimental Botany 66(18): 5441–5452.

Newman, E.I. 1969. Resistance to water flow in soil and plant. I. Soil resistance in relation to amounts of root: theoretical estimates. Journal of Applied Ecology 6: 1–12.

Pagès, L., C. Bécel, H. Boukcim, D. Moreau, C. Nguyen, and A.-S. Voisin. 2014. Calibration and evaluation of ArchiSimple, a simple model of root system architecture. Ecological Modelling 290: 76–84.

Pagès, L. 2011. Links between root developmental traits and foraging performance. Plant, Cell and Environment 34: 1749–1760.

Pagès, L., G. Vercambre, J.-L. Drouet, F. Lecompte, C. Collet, and J. Le Bot. 2004. Root Typ: a generic model to depict and analyse the root system architecture. Plant and Soil 258(1): 103–119.

Poorter, H., F. Fiorani, R. Pieruschka, T. Wojciechowski, W.H. Putten, M. Kleyer, U. Schurr, and J. Postma. 2016. Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. New Phytologist 212(4): 838–855.

Postma, J.A., C. Kuppe, M.R. Owen, N. Mellor, M. Griffiths, M.J. Bennett, J.P. Lynch, and M. Watt. 2017. OpenSimRoot: widening the scope and application of root architectural models. New Phytologist 215(3): 1274–1286.

Postma, J.A., and J.P. Lynch. 2011. Root cortical aerenchyma enhances the growth of maize on soils with suboptimal availability of nitrogen, phosphorus, and potassium. Plant Physiology 156(3): 1190–1201.

Schnepf, A., D. Leitner, M. Landl, G. Lobet, T.H. Mai, S. Morandage, C. Sheng, M. Zoerner, J. Vanderborght, and H. Vereecken. 2018. CRootBox: a structural - functional modelling framework for root systems. Annals of Botany, mcx221.

Scott, D. W. 1979. On optimal and data-based histograms. Biometrika 66: 605–610.

Vadez, V. 2014. Root hydraulics: the forgotten side of roots in drought adaptation. Field Crops Research 165: 15–24.

Van Noordwijk, M., J. Floris, and A. De Jager. 1985. Sampling schemes for estimating root density distribution in cropped fields. Netherlands Journal of Agriculture Science 33: 241–262.

Wojciechowski, T., M.J. Gooding, L. Ramsay, and P.J. Gregory. 2009. The effects of dwarfing genes on seedling root growth of wheat. Journal of Experimental Botany 60(9): 2565–2573.

Zhang, B. G., De Reffye, P., Liu, L., Kang, M. Z., Li, B. G. (2003): Analysis and modeling of the root system architecture of winter wheat seedlings, in: Hu Bao-Gang, J. M. (ed.): Plant Growth Modeling and Applications: Proceedings-PMA03. Presented at the International Symposium on Plant Growth Modeling, Simulation, Visualization and their Applications (PMA03), Tsinghua University Press/Springer, Beijng/Heidelberg, pp. 321–328.

Zuo, Q., F. Jie, R. Zhang, and L. Meng. 2004. A generalized function of wheat's root length density distributions. Vadose Zone Journal 3(1): 271–277.