Overcoming the challenges to enhancing experimental plant biology with computational modeling

Renee Dale¹§, Scott Oswald², Amogh Jaliha³, Mary-Francis LaPorte⁴, Daniel M. Fletcher⁵, Allen Hubbard¹, Shin-Han Shiu⁶, Andrew David Lyle Nelson⁷, Alexander Bucksch^{2,8,9}

¹ Donald Danforth Plant Science Center, St. Louis, MO

² Warnell School of Forestry and Natural Resources and Institute of Bioinformatics, University of Georgia, Athens, GA, USA

³ Department of Systems Biology, Harvard Medical School, Boston, MA

⁴ University of California, Davis, One Shields Avenue, Davis, CA

⁵ Bioengineering Sciences Research Group, Department of Mechanical Engineering, School of Engineering, Faculty of Engineering and Physical Sciences, University of Southampton, Southampton, UK

⁶ Department of Plant Biology and Department of Computational Mathematics, Science, and Engineering, Michigan State University, East Lansing MI

⁷ Boyce-Thompson Institute, Cornell University, Ithaca, NY

⁸ Department of Plant Biology, University of Georgia, Athens, GA

⁹ Institute of Bioinformatics, University of Georgia, , Athens, GA

§Corresponding author: rdale825@gmail.com

Abstract

The study of complex biological systems necessitates computational modeling approaches that are currently underutilized in plant biology. Many plant biologists have trouble identifying or adopting modeling methods to their research, particularly mechanistic mathematical modeling. Here we address challenges that limit the use of computational modeling methods, particularly mechanistic mathematical modeling. We divide computational modeling techniques into either pattern models (e.g., bioinformatics, machine learning, or morphology) or mechanistic mathematical models (e.g., biochemical reactions, biophysics, or population models), which both contribute to plant biology research at different scales to answer different research questions. We present arguments and recommendations for the increased adoption of modeling by plant biologists interested in incorporating more modeling into their research programs. As some researchers find math and quantitative methods to be an obstacle to modeling, we provide suggestions for easy-to-use tools for non-specialists and for collaboration with specialists. This may especially be the case for mechanistic mathematical modeling, and we spend some extra time discussing this. Through a more thorough appreciation and awareness of the power of different kinds of modeling in plant biology, we hope to facilitate interdisciplinary, transformative research.

Introduction

Generating knowledge requires the integration and contextualization of information: ‘A collection of facts is no more a science than a heap of stones is a house’ (Henri Poincare). The increasing availability of data provides opportunities as well as challenges to integrate information and properly describe complex biological systems. **Mathematical modeling** is the process of describing complex systems in a logically consistent and explicit manner using a quantitative framework [24]. Such models can generate testable hypotheses by relating possible mechanisms and relationships to observable, measurable phenomena [14]. In addition, models are used to identify non-intuitive relationships, emergent properties, and the conditions under which phenomena arise. In the first half of the paper, we address questions that plant biologists may have about modeling (Questions 1-4), followed by challenges to overcome hurdles (Challenge 1-5). Here, we begin by dividing the field of mathematical modeling in plant biology into two categories: pattern-finding and mechanistic mathematical models (Question 1). We then address how these types of models are used in different subfields of plant biology, and how pattern and mechanistic mathematical models complement each

FIGURE 1. Pattern and mechanistic models approach the same problem in different ways, producing different inferences. (a) A pattern model’s analysis of data might show that generally y and x increase proportional to one another. (b) This type of pattern discovery produces hypotheses about the mechanisms involved, but might not include the action of another component, z , known to be involved in the system. (c) Mechanistic mathematical models represent hypotheses about the putative mechanisms driving this process inclusive of component z . Through analysis of the mechanistic mathematical models, we can predict patterns expected under the different possible regimes.

other (Question 2) [13, 38]. Then we further describe the scientific value of modeling (Question 3). We then specifically focus on modeling approaches that are under-used in plant biology (Question 4). In the second half, we identify the current challenges and potential solutions to broadening engagement with models in plant biology, such as the required expertise and the difficulty finding modeling collaborators.

Question 1: What kinds of models are there?

To facilitate communication, we divide computational models roughly by their utility to plant biology - to study patterns or mechanisms. **Pattern models** test hypotheses about spatial, temporal, or relational patterns between system components (e.g., individual plants, proteins, genes). The mathematical representation of these hypotheses is based on assumptions about the data and statistical properties (such as regulatory network topology [4] or appropriate probability distributions for phenotypic data [67]). Thus, pattern models are typically more ‘data-driven’, i.e., involving finding patterns from the data. Pattern models draw from many disciplines such as bioinformatics, statistics, and machine learning [11]. Many areas of plant biology are studied with pattern models, including the development of genome annotations, phenomics, proteomics, and metabolomics. Big data problems are often addressed using methods such as dimension reduction (e.g., clustering of expression data), latent feature extraction, or machine learning (e.g., neural networks) [107]. Spatially-derived patterns, such as plant anatomical structures, are typically addressed using topology and geometry [60]. The identified patterns (e.g. correlation between x and y in Fig. 1) constrain the set of possible hypotheses about mechanistic relationships that can explain these observed patterns.

Mechanistic¹ mathematical models describe the underlying chemical, biophysical, and mathematical properties within a biological system to predict and understand its behavior mechanistically [17]. Examples of some well-known mechanistic relationships include density-dependent degradation that produces exponential decay; the law of mass-action in biochemical kinetics; and logistic population growth. Mechanistic

¹The term “mechanistic” is often used to describe models that incorporate a plant’s physiological and regulatory mechanisms [54, 58, 59, 57]. This definition of “mechanistic” is distinct from the concept of mathematical abstraction discussed in this review. These two usages of the term are not mutually exclusive, however, as some models use mechanistic mathematical models related to plant regulatory mechanisms [30, 32].

mathematical models are descriptions of real systems but must balance realism with parsimony. Parsimony refers to the simplest but necessary core processes and components (e.g., Occam's razor) - itself a knowledge-generating process. Parsimonious models permit the study of relationships between the system's hypothesized structure and the resulting behavior of the system [3]. Fully realistic models are rarely possible, given the number of biological unknowns, and present computational challenges. For pattern models, parsimony is not always an issue. Some statistical approaches may penalize high-dimensional models, but other approaches (such as neural nets) may use thousands of parameters.

Many mechanistic mathematical models are ordinary differential equations (ODEs, Fig. 1). In essence, these models specify how components change with respect to time or space, such as biochemical reactions changing the concentration of proteins. The reactions between components are controlled by one or more rate parameters. These parameters represent the strength and directionality of an interaction or reaction, and may be estimated from data or literature. In addition to specific measurements, we can compare model predictions to our conceptual understanding of how the system works. Different mathematical formulations can be used to describe different biological properties, and affect how the inputs influence the model components.

Mechanistic mathematical models permit the rigorous study of our hypotheses about phenomena without data. For example, in Fig. 1, a mechanistic mathematical model could predict the impact of z suppressing, promoting, or having no effect on x over a suite of concentrations of x , y , and z . Such an exercise may allow for the elimination of possibilities based on current understanding of the system before data are collected - even guiding the experimental design [15]. Mechanistic mathematical models have yet to reach their full potential in plant biology [25]. This is at least partly due to the challenges associated with the lack of quantitative education in biology curriculum [61] and communicating mathematical representations of the models to biologists [62].

Question 2: What kinds of plant science research have utilized modeling approaches?

While pattern and mechanistic mathematical models complement each other, there are far fewer mechanistic mathematical models being used in plant biology (with a few exceptions). Several limitations to their adoption exist - but before we address these issues, we will establish why mechanistic mathematical modeling is relevant to you and your research.

Gene expression Pattern models are widely used in plant science to study genetics and gene expression. These models exploit statistical detection of patterns, often through analysis of variability, combined with computational algorithms that allow their application to large datasets across genotype and time. Currently, one of the most abundant types of data is from RNA sequencing (RNA-seq) approaches. RNA-seq is used to measure transcript abundance at a genome-wide scale, examine degrading RNAs, RNA structure, post-transcriptional modifications, and small RNA populations. Software such as DESeq2 deploy general linearized modeling approaches, often utilizing a negative binomial distribution, to identify genes whose expression changes under the influence of a treatment condition (DESeq2)[119]. Pattern modeling can integrate molecular (e.g., transcript abundance) and physiological phenotypes to predict causal genes underlying a trait of interest through the identification of correlations. For example, transcriptome-wide association studies (TWAS) showed that altered transcript abundance explains half of the variation in a number of metabolic and agricultural traits in maize [73]. In addition, pattern models have been used to identify genes that influence phenotypes such as yield through their impact on the metabolome using metabolomics QTL (mQTL) [74]. Identifying functionally correlated transcripts from small populations of samples, or time series data, are typically performed using pattern modeling approaches such as weighted gene co-expression analyses (WGCNA; [75]), or circadian aware statistical models such as JTK.Cycle [76]. In the realm of single-cell informatics, statistical models such as Seurat or Monocle allow the tracking of cells along development without a priori knowledge of the specific transcripts that define those processes.

In the analysis of gene expression, pattern models typically look for linear relationships between variation in gene expression across a putative driver of that variation, such as different genotypes. However, the underlying processes that drive plant adaptation and behavior are very nonlinear, and statistical approaches that focus on correlations are limited in their discovery ability [24]. Besides, correlation in pattern models is not causation. Mechanistic mathematical models then come into focus as useful to understand the processes that may be driving what we observe. For example, in a mechanistic mathematical model, developmental timing stochasticity explains ‘noise’ and patterns of gene expression in *Arabidopsis* roots [28]. This work is a nice example of how patterns and mechanisms inform each other, and we anticipate many more discoveries of this type thanks to the interplay between these models in the future.

Gene regulatory networks Predicting gene regulatory networks (GRNs) is a core interest in plant systems biology [106]. Biological responses to internal and external signals are mediated by transcription factors (TFs), some of which regulate the expression of hundreds of genes [72]. The structure and dynamics of TF-gene and TF-TF interactions control diverse biological processes ranging from spatial patterning in tissues [63], to stress responses [64]. Due to the sheer number of interacting components, TF-gene interactions are often represented as directed networks (GRNs). The past decades have seen numerous pattern modeling approaches for inferring GRNs (GENIE3, etc, best performers of DREAM4) from a variety of sequencing data (gene expression data from RNAseq, TF occupancy data such ATACseq, etc). Although inferred GRNs have significantly improved our understanding of how plant gene expression is regulated, these GRNs are static and thus limited in providing mechanistic insight into the biological process itself. Static networks cannot be used to explore the temporal dynamics of processes, and fail to capture the interactions between GRN components.

GRNs have been successfully implemented beyond static representations through the incorporation of mechanisms. Mechanistic mathematical models can be generated from data-focused pattern modeling techniques, and these models in turn predict patterns that can be validated [12]. The mechanistic mathematical model represents the GRN as a dynamic network which can be simulated by altering the state (Boolean ON or OFF, i.e., bound or not bound) of each TF in the network - an approach that can accurately capture TF regulatory mechanisms [29, 12]. In this manner, mechanistic networks can provide insight into various network behaviors and cellular decision-making. Mechanistic mathematical models can also be expanded to include metabolomic components of regulatory networks, such as the Boolean network model of the ABA drought stress regulatory network [29]. This approach requires extensive curation of genetic and metabolomic activity, but produces a system that predicts a wide variety of mutants on the network behavior.

Signal transduction pathways Mechanistic mathematical models are popular in this area of plant biology. At this scale, plant biologists are more able to collect temporal data with sufficient time resolution to capture the dynamics of system components. At larger scales, from in vivo tissue to organs or whole plants, this may require many sampling points and data types that push the boundaries of existing technologies. Pattern models are rarely applied at this scale of plant biology, often asking questions about spatio-temporal gene expression or regulation patterns [41], or developmental patterning [36]. Models of cellular processes include circadian clock and signaling [31], the cell cycle [30], gene expression [28], development and cell fate [35], membrane batteries [32], photosynthesis [34], and carbon flux through metabolic pathways [108].

Physiology Dynamic processes are the key phenomena of interest in plant physiology. From how water moves throughout a plant to how plants grow, plant physiology is concerned with the flow and change of matter and energy throughout the plant body. Mathematical modeling is necessary to describe these processes precisely and in detail, thus modeling is popular in this area of plant biology.

The regulation of stomatal aperture is an excellent example of mechanistic modeling in plant physiology. Stomatal aperture controls the rate of carbon dioxide assimilation (and therefore photosynthesis) but also controls the rate of transpiration (and therefore plant water balance). Since these processes ultimately determine the productivity and water use in crops and forests alike, mechanistic quantitative descriptions provided

by mathematical models are necessary for agriculture and climate forecasting. The regulation of stomatal aperture is also a microcosm of approaches to plant physiological modeling, ranging from the phenomenological (e.g., [79, 80]) to the biochemical/reductionist ([81]) to teleonomic/non-reductionist [82, 83, 84, 85, 86].

A variety of mathematical models for many plant physiological processes have been proposed such as leaf and canopy photosynthesis [88], xylem hydraulics [87], phloem translocation [89], growth morphology [90, 91, 92], carbon allocation dynamics [93, 94, 95, 96], genotype-to-phenotype mapping [77], and carbon-to-nitrogen allocation [97, 98, 99]. In addition to physiological processes, these examples range in biological scale, computational complexity, and mathematical sophistication; while commonalities exist along each of these axes, general statements are difficult. However, this diversity suggests a research program aimed at synthesis and provides an excellent source of ideas for that goal.

Shape and morphology A variety of mathematical techniques from topology and geometry are used to describe plant shape and exploit the analytically common or distinguishing characteristics of shapes as phenotypic traits [38, 37]. Technically, morphological modeling is realized through image processing as a means to extract plant geometry, segmentation or computer simulation to characterize relations between elements like connectivity and hierarchy of branches, arrangements of cells in a space or location of molecules. The field of morphological modeling seeks to understand how underlying mechanisms, including gene regulatory networks, cellular signaling, organ signaling, and biophysical limitations, interact with physical growth processes, and how this ultimately produces the overall size and shape of different plant organs [6, 9, 38].

Persistent homology is a topological pattern modeling technique that describes a relation between plant morphology and a known expanding mathematical function. For example, a circle that continuously increases its diameter from the center of mass of a leaf outline can record the diameters at which serration of the leaf begins (birth) and ends (death) by tracking the intersections between leaf outline and the circle. In that way, subtle differences in the regularity of serration can be detected and potentially linked to genes controlling the serration pattern [37]. Similarly, the same technique can be used to quantify the branching complexity of root systems in 2D images by recording loops in the skeleton of the 2D projection. The difference between birth and death diameters allows for insight into size variation within loops and therefore summarizes branching frequencies and root density distribution within the root system in one mathematical construct.

Mechanistic mathematical models can enhance the information content and prediction of shape development. For example, the FiberWalk [38] characterizes the interaction between elongation and lateral expansion processes of tip-driven growth of a branch. The model predicts that tip-driven growth does not result in an equally thick branch and can not reach all spatial locations, providing mechanistic interpretation of some of the observed variation in measured phenotyping data. Both models and segmented images of plant geometry can be used in mechanistic models of plant functions. This approach is often utilized in root-soil models where the geometry of the root systems are hypothesized to play an important role in the root function (e.g., water and nutrient acquisition and stability) [109]. For example, in the FiberWalk model, branching was found to be a necessary process to optimize nutrient and water uptake below ground [38].

Root-soil models Understanding the structure and growth of roots is important for improving plant productivity. However, the difficulty involved in imaging roots in opaque soil motivates mechanistic mathematical modeling of root growth and the resulting root architecture [110]. These models (often called ‘root-soil’ models) need both a mechanistic description of plant and soil processes to understand the function of root systems. Root-soil models are a good example of mechanistic mathematical models applied across scales (e.g., root branching and the biophysical processes involved in water uptake), as well as the seamless transition from the pattern modeling approach of morphology to mechanistic mathematical modeling.

Factors such as water flow in the xylem, transpiration, and diurnal rhythm often play an important role in root-soil models [111, 112]. For example, if water is stored throughout the depth of the soil, deeper rooting

growth patterns are preferable; while if the soil has a low water-retention capacity, dense and shallow rooting is preferable [113, 114]. Roots can be represented by the root length/surface density (unit length/surface of root per unit volume of soil) as a function of soil depth and time [115, 116] or image-resolved geometries [117]. The function of root, root hairs and soil aggregate geometries can be studied using image-based modeling (a mechanistic approach) using high-resolution 3D imaging of roots in soil, typically X-ray computed tomography. For example, an image-based model found that root hairs and the root contributed equally to phosphorus uptake due to the larger surface area of the root compared to the root hairs. Image-based modeling can complement root imaging studies by solving the mechanistic mathematical model on the image-derived computational mesh, and comparing model predictions to morphological measurements of the root structure [118]. Root systems which had root tips in close proximity obtained the most additional phosphorus uptake due to organic-acid exudation. In summary, mechanistic mathematical models are also powerful vehicles to incorporate multi-scale processes, heterogeneous data such as soil, and complex geometries and are a future direction of focus for the field [38, 9, 7, 117]. Additionally, mechanistic models can be coupled with imaging studies and growth models to link observed plant structure to underlying function.

Whole plant and agronomic traits Crop models (CM) attempt to describe the development, physiology, yield, and agronomic qualities of crop plants, based on genetics, environment, and management. CMs are used by geneticists and breeders to understand the impacts of genotype and environment on traits such as yields, pathogen resistance, and agronomic quality, or to further the understanding and experimental direction for a crop plant of interest [51]. CMs often incorporate a variety of inputs, including nutrient availability, radiation, weather, genetic influences on growth, influences from pests and pathogens, and/or field management practices [51, 54, 52]. To synthesize these complex inputs into a cogent model, crop modelers utilize both mechanistic and pattern models.

CM are unique in that they are neither purely patternistic nor mechanistic, often integrating both. Typically, mechanistic mathematical models are incorporated as “sub-models” of a compartment (such as weather patterns or photosynthesis) within a larger empirical modeling structure, often including pattern modeling components. For example, DSSAT (Decision Support System for Agrotechnology Transfer) models simulate crop growth by utilizing mathematical representations of soil and weather relations alongside empirical findings for specific crops’ growth habits (Jones 2003). Thinking about CM may be useful to experimentalists learning about mechanistic mathematical models as an “exception that clarifies the rule”. In the future, we expect the field of CM to become increasingly mechanistic, particularly as computational limitations decrease. The *crops in silico* project has begun to visualize and simulate biological processes, from the molecular to ecosystem-scale [55]. These efforts have the potential of producing fully mechanistic mathematical models, which could inform further experimentation and research directions. Since the two types of models we lay out here are a spectrum, rather than a dichotomy, considering the way that CMs integrate both areas of modeling may be helpful to clarify the conceptual differences.

Question 3: What can I get from a model?

Hopefully you now see some intersection and value in pattern and mechanistic mathematical models with your research. Even so, models are extra, potentially new work involving learning coding, mathematics, and other quantitative theories - what do they bring to the table in general? We argue that mechanistic mathematical models are not only the natural next step to pattern model discoveries widely used already, but also function in a unique manner to advance plant science for four reasons.

- (1) **Abstraction of complex systems to produce tractable problems** We begin with a list of facts and information. This gets reduced and simplified depending on the research question. This abstraction process can be helpful to enhance our understanding of biology, addressing questions such as the minimum required components to produce a given phenomenon including feedback, oscillatory behavior, or spatial patterning. For example, a mechanistic mathematical model of auxin signaling in the formation of root nodules predicted ‘signature’ patterns that allow experimental

discrimination between the possible underlying mechanisms driving the behavior [71].

- (2) **Predicting emergent phenomena** The interesting parts of a system are when you begin to observe unexpected behavior. Such behavior helps us identify the significance of the roles of specific components within a system. Theoretical tools can be applied to mechanistic mathematical models to allow us to make claims about qualitative and emergent behaviors of a system. For example, bifurcation analysis predicts previously unknown protective relationships between pathogens in a model of disease transmission [2] or how precipitation regimes give rise to distinct landscape vegetation patterns [100]. Mathematical phenomena like switches, bi-stability, and attractors [70, 69] may produce additional emergent behaviors that otherwise may go unnoticed with standard experimental exploration of the stimulus space. Additional analytical or numerical study can predict 'breaking point' or 'unrealistic' behavior. If a predicted 'breaking point' is not observed experimentally, the model's representation needs to be re-evaluated. This allows us to avoid wondering if "maybe it just wasn't enough of a [stimulus]" when designing experiments.
- (3) **Suggesting mechanisms not present in our intuition of a system** After formulation of a model describing a system, we may notice that it critically disagrees with our observations. In this case, we may question the suitability of the pattern model for the data without considering the disagreement biologically informative. On the other hand, disagreement between a mechanistic mathematical model and the data often suggests our understanding of the system may be wrong. Mechanistic mathematical models are quantitative representations of our hypotheses. Disagreement between mechanistic mathematical models and data may also predict the existence of relationships not previously considered as critical to producing the phenomena or dynamics of interest, or the mathematical representation is not appropriate. The back-and-forth between quantifying our understanding via mechanistic mathematical models and assessing agreement with data has the potential to produce new biology, new mathematics, and new mathematical biology questions to be pursued.
- (4) **Integrating knowledge and understanding across system scales** Experimentation reveals how a particular component interacts with other components in the system. Pattern models can reveal these interactions, while mechanistic mathematical models can test them. One of the strongest benefits of mechanistic mathematical modeling is the ability to incorporate multi-disciplinary concepts, such as chemistry [39, 81], biophysics [40, 71, 32], and multi-scale processes [42]. Biological systems are necessarily controlled by chemical and physical processes, and in certain cases these effects should not be ignored.

Question 4: Why is mechanistic mathematical modeling under-utilized in plant biology?

The biggest challenge to the wider adoption of mechanistic mathematical modeling in plant biology is implementation. Indeed, it is often challenging for non-modelers to specify a modeling approach, let alone develop the necessary models. Experience is needed to propose a minimal model of the system, identify the appropriate experimental design, choose an appropriate mathematical representation, and carry out computational and mathematical analysis to study the resulting model.

Mechanistic mathematical models are usually specific Pattern models can often be useful as "black-boxes" (e.g., an input of data into a pattern model, an output of a p-value). However, mechanistic mathematical models are typically very specific. While mechanistic mathematical models excel at making predictions for a variety of contexts, analysis of a given data set often requires modifications. Their utility results from synthesizing biological concepts into a coherent whole and applying them to specific phenomena or experiments. Mechanistic model development requires an understanding of both the biological system and the mathematics; pattern models can be developed for many applications since correlations in data exist

independent of what the data represent. For example, an RNA-seq approach could be applied to any species, for any environmental condition to understand gene expression patterns. A mechanistic mathematical model would need to be specific to the TFs and genes of interest; further, environmental stressors cause changes in different response pathways, necessitating completely different models.

Model development requires math (to some extent) Mechanistic mathematical models require math, which may be intimidating - whether you are writing one yourself, or trying to collaborate with a modeler. A good background to mechanistic mathematical modeling includes understanding the theoretical basis as well as its practical relevance to plant biology and their implementation and validation. For example, ODE modeling uses mathematical biology theory such as mass-action kinetics or standard mechanistic equations mentioned earlier in Question 1; mathematical concepts from calculus and differential equations; methods to simulate and solve equations; and computational methods to estimate parameters, from least-squares to Bayesian and machine-learning approaches.

Collaborating with mathematicians Traditionally, mathematicians who developed mechanistic mathematical models were experts in a field of theoretical math. This means that biologists seeking to develop mechanistic mathematical models for their research needed a deep understanding of modeling for a productive collaborative discussion to take place. Alternatively, biological problems would have to reach the ears of applied mathematicians, who then sought out biologists. Fortunately, we now have specialists in computational plant biology, as well as mathematical modelers working on similar phenomena in other biological systems. This greatly reduces, although does not eliminate, collaborative issues.

Challenge 1: Modeling when you don't like math

It's an old stereotype that people go into the field of biology because they don't like math [65]. Rest assured - you can still model without doing math. In some cases, models can be developed using software for a wide array of biological systems without an in-depth knowledge in the underlying mathematics, including biochemical questions (COPASI [103]); signaling, cellular, and multicellular questions (VCell [101] and SBML [102]); and spatial and ecological questions (LANDIS II[104]). These tools automatically translate diagrams and rules into equations, with anywhere from minimal to high levels of coding required. Tools are also developed to study special systems in plant biology, such as stomatal regulation [81]. While research questions often still require the attention of a modeler, these approaches would certainly help facilitate conversation with, if not totally suffice as the model.

Challenge 2: Finding and collaborating with modelers

To facilitate collaboration for those cases where more complex analysis is required, we recognize the importance of the personal connection. Collaborative incubators and workshops have increasingly sprung up to meet this need, such as Finding Your Inner Modeler (FYIM), Probability Meets Biology [45], Quantitative Cell Biology network (QCB [44]), and NIMBioS workshops [43]. However, more work is needed. Math can be scary, and we need human connection. To partially address this we are developing a collaborative website, www.initmathbio.com. This website works in conjunction with an open database of participants at previously held collaborative workshops we have held. We hope that with this website, you will be able to describe your problem, obtain feedback from subject matter experts, and find collaborators to jump-start your modeling. Some aspects of collaboration are particularly challenging, and we offer the following suggestions:

- (1) **Don't assume anything is not important** The experimental assumptions and methods are often just as (if not more) important than the system itself. Models often have to reflect the experimental design as well as the biology of interest [105].
- (2) **Specify research questions and their impact** Why do you want to model your system? What is currently unknown, and what significance does that have to the field? Often experimental biologists want a model for vague reasons (e.g., 'surely I have enough data to model'). Hopefully, after reading

FIGURE 2. **Example of using VCell.** A conceptual model of plant mycorrhizal trade is developed (A). A model of this system can be developed in VCell [101] through a graphical interface (B) and text-based descriptions of rules, such as reactions and movement. Reproduced from [33].

this paper, you are now aware that modeling comprises a vast array of approaches; a collaboration will be more efficient if a model can be contextualized.

- (3) **Be patient** The interaction between modelers and experimentalists is a learning process that involves both parties. Modelers must develop domain knowledge relevant to the biological question, while experimentalists need to get familiar with the abstract thinking (simplification) in the modeling approach. Many conversations will be required before model development begins.

Challenge 3: Appreciating how pattern and mechanistic mathematical models fit into the scientific method

Modeling should be a back and forth between model and experiment, and an iterative improvement over previous models in order to answer a question [5, 4, 3]. This integrative process is called the modeling cycle, and mirrors the scientific method (hypothesis, experiment, evaluation, repeat). The modeling cycle starts with composing a preliminary model of the phenomena of interest (“hypothesis”). The model may be a network of components with interactions based on scientific theory, existing data, or an existing model. The model is compared to experimental data, or used to predict experimental designs where certain outcomes will occur (“experiment”). The resulting model can then be used to adjust our experimental designs to fill knowledge gaps (“evaluation”). The back and forth process between model predictions, in silico simulations, and experimentation produces gradually improved models and depth of biological inference that lets the utility of modeling shine [17, 21, 23, 25, 5, 1].

FIGURE 3. **Pattern (L) and mechanistic mathematical models (R) have different strengths that fit into the scientific method.** Pattern models test predicted patterns based on biological theory. Observed patterns then allow us to form a more detailed theory on possible mechanisms driving these patterns. Mechanistic mathematical models allow us to identify ways to discriminate between these possibilities. Once validated, the biological theory is updated further, and new patterns are predicted.

Rather than thinking of mathematical models as black boxes that data is shoved into, plant biologists of the future need to “move seamlessly between computational and cell biology” to understand how models predict results, drive design, and produce hypotheses [22]. This is challenging due to social and technical difficulties associated with quantitative proficiency. Fortunately, it has been shown that math appreciation increases with its utilization [18, 2]. Recent emphasis on integrative and translational research and large collaborative groups or hiring clusters facilitate the collaborative, “non-specialized” nature of modern science and facilitate those transitions.

Challenge 4. Consulting modelers before experiments take place

One of the most under utilized benefits of models is their ability to predict interesting behavior based on a preliminary model. Although this collaborative approach necessitates additional upfront work, or the willingness of the experimental biologist to get their hands dirty with math, the outcome is far preferable to an experiment that won’t let us fully interrogate the patterns or mechanisms in question.

One approach to implementing this successfully is *model-based experimental design*. Designing an experiment that will facilitate modeling and maximize its inferential power isn’t always intuitive, and we recommend consulting a modeler during the design process to ensure the model provides insight to the research question [15]. Other considerations of experimental design include how the existing model can be improved, via its structure, parameter estimates, or assumptions. There are different methods of evaluating its quality - such as frequentist or Bayesian statistical approaches [47], control theory [50], optimization theory [49],

sensitivity analyses [47, 48] - that may affect the amount or type of data required.

Models don't need a ton of data to be useful - but they need the appropriate data. Sometimes what is a traditional, convenient, or intuitive design to an experimental biologist is not appropriate or sufficient for the modeling approach. For example, when addressing questions of how a range of a stimulus impacts behavior, it would be better to use a model to determine where interesting or limiting behavior might occur. It is common for modelers to be humorously critiqued for asking for impossible data - communication is required to establish the happy medium between the two perspectives and maximize our science. If a model is sufficiently precise it can describe the relationship that will appear in the data we do have, rather than the data we wish we had.

Challenge 5: Beyond specialization: plant computational biology as a discipline

The issue of improperly designed and implemented experiments is a well-known problem in statistics. Far from being an esoteric concern, improper experimental design limits statistical power and depth of inference. Even scientists who are careful with their analyses may run into problems. If an experiment is poorly planned or executed, computational analyses (especially toolboxes or software) will often spit out something. Although mechanistic mathematical models are rarely applied as black boxes, they can be misused in other ways. The quality of a model depends on the practical implications of those flaws for prediction, inference, or decision making. We need plant biologists to be able to evaluate the purpose, utility, and basic practices involved in modeling.

Conclusion

Computational thinking is a fundamental skill for plant biologists [26, 27]. It complements the theoretical nature of biology and how we understand how things work through the process of abstraction [27]. With education and increased access to computational resources, mathematical and computational methods will become more common throughout plant biology. The field of plant computational biology meets this need, where applied mathematical biologists and computational biologists are experts in both mathematical and computational tools, and their applications to plant biology. We urge plant biologists interested in enhancing their research with computational modeling to meet our challenges, appreciate the science and the specialist nature of modeling, and start collaborative conversations with patience.

Acknowledgements The authors thank Dr Belinda Akpa and Dr Carolyn Rassmussen for conversations at the PlantBiology 2019 and 2020 conference workshops on computational plant biology led by us. We thank the American Society of Plant Biology for hosting the workshops. We also thank Dr Ingo Dreyer and Dr Ivan Baxter for their comments on the manuscript.

Funding R.D. supported by DE-SC0018277 and IOS – 1638507. DE-SC0021110 to M-F.L. NSF IOS-1546617, DEB-1655386, DGE-1828149 and U.S. Department of Energy Great Lakes Bioenergy Research Center BER DE-SC0018409 to S.-H.S. NSF IOS-1758532, NSF IOS-2023310 to A.D.L.N.

References

- [1] Tyson, J. Bringing cartoons to life. *Nature* 445, 823 (2007). <https://doi.org/10.1038/445823a>
- [2] Chen, M. M., Scott, S. M. & Stevens, J. D. Technology as a tool in teaching quantitative biology at the secondary and undergraduate levels: a review. *Letters in Biomathematics* 5, 30–48 (2018).
- [3] Gómez-Mourelto, P. & Ginovart, M. The differential equation counterpart of an individual-based model for yeast population growth. *Computers & Mathematics with Applications* 58, 1360–1369 (2009).
- [4] Tyson, J. J. & Novák, B. Functional Motifs in Biochemical Reaction Networks. *Annu. Rev. Phys. Chem.* 61, 219–240 (2010).
- [5] Mogilner, A., Wollman, R. & Marshall, W. F. Quantitative Modeling in Cell Biology: What Is It Good for? *Developmental Cell* 11, 279–287 (2006).
- [6] Chickarmane, V., Roeder, A.H.K., Tarr, P.T., Cunha, A., Tobin, C., and Meyerowitz, E.M. (2010). Computational Morphodynamics: A Modeling Framework to Understand Plant Growth. *Annu. Rev. Plant Biol.* 61, 65–87.

- [7] Roeder, A.H.K., Tarr, P.T., Tobin, C., Zhang, X., Chickarmane, V., Cunha, A., and Meyerowitz, E.M. (2011). Computational morphodynamics of plants: integrating development over space and time. *Nat Rev Mol Cell Biol* 12, 265–273.
- [8] Jönsson, H., Gruel, J., Krupinski, P., and Troein, C. (2012). On evaluating models in Computational Morphodynamics. *Current Opinion in Plant Biology* 15, 103–110.
- [9] Hong, L., Dumond, M., Zhu, M., Tsugawa, S., Li, C.-B., Boudaoud, A., Hamant, O., and Roeder, A.H.K. (2018). Heterogeneity and Robustness in Plant Morphogenesis: From Cells to Organs. *Annu. Rev. Plant Biol.* 69, 469–495.
- [10] Bucksch, A., Atta-Boateng, A., Azihou, A.F., Battogtokh, D., Baumgartner, A., Binder, B.M., Braybrook, S.A., Chang, C., Coneva, V., DeWitt, T.J., et al. (2017). Morphological Plant Modeling: Unleashing Geometric and Topological Potential within the Plant Sciences. *Front. Plant Sci.* 8, 900.
- [11] Zakharova, L., Meyer, K.M., and Seifan, M. (2019). Trait-based modelling in ecology: lessons from two decades of research (PeerJ Preprints).
- [12] Pratapa, A., Jalihal, A.P., Law, J.N. et al. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nat Methods* 17, 147–154 (2020). <https://doi.org/10.1038/s41592-019-0690-6>
- [13] Passot, S., Couvreur, V., Meunier, F., Draye, X., Javaux, M., Leitner, D., Pagès, L., Schnepf, A., Vanderborght, J., and Lobet, G. (2019). Connecting the dots between computational tools to analyse soil–root water relations. *Journal of Experimental Botany* 70, 2345–2357.
- [14] Bennett, A.E., Preedy, K., Golubski, A., Umbanhowar, J., Borrett, S.R., Byrne, L., Apostol, K., Bever, J.D., Biederman, L., Classen, A.T., et al. (2019). Beyond the black box: promoting mathematical collaborations for elucidating interactions in soil ecology. *Ecosphere* 10.
- [15] Braniff, N., and Ingalls, B. (2018). New opportunities for optimal design of dynamic experiments in systems and synthetic biology. *Current Opinion in Systems Biology* 9, 42–48.
- [16] Drubin, D.G., and Oster, G. (2010). Experimentalist Meets Theoretician: A Tale of Two Scientific Cultures. *MBoc* 21, 2099–2101.
- [17] Keurentjes, J.J.B., Angenent, G.C., Dicke, M., Santos, V.A.P.M.D., Molenaar, J., van der Putten, W.H., de Ruiter, P.C., Struik, P.C., and Thomma, B.P.H.J. (2011). Redefining plant systems biology: from cell to ecosystem. *Trends in Plant Science* 16, 183–190.
- [18] Marsteller, P. (2010). Beyond BIO2010 : Integrating Biology and Mathematics: Collaborations, Challenges, and Opportunities. *LSE* 9, 141–142.
- [19] Miller, L.A., and Alben, S. (2012). Interfacing Mathematics and Biology: A Discussion on Training, Research, Collaboration, and Funding. *Integrative and Comparative Biology* 52, 616–621.
- [20] Brodland, G.W. (2015). How computational models can help unlock biological systems. *Seminars in Cell & Developmental Biology* 47–48, 62–73.
- [21] Ratushny, A.V., Ramsey, S.A., and Aitchison, J.D. (2011). Mathematical Modeling of Biomolecular Network Dynamics. In *Network Biology*, G. Cagney, and A. Emili, eds. (Totowa, NJ: Humana Press), pp. 415–433.
- [22] Short, B. (2009). Cell biologists expand their networks. *J Cell Biol* 186, 305–311.
- [23] Long, S.P. (2019). Making our plant modelling community more than the sum of its parts: a personal perspective. *In Silico Plants* 1.
- [24] Nijhout, H.F., Best, J.A., and Reed, M.C. (2015). Using mathematical models to understand metabolism, genes, and disease. *BMC Biol* 13, 79.
- [25] Holzheu, P., and Kummer, U. (2020). Computational systems biology of cellular processes in *Arabidopsis thaliana*: an overview. *Cell. Mol. Life Sci.* 77, 433–440.
- [26] Schatz, M.C. (2012). Computational thinking in the era of big data biology. *Genome Biol* 13, 177.
- [27] Wing JM: Computational Thinking [<http://www.cs.cmu.edu/~wing/publications/Wing06.pdf>]
- [28] Greenwood, M., Domijan, M., Gould, P. D., Hall, A. J. W. & Locke, J. C. W. Coordinated circadian timing through the integration of local inputs in *Arabidopsis thaliana*. *PLoS Biol* 17, e3000407 (2019).
- [29] Albert, R. et al. A new discrete dynamic model of ABA-induced stomatal closure predicts key feedback loops. *PLoS biology* 15, e2003451 (2017).
- [30] Roodbarkelari, F. et al. CULLIN 4-RING FINGER-LIGASE plays a key role in the control of endoreplication cycles in *Arabidopsis trichomes*. *Proceedings of the National Academy of Sciences* 107, 15275–15280 (2010).
- [31] Grima, R. et al. Insight into nuclear body formation of phytochromes through stochastic modelling and experiment. *Physical Biology* 15, 056003 (2018).
- [32] Dreyer, I., Spitz, O., Kanonenberg, K., Montag, K., Handrich, M.R., Ahmad, S., Schott-Verdugo, S., Navarro-Retamal, C., Rubio-Meléndez, M.E., Gomez-Porras, J.L., Riedelsberger, J., Molina-Montenegro, M.A., Succurro, A., Zuccaro, A., Gould, S.B., Bauer, P., Schmitt, L. and Gohlke, H. (2019), Nutrient exchange in arbuscular mycorrhizal symbiosis from a thermodynamic point of view. *New Phytol*, 222: 1043-1053. <https://doi.org/10.1111/nph.15646>
- [33] Schott S, Valdebenito B, Bustos D, Gomez-Porras JL, Sharma T, Dreyer I. Cooperation through Competition-Dynamics and Microeconomics of a Minimal Nutrient Trade System in Arbuscular Mycorrhizal Symbiosis. *Front Plant Sci.* 2016;7:912. Published 2016 Jun 27. doi:10.3389/fpls.2016.00912
- [34] BRIAN D. HAHN, A Mathematical Model of Photorespiration and Photosynthesis, *Annals of Botany*, Volume 60, Issue 2, August 1987, Pages 157–169, <https://doi.org/10.1093/oxfordjournals.aob.a087432>

- [35] van Berkel, K., de Boer, R. J., Scheres, B. & ten Tusscher, K. Polar auxin transport: models and mechanisms. *Development* 140, 2253–2268 (2013).
- [36] Di Mambro, R. et al. Auxin minimum triggers the developmental switch from cell division to cell differentiation in the Arabidopsis root. *Proc Natl Acad Sci USA* 114, E7641–E7649 (2017).
- [37] Mao Li, Margaret H. Frank, Viktoriya Coneva, Washington Mio, Daniel H. Chitwood, Christopher N. Topp. (2018). The Persistent Homology Mathematical Framework Provides Enhanced Genotype-to-Phenotype Associations for Plant Morphology. *Plant Physiology* 177(4) 1382-1395; DOI: 10.1104/pp.18.00104
- [38] Bucksch, A., Turk, G. & Weitz, J. S. The Fiber Walk: A Model of Tip-Driven Growth with Lateral Expansion. *PLoS ONE* 9, e85585 (2014).
- [39] Dale, R. & Kato, N. Truly quantitative analysis of the firefly luciferase complementation assay. *Current Plant Biology* 5, 57–64 (2016).
- [40] Amiri, K. P., Kalish, A. & Mukherji, S. Robust organelle size control via bursty growth. <http://biorxiv.org/lookup/doi/10.1101/789453> (2019) doi:10.1101/789453.
- [41] Geng, Y. et al. A Spatio-Temporal Understanding of Growth Regulation during the Salt Stress Response in Arabidopsis. *The Plant Cell* 25, 2132–2154 (2013).
- [42] Feller, C. et al. Mathematical Modeling of the Dynamics of Shoot-Root Interactions and Resource Partitioning in Plant Growth. *PLoS ONE* 10, e0127905 (2015).
- [43] NIMBioS Investigative Workshops.
- [44] Workshops — QCBNet.
- [45] Probability meets biology.
- [46] Finding Your Inner Modeler: An NSF Workshop – Year 3. <https://kawai.phy.uab.edu/fyim3/>.
- [47] Barnes, C. P., Silk, D. & Stumpf, M. P. H. Bayesian design strategies for synthetic biology. *Interface Focus* 1, 895–908 (2011).
- [48] Heinemann, T. & Raue, A. Model calibration and uncertainty analysis in signaling networks. *Current Opinion in Biotechnology* 39, 143–149 (2016).
- [49] Wang, Y., Zhang, X.-S. & Chen, L. Optimization meets systems biology. *BMC Syst Biol* 4, S1 (2010).
- [50] Thomas, P.J., Olufsen, M., Sepulchre, R. et al. Control theory in biology and medicine. *Biol Cybern* 113, 1–6 (2019). <https://doi.org/10.1007/s00422-018-00791-5>
- [51] Asseng, S., Y. Zhu, B. Basso, T. Wilson, and D. Cammarano. “Simulation Modeling: Applications in Cropping Systems.” In *Encyclopedia of Agriculture and Food Systems*, edited by Neal K. Van Alfen, 102–12. Oxford: Academic Press, 2014. <https://doi.org/10.1016/B978-0-444-52512-3.00233-3>.
- [52] Donatelli, M., R.D. Magarey, S. Bregaglio, L. Willocquet, J.P.M. Whish, and S. Savary. “Modelling the Impacts of Pests and Diseases on Agricultural Systems.” *Agricultural Systems* 155 (July 1, 2017): 213–24. <https://doi.org/10.1016/j.agry.2017.01.019>.
- [53] Hammer, Graeme, Charlie Messina, Alex Wu, and Mark Cooper. “Biological Reality and Parsimony in Crop Models—Why We Need Both in Crop Improvement!” In *Silico Plants* 1, no. diz010 (January 1, 2019). <https://doi.org/10.1093/insilicoplants/diz010>.
- [54] Jones, J.W, G Hoogenboom, C.H Porter, K.J Boote, W.D Batchelor, L.A Hunt, P.W Wilkens, U Singh, A.J Gijsman, and J.T Ritchie. “The DSSAT Cropping System Model.” *Modelling Cropping Systems: Science, Software and Applications* 18, no. 3 (January 1, 2003): 235–65. [https://doi.org/10.1016/S1161-0301\(02\)00107-7](https://doi.org/10.1016/S1161-0301(02)00107-7).
- [55] Marshall-Colon, Amy, Stephen P. Long, Douglas K. Allen, Gabrielle Allen, Daniel A. Beard, Bedrich Benes, Susanne von Caemmerer et al. “Crops In Silico: Generating Virtual Crops Using an Integrative and Multi-Scale Modeling Platform.” *Frontiers in Plant Science* 8 (2017): 786. <https://doi.org/10.3389/fpls.2017.00786>.
- [56] in silico plants <https://academic.oup.com/insilicoplants>
- [57] Spitters, C.J.T. “Crop Growth Models: Their Usefulness and Limitations.” *Acta Hort.* 267 (1990): 349-368. <https://doi.org/10.17660/ActaHortic.1990.267.42>
- [58] Keating, B. A.; Carberry, P. S.; Hammer, G. L.; Probert, M. E.; Robertson, M. J.; Holzworth, D.; Huth, N. I.; Hargreaves, J. N. G.; Meinke, H.; Hochman, Z.; McLean, G.; Verburg, K.; Snow, V.; Dimes, J. P.; Silburn, M.; Wang, E.; Brown, S.; Bristow, K. L.; Asseng, S.; Chapman, S.; McCown, R. L.; Freebairn, D. M.; Smith, C. J. An Overview of APSIM, a Model Designed for Farming Systems Simulation. *European Journal of Agronomy* 2003, 18 (3), 267–288. [https://doi.org/10.1016/S1161-0301\(02\)00108-9](https://doi.org/10.1016/S1161-0301(02)00108-9).
- [59] Estes, L.D., Bradley, B.A., Beukes, H., Hole, D.G., Lau, M., Oppenheimer, M.G., Schulze, R., Tadross, M.A. and Turner, W.R. (2013), Crop model intercomparison. *Global Ecology and Biogeography*, 22: 1007-1018. <https://doi.org/10.1111/geb.12034>
- [60] Amézquita, EJ, Quigley, MY, Ophelders, T, Munch, E, Chitwood, DH. The shape of things to come: Topological data analysis and biology, from molecules to organisms. *Developmental Dynamics*. 2020; 249: 816– 833. <https://doi.org/10.1002/dvdy.175>
- [61] Bialek, W. (2004). Introductory Science and Mathematics Education for 21st-Century Biologists. *Science*, 303(5659), 788–790. doi:10.1126/science.1095480
- [62] Fawcett, T. W. & Higginson, A. D. Heavy use of equations impedes communication among biologists. *Proceedings of the National Academy of Sciences* 109, 11735–11739 (2012).

- [63] Adrian J, Chang J, Ballenger CE, Bargmann BO, Alassimone J, Davies KA, Lau OS, Matos JL, Hachez C, Lanctot A, Vatén A, Birnbaum KD, Bergmann DC. Transcriptome dynamics of the stomatal lineage: birth, amplification, and termination of a self-renewing population. *Dev Cell*. 2015 Apr 6;33(1):107-18. doi: 10.1016/j.devcel.2015.01.025. PMID: 25850675; PMCID: PMC4390738.
- [64] Song, L. et al. A transcription factor hierarchy defines an environmental stress response network. *Science* 354, 1550–1550 (2016).
- [65] Wachsmuth LP, Runyon CR, Drake JM, Dolan EL. Do Biology Students Really Hate Math? Empirical Insights into Undergraduate Life Science Majors' Emotions about Mathematics. *CBE Life Sci Educ*. 2017 Fall;16(3):ar49. doi: 10.1187/cbe.16-08-0248. PMID: 28798211; PMCID: PMC5589429.
- [66] Tyson, J. J., Laomettachit, T. & Kraikivski, P. Modeling the dynamic behavior of biochemical regulatory networks. *Journal of Theoretical Biology* 462, 514–527 (2019).
- [67] Kirkpatrick RM, Neale MC. Applying Multivariate Discrete Distributions to Genetically Informative Count Data. *Behav Genet*. 2016 Mar;46(2):252-68. doi: 10.1007/s10519-015-9757-z. Epub 2015 Oct 24. PMID: 26497008; PMCID: PMC4752908.
- [68] Orth, J. D., Thiele, I. & Palsson, B. Ø. What is flux balance analysis? *Nature Biotechnology* 28, 245–248 (2010).
- [69] Saadatpour, A. & Albert, R. A comparative study of qualitative and quantitative dynamic models of biological regulatory networks. *EPJ Nonlinear Biomedical Physics* 4, 5 (2016).
- [70] Rata, S. et al. Two Interlinked Bistable Switches Govern Mitotic Control in Mammalian Cells. *Current Biology* 28, 3824–3832.e6 (2018).
- [71] Deinum, E. E., Geurts, R., Bisseling, T. & Mulder, B. M. Modeling a Cortical Auxin Maximum for Nodulation: Different Signatures of Potential Strategies. *Front. Plant Sci*. 3, (2012).
- [72] Bilu, Y., Barkai, N. The design of transcription-factor binding sites is affected by combinatorial regulation. *Genome Biol* 6, R103 (2005). <https://doi.org/10.1186/gb-2005-6-12-r103>
- [73] Kremling KAG, Diepenbrock CH, Gore MA, Buckler ES, Bandillo NB. Transcriptome-Wide Association Supplements Genome-Wide Association in *Zea mays*. G3 (Bethesda). 2019;9(9):3023-3033. Published 2019 Sep 4. doi:10.1534/g3.119.400549
- [74] Wei, J., Wang, A., Li, R. et al. Metabolome-wide association studies for agronomic traits of rice. *Heredity* 120, 342–355 (2018). <https://doi.org/10.1038/s41437-017-0032-3>
- [75] Langfelder, P., Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559 (2008). <https://doi.org/10.1186/1471-2105-9-559>
- [76] Hughes ME, Hogenesch JB, Kornacker K. JTK-CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J Biol Rhythms*. 2010 Oct;25(5):372-80. doi: 10.1177/0748730410379711. PMID: 20876817; PMCID: PMC3119870.
- [77] Diane R Wang, Carmela R Guadagno, Xiaowei Mao, D Scott Mackay, Jonathan R Pleban, Robert L Baker, Cynthia Weinig, Jean-Luc Jannink, Brent E Ewers, A framework for genomics-informed ecophysiological modeling in plants, *Journal of Experimental Botany*, Volume 70, Issue 9, 15 April 2019, Pages 2561–2574, <https://doi.org/10.1093/jxb/erz090>
- [78] Buckley, T. N. Modeling Stomatal Conductance. *Plant Physiol*. 174, 572–582 (2017).
- [79] Jarvis PG (1976) The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philos Trans R Soc Lond B Biol Sci* 273: 593–610
- [80] Ball JT, Woodrow IE, Berry JA (1987) A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In J Biggens, ed, *Progress in Photosynthesis Research*. Martinus Nijhoff Publishers, The Netherlands, pp 221–224
- [81] Hills A, Chen Z-H, Amtmann A, Blatt MR, Lew VL (2012) OnGuard, a computational platform for quantitative kinetic modeling of guard cell physiology. *Plant Physiol* 159: 1026–1042
- [82] Cowan IR, Farquhar GD (1977) Stomatal function in relation to leaf metabolism and environment. *Symp Soc Exp Biol* 31: 471–505
- [83] Manzoni S, Vico G, Porporato A, Katul G (2013) Biological constraints on water transport in the soil–plant–atmosphere system. *Adv Water Resour* 51: 292–304
- [84] Wolf A, Anderegg WR, Pacala SW (2016) Optimal stomatal behavior with competition for water and risk of hydraulic impairment. *Proc Natl Acad Sci USA* 113: E7222–E7230
- [85] Sperry JS, Venturas MD, Anderegg WR, Mencuccini M, Mackay DS, Wang Y, Love DM (2017) Predicting stomatal responses to the environment from the optimization of photosynthetic gain and hydraulic cost. *Plant Cell Environ* doi:10.1111/pce.12852
- [86] Mrad, A. et al. A Dynamic Optimality Principle for Water Use Strategies Explains Isohydic to Anisohydic Plant Responses to Drought. *Front. For. Glob. Change* 2, 49 (2019).
- [87] Mrad, A., Domec, J.-C., Huang, C.-W., Lens, F. & Katul, G. A network model links wood anatomy to xylem tissue hydraulic behaviour and vulnerability to cavitation: Model links wood anatomy to plant hydraulics. *Plant Cell Environ* 41, 2718–2730 (2018).
- [88] Hikosaka, K. Optimality of nitrogen distribution among leaves in plant canopies. *J Plant Res* 129, 299–311 (2016). <https://doi.org/10.1007/s10265-016-0824-1>

- [89] Stanfield, RC, Schulte, PJ, Randolph, KE, Hacke, UG. Computational models evaluating the impact of sieve plates and radial water exchange on phloem pressure gradients. *Plant Cell Environ.* 2019; 42: 466– 479. <https://doi.org/10.1111/pce.13414>
- [90] Przemysław Prusinkiewicz, Pierre Barbier de Reuille, Constraints of space in plant development, *Journal of Experimental Botany*, Volume 61, Issue 8, May 2010, Pages 2117–2129, <https://doi.org/10.1093/jxb/erq081>
- [91] Prusinkiewicz, P. and Runions, A. (2012), Computational models of plant development and form. *New Phytologist*, 193: 549–569. <https://doi.org/10.1111/j.1469-8137.2011.04009.x>
- [92] Sievänen, R., Godin, C., DeJong, T. M., & Nikinmaa, E. (2014). Functional-structural plant models: a growing paradigm for plant studies. *Annals of botany*, 114(4), 599–603. <https://doi.org/10.1093/aob/mcu175>
- [93] Le Roux, X., Bariac, T., Sinoquet, H., Genty, B., Piel, C., Mariotti, A., Girardin, C. and Richard, P. (2001), Spatial distribution of leaf water-use efficiency and carbon isotope discrimination within an isolated tree crown. *Plant, Cell & Environment*, 24: 1021–1032. <https://doi.org/10.1046/j.0016-8025.2001.00756.x>
- [94] Oskar Franklin, Jacob Johansson, Roderick C. Dewar, Ulf Dieckmann, Ross E. McMurtrie, Åke Brännström, Ray Dybzinski, Modeling carbon allocation in trees: a search for principles, *Tree Physiology*, Volume 32, Issue 6, June 2012, Pages 648–666, <https://doi.org/10.1093/treephys/tpz138>
- [95] De Kauwe, M.G., Medlyn, B.E., Zaehle, S., Walker, A.P., Dietze, M.C., Wang, Y.-P., Luo, Y., Jain, A.K., El-Masri, B., Hickler, T., Wärlind, D., Weng, E., Parton, W.J., Thornton, P.E., Wang, S., Prentice, I.C., Asao, S., Smith, B., McCarthy, H.R., Iversen, C.M., Hanson, P.J., Warren, J.M., Oren, R. and Norby, R.J. (2014), Where does the carbon go? A model–data intercomparison of vegetation carbon allocation and turnover processes at two temperate forest free-air CO₂ enrichment sites. *New Phytol*, 203: 883–899. <https://doi.org/10.1111/nph.12847>
- [96] Katarína Merganičová, Ján Merganič, Aleksi Lehtonen, Giorgio Vacchiano, Maša Zorana Ostrogović Sever, Andrey L D Augustynczyk, Rüdiger Grote, Ina Kyselová, Annikki Mäkelä, Rasoul Yousefpour, Jan Krejza, Alessio Collalti, Christopher P O Reyer, Forest carbon allocation modelling under climate change, *Tree Physiology*, Volume 39, Issue 12, December 2019, Pages 1937–1960, <https://doi.org/10.1093/treephys/tpz105>
- [97] Chen, J.L., Reynolds, J.F., Harley, P.C. et al. Coordination theory of leaf nitrogen distribution in a canopy. *Oecologia* 93, 63–69 (1993). <https://doi.org/10.1007/BF00321192>
- [98] Dybzinski R, Farrior C, Wolf A, Reich PB, Pacala SW. Evolutionarily stable strategy carbon allocation to foliage, wood, and fine roots in trees competing for light and nitrogen: an analytically tractable, individual-based model and quantitative comparisons to data. *Am Nat.* 2011 Feb;177(2):153–66. doi: 10.1086/657992. PMID: 21460552.
- [99] Barillot, R., Chambon, C., & Andrieu, B. (2016). CN-Wheat, a functional-structural model of carbon and nitrogen metabolism in wheat culms after anthesis. I. Model description. *Annals of botany*, 118(5), 997–1013. <https://doi.org/10.1093/aob/mcw143>
- [100] Tarnita, C. E. et al. A theoretical foundation for multi-scale regular vegetation patterns. *Nature* 541, 398–401 (2017).
- [101] <https://vcell.org/>
- [102] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden, A. A. Cuellar, S. Dronov, E. D. Gilles, M. Ginkel, V. Gor, I. I. Goryanin, W. J. Hedley, T. C. Hodgman, J.-H. Hofmeyr, P. J. Hunter, N. S. Juty, J. L. Kasberger, A. Kremling, U. Kummer, N. Le Novère, L. M. Loew, D. Lucio, P. Mendes, E. Minch, E. D. Mjolsness, Y. Nakayama, M. R. Nelson, P. F. Nielsen, T. Sakurada, J. C. Schaff, B. E. Shapiro, T. S. Shimizu, H. D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, J. Wang, and the rest of the SBML Forum., The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models, *Bioinformatics*, Volume 19, Issue 4, 1 March 2003, Pages 524–531, <https://doi.org/10.1093/bioinformatics/btg015>
- [103] <http://copasi.org/>
- [104] <http://www.landis-ii.org/>
- [105] Dale, R., Ohmuro-Matsuyama, Y., Ueda, H. & Kato, N. Mathematical Model of the Firefly Luciferase Complementation Assay Reveals a Non-Linear Relationship between the Detected Luminescence and the Affinity of the Protein Pair Being Analyzed. *PLOS ONE* 11, e0148256 (2016).
- [106] Haque, S., Ahmad, J. S., Clark, N. M., Williams, C. M. & Sozzani, R. Computational prediction of gene regulatory networks in plant growth and development. *Current Opinion in Plant Biology* 47, 96–105 (2019).
- [107] Hériché, J.-K., Alexander, S., and Ellenberg, J. (2019). Integrating Imaging and Omics: Computational Methods and Challenges. *Annu. Rev. Biomed. Data Sci.* 2, 175–197.
- [108] Allen DK, Libourel IG, Shachar-Hill Y. Metabolic flux analysis in plants: coping with complexity. *Plant Cell Environ.* 2009 Sep;32(9):1241–57. doi: 10.1111/j.1365-3040.2009.01992.x. Epub 2009 Apr 22. PMID: 19422611.
- [109] Dunbabin, V.M., Postma, J.A., Schnepf, A., Pagès, L., Javaux, M., Wu, L., Leitner, D., Chen, Y.L., Rengel, Z. and Diggle, A.J., 2013. Modelling root–soil interactions using three–dimensional models of root growth, architecture and function. *Plant and soil*, 372(1), pp.93–124
- [110] Schnepf, A., Leitner, D., Landl, M., Lobet, G., Mai, T.H., Morandage, S., Sheng, C., Zörner, M., Vanderborght, J. and Vereecken, H., 2018. CRootBox: a structural–functional modelling framework for root systems. *Annals of botany*, 121(5), pp.1033–1053.
- [111] Hayat, F., Zarebanadkouki, M., Ahmed, M.A., Buecherl, T. and Carminati, A., 2020. Quantification of hydraulic redistribution in maize roots using neutron radiography. *Vadose Zone Journal*, 19(1), p.e20084.

- [112] Schnepf, A., Leitner, D. and Klepsch, S., 2012. Modeling phosphorus uptake by a growing and exuding root system. *Vadose Zone Journal*, 11(3), pp.vzj2012-0001.
- [113] Leitner, D., Meunier, F., Bodner, G., Javaux, M. and Schnepf, A., 2014. Impact of contrasted maize root traits at flowering on water stress tolerance—a simulation study. *Field crops research*, 165, pp.125-137.
- [114] Tron, S., Bodner, G., Laio, F., Ridolfi, L. and Leitner, D., 2015. Can diversity in root architecture explain plant water use efficiency? A modeling study. *Ecological modelling*, 312, pp.200-210.
- [115] Fletcher, D.M., Ruiz, S.A., Dias, T., Chadwick, D.R., Jones, D.L. and Roose, T., 2021. Precipitation-optimised targeting of nitrogen fertilisers in a model maize cropping system. *Science of the Total Environment*, 756, p.144051.
- [116] Ruiz, S., Koebernick, N., Duncan, S., Fletcher, D.M., Scotson, C., Boghi, A., Marin, M., Bengough, A.G., George, T.S., Brown, L.K. and Hallett, P.D., 2020. Significance of root hairs at the field scale—modelling root water and phosphorus uptake under different field conditions. *Plant and soil*, 447(1), pp.281-304.
- [117] Ruiz, S.A., McKay Fletcher, D., Williams, K. and Roose, T., 2020. Review of plant-soil modelling: root growth, nutrient and water transport/uptake, and mechanics. *Annual Plant Reviews online*.
- [118] McKay Fletcher, D.M., Ruiz, S., Dias, T., Petroselli, C. and Roose, T., 2020. Linking root structure to functionality: the impact of root system architecture on citrate-enhanced phosphate uptake. *New Phytologist*, 227(2), pp.376-391.
- [119] Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550 (2014).