HAL open science

Inferring the structure and dynamics of tropical rain forests with individual-based forest growth models

Fabian Jörg Fischer

▶ To cite this version:

Fabian Jörg Fischer. Inferring the structure and dynamics of tropical rain forests with individual-based forest growth models. Earth Sciences. Université Paul Sabatier - Toulouse III, 2019. English. NNT : 2019TOU30229 . tel-02930242

HAL Id: tel-02930242

https://theses.hal.science/tel-02930242v1

Submitted on 4 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Fabian FISCHER

Le 17 décembre 2019

Inférence de la structure et dynamique des forêts tropicales humides avec un modèle individu-centré

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries**

Spécialité : **Ecologie, biodiversité et évolution**

Unité de recherche :
EDB - Evolution et Diversité Biologique

Thèse dirigée par
Jérôme CHAVE

Jury

M. Georges Kunstler, Rapporteur
Mme Anja Rammig, Rapporteure
M. Martin Herold, Examinateur
M. Mathias Disney, Examinateur
Mme Rosie Fisher, Examinatrice
M. Cédric Véga, Examinateur
M. Christophe Thébaud, Examinateur
M. Jérôme CHAVE, Directeur de thèse

Acknowledgements

I would like to thank all the people that have accompanied me, supported me and made me laugh – and laughed at my jokes – throughout the past three and a half years.

Thank you, first of all, to my PhD committee who has provided important comments and corrections to my initial ideas. Thank you to the reviewers of this manuscript and the jury to whom I am deeply grateful for their efforts, and thank you to my funding organisation, the Studienstiftung des Deutschen Volkes, who have generously sponsored my academic education.

More than anyone, I would like to thank my supervisor, Jérôme, who took me on as a PhD student, knowing me from little more than a skype session and a CV, and then went on to provide me both with the freedom to develop my own ideas and a critical look to shape my ecological thinking. Thank you for your rapid responses to my messages, from all continents and at any time, the quick and thorough readings of my texts, and the financial backing to carry out my project, from attending conferences to doing fieldwork in French Guiana. Many thanks, it has been a great experience!

I would also like to thank all my colleagues, who have been more than colleagues. Thank you to my office mates, Delphine, Uxue and Jordi, whom it has always been a great pleasure to be working next to – even though not as often as we might have hoped for. Then, of course, an enormous thanks to the other PhD students and scientists, Jade, Kévin, Marine, Antoine, Max, Eglantine, Maëva, Luana, Isa, Séb, Félix, Sandra, Jan, Céline, Quentin, Alex, Juan, Jessica, Guillaume, Jérôme, Amaia. Thanks to you I have felt very much at home in Toulouse and at the lab! Also, a big shout-out to the other vegetation aficionados from our lab, Julian, Tony, Nico, E-Ping, and those who might be far away, but who are practically still here, Emilie and Isabelle! Thanks for all the times we were working together in French Guiana, going running, or discussing science and politics! Thanks to all those as well that have been on field missions with me or otherwise helped my scientific education, Andrew, Bruno, Grég, Andreas, to the workers who have carried out meticulous field censuses and botanical identification, and to our administration who have always been patient and helped me more than once when I did not fully comprehend the intricacies of bureaucracy. Big thanks also to Jan, who made sure this thesis turned from a Toulouse-generated pdf into a printed-out manuscript in Munich.

A very special thank you is due to my family – first of all to my grandparents, Xaver and Resi, who have not lived long enough to see this project come to fruition, but whose love for the forests very much lives on. To my aunt, Hermine, who has always encouraged me to be outdoors and whose values still shape my life. To my parents, Arthur and Isolde, whose steady and unquestioning support has made my whole academic career possible in the first place. An acknowledgement section in a PhD introduction does not do justice to what you have done for me and continue to do for me. And finally, I would like to thank my girlfriend, Netsanet, with all my heart, for having supported me throughout this project. You have had to bear the daily ups and downs that are part of a PhD project and you have showed great patience with me in the last weeks and months of my PhD, when algorithms had to be rerun and models put on servers even during our holidays. Your presence and love has been fundamental for making this possible!

Thank you!

Table of Contents

Table of contents **1**

Introduction **3**
A. Forests in an age of planetary challenges 3
B. Towards a predictive ecology: Functional traits, remote sensing and the integrative power of vegetation modelling 12
C. From regression to the mean to individual-based models: The importance of the individual 25

Objectives **30**

Chapter 1: Improving plant allometry by fusing forest models and remote sensing **60**

Chapter 2: A new method to infer forest structure and tree allometry from airborne laser scanning and forest inventories **69**

Chapter 3: Calibrating the short-term dynamics of the TROLL individual-based model in an old-growth tropical forest **113**

Chapter 4: The importance of considering inter-individual variation in allometric and functional traits in an individual-based forest model **157**

Chapter 5: Global patterns and evolutionary trends in wood density **189**

Discussion **236**
A. Individual-level variation and a predictive ecology at global scale 236
B. Future improvements for predictive modelling 239
C. The integration of predictive science with society 241

INTRODUCTION

Throw up a handful of feathers, and all must fall to the ground according to definite laws; but how simple is this problem compared to the action and reaction of the innumerable plants and animals which have determined, in the course of centuries, the proportional numbers and kinds of trees now growing on the old Indian ruins!

Charles Darwin in *On the Origin of Species (1859)*

A. Forests in an age of planetary challenges

The significance of forests

Ever since Gilgamesh set out on his journey to the Cedar Forests in Lebanon (George, 2003) – and presumably much earlier –, human beings have shown a deep fascination for forests. Trees and forests have been variously worshipped in sacred groves or mythologized as the center of the universe (Sturlson, 2005). They are the place of legendary battles[1] and philosophical thought experiments, have been feared for the creatures that might emerge from them, and turned into potent metaphors that continue to inform our thinking[2]. In 18th and 19th century Europe, under the influence of burgeoning industrialization and secularisation, Romantic poets, in particular, rediscovered them both as save havens from emerging technology and the uncanny locations where humans could encounter their unconscious (Cox, 1985; Pensel, 2019). During the same period, early scientists and explorers such as Alexander von Humboldt set out on their journeys to discover the tropical rainforests of South America, spurring a new fascination for the biodiversity and complexity of nature (von Humboldt &

[1] For instance, the Teutoburg Forest where Roman legions and Germanic tribes clashed, and Sherwood Forest, the hiding place of Robin Hood.

[2] This ranges from "deeply rooted" ideas, over proto-scientific knowledge collections that were known as forests, such as Francis Bacon's *Sylva Sylvarum* (De Bruyn, 2001; Rusu & Lüthy, 2017), to modern-day algorithms ("random forest").

Bonpland, 1814). These ideas continue to resonate until today, where forests are still regarded as refuges from civilisation and their destruction has become one of the most vivid symbols of humanity's impact on the Earth.

The fascination with forests, aesthetic and beyond, is deeply connected with the vital role forests have played and continue to play for human existence. It is not by accident that, already in the earliest texts we know of, they are linked to mankind's technological rebellion against the gods. Gilgamesh's infraction consists in cutting down the cedar trees. In Greek and Roman mythology, it is only in the paradisical Golden Age of mankind that "[n]o pine had yet, on its high mountain felled / Descended to the sea to find strange lands / Afar;" (Ovid, 2009). Throughout history, human beings have relied on functioning forest ecosystems, either directly – harvesting them for fuelwood and timber, the building of ships and accommodation – or indirectly – as sources of food and water, shelter from natural and human-made hazards, and for their mental health (Bratman *et al.*, 2012; Vira *et al.*, 2015).

In return – and as a result of their importance as resources –, human beings have always influenced forests. Forests across the world have been deeply and lastingly transformed by human beings, both in their extent and in their composition for thousands of years (Thompson *et al.*, 2013; Roberts *et al.*, 2018; Odonne *et al.*, 2019). In recent times, this influence has become even more pronounced, with large-scale destruction of forests looming over biomes across the world, threatening their biodiversity (Barlow *et al.*, 2016) and the many services they provide. And while humans have always transformed forests and often at regional scales, they are now transforming forests globally, rendering changes to the whole Earth-system a distinct possibility.

It is here, at these large scales, that forests and trees are directly linked to the future of humanity. First, they greatly influence the hydrological cycle, can change local and regional climates and thus, in some places, make human existence possible in the first place (Ellison *et al.*, 2017). Second, since plants, through photosynthesis, assimilate carbon and store it in their tissue, the biosphere in general and forests in particular represent a large terrestrial carbon sink (Pan *et al.*, 2011). They are estimated to take up around 30% of yearly carbon emissions (Le Quéré *et al.*, 2009; Quéré *et al.*, 2018), providing an essential buffer to anthropogenic carbon emissions and slowing down a changing climate. But what exactly becomes of the biosphere under further anthropogenic climate change, and how this will feed back into climate change itself, is an unresolved question (Bonan, 2008).

It is therefore essential to develop approaches and methods to better understand forest dynamics, acquire the data sets to test our hypotheses and develop tools that can transform theoretical knowledge into practical predictions for the management of forests at global scales (Millar *et al.*, 2007).

Climate change – a planetary challenge

The overarching challenge in managing and understanding the future of the world's forests is, first and foremost, that they are embedded in a global system that is in itself complex. Therefore, the future of forests cannot be extricated from the larger systematic changes that the Earth system as a whole is currently undergoing as a result of climate change.

The source of current climate change is well known: an increase in anthropogenic carbon emissions, leading to an increase in atmospheric carbon levels

from a preindustrial level of 280ppm to more than 400ppm in recent years[3]. Equally well-known is the basic mechanism: greenhouse gas molecules such as carbon dioxide vibrate in ways that their energy spacings correspond to the frequency of infrared radiation, but not the visible spectrum, and thus capture and reemit a large portion of energy back to the Earth's surface.

But the resultant anthropogenic climate change confronts human society with a planetary-scale challenge whose solution lies beyond nation states and whose impact will likely persist for centuries or even milennia (IPCC, 2014). Over the past decades, a concerted effort has been made to quantify and monitor the current state of the Earth and its ecosystems (Goetz *et al.*, 2009), define limits beyond which the Earth's systems might undergo irreversible change (Rockström *et al.*, 2009; Lenton, 2011) and to provide technological (Keith, 2009; Praetorius & Schumacher, 2009) and non-technological solutions (Jackson *et al.*, 2008) to offset carbon emissions or adapt to climate change (Smith *et al.*, 2011). Prediction, however, is at the heart of most scientific activity – prediction both in the sense of predicting current patterns and, more importantly for climate change, in the sense of forecasting future patterns. For science to support political decision making, it needs to provide scenarios for future states of the Earth system (Meehl *et al.*, 2002), and it is exactly here that climate change poses one of its biggest challenges.

Climate dynamics are complex in the sense that they include many nonlinear effects, feedbacks and subsystems that respond at different timescales (Colman & McAvaney, 2009). Some of these feedbacks can be predicted with high confidence, such as surface albedo decreasing with decreasing ice extents and, in turn, reinforcing

[3] Source: NOAA Earth System Laboratory, Global Monitoring Division: https://www.esrl.noaa.gov/gmd/ccgg/trends/global.html, last accessed on October 12, 2019.

temperature increases (Curry *et al.*, 1995). Others such as aerosol-cloud interaction are highly uncertain (IPCC, 2014). While there is consensus that catastrophic whole-Earth system changes, such as a runaway greenhouse effect on Venus, are highly unlikely due to anthropogenic forcing (Goldblatt & Watson, 2012), there is considerable uncertainty about so-called "tipping points", abrupt system shifts brought about, for example, by changes in oceanic circulation, a rapid loss of icesheets or vegetation diebacks (Schellnhuber, 2009; Lenton, 2011) that could potentially alter a vast part of Earth's ecosystems and have catastrophic consequences for human society (cf. Figure 1 for a geographic overview).

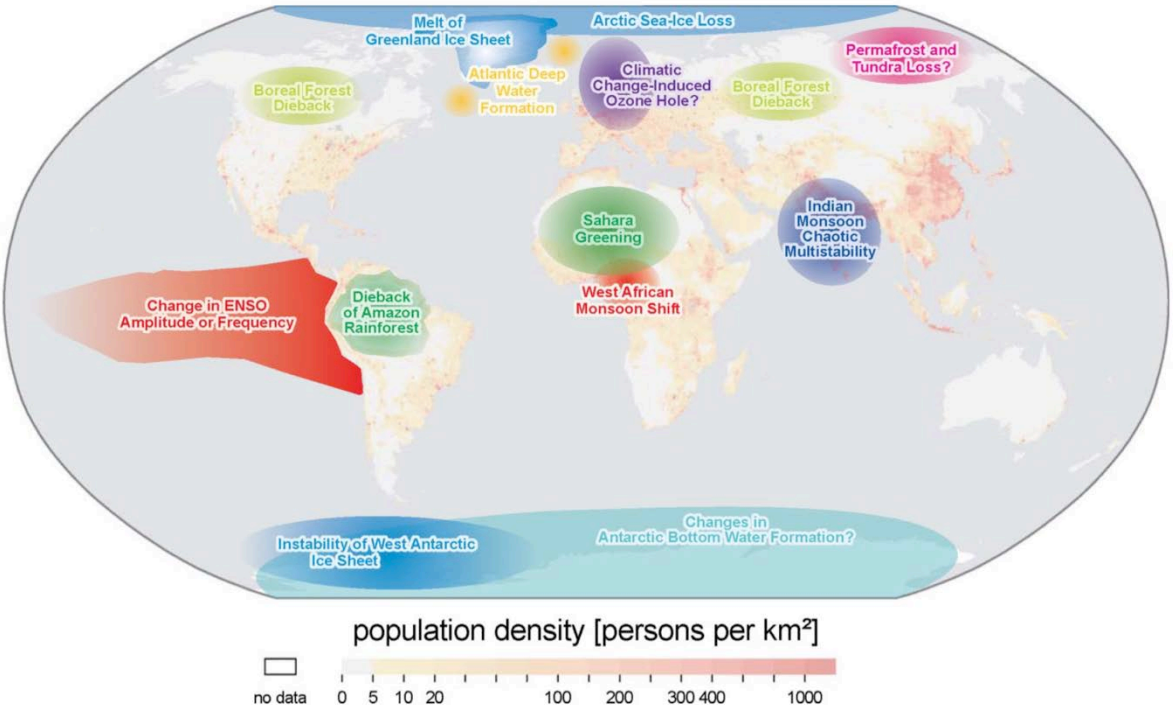


Figure 1: Map of potential tipping points due to climatic change in the 21st century. Shown are large-scale shifts of some of the Earth's subsystems that could occur in the 21st century, overlaid on a map of population density and modified with a question mark when particularly unlikely. The figure is taken from Lenton *et al.*, 2008. It should be noted that the likelihood of some of the shown tipping points has been reassessed more recently, and that the "Amazon dieback", for example, is now estimated to have much lower probability than initial model estimates suggested (Malhi *et al.*, 2009; Rammig *et al.*, 2010; Good *et al.*, 2013).

But even if tipping points are avoided, climate change will have and already has important consequences for the biosphere. Changes in global temperatures and weather variability impact the suitability of habitats, with ecosystems and organisms responding in various ways, through spatial and temporal shifts, adaptations (Bell & Gonzalez, 2009) or individual-level plasticity. Not all responses affect biodiversity negatively (Bellard et al., 2012), but they likely lead to changes in global ecosystem functioning (cf. Figure 2), and further cascading effects on the Earth system. Ecological regime shifts, for example, have the potential to substantially endanger human livelihoods, affect food supplies and increase the prevalence of diseases worldwide (Godfray *et al.*, 2010; Altizer *et al.*, 2013; Scheffers *et al.*, 2016). This, in turn, increases the likelihood of further rapid changes, such as mass migrations and armed conflict (Reuveny, 2007; Kelley *et al.*, 2015).

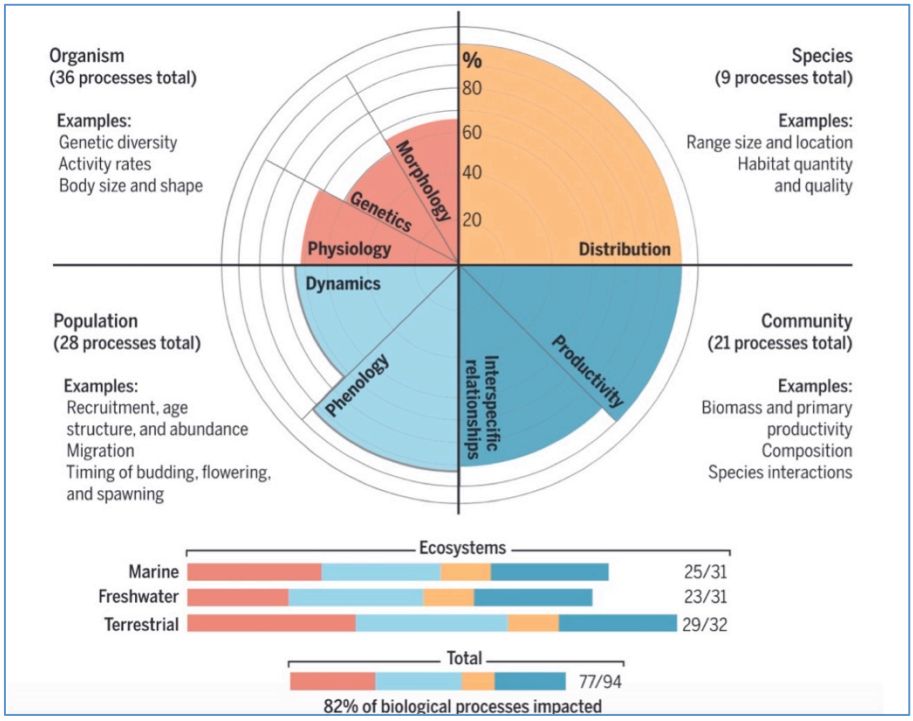


Figure 2: Impact of climate change on ecosystem processes. This figure shows that out of 94 processes identified in biological systems, 77 were impacted by climate change. Figure taken from Scheffers *et al.*, 2016 .

The uncertainty of vegetation dynamics

While high uncertainty is thus a feature of climate change as a whole, it is a particularly prominent property of forest ecosystems. How exactly forests, and tropical forests in particular, with their high levels of biodiversity and large networks of organisms, will respond to further climate change is one of the great challenges to tackle.

On the one hand, a number of factors reinforce carbon sequestration in forest ecosystems, such as increasing carbon dioxide concentrations that lead to a carbon fertilization effect (Zhu *et al.*, 2016), increasing temperatures that result in higher metabolic activities of plants (Dusenge *et al.*, 2019), and changes in phenology that lead to extended growth periods for plants in temperate and boreal regions (Cleland *et al.*, 2007). On the other hand, climate change might not only result in higher assimilation rates, but also considerably increase respiration and biomass turnover – with photosynthetic activity decoupled from carbon sequestration (Malhi, 2012; Fatichi *et al.*, 2014). Furthermore, stronger variability in precipitation and temperature patterns and more extreme events, such as longer and more extreme droughts, might increase tree mortality or suppress tree growth (Phillips *et al.*, 2009; Bonal *et al.*, 2016), or even result in vegetation diebacks and a transition of ecosystems to new quasi-equilibrium states (Malhi *et al.*, 2009).

There is a particular risk when self-amplification and nonlinear feedbacks are involved. Since forest ecosystems feed back into local climates through evapotranspiration (Salati *et al.*, 1979; Eltahir & Bras, 1994; Moreira *et al.*, 1997), increasing droughts leading to tree mortality in tropical forests can in turn amplify the risk of further droughts and further tree mortality (Zemp *et al.*, 2017). One particular scenario that has been evoked in this context is the transition from tropical rainforest to

savanna – or from savanna to forest – due to changes in precipitation levels (Lobo Sternberg, 2001; Hirota *et al.*, 2011).

Empirical evidence suggests that, in the past decades, tropical forest ecoystems such as the Amazon or African rain forests have not undergone such drastic changes, but rather had steady carbon accumulation rates of ca. 0.5 MgC $ha^{-1} yr^{-1}$ (Phillips *et al.*, 2008; Lewis *et al.*, 2009), offsetting losses through deforestation and land-use changes by increased growth (Gaubert *et al.*, 2019). Furthermore, tropical forests seem to have large restoration potentials, with secondary forest recovery buffering losses due to deforestation elsewhere (Poorter *et al.*, 2016).

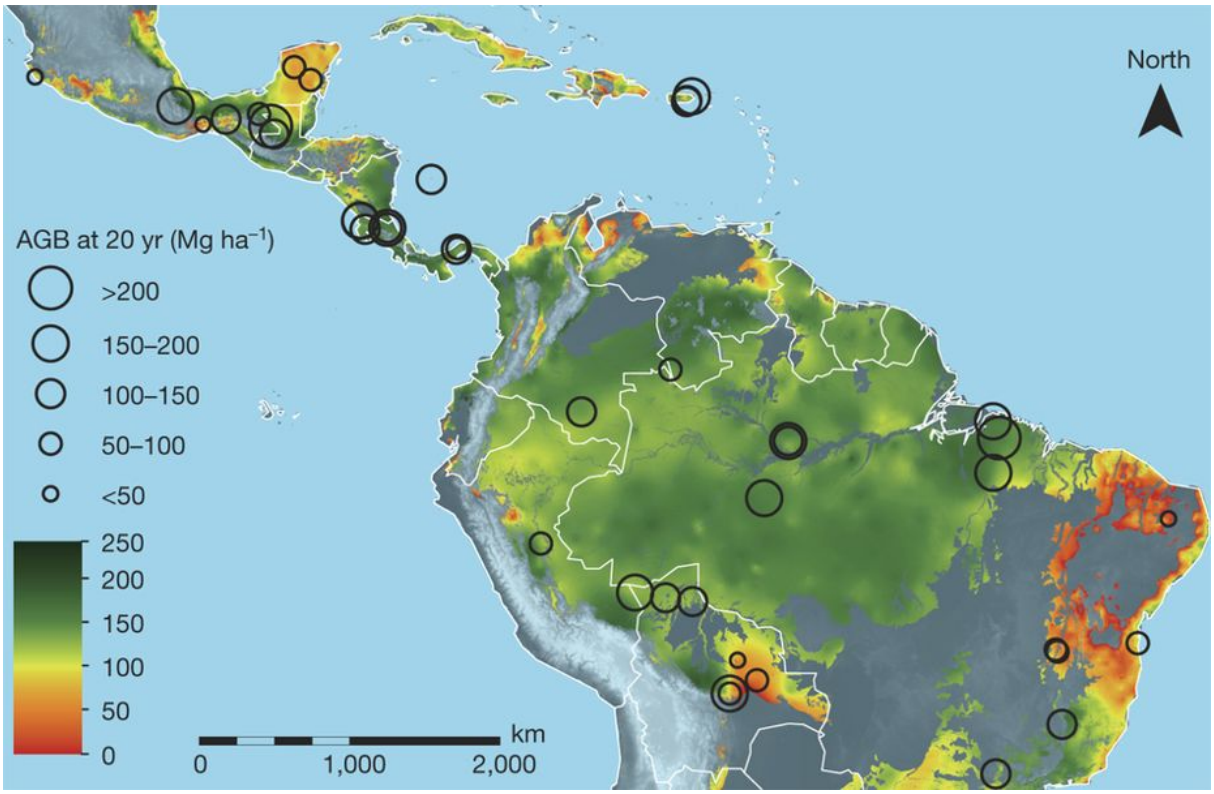


Figure 3: Potential aboveground biomass accumulation due to secondary forest growth in the Neotropics. The map shows the potential accumulation of aboveground biomass over a time span of 20 years due to lowland secondary forest growth, based on 44 study sites (indicated by circles, scaled according to their carbon accumulation rates). From Poorter *et al.*, 2016.

However, to what extent these trends and potentials will persist is unclear. Evidence has also emerged that the sink potential of tropical forests is slowing down (Brienen *et al.*, 2015), and there is considerable uncertainty regarding a potential transition from carbon sink to source (Baccini *et al.*, 2017). While a uniform response from a system such as the Amazon is generally unlikely, continued deforestation and land-use change will have strong impacts regionally (Lewis *et al.*, 2015; Houghton & Nassikas, 2017) and increase the risk of long-term, gradual degradation, mainly due to changes in hydrology that could result in further biomass losses and a shift from rain to dry forests (Malhi *et al.*, 2009; Rammig *et al.*, 2010; Levine *et al.*, 2016).

The challenge to correctly assess feedbacks and predict the future of our forests, is enormous, and does not only rely on a good understanding of biogeochemical processes, but also the demographics and evolutionary dynamics that shape and are shaped by these processes (Aitken *et al.*, 2008; Fisher *et al.*, 2018). This brings us to the initial quote by Charles Darwin. While Darwin might have been very generous in his assessment of the predictive power of physics, the problem of predictability in ecology and evolution has become even more accute in the light of climate change and global forest losses. Knowing that we fundamentally rely on functioning forests, their contribution to hydrological cycles and global carbon stocks, the overarching question that poses itself is: How can we understand, predict and manage forests globally? And how, in particular, can we forecast the future of tropical rainforests with their high biodiversity and complex community dynamics?

B. Towards a predictive ecology: Functional traits, remote sensing and the integrative power of vegetation modelling

Prediction in ecology

While it is impossible to define such a complex, social endeavour as scientific inquiry in a single, overarching concept, prediction has always been one of its main tenets. Together with explanation and the continuous process of methodical data acquisition and hypothesis testing – linking empirical observations through rules –, prediction – using those rules to make inferences about what is or about what is yet to come – could be said to constitute the sciences' "family resemblance" (Wittgenstein, 1953). While some have held that explanation can ultimately be reduced to prediction – i.e. a good scientific explanation is one that is successful in predicting patterns (Peirce, 1878; Dewey, 1903; Houlahan *et al.*, 2017) –, others have pointed out that scientific theories, even though they are often used for predictive purposes, are more than instruments (Popper, 1963).

Irrespective of the larger philosophical implications, however, there seems to be a broad underlying agreement that scientific inquiry is a formalization and sophistication of everyday inquisitiveness, i.e "enlightened common sense" (Popper, 1972), and heavily revolves around gathering information to solve problems. If one conceives of biological adapations encoded in organisms as information (Davies, 2019), this practical aspect of science could even be regarded as an extension of the trial-and-error process of life in general (Dewey, 1903; Popper, 1990). Broadly, the process by which a crow finds out how to crack the shell of a walnut is mirrored in a physicist's

attempt to excite molecules or the ecologist's attempt to predict patterns of biodiversity[4].

In this broad sense of the word, studies of ecology and evolution have always been predictive. Even when experimentation has to be replaced by observational studies, we hypothesize that relationships and patterns between variables found at one specific site will be similar at another site – i.e. transferable (Wenger & Olden, 2012) –, and if they are not, that we can ameliorate our original model to incorporate the deviation into future hypotheses. We expect, for example, that the rules that govern bird speciation on islands will share some general characteristics with those that have been found for Darwin's finches on Galapagos (Grant, 1996), or that patterns of species abundances at one site will be repeated at another site – not with the exact same parameters or species in either case, but based on some basic principles that do not change.

At the same time, the high complexity of ecosystems and the many variables involved, most of them difficult to measure – e.g., the behavior of animals, the development of trees across decades –, render the falsification of hypotheses complicated. While there have been, for example, experimental manipulations of ecosystems to assess the consequences of climatic change (Brando *et al.*, 2008; Norby & Zak, 2011), traditional hypothetico-deductive methods quickly reach their limits when addressing systems such as tropical rainforests that evolve on global scales and over timescales larger than any single researcher's lifetime. Furthermore, if a particular hypothesis cannot be supported in ecology, alternative explanations are legion (Quinn & Dunham, 1983). While this might explain why ecology has often focussed on broad patterns instead of quantitative predictions and why there are wider issues in

[4] The physicist and ecologist presumably have stricter protocols than the crow.

replication in the social and life sciences (Earp & Trafimow, 2015), there is a need to produce predictions and subject theories to rigorous testing (Mouquet *et al.*, 2015; Houlahan *et al.*, 2017).

In the past two decades, ecology has, however, seen a number of developments that hold great promise to enhance its predictive capabilities. These developments comprise the study of functional traits as a mechanistic link between species and their environment, the deployment of remote sensing techniques at unprecedented scales to discover patterns across the globe, and the development of dynamic vegetation models that can serve as integrators for various data sources and put the reliability of ecological knowledge to the test.

Trait-based ecology and scaling laws

Traits are properties of individual organisms at the phenotype level, and linked to an organism's evolutionary history through its genes. Many of these traits, due to their evolutionary origins, play a particular functional role for the organism, and when these functional roles can be linked to overall ecosystem functioning, a framework emerges that links biodiversity and community ecology with ecosystem functioning (Grime, 1997), comprises evolutionary history (Reich *et al.*, 2003) and allows for predictions across different environmental conditions (Reich *et al.*, 1997).

With regard to plant traits it has, for example, been shown that they strongly impact the competitive abilities of plants, that they can promote coexistence through functional differentiation, and that they thus are intrinsically linked to the vital rates of organisms (Poorter *et al.*, 2008; Hérault *et al.*, 2011; Kraft *et al.*, 2015; Kunstler *et al.*, 2016; Visser *et al.*, 2016). Although variability around predicted relationships is often high, thus substantially reducing the precision of prediction (Paine *et al.*, 2015; Poorter

et al., 2018), functional traits have proved an important step towards a predictive ecology across scales (Lavorel & Garnier, 2002).

In particular, relying on large global data sets (Kattge *et al.*, 2011), trait-based ecology has identified fundamental trade-offs between various leaf and wood traits, leading to the hypothesis of a leaf economics spectrum (Wright *et al.*, 2004) and a wood economics spectrum (Chave *et al.*, 2009). Among the trade-offs found are, for example, a positive correlation between leaf mass per area and leaf lifespan and a negative correlation between leaf mass per area and photosynthetic capacity – meaning that plants either invest in long-lived and expensive leaves with lower photosynthetic capacity, or in short-lived, but cheap and photosynthetically more productive leaves, thus aligning on a "slow-fast" continuum of plant functioning (Reich, 2014). Similar trade-offs, such as the trade-off between high mortality risks and fast growth at low wood densities (compared to slow growth, but low mortality) can be found for wood density. In a recent global synthesis, further comprising traits such as adult height and diaspore mass, Díaz *et al.* (2016) showed that there were clear trade-offs in investment, suggesting strong evolutionary limitations in the way plants invest the carbon they acquire.

In this light, trait-based approaches can also be seen as part of a larger program in ecology that is interested in the scaling relationships within and between organisms (Jarvis, 1995). Given the relative rarity of isometric scaling (i.e. the preservation of proportions, but see Reich *et al.*, 2006), these scaling relationships are also often simply called allometries. Allometries have been found empirically for a wide number of plant traits and dimensions, such as the scaling of leaf area with sapwood area (Vertessy *et al.*, 1995), of crown dimensions with stem diameter (Jucker *et al.*, 2017) and above-ground biomass and stem diameter (Chave *et al.*, 2014). Given that stem diameter

measurements are routinely measured in ground inventories, but biomass estimates could only be directly accessed through destructive sampling, the latter allometry in particular has served as an important tool to predict above ground biomass at global scales.

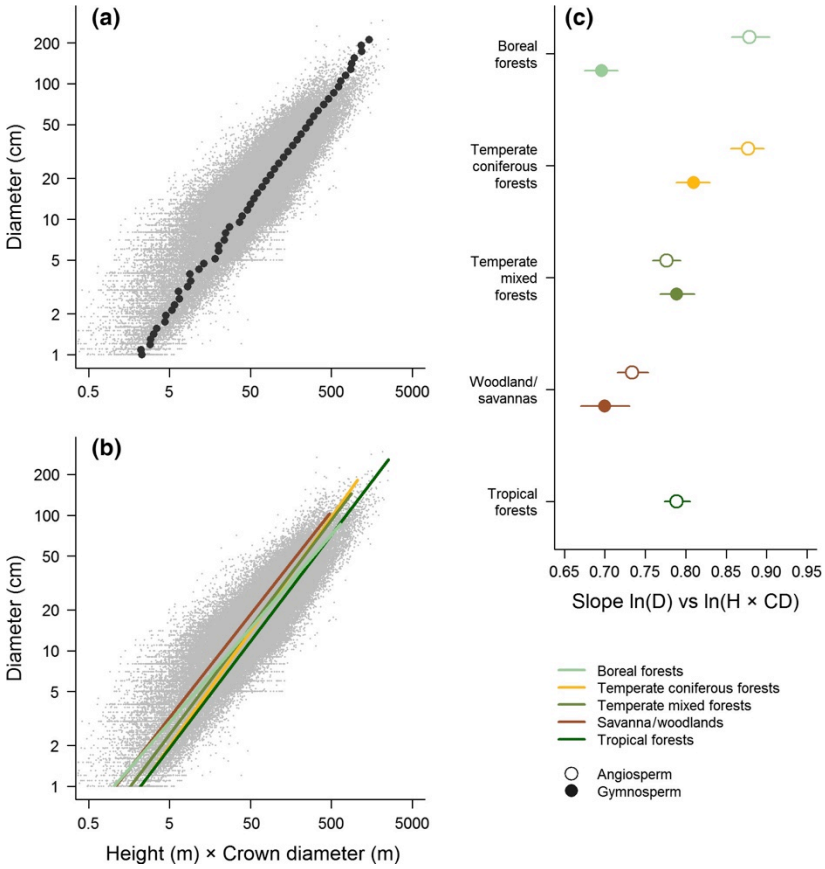


Figure 4: Global scaling relationship between crown dimensions and stem diameter. Shown is the relationship (on logscales) between the product of tree height and crown diameter with stem diameter for different forest types. Panel a) shows raw data and the means for each diameter class, panel b) separate relationships for different forest types and panel c) the corresponding slope parameters. From Jucker *et al.*, 2017.

What makes scaling relationshps an important part of predictive ecology is that they reflect physiological and evolutionary constraints of organisms (Niklas, 1994) and could thus reflect general underlying laws. Metabolic scaling theory has developed a framework to explain a wide number of scaling relationships found in animals and

plants through the branching patterns of vessels and physiological constraints on fluid flow (West *et al.*, 1999; Enquist *et al.*, 2003). And while there are notable deviations from metabolic scaling (Reich *et al.*, 2006; Enquist *et al.*, 2007; Coomes *et al.*, 2011; Price *et al.*, 2012), the generality of scaling rules and the use of testable predictions have contributed strongly to a more predictive ecological research agenda.

Remote scaling as a global observatory

While trait-based research has focussed on the properties of individual organisms that are measured through intense ground work in many, but dispersed, sampling sites, an ecology considering processes at global scales increasingly needs data at global scales. In many ways, the challenges encountered here mirror those of astronomy: systems operating at spatial and temporal scales that can rarely be accessed by human beings, a resulting lack of experimental opportunities and the need to rely on observational evidence. As a consequence, some of ecology's approaches also mirror closely those of astronomy, especially in the use of so-called remote sensing techniques.

Remote sensing, in the broadest sense of the word, is the inference of object properties without physical contact. It is based on the idea that the frequency distribution of a radiative signal is shaped by the objects that have emitted or reflected it and the media through which the radiation has passed on the way to the receiver. Where astrophysics infers the composition of stars through absorption lines associated with elements or molecules, ecology can use the spectral signature of signals to infer the existence of vegetation, its density and, ideally, its chemical composition. Remote sensing in ecology takes both active forms – probing objects and organisms of interest with a source of radiation – and passive forms – using already existing sources of radiation such as sunlight or thermal radiation –, and can then be used to gather

information either about individual organisms ("direct") or about these organisms' environment ("indirect) (Turner *et al.*, 2003).

The use of remote sensing in ecology dates back at least to the early 1970s when researchers started to use data from the Landsat satellite to construct vegetation indices (Rouse *et al.*, 1973; Tucker, 1979). The Normalized Difference Vegetation Index (NDVI), based on the different absorption and reflectance properties of photosynthetic tissues in the near infrared region and the visible part of the light spectrum, has proved a valuable source for ecological studies ever since (Pettorelli *et al.*, 2011). Compared to early uses of remote sensing, we nowadays possess a much wider arsenal of remote sensing sources – from the optical part of the spectrum, over infrared to radio waves – that combine to form a large global observatory of the Earth system and its biosphere, or "flux towers in the sky" (Schimel & Schneider, 2019). This has allowed for a wide variety of ecological questions and challenges to be tackled, including wetland methane emissions (Bloom *et al.*, 2010), estimates of biomass (Le Toan *et al.*, 1992; Saatchi *et al.*, 2011), primary productivity (Frankenberg *et al.*, 2011), leaf phenology (Richardson *et al.*, 2009), the detection of invasive species (Asner *et al.*, 2008) and biodiversity predictions (Gillespie *et al.*, 2008), to name but a few.

In forest ecology, lidar ("light detection and ranging"), or the use of lasers in the visible or infrared spectrum, has proved a particularly powerful tool. Lidar systems measure the distance between the emitter and the object that is scanned and thus can estimate the position of the scanned objects, with the resulting data typically obtained either as full waveform data or as discretized returns (cf. Figure 5). On one end of the scale, its terrestrial version (TLS) allows for detailed reconstructions of forests tree by tree (Calders *et al.*, 2018) that can, in turn, be passed on to further studies, such as

susceptibility to wind (Jackson *et al.*, 2019), precise biomass estimation (Disney *et al.*, 2018) or to supplement and aid forest inventories (Bauwens *et al.*, 2016).

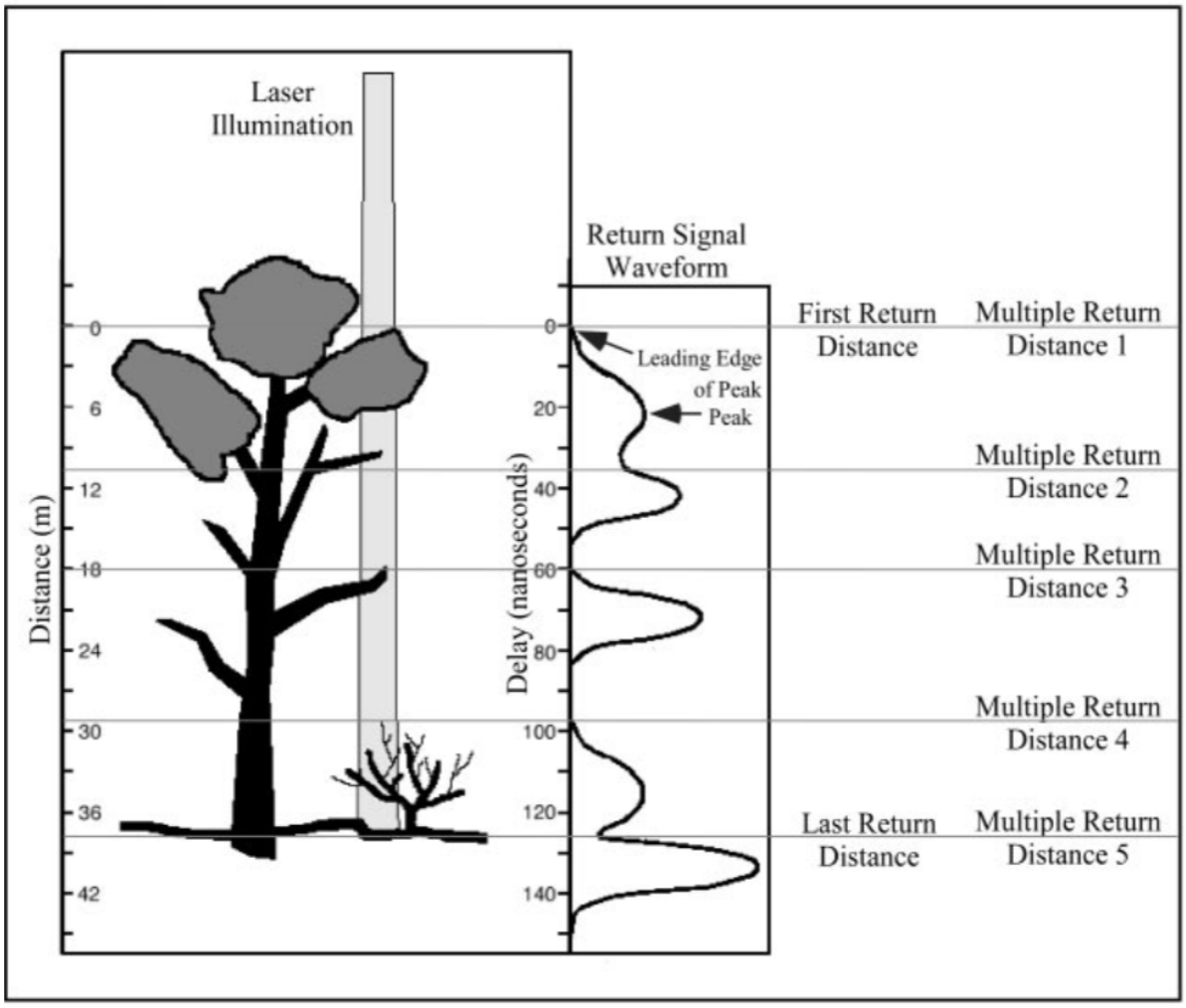


Figure 5: Illustration of airborne lidar scanning. Shown is an illustration of vegetation being scanned by an airborne lidar instrument, as well as the lidar waveform that would be recorded. Also shown is the concept of translating the waveform into discrete returns by identifying the leading edges of peaks in the signal. Figure from Lefsky *et al.*, 2002.

When mounted on airplanes (so-called airborne lidar or ALS), on the other hand, lidar scans can be carried out across thousands of hectares and then be used to infer detailed topographic information (Höfle & Rutzinger, 2011), delineate individual trees across forest types (Morsdorf *et al.*, 2004; Vega *et al.*, 2014; Ferraz *et al.*, 2016) and estimate the underlying biomass, with a wide range of techniques available (Asner & Mascaro,

2014; Coomes *et al.*, 2017). Finally, this approach can also be extended to satellite-based observations (spaceborne lidar, SLS), allowing to create estimates of forest height on regional to global scales (Rosette *et al.*, 2008; Simard *et al.*, 2011). NASA's recently launched GEDI system onboard the international space station, for example, will provide high resolution spaceborne data and thus considerable improve estimates of forest structure globally (Qi & Dubayah, 2016).

Together with other future missions, such as the radar-based BIOMASS mission (Le Toan *et al.*, 2011), and supported by a strong network of field data (Chave *et al.*, 2019), remote sensing provides the spatial extent in data that is needed to make and validate predictions at global scales.

Dynamic vegetation models

Trait-based approaches and remote sensing represent important and complementary steps in the development of testable predictions about forest ecosystems. It remains, however, a challenge to link them to each other (Homolová *et al.*, 2013; Antonarakis *et al.*, 2014) and to a wider body of ecological theory, such as demographic processes (Salguero-Gómez *et al.*, 2018). While allometries, in particular, have served as important tools to translate between ground and remote sensing observations (Asner & Mascaro, 2014; Jucker *et al.*, 2017), further techniques are needed that can easily translate between various data types and approaches, take into account both dynamic changes and spatial extent and synthesize knowledge across scales. And this is where the great power of vegetation models lies: They bring together knowledge from various domains of ecology to build a virtual version of the observed ecoystems and thus both identify knowledge gaps in current ecological theory and derive predictions for the future (Shugart *et al.*, 2015; Schimel & Schneider, 2019).

Vegetation models have a comparatively long history. The growth and yield tables employed by German foresters in the late 18th century are often regarded as the beginning of forest growth modelling (Vanclay, 1994; Pretzsch, 2009). With the objective to predict the yield of forest stands (i.e. the expected volume increments), foresters collected data over years and extrapolated to future conditions, initially using graphical methods and later with more sophisticated equations and models (Peng, 2000). Forest models have thus, from the very beginning, been tied closely to predictive aims and were designed to help in decision-making processes. According to the respective aims of practitioners, forest models could vary widely in their complexity, including differential equations (Garcia, 1983), cellular automata (Karafyllidis & Thanailakis, 1997) or complex process-based models of later generations (Friend *et al.*, 1993).

While the exact terminology of vegetation modelling varies (Porté & Bartelink, 2002), two important steps in the development of the discipline have been process-based ecoystem models on the one hand (McMurtrie & Wolf, 1983; Running & Coughlan, 1988; Running & Gower, 1991) and so-called gap models on the other hand (Botkin *et al.*, 1972; Shugart, & West, 1980). The former were developed with the emergence of remote sensing data in mind and aimed at representing ecosystem processes, such as carbon, water and nutrient cycling, mechanistically through the explicit calculation of processes such as photosynthesis, respiration, and evapotranspiration. While they strived for high accuracy concerning exchanges of mass and energy, they relied, however, on minimalistic information regarding species and their ecology, often representing the whole forest canopy as one "big leaf" (Monteith, 1981; Running & Coughlan, 1988). Gap models, on the other hand, were designed to

mirror the mosaic dynamics of forest growth, simulating individual tree growth and vertical competition for light on small patches (typically 100 to 1000m^2) that were decoupled from each other and could thus represent different successional states. While gap models successfully simulated the dynamics of a wide variety of ecosystems, became more sophisticated over time, e.g. through explicit tree representations (Pacala *et al.*, 1996; DeAngelis & Mooij, 2005) and continue to inform models (Fischer *et al.*, 2016, cf. Figure 6), they have often had the drawback of using highly aggregated, empirical relationships to calculate processes such as mortality or even tree growth (Bugmann, 2001).

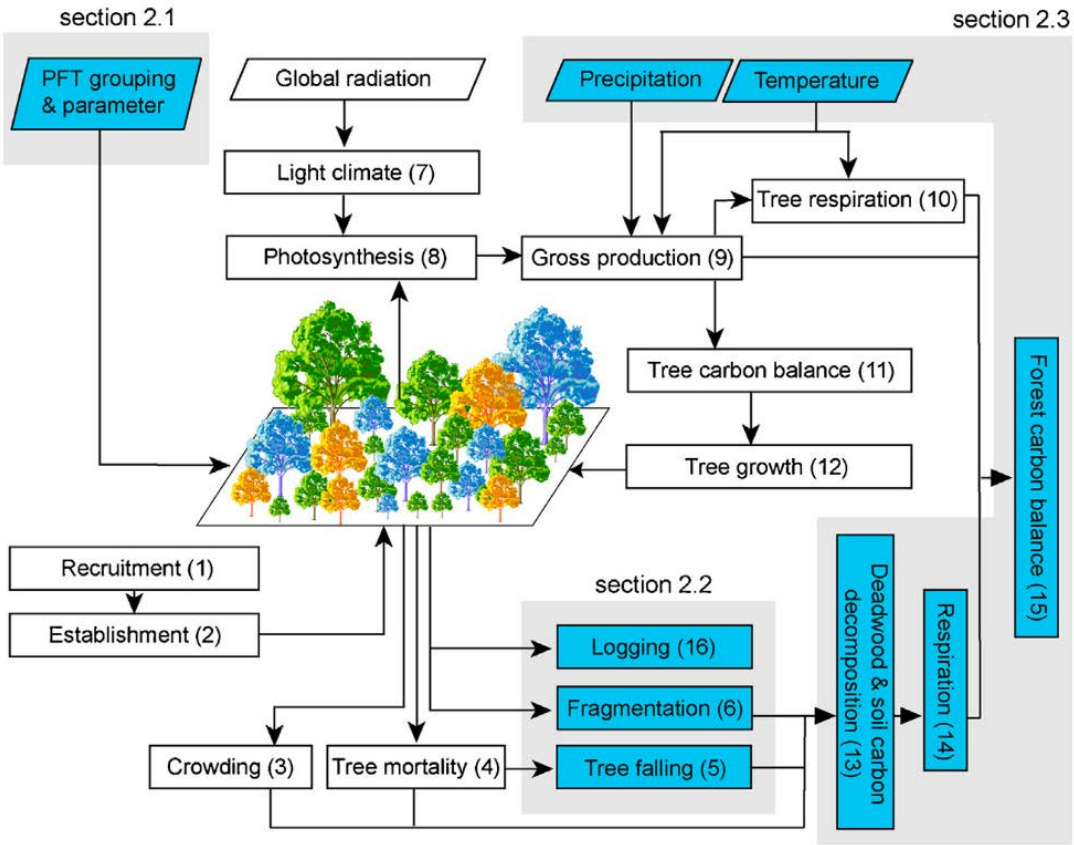


Figure 6: Conceptual representation of the FORMIND gap model. The figure shows the basic processes modelled in the forest gap model FORMIND in white, as well as additional simulated processes in blue. Parallelograms describe model input. Numbers behind the respective processes indicate the sequence in which they are carried out. Sections refer to the publication this figure has been taken from (Fischer *et al.*, 2016).

However, since the 1990s, a growing synthesis between the two model types has begun to emerge. The first Dynamic Global Vegetation Models (DGVM) such as HYBRID (Friend *et al.*, 1997) or the Integrated Biosphere Simulator IBIS (Foley *et al.*, 1996) combined the simulation of light competition between trees known from gap models with the fundamental ecosystem functioning known from process-based models and were then coupled to General Circulation Models (GCMs). While the resulting climate and carbon cycle projections were often highly diverging, with the tropics and Amazonia a particular source of uncertainty (Sitch *et al.*, 2008), DGVMs thus presented a unique opportunity to forecast the future of the Earth's system, analyze atmosphere-biosphere feedbacks and assess the effects on the global carbon cycle.

Of particular concern in the development of DGVMs has been the use of so-called PFTs – plant functional types – that group together a wide range of species into entities that are supposed to have similar plant functioning (Purves & Pacala, 2008). While often necessary to reduce computational demands, PFTs can act as hidden model-tuning parameters (Scheiter *et al.*, 2013) and most likely fail to capture the complex dynamics of forest ecosystems. The large biodiversity of tropical forest vegetation has, for example, often been subdivided into only two very broad categories – tropical broadleaf evergreen and tropical broadleaf deciduous trees (Foley *et al.*, 1996; Sitch *et al.*, 2003; Krinner *et al.*, 2005). The question of whether this representation is adequate feeds back into one of the fundamental ecological questions, i.e. how biodiversity affects ecosystems (Sutherland *et al.*, 2013) and the particular role of rare and extreme trait combinations in the functioning of ecosystems (Ter Steege *et al.*, 2013; Violle *et al.*, 2017).

On most of the issues plaguing early-generation DGVMs (Prentice *et al.*, 2007), however, considerable progress has been made in the past years. Advances include the integration of demographic processes (Medvigy *et al.*, 2009; Fisher *et al.*, 2018), the multiplication of plant functional types or their conversion into continuous trait distributions (Pavlick *et al.*, 2013; Fyllas *et al.*, 2014; Sakschewski *et al.*, 2015), and new methods to rapidly translate from individual-based dynamics to large-scale ecosystem functioning (Purves *et al.*, 2008; Strigul *et al.*, 2008). Furthermore, modern successors to early gap models are now used at large scales to provide estimates of biomass (Rödig *et al.*, 2017) and analyze biodiversity-productivity relationships (Bohn & Huth, 2017). With their ability to thus harmonize trait-based research, assimilate remote sensing data and translate them into predictions of ecosystem functioning, dynamic vegetation models have become an essential part of predictive ecology and climate science.

C. From regression to the mean to individual-based models: The importance of the individual

Despite the advances in trait-based approaches, the large-scale data provided by remote sensing and the integrative potential of next-generation DGVMs, one largely unresolved issue remains how to exactly incorporate individual organisms such as trees into our predictive tools.

On the one hand, individuals could be considered one of the most basic, if not the most basic unit of ecology (Railsback, 2001; Begon *et al.*, 2005). While there has been considerable debate about what consitutes the units of selection in evolution (Brandon & Burian, 1986), many fundamental processes in ecology occur between individuals, including such diverse interactions as competition, facilitation, predator-prey relationships and pollination. Whether a tree can grow and survive in a particular environment, will depend on the biotic environment as much as on the abiotic environment. And while some aspects of the biotic environment are mediated by evolutionary history beyond the individual, i.e. what competitors the tree will encounter or what mycorrhizal funghi it will associate with, there is clear empirical evidence that an individual tree's growth is shaped by its neighbors and in return shapes its neighbors' growth (Wright *et al.*, 2014; Chen *et al.*, 2016; Williams *et al.*, 2017), with important consequences for community ecology (Chesson, 1986). This importance of the local environment and interactions between individuals translates also to ecological research. Given that individuals shape each other, a large number of ecological experiments and field measurements have been, by necessity, carried out at the individual level.

At the same time, individuals have rarely found their way into dynamic vegetation models – or as "average individuals" –, and how exactly individual-based measurements should be assimilated, has been an open question (Purves & Pacala, 2008). In a similar vein, trait-based ecology has often focussed on average traits, although trait-fitness relationships are likely mediated by individuals (Shipley *et al.*, 2016).

In part, this lack of consideration for individual variation is due to practicality: remote sensing data are typically not obtained at the individual level, averaging procedures greatly reduce computational efforts in models, and sampling efforts are much lower in trait-based studies when species means can be used. More specifically, if individual-level variation is discounted, large data bases can be used to supply missing trait data for species means (Zanne *et al.*, 2009; Kattge *et al.*, 2011).

Another reason for the neglect of individual-level variation might, however, also be found in the particularities of knowledge acquisition and the scientific method, i.e. the tendency towards generality and simplicity. The idea that the basic units of cognition, i.e. what is individual to objects and beings, cannot be known and is always averaged out in knowledge production ("Individuum est ineffabile") can be traced back to Aristotle (Aristotle, 1963). It is in this sense that science often relies on generalisations. Its most popular statistical model – ordinary least squares regression –, summarizes variation around the predicted mean as error. Ecological variation that cannot be accounted for is thus typically summarized as "unexplained variation".

Science's tendency towards generalisation has been further reinforced by a drive towards parsimony. The underlying principle, known as Occam's razor, states that "the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience." (Einstein, 1934). If a simpler explanation is available, that

nonetheless still describes the system well, it is generally preferred to one that requires many variables and assumptions. Individual variation, in short, should thus not be included, if our understanding does not require it.

A key point of the principle of parsimony is, however, that simplicity is justified only as long as it is adequate to the system that is studied, i.e. a model that is simple, but cannot predict or reproduce empirical patterns, is not a good model (Houlahan *et al.*, 2017). When all models are wrong and only some useful (Box, 1979), abstractions need to be judged by their predictive power and usefulness. While correlative studies might thus profit from applying parsimonious principles to distinguish measurement noise from real variation, this does not necessarily hold true for mechanistic modelling (Coelho *et al.*, 2019).

In particular, there are several ways in which individual-level information is mechanistically important for ecology in general and the prediction of vegetation dynamics in particular. Individual (or intraspecific) variation often dominates trait variation and can qualitatively alter ecological dynamics in the presence of nonlinearity (Bolnick *et al.*, 2003, 2011). Since most trait-trait relationships and allometric scaling laws are described on logarithmic scales, the prediction from the mean of several trait values is not the same as the mean of several predictions from trait values (cf. also Figure 7). In particular, if we assume ecosystems to be complex, i.e. with emerging dynamics, then small alterations in underlying distributions could have strong repercussions on the predicted dynamics. Furthermore, plasticity in plant traits and dimensions has consistently be shown to affect forest structure and functioning (Longuetaud *et al.*, 2013; Jucker *et al.*, 2015), and large parts of biodiversity might be explained by high-dimensional variation between individuals (Clark *et al.*, 2010).

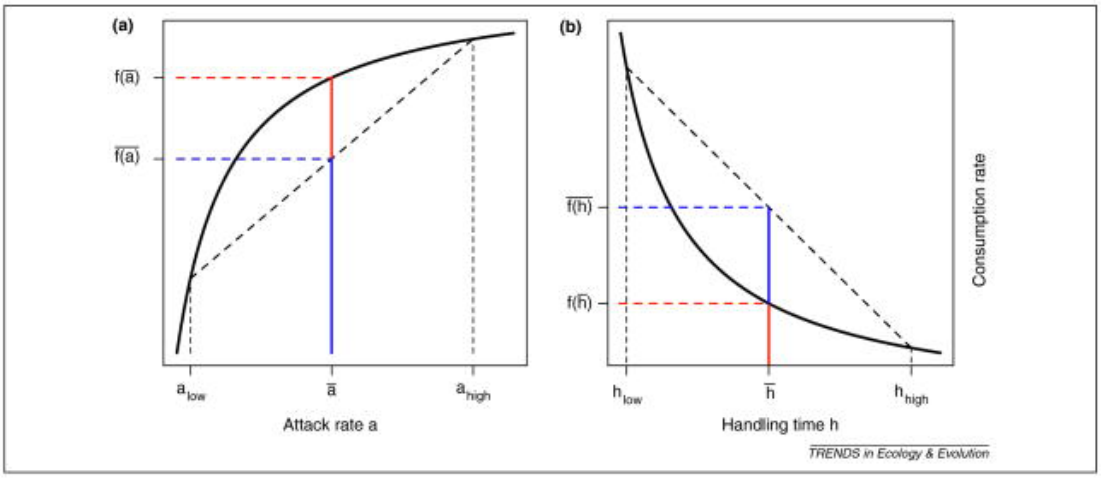


Figure 7: Jensen's inequality in ecology. Demonstration of the effect of individual variation in predator attack rate and handling time on feeding rate, when feeding rate is a convex function of attack rate and a concave function of handling time. Assuming dimorphic populations, i.e. with individuals that have either high or low attack rates (or handling times), average feeding rates are shown in blue. In red are shown the feeding rates of the average individual. This illustrates how using mean individuals in ecological models could produce wrong predictions despite relying on the same underlying relationship between attacking rates (or handling times) and feeding rates. From Bolnick *et al.*, 2011.

It is in this context that individual-based modelling (IBM) approaches are a highly promising tool for the future of predictive ecology. Originating from the gap-modelling philosophy, individual-based models of forest ecoystems have been developed early on. Simulating the dynamics of forests tree by tree (DeAngelis & Mooij, 2005), often in spatially explicit ways (Pacala *et al.*, 1996; Chave, 1999; Maréchaux & Chave, 2017), individual-based models were thus able to simulate ecosystem dynamics bottom-up and across scales. This allows not only for the simulation of interactions between individuals, the explicit integration of ecosystem functioning and biodiversity (Grimm *et al.*, 2017) or the mechanistic representation of mortality events such as treefall, but also for a highly flexible integration of data, from ground data up to remotely sensed canopy data (Shugart *et al.*, 2015; Knapp *et al.*, 2018).

And while the incorporation of individuals into dynamic vegetation models and variability around traits might make models more complex, it also allows for a wider variety of tests that can ensure that dynamics are rendered accurately. Pattern-based modelling, in particular, i.e. the repeated validation of submodels, can substantially increase the confidence in model predictions (Grimm *et al.*, 2005; Grimm & Railsback, 2012).

Summary: Towards large-scale predictions of forest dynamics

In summary, forest ecosystems, and particularly tropical rainforests are complex adaptive systems (Levin, 1998) that are essential to human life and that emerge from the competitive – and facilitating (Brooker *et al.*, 2008) – interactions of plant organisms. These organisms are modular, plastically react to their environment, develop over decades or centuries and are embedded in complex ecological and evolutionary processes. To better predict their future and inform human policy making, approaches are needed to adequately simulate these emerging dynamics. Coupled with remote sensing and the powers of modern computers, individual-based approaches offer a highly promising answer to this challenge, either by informing or even becoming part of DGVMs (Sato *et al.*, 2007; Smith *et al.*, 2014; Shugart *et al.*, 2018), thus opening up avenues for the large-scale prediction of forest growth from individual tree organisms at global scales.

OBJECTIVES

The ultimate aim of the work at hand was to further our understanding of the complex ecology and functioning of forests, specifically tropical rainforests – in the hope that, one day, we can predict the future of these ecosystems with reasonable accuracy and precision. At the heart of the effort lies the continued development of an individual-based forest growth model, TROLL, whose code had been first written 20 years ago (Chave, 1999) and which has been comprehensively updated more recently to simulate a tropical rainforest in French Guiana (Maréchaux & Chave, 2017).

While this most recent version of TROLL has shown good correspondence to a number of important metrics of forest structure and dynamics at the study site – aboveground biomass, primary productivity, successional dynamics (Maréchaux & Chave, 2017) –, it was the aim of this PhD project to further increase the realism and the predictive potential of the model to lay the ground-work for large-scale predictions. Most important in this regard was the assimilation of various data sources into TROLL, both for calibration and validation purposes, with a special focus on airborne laser scanning (ALS) and its potential to extend predictions across several thousands of hectares of forest. Figure 8 illustrates how continuous improvements in methodology in this PhD work have improved the match between TROLL canopies and ALS-derived canopy height estimates. This successful integration could serve as a blueprint for the integration of futher remote sensing sources, both terrestrial and space-borne, and, one day, the extension of TROLL to regional, if not global, scales.

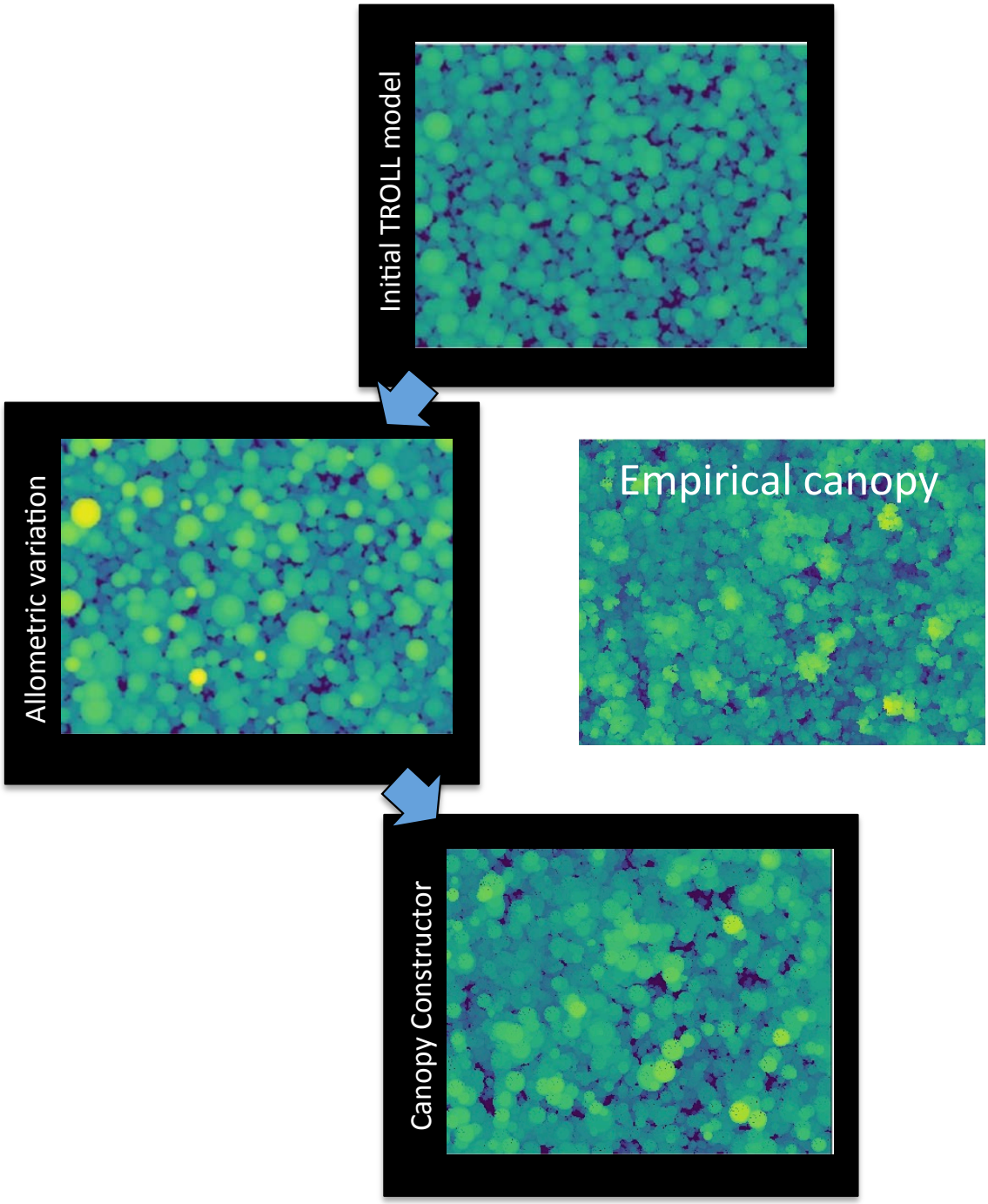


Figure 8: Improvements in the representation of empirically observed canopy height with the TROLL forest growth model. The above graph shows how TROLL-based rendering of empirical forest canopies have evolved from an initial simulation (upper panel) with little variation in tree height, over the inclusion of variation around allometric means (middle left) up to the spatially explicit fitting with the Canopy Constructor (bottom panel), as laid out in Chapter 2.

Throughout this work, individual-based model-building with TROLL invariably posed a number of challenges. First and foremost, since individual-based models simulate the dynamics of ecosystems bottom-up, they ideally give rise to similar emergent behavior as natural systems (Railsback, 2001; DeAngelis & Grimm, 2014). In practical terms this meant, however, that the causes of divergences from empirical patterns were often difficult to identify in a straightforward way. Second, individual-based models are often parameter-rich mechanistic models, including both explicit (tuned parameters) and implicit (empirical formulae) parameterisations and thus inherit the problems of other complex ecosystem models. A particular issue is equifinality, the potential to generate similar patterns as found empirically, but without simulating the underlying processes adequately (Beven & Freer, 2001). As a consequence, a delicate balance had to be maintained between mirroring natural complexity and keeping the model understandable and suitable for the purposes of prediction (Levins, 1966).

When integrating TROLL with both field data and airborne lidar scans, a major challenge emerged in how to constrain parameters to which the model was highly sensitive, but for which sufficient field data was lacking. This concerned, for example, allometric laws relating trunk diameter and crown extent, biomass allocation rules and mortality-related parameters. In particular, the question emerged of how to do this in an efficient way, considering the high computational demands of the spatially explicit TROLL simulations. How could we, for example, in a pratical way, decide which deviations from empirical canopies were due to useful model abstractions (i.e. crown geometry), and which deviations represented fundamental misrepresentations (treefall dynamics)? Did the model impose sufficient constraints to infer scaling relationships between tree properties? And how could this be extended to spatial scales relevant to

research on climate change? Much of this thinking has informed Chapter 1 and was published as a *Tansley Insight* (Fischer *et al.*, 2019).

What emerged from a continued pursuit of these questions was the idea of a divide-and-conquer approach, inspired by pattern-oriented modelling (Grimm *et al.*, 2005). Pattern-oriented modelling means that, if an individual-based model indeed simulates a natural system bottom-up, then it should also mirror the natural system across all levels of representation, and validation should be carried out at every scale. With this approach in mind, instead of fitting all of TROLL's parameters at the same time, the idea was to split the inference of forest structure and dynamics into two steps. The first step asked: What is the best representation of forest structure that we can create of an empirically observed canopy? Could we reconstruct a tree configuration that conforms to TROLL principles and that fits empirical data, having only a field inventory and an airborne lidar scan at our disposal? Could we potentially extend this to a larger area, e.g. the area covered by the whole lidar scan? The result was the so-called Canopy Constructor algorithm which shared some broad characteristics with recent developments by the FORMIND modelling group (Taubert *et al.*, 2015; Bohn & Huth, 2017), but would be a much more general tool to create geometric representations of tropical forests and to infer forest biomass or tree abundances at large scales (Chapter 2).

The second step of the divide-and-conquer approach was the translation of a static forest structure into a dynamical ensemble of growing and dying trees. Now that the Canopy Constructor provided a means to infer static allometries, it was necessary to link it with the process-based approach of TROLL. The most fundamental challenge here

was to translate an ensemble of trees that corresponded well to a lidar scan geometrically into a biologically viable old-growth forest and then to ensure that they would continue to coexist in a stable way. To improve stability and transferability, TROLL was updated to include intra-specific variation, crown plasticity and an improved representation of photosynthetic dynamics. Based on this work, the realism of TROLL with regard to empirical patterns could be assessed, and a platform was built for the future prediction of vegetation dynamics directly from remotely sensed old-growth forests (Chapter 3).

Once the model was calibrated, this new version of TROLL then served as basis to tackle a crucial ecological question, namely: How does inter-individual variation in traits and allometries and the plasticity of plants to environmental conditions influence whole-ecosystem properties in the old-growth forests of French Guiana? In ecological systems, non-linear responses between organisms are common (Bolnick *et al.*, 2011), and in forests in particular, individual variation has been shown to greatly impact whole-stand aboveground biomass (Pretzsch, 2014). This study thus served to make use of the recently developed improvements of TROLL to explore the relationship between variation and functioning more thoroughly (Chapter 4).

Finally, albeit not at the center of the thesis project at hand, plants traits have continually played an important role in it, be it for the inference of tree demography in TROLL, or as in the case of wood density, for the estimation of aboveground biomass with the Canopy Constructor. To apply both the Canopy Constructor and TROLL at large spatial and temporal scales, good collections of traits and a good understanding of the eco-evolutionary dynamics that shape them will be a prerequisite. It is in this context

that this PhD work zoomed out again, from the narrower questions of model-calibration and individual variation in the Guianas to the wider patterns of wood density variation. Over the course of two years, a new global data base of wood densities was created, to update a previous collection (Chave *et al.*, 2009; Zanne *et al.*, 2009), and to explore both the evolutionary divergences and global distribution of wood density (Chapter 5).

LITERATURE

Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S. 2008. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* **1**: 95–111.

Altizer S, Ostfeld RS, Johnson PTJ, Kutz S, Harvell CD. 2013. Climate change and infectious diseases: From evidence to a predictive framework. *Science* **341**: 514–519.

Antonarakis AS, Munger JW, Moorcroft PR. 2014. Imaging spectroscopy- and lidar-derived estimates of canopy composition and structure to improve predictions of forest carbon fluxes and ecosystem dynamics. *Geophysical Research Letters* **41**: 2535–2542.

Aristotle. 1963. *Categories and De Interpretatione. Translated with notes* (JD Ackrill, Ed.). Clarendon Press.

Asner GP, Jones MO, Martin RE, Knapp DE, Hughes RF. 2008. Remote sensing of native and invasive species in Hawaiian forests. *Remote Sensing of Environment* **112**: 1912–1926.

Asner GP, Mascaro J. 2014. Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. *Remote Sensing of Environment* **140**: 614–624.

Baccini A, Walker W, Carvalho L, Farina M, Sulla-Menashe D, Houghton RA. 2017. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science* **358**: 230–234.

Barlow J, Lennox GD, Ferreira J, Berenguer E, Lees AC, Nally R Mac, Thomson JR, Ferraz SFDB, Louzada J, Oliveira VHF, *et al.* 2016. Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. *Nature* **535**: 144–147.

Bauwens S, Bartholomeus H, Calders K, Lejeune P. 2016. Forest inventory with

terrestrial LiDAR: A comparison of static and hand-held mobile laser scanning. *Forests* **7**.

Begon M, Townsend CR, Harper JL. 2005. *Ecology: From Individuals to Ecosystems, 4th Edition*. Wiley-Blackwell.

Bell G, Gonzalez A. 2009. Evolutionary rescue can prevent extinction following environmental change. *Ecology Letters* **12**: 942–948.

Bellard C, Bertelsmeier C, Leadley P, Thuiller W, Courchamp F. 2012. Impacts of climate change on the future of biodiversity. *Ecology letters* **15**: 365–377.

Beven K, Freer J. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* **249**: 11–29.

Bloom AA, Palmer PI, Fraser A, David SR, Frankenberg C. 2010. Large-scale controls of methanogenesis inferred from methane and gravity spaceborne data. *Science* **327**: 322–325.

Bohn FJ, Huth A. 2017. The importance of forest structure to biodiversity-productivity relationships. *Royal Society Open Science* **4**: 160521.

Bolnick DI, Amarasekare P, Araújo MS, Bürger R, Levine JM, Novak M, Rudolf VHW, Schreiber SJ, Urban MC, Vasseur DA. 2011. Why intraspecific trait variation matters in community ecology. *Trends in Ecology and Evolution* **26**: 183–192.

Bolnick DI, Svanbäck R, Fordyce JA, Yang LH, Davis JM, Hulsey CD, Forister ML. 2003. The ecology of individuals: Incidence and implications of individual specialization. *American Naturalist* **161**: 1–28.

Bonal D, Burban B, Stahl C, Wagner F, Hérault B. 2016. The response of tropical rainforests to drought—lessons from recent research and future prospects. *Annals of Forest Science* **73**: 27–44.

Bonan GB. **2008**. Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science* **320**: 1444–1449.

Botkin DB, Janak JF, Wallis JR. **1972**. Some ecological consequences of a computer model of forest growth. *The Journal of Ecology* **60**: 849–872.

Box GEP. **1979**. Robustness in the Strategy of Scientific Model Building. In: Robustness in Statistics. 201–236.

Brando PM, Nepstad DC, Davidson EA, Trumbore SE, Ray D, Camargo P. **2008**. Drought effects on litterfall, wood production and belowground carbon cycling in an Amazon forest: Results of a throughfall reduction experiment. In: Philosophical Transactions of the Royal Society B: Biological Sciences. 1839–1848.

Brandon RN, Burian R. **1986**. *Genes, Organisms, Populations: Controversies Over the Units of Selection*. Bradford.

Bratman GN, Hamilton JP, Daily GC. **2012**. The impacts of nature experience on human cognitive function and mental health. *Annals of the New York Academy of Sciences* **1249**: 118–136.

Brienen RJW, Phillips OL, Feldpausch TR, Gloor E, Baker TR, Lloyd J, Lopez-Gonzalez G, Monteagudo-Mendoza A, Malhi Y, Lewis SL, *et al.* 2015. Long-term decline of the Amazon carbon sink. *Nature* **519**: 344–348.

Brooker RW, Maestre FT, Callaway RM, Lortie CL, Cavieres LA, Kunstler G, Liancourt P, Tielbörger K, Travis JMJ, Anthelme F, *et al.* 2008. Facilitation in plant communities: The past, the present, and the future. *Journal of Ecology* **96**: 18–34.

De Bruyn F. **2001**. The Classical Silva and the Generic Development of Scientific Writing in Seventeenth-Century England. *New Literary History* **32**: 347–373.

Bugmann H. **2001**. A review of forest gap models. *Climatic Change* **51**: 259–305.

Calders K, Origo N, Burt A, Disney M, Nightingale J, Raumonen P, Åkerblom M,

Malhi Y, Lewis P. 2018. Realistic forest stand reconstruction from terrestrial LiDAR for radiative transfer modelling. *Remote Sensing* **10**: 933.

Chave J. 1999. Study of structural, successional and spatial patterns in tropical rain forests using TROLL, a spatially explicit forest model. *Ecological Modelling* **124**: 233–254.

Chave J, Coomes D, Jansen S, Lewis SL, Swenson NG, Zanne AE. 2009. Towards a worldwide wood economics spectrum. *Ecology Letters* **12**: 351–366.

Chave J, Davies SJ, Phillips OL, Lewis SL, Sist P, Schepaschenko D, Armston J, Baker TR, Coomes D, Disney M, *et al.* 2019. Ground Data are Essential for Biomass Remote Sensing Missions. *Surveys in Geophysics* **40**: 863–880.

Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WBC, Duque A, Eid T, Fearnside PM, Goodman RC, *et al.* 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology* **20**: 3177–3190.

Chen Y, Wright SJ, Muller-Landau HC, Hubbell SP, Wang Y, Yu S. 2016. Positive effects of neighborhood complementarity on tree growth in a Neotropical forest. *Ecology* **97**: 776–785.

Chesson PL. 1986. Environmental variation and the coexistence of species. *Community Ecology*: 240–256.

Clark JS, Bell D, Chu C, Courbaud B, Dietze M, Hersh M, Hillerislambers J, Ibáñez I, Ladeau S, McMahon S, *et al.* 2010. High-dimensional coexistence based on individual variation: A synthesis of evidence. *Ecological Monographs* **80**: 569–608.

Cleland EE, Chuine I, Menzel A, Mooney HA, Schwartz MD. 2007. Shifting plant phenology in response to global change. *Trends in Ecology and Evolution* **22**: 357–365.

Coelho MTP, Diniz-Filho JA, Rangel TF. 2019. A parsimonious view of the parsimony

principle in ecology and evolution. *Ecography* **42**: 968–976.

Colman R, McAvaney B. **2009**. Climate feedbacks under a very broad range of forcing. *Geophysical Research Letters* **36**: L01702.

Coomes DA, Asner GP, Lewis SL, Dalponte M, Phillips OL, Qie L, Nilus R, Phua M-H, Banin LF, Burslem DFRP, *et al.* 2017. Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asian forests from airborne laser scanning data. *Remote Sensing of Environment* **194**: 77–88.

Coomes DA, Lines ER, Allen RB. **2011**. Moving on from Metabolic Scaling Theory: Hierarchical models of tree growth and asymmetric competition for light. *Journal of Ecology* **99**: 748–756.

Cox TR. **1985**. Americans and Their Forests: Romanticism, Progress, and Science in the Late Nineteenth Century. *Forest & Conservation History* **29**: 156–168.

Curry JA, Schramm JL, Ebert EE. **1995**. Sea ice-albedo climate feedback mechanism. *Journal of Climate* **8**: 240–247.

Davies P. **2019**. *The Demon in the Machine*. Allen Lane, an imprint of Penguin Books.

DeAngelis DL, Grimm V. **2014**. Individual-based models in ecology after four decades. *F1000Prime Reports* **6**: 39.

DeAngelis DL, Mooij WM. **2005**. Individual-based modeling of ecological and evolutionary processes. *Annual Review of Ecology, Evolution, and Systematics* **36**: 147–168.

Dewey J. **1903**. *Studies in Logical Theory*. American Mathematical Society.

Díaz S, Kattge J, Cornelissen JHC, Wright IJ, Lavorel S, Dray S, Reu B, Kleyer M, Wirth C, Colin Prentice I, *et al.* 2016. The global spectrum of plant form and function. *Nature* **529**: 167–171.

Disney MI, Boni Vicari M, Burt A, Calders K, Lewis SL, Raumonen P, Wilkes P. **2018**.

Weighing trees with lasers: Advances, challenges and opportunities. *Interface Focus* **8**: 20170048.

Dusenge ME, Duarte AG, Way DA. **2019**. Plant carbon metabolism and climate change: elevated CO2 and temperature impacts on photosynthesis, photorespiration and respiration. *New Phytologist* **221**: 32–49.

Earp BD, Trafimow D. **2015**. Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology* **6**: 621.

Einstein A. **1934**. On the Method of Theoretical Physics. *Philosophy of Science* **1**: 163–169.

Ellison D, Morris CE, Locatelli B, Sheil D, Cohen J, Murdiyarso D, Gutierrez V, Noordwijk M van, Creed IF, Pokorny J, *et al.* 2017. Trees, forests and water: Cool insights for a hot world. *Global Environmental Change* **43**: 51–61.

Eltahir EAB, Bras RL. **1994**. Precipitation recycling in the Amazon basin. *Quarterly Journal of the Royal Meteorological Society* **120**: 861–880.

Enquist BJ, Allen AP, Brown JH, Gillooly JF, Kerkhoff AJ, Niklas KJ, Price CA, West GB. **2007**. Does the exception prove the rule? *Nature* **445**: E9–E10.

Enquist BJ, Economo EP, Huxman TE, Allen AP, Ignace DD, Gillooly JF. **2003**. Scaling metabolism from organisms to ecosystems. *Nature* **423**: 639–642.

Fatichi S, Leuzinger S, Körner C. **2014**. Moving beyond photosynthesis: from carbon source to sink-driven vegetation modeling. *New Phytologist* **201**: 1086–1095.

Ferraz A, Saatchi S, Mallet C, Meyer V. **2016**. Lidar detection of individual tree size in tropical forests. *Remote Sensing of Environment* **183**: 318–333.

Fischer R, Bohn F, Dantas de Paula M, Dislich C, Groeneveld J, Gutiérrez AG, Kazmierczak M, Knapp N, Lehmann S, Paulick S, *et al.* 2016. Lessons learned from applying a forest gap model to understand ecosystem and carbon dynamics of complex

tropical forests. *Ecological Modelling* **326**: 124–133.

Fischer FJ, Maréchaux I, Chave J. **2019**. Improving plant allometry by fusing forest models and remote sensing. *New Phytologist* **223**: 1159–1165.

Fisher RA, Koven CD, Anderegg WRL, Christoffersen BO, Dietze MC, Farrior CE, Holm JA, Hurtt GC, Knox RG, Lawrence PJ, *et al.* 2018. Vegetation demographics in Earth System Models: A review of progress and priorities. *Global Change Biology* **24**: 35–54.

Foley JA, Prentice IC, Ramankutty N, Levis S, Pollard D, Sitch S, Haxeltine A. **1996**. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. *Global Biogeochemical Cycles* **10**: 603–628.

Frankenberg C, Fisher JB, Worden J, Badgley G, Saatchi SS, Lee JE, Toon GC, Butz A, Jung M, Kuze A, *et al.* 2011. New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters* **38**: L17706.

Friend AD, Schugart HH, Running SW. **1993**. A Physiology-Based Gap Model of Forest Dynamics. *Ecology* **74**: 792–797.

Friend AD, Stevens AK, Knox RG, Cannell MGR. **1997**. A process-based, terrestrial biosphere model of ecosystem dynamics (Hybrid v3.0). *Ecological Modelling* **95**: 249–287.

Fyllas NM, Gloor E, Mercado LM, Sitch S, Quesada CA, Domingues TF, Galbraith DR, Torre-Lezama A, Vilanova E, Ramírez-Angulo H, *et al.* 2014. Analysing Amazonian forest productivity using a new individual and trait-based model (TFS v.1). *Geoscientific Model Development* **7**: 1251–1269.

Garcia O. **1983**. A Stochastic Differential Equation Model for the Height Growth of Forest Stands. *Biometrics* **39**: 1059.

Gaubert B, Stephens BB, Basu S, Chevallier F, Deng F, Kort EA, Patra PK, Peters W, Rödenbeck C, Saeki T, *et al.* 2019. Global atmospheric CO2 inverse models converging on neutral tropical land exchange, but disagreeing on fossil fuel and atmospheric growth rate. *Biogeosciences* **16**: 117–134.

George AR. **2003**. *The Babylonian Gilgamesh Epic. Introduction, Critical Edition and Cuneiform Texts*. Oxford University Press.

Gillespie TW, Foody GM, Rocchini D, Giorgi AP, Saatchi S. **2008**. Measuring and modelling biodiversity from space. *Progress in Physical Geography* **32**: 203–221.

Godfray HCJ, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C. **2010**. Food security: The challenge of feeding 9 billion people. *Science* **327**: 812–818.

Goetz SJ, Baccini A, Laporte NT, Johns T, Walker W, Kellndorfer J, Houghton RA, Sun M. **2009**. Mapping and monitoring carbon stocks with satellite observations: A comparison of methods. *Carbon Balance and Management* **4**: 2.

Goldblatt C, Watson AJ. **2012**. The runaway greenhouse: Implications for future climate change, geoengineering and planetary atmospheres. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **370**: 4197–4216.

Good P, Jones C, Lowe J, Betts R, Gedney N. **2013**. Comparing tropical forest projections from two generations of hadley centre earth system models, HadGEM2-ES and HadCM3LC. *Journal of Climate* **26**: 495–511.

Grant PR. **1996**. Speciation and hybridization in island birds. *Philosophical Transactions of the Royal Society B: Biological Sciences* **351**: 765–772.

Grime JP. **1997**. Biodiversity and ecosystem function: The debate deepens. *Science* **277**: 1260–1261.

Grimm V, Ayllón D, Railsback SF. **2017**. Next-Generation Individual-Based Models

Integrate Biodiversity and Ecosystems: Yes We Can, and Yes We Must. *Ecosystems* **20**: 229–236.

Grimm V, Railsback SF. **2012**. Pattern-oriented modelling: a 'multi-scope' for predictive systems ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**: 298–310.

Grimm V, Revilla E, Berger U, Jeltsch F, Mooij WM, Railsback SF, Thulke HH, Weiner J, Wiegand T, DeAngelis DL. **2005**. Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science* **310**: 987–991.

Hérault B, Bachelot B, Poorter L, Rossi V, Bongers F, Chave J, Paine CET, Wagner F, Baraloto C. **2011**. Functional traits shape ontogenetic growth trajectories of rain forest tree species. *Journal of Ecology* **99**: 1431–1440.

Hirota M, Holmgren M, Van Nes EH, Scheffer M. **2011**. Global resilience of tropical forest and savanna to critical transitions. *Science* **334**: 232–235.

Höfle B, Rutzinger M. **2011**. Topographic airborne LiDAR in geomorphology: A technological perspective. *Zeitschrift fur Geomorphologie* **55**: 1–29.

Homolová L, Malenovský Z, Clevers JGPW, García-Santos G, Schaepman ME. **2013**. Review of optical-based remote sensing for plant trait mapping. *Ecological Complexity* **15**: 1–16.

Houghton RA, Nassikas AA. **2017**. Global and regional fluxes of carbon from land use and land cover change 1850–2015. *Global Biogeochemical Cycles* **31**: 456–472.

Houlahan JE, McKinney ST, Anderson TM, McGill BJ. **2017**. The priority of prediction in ecological understanding. *Oikos* **126**: 1–7.

von Humboldt A, Bonpland A. **1814**. *Personal Narrative of Travels to the Equinoctial Regions of the New Continent, during the years 1799-1804. Vol. 1* (HM Williams, Ed.).

IPCC. **2014**. *IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working*

Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Jackson RB, Randerson JT, Canadell JG, Anderson RG, Avissar R, Baldocchi DD, Bonan GB, Caldeira K, Diffenbaugh NS, Field CB, *et al.* 2008. Protecting climate with forests. *Environmental Research Letters* **3**: 044006.

Jackson T, Shenkin A, Wellpott A, Calders K, Origo N, Disney M, Burt A, Raumonen P, Gardiner B, Herold M, *et al.* 2019. Finite element analysis of trees in the wind based on terrestrial laser scanning data. *Agricultural and Forest Meteorology* **265**: 137–144.

Jarvis PG. **1995**. Scaling processes and problems. *Plant, Cell & Environment* **18**: 1079–1089.

Jucker T, Bouriaud O, Coomes DA. **2015**. Crown plasticity enables trees to optimize canopy packing in mixed-species forests. *Functional Ecology* **29**: 1078–1086.

Jucker T, Caspersen J, Chave J, Antin C, Barbier N, Bongers F, Dalponte M, van Ewijk KY, Forrester DI, Haeni M, *et al.* 2017. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Global Change Biology* **23**: 177–190.

Karafyllidis I, Thanailakis A. **1997**. A model for predicting forest fire spreading using cellular automata. *Ecological Modelling* **99**: 87–97.

Kattge J, Díaz S, Lavorel S, Prentice IC, Leadley P, Bönisch G, Garnier E, Westoby M, Reich PB, Wright IJ, *et al.* 2011. TRY - a global database of plant traits. *Global Change Biology* **17**: 2905–2935.

Keith DW. **2009**. Why capture CO2 from the atmosphere? *Science* **325**: 1654–1655.

Kelley CP, Mohtadi S, Cane MA, Seager R, Kushnir Y. **2015**. Climate change in the Fertile Crescent and implications of the recent Syrian drought. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 3241–3246.

Knapp N, Fischer R, Huth A. 2018. Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states. *Remote Sensing of Environment* **205**: 199–209.

Kraft NJB, Godoy O, Levine JM. 2015. Plant functional traits and the multidimensional nature of species coexistence. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 797–802.

Krinner G, Viovy N, de Noblet-Ducoudré N, Ogée J, Polcher J, Friedlingstein P, Ciais P, Sitch S, Prentice IC. 2005. A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*.

Kunstler G, Falster D, Coomes DA, Hui F, Kooyman RM, Laughlin DC, Poorter L, Vanderwel M, Vieilledent G, Wright SJ, *et al.* 2016. Plant functional traits have globally consistent effects on competition. *Nature* **529**: 204–207.

Lavorel S, Garnier E. 2002. Predicting changes in community composition and ecosystem functioning from plant traits: Revisiting the Holy Grail. *Functional Ecology* **16**: 545–556.

Lefsky MA, Cohen WB, Parker GG, Harding DJ. 2002. Lidar Remote Sensing for Ecosystem Studies. *BioScience* **52**: 19.

Lenton TM. 2011. Early warning of climate tipping points. *Nature Climate Change* **1**: 201–209.

Lenton TM, Held H, Kriegler E, Hall JW, Lucht W, Rahmstorf S, Schellnhuber HJ. 2008. Tipping elements in the Earth's climate system. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 1786–1793.

Levin SA. 1998. Ecosystems and the biosphere as complex adaptive systems. *Ecosystems* **1**: 431–436.

Levine NM, Zhang K, Longo M, Baccini A, Phillips OL, Lewis SL, Alvarez-Dávila E, De

Andrade ACS, Brienen RJW, Erwin TL, *et al.* 2016. Ecosystem heterogeneity determines the ecological resilience of the Amazon to climate change. *Proceedings of the National Academy of Sciences of the United States of America* **113**: 793–797.

Levins R. **1966**. The strategy of model building in population biology. *American Naturalist* **54**: 421–431.

Lewis SL, Edwards DP, Galbraith D. **2015**. Increasing human dominance of tropical forests. *Science* **349**: 827–832.

Lewis SL, Lopez-Gonzalez G, Sonké B, Affum-Baffoe K, Baker TR, Ojo LO, Phillips OL, Reitsma JM, White L, Comiskey JA, *et al.* 2009. Increasing carbon storage in intact African tropical forests. *Nature* **457**: 1003–1006.

Lobo Sternberg LDS. **2001**. Savanna-forest hysteresis in the tropics. *Global Ecology and Biogeography* **10**: 369–378.

Longuetaud F, Piboule A, Wernsdörfer H, Collet C. **2013**. Crown plasticity reduces inter-tree competition in a mixed broadleaved forest. *European Journal of Forest Research* **132**: 621–634.

Malhi Y. **2012**. The productivity, metabolism and carbon cycle of tropical forest vegetation. *Journal of Ecology* **100**: 65–75.

Malhi Y, Aragão LEOC, Galbraith D, Huntingford C, Fisher R, Zelazowski P, Sitch S, McSweeney C, Meir P. **2009**. Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 20610–20615.

Maréchaux I, Chave J. **2017**. An individual-based forest model to jointly simulate carbon and tree diversity in Amazonia: description and applications. *Ecological Monographs* **87**: 632–664.

McMurtrie R, Wolf L. **1983**. A Model of Competition between Trees and Grass for

Radiation, Water and Nutrients. *Annals of Botany* **52**: 449–458.

Medvigy D, Wofsy SC, Munger JW, Hollinger DY, Moorcroft PR. 2009. Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2. *Journal of Geophysical Research: Biogeosciences* **114**: G01002.

Meehl GA, Boer GJ, Covey C, Latif M, Stouffer RJ. 2002. The Coupled Model Intercomparison Project (CMIP). *Bulletin of the American Meteorological Society* **81**: 313–318.

Millar CI, Stephenson NL, Stephens SL. 2007. Climate change and forests of the future: Managing in the face of uncertainty. *Ecological Applications* **17**: 2145–2151.

Monteith JL. 1981. Evaporation and surface temperature. *Quarterly Journal of the Royal Meteorological Society* **107**: 1–27.

Moreira M, Sternberg L, Martinelli L, Victoria R, Barbosa E, Bonates L, Nepstad D. 1997. Contribution of transpiration to forest ambient vapour based on isotopic measurements. *Global Change Biology* **3**: 439–450.

Morsdorf F, Meier E, Kötz B, Itten KI, Dobbertin M, Allgöwer B. 2004. LIDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. In: Remote Sensing of Environment. 353–362.

Mouquet N, Lagadeuc Y, Devictor V, Doyen L, Duputié A, Eveillard D, Faure D, Garnier E, Gimenez O, Huneman P, *et al.* 2015. Predictive ecology in a changing world. *Journal of Applied Ecology* **52**: 1293–1310.

Niklas KJ. 1994. Morphological evolution through complex domains of fitness. *Proceedings of the National Academy of Sciences* **91**: 6772–6779.

Norby RJ, Zak DR. 2011. Ecological Lessons from Free-Air CO 2 Enrichment (FACE) Experiments . *Annual Review of Ecology, Evolution, and Systematics* **42**: 181–203.

Odonne G, Bel M, Burst M, Brunaux O, Bruno M, Dambrine E, Davy D, Desprez M,

Engel J, Ferry B, *et al.* 2019. Long-term influence of early human occupations on current forests of the Guiana Shield. *Ecology* **100**: e02806.

Ovid. **2009**. *Metamorphoses* (AD Melville and EJ Kenney, Eds.). Oxford University Press.

Pacala SW, Canham CD, Saponara J, Silander JA, Kobe RK, Ribbens E. **1996**. Forest models defined by field measurements: estimation, error analysis and dynamics. *Ecological Monographs* **66**: 1–43.

Paine CET, Amissah L, Auge H, Baraloto C, Baruffol M, Bourland N, Bruelheide H, Daïnou K, de Gouvenain RC, Doucet JL, *et al.* 2015. Globally, functional traits are weak predictors of juvenile tree growth, and we do not know why. *Journal of Ecology* **103**: 978–989.

Pan Y, Birdsey RA, Fang J, Houghton R, Kauppi PE, Kurz WA, Phillips OL, Shvidenko A, Lewis SL, Canadell JG, *et al.* 2011. A Large and Persistent Carbon Sink in the World's Forests. *Science* **333**: 988 LP – 993.

Pavlick R, Drewry DT, Bohn K, Reu B, Kleidon A. **2013**. The Jena Diversity-Dynamic Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs. *Biogeosciences* **10**: 4137–4177.

Peirce CS. **1878**. How to Make Our Ideas Clear. *Popular Science Monthly* **12**: 286–302.

Peng C. **2000**. Growth and yield models for uneven-aged stands: Past, present and future. *Forest Ecology and Management* **132**: 259–279.

Pensel D. **2019**. '...tiefer in den Wald hinein'. Zum Raum des Unbewussten in der Literatur des 19. Jahrhunderts. In: Frenzel M, Geist K, Oberrauch C, eds. Ein Ort, viel Raum(theorie)?: Imaginationen gleicher Räume und Orte in Literatur und Film. 163–198.

Pettorelli N, Ryan S, Mueller T, Bunnefeld N, Jedrzejewska B, Lima M, Kausrud K.

2011. The Normalized Difference Vegetation Index (NDVI): Unforeseen successes in animal ecology. *Climate Research* **46**: 15–27.

Phillips OL, Aragão LEOC, Lewis SL, Fisher JB, Lloyd J, López-González G, Malhi Y, Monteagudo A, Peacock J, Quesada CA, *et al.* 2009. Drought sensitivity of the amazon rainforest. *Science* **323**: 1344–1347.

Phillips OL, Lewis SL, Baker TR, Chao KJ, Higuchi N. 2008. The changing Amazon forest. In: Philosophical Transactions of the Royal Society B: Biological Sciences. 1819–1827.

Poorter L, Bongers F, Aide TM, Almeyda Zambrano AM, Balvanera P, Becknell JM, Boukili V, Brancalion PHS, Broadbent EN, Chazdon RL, *et al.* 2016. Biomass resilience of Neotropical secondary forests. *Nature* **530**: 211–214.

Poorter L, Castilho C V., Schietti J, Oliveira RS, Costa FRC. 2018. Can traits predict individual growth performance? A test in a hyperdiverse tropical forest. *New Phytologist* **219**: 109–121.

Poorter L, Wright SJ, Paz H, Ackerly DD, Condit R, Ibarra-Manríquez G, Harms KE, Licona JC, Martínez-Ramos M, Mazer SJ, *et al.* 2008. Are functional traits good predictors of demographic rates? Evidence from five neotropical forests. *Ecology* **89**: 1908–1920.

Popper K. 1963. *Conjectures and Refutations*. Routledge & Kegan Paul.

Popper KR. 1972. *Objective Knowledge*. Oxford, Clarendon Press.

Popper K. 1990. *A World of Propensities*. Thoemmes.

Porté A, Bartelink HH. 2002. Modelling mixed forest growth: A review of models for forest management. *Ecological Modelling* **150**: 141–188.

Praetorius B, Schumacher K. 2009. Greenhouse gas mitigation in a carbon constrained world: The role of carbon capture and storage. *Energy Policy* **37**: 5081–

5093.

Prentice IC, Bondeau A, Cramer W, Harrison SP, Hickler T, Lucht W, Sitch S, Smith B, Sykes MT. **2007**. Dynamic Global Vegetation Modeling: Quantifying Terrestrial Ecosystem Responses to Large-Scale Environmental Change. In: Canadell JG, Pataki DE, Pitelka LF, eds. Terrestrial Ecosystems in a Changing World. Berlin, Heidelberg: Springer Berlin Heidelberg, 175–192.

Pretzsch H. **2009**. *Forest Dynamics, Growth, and Yield*. Springer.

Pretzsch H. **2014**. Canopy space filling and tree crown morphology in mixed-species stands compared with monocultures. *Forest Ecology and Management* **327**: 251–264.

Price CA, Weitz JS, Savage VM, Stegen J, Clarke A, Coomes DA, Dodds PS, Etienne RS, Kerkhoff AJ, Mcculloh K, *et al.* 2012. Testing the metabolic theory of ecology. *Ecology Letters* **15**: 1465–1474.

Purves DW, Lichstein JW, Strigul N, Pacala SW. **2008**. Predicting and understanding forest dynamics using a simple tractable model. *Proceedings of the National Academy of Sciences* **105**: 17018–17022.

Purves D, Pacala S. **2008**. Predictive models of forest dynamics. *Science* **320**: 1452–1453.

Qi W, Dubayah RO. **2016**. Combining Tandem-X InSAR and simulated GEDI lidar observations for forest structure mapping. *Remote Sensing of Environment* **187**: 253–266.

Quéré C, Andrew R, Friedlingstein P, Sitch S, Hauck J, Pongratz J, Pickers P, Ivar Korsbakken J, Peters G, Canadell J, *et al.* 2018. Global Carbon Budget 2018. *Earth System Science Data* **10**: 2141–2194.

Le Quéré C, Raupach MR, Canadell JG, Marland G, Bopp L, Ciais P, Conway TJ, Doney SC, Feely RA, Foster P, *et al.* 2009. Trends in the sources and sinks of carbon

dioxide. *Nature Geoscience* **2**: 831–836.

Quinn JF, Dunham AE. **1983**. On hypothesis testing in ecology and evolution. *American Naturalist* **122**: 602–617.

Railsback SF. **2001**. Concepts from complex adaptive systems as a framework for individual-based modelling. *Ecological Modelling* **139**: 47–62.

Rammig A, Jupp T, Thonicke K, Tietjen B, Heinke J, Ostberg S, Lucht W, Cramer W, Cox P. **2010**. Estimating the risk of Amazonian forest dieback. *New Phytologist* **187**: 694–706.

Reich PB. **2014**. The world-wide 'fast-slow' plant economics spectrum: A traits manifesto. *Journal of Ecology* **102**: 275–301.

Reich PB, Tjoelker MG, Machado JL, Oleksyn J. **2006**. Universal scaling of respiratory metabolism, size and nitrogen in plants. *Nature* **439**: 457–461.

Reich PB, Walters MB, Ellsworth DS. **1997**. From tropics to tundra: Global convergence in plant functioning. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 13730–13734.

Reich PB, Wright IJ, Cavender-Bares J, Craine JM, Oleksyn J, Westoby M, Walters MB. **2003**. The evolution of plant functional variation: Traits, spectra, and strategies. *International Journal of Plant Sciences* **164**: S143–S164.

Reuveny R. **2007**. Climate change-induced migration and violent conflict. *Political Geography* **26**: 656–673.

Richardson AD, Braswell BH, Hollinger DY, Jenkins JP, Ollinger S V. 2009. Near-surface remote sensing of spatial and temporal variation in canopy phenology. *Ecological Applications* **19**: 1417–1428.

Roberts N, Fyfe RM, Woodbridge J, Gaillard MJ, Davis BAS, Kaplan JO, Marquer L, Mazier F, Nielsen AB, Sugita S, *et al.* 2018. Europe's lost forests: A pollen-based

synthesis for the last 11,000 years. *Scientific Reports* **8**: 716.

Rockström J, Steffen W, Noone K, Persson A, Chapin FS, Lambin EF, Lenton TM, Scheffer M, Folke C, Schellnhuber HJ, *et al.* 2009. A safe operating space for humanity. *Nature* **461**: 472–5.

Rödig E, Cuntz M, Heinke J, Rammig A, Huth A. 2017. Spatial heterogeneity of biomass and forest structure of the Amazon rain forest: Linking remote sensing, forest modelling and field inventory. *Global Ecology and Biogeography* **26**: 1292–1302.

Rosette JAB, North PRJ, Suárez JC. 2008. Vegetation height estimates for a mixed temperate forest using satellite laser altimetry. In: International Journal of Remote Sensing. 1475–1493.

Rouse JW, Hass RH, Schell JA, Deering DW. 1973. Monitoring vegetation systems in the great plains with ERTS. In: Third Earth Resources Technology Satellite (ERTS) symposium. 309–317.

Running SW, Coughlan JC. 1988. A general model of forest ecosystem processes for regional applications I. Hydrologic balance, canopy gas exchange and primary production processes. *Ecological Modelling* **42**: 125–154.

Running SW, Gower ST. 1991. FOREST-BGC, A general model of forest ecosystem processes for regional applications. II. Dynamic carbon allocation and nitrogen budgets. *Tree Physiology* **9**: 147–160.

Rusu DC, Lüthy C. 2017. Extracts from a paper laboratory: the nature of Francis Bacon's Sylva sylvarum. *Intellectual History Review* **27**: 171–202.

Saatchi SS, Harris NL, Brown S, Lefsky M, Mitchard ETA, Salas W, Zutta BR, Buermann W, Lewis SL, Hagen S, *et al.* 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 9899–9904.

Sakschewski B, von Bloh W, Boit A, Rammig A, Kattge J, Poorter L, Peñuelas J, Thonicke K. 2015. Leaf and stem economics spectra drive diversity of functional plant traits in a dynamic global vegetation model. *Global Change Biology* **21**: 2711–2725.

Salati E, Dall'Olio A, Matsui E, Gat JR. 1979. Recycling of water in the Amazon Basin: An isotopic study. *Water Resources Research* **15**: 1250–1258.

Salguero-Gómez R, Violle C, Gimenez O, Childs D. 2018. Delivering the promises of trait-based approaches to the needs of demographic approaches, and vice versa. *Functional Ecology* **32**: 1424–1435.

Sato H, Itoh A, Kohyama T. 2007. SEIB-DGVM: A new Dynamic Global Vegetation Model using a spatially explicit individual-based approach. *Ecological Modelling* **200**: 279–307.

Scheffers BR, De Meester L, Bridge TCL, Hoffmann AA, Pandolfi JM, Corlett RT, Butchart SHM, Pearce-Kelly P, Kovacs KM, Dudgeon D, *et al.* 2016. The broad footprint of climate change from genes to biomes to people. *Science* **354**: aaf7671.

Scheiter S, Langan L, Higgins SI. 2013. Next-generation dynamic global vegetation models: Learning from community ecology. *New Phytologist* **198**: 957–969.

Schellnhuber HJ. 2009. Tipping elements in the Earth System. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 20561–20563.

Schimel D, Schneider FD. 2019. Flux towers in the sky: global ecology from space. *New Phytologist* **224**: 570–584.

Shipley B, De Bello F, Cornelissen JHC, Laliberté E, Laughlin DC, Reich PB. 2016. Reinforcing loose foundation stones in trait-based plant ecology. *Oecologia* **180**: 923–931.

Shugart, HH, West DC. 1980. Forest Succession Models. *BioScience* **30**: 308–313.

Shugart HH, Asner GP, Fischer R, Huth A, Knapp N, Le Toan T, Shuman JK. 2015.

Computer and remote-sensing infrastructure to enhance large-scale testing of individual-based forest models. *Frontiers in Ecology and the Environment* **13**: 503–511.

Shugart HH, Wang B, Fischer R, Ma J, Fang J, Yan X, Huth A, Armstrong AH. 2018. Gap models and their individual-based relatives in the assessment of the consequences of global change. *Environmental Research Letters* **13**: 033001.

Simard M, Pinto N, Fisher JB, Baccini A. 2011. Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences* **116**: G04021.

Sitch S, Huntingford C, Gedney N, Levy PE, Lomas M, Piao SL, Betts R, Ciais P, Cox P, Friedlingstein P, *et al.* 2008. Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). *Global Change Biology* **14**: 2015–2039.

Sitch S, Smith B, Prentice IC, Arneth A, Bondeau A, Cramer W, Kaplan JO, Levis S, Lucht W, Sykes MT, *et al.* 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology* **9**: 161–185.

Smith MS, Horrocks L, Harvey A, Hamilton C. 2011. Rethinking adaptation for a 4°C world. In: Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 196–216.

Smith B, Wärlind D, Arneth A, Hickler T, Leadley P, Siltberg J, Zaehle S. 2014. Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences* **11**: 2027–2054.

Ter Steege H, Pitman NCA, Sabatier D, Baraloto C, Salomão RP, Guevara JE, Phillips OL, Castilho C V., Magnusson WE, Molino JF, *et al.* 2013. Hyperdominance in the Amazonian tree flora. *Science* **342**: 1243092.

Strigul N, Pristinski D, Purves D, Dushoff J, Pacala S. 2008. Scaling from trees to

forests: Tractable macroscopic equations for forest dynamics. *Ecological Monographs* **78**: 523–545.

Sturlson S. **2005**. *The Prose Edda* (J Byock, Ed.). Penguin Classics.

Sutherland WJ, Freckleton RP, Godfray HCJ, Beissinger SR, Benton T, Cameron DD, Carmel Y, Coomes DA, Coulson T, Emmerson MC, *et al.* 2013. Identification of 100 fundamental ecological questions. *Journal of Ecology* **101**: 58–67.

Taubert F, Jahn MW, Dobner H-J, Wiegand T, Huth A. **2015**. The structure of tropical forests and sphere packings. *Proceedings of the National Academy of Sciences* **112**: 15125–15129.

Thompson JR, Carpenter DN, Cogbill C V., Foster DR. **2013**. Four Centuries of Change in Northeastern United States Forests. *PLoS ONE* **8**: e72540.

Le Toan T, Beaudoin A, Riom J, Guyon D. **1992**. Relating Forest Biomass to SAR Data. *IEEE Transactions on Geoscience and Remote Sensing* **30**: 403–411.

Le Toan T, Quegan S, Davidson MWJ, Balzter H, Paillou P, Papathanassiou K, Plummer S, Rocca F, Saatchi S, Shugart H, *et al.* 2011. The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. *Remote Sensing of Environment* **115**: 2850–2860.

Tucker CJ. **1979**. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* **8**: 127–150.

Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M. **2003**. Remote sensing for biodiversity science and conservation. *Trends in Ecology and Evolution* **18**: 306–314.

Vanclay JK. **1994**. Modelling forest growth and yield: applications to mixed tropical forests. *Modelling forest growth and yield: applications to mixed tropical forests*: 537.

Vega C, Hamrouni A, El Mokhtari A, Morel M, Bock J, Renaud JP, Bouvier M,

Durrieue S. 2014. PTrees: A point-based approach to forest tree extractionfrom lidar data. *International Journal of Applied Earth Observation and Geoinformation* **33**: 98–108.

Vertessy RA, Benyon RG, O'Sullivan SK, Gribben PR. 1995. Relationships between stem diameter, sapwood area, leaf area and transpiration in a young mountain ash forest. *Tree Physiology* **15**: 559–567.

Violle C, Thuiller W, Mouquet N, Munoz F, Kraft NJB, Cadotte MW, Livingstone SW, Mouillot D. 2017. Functional Rarity: The Ecology of Outliers. *Trends in Ecology and Evolution* **32**: 356–367.

Vira B, Wildburger C, Mansourian S. 2015. *Forests , Trees and Landscapes for Food Security and Nutrition. A Global Assessment Report.*

Visser MD, Bruijning M, Wright SJ, Muller-Landau HC, Jongejans E, Comita LS, de Kroon H. 2016. Functional traits as predictors of vital rates across the life cycle of tropical trees. *Functional Ecology* **30**: 168–180.

Wenger SJ, Olden JD. 2012. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods in Ecology and Evolution* **3**: 260–267.

West GB, Brown JH, Enquist BJ. 1999. A general model for the structure and allometry of plant vascular systems. *Nature* **400**: 664–667.

Williams LJ, Paquette A, Cavender-Bares J, Messier C, Reich PB. 2017. Spatial complementarity in tree crowns explains overyielding in species mixtures. *Nature Ecology and Evolution* **1**: 0063.

Wittgenstein L. 1953. *Philosophical Investigations* (GEM Anscombe, Ed.).

Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F, Cavender-Bares J, Chapin T, Cornellssen JHC, Diemer M, *et al.* 2004. The worldwide leaf economics spectrum. *Nature* **428**: 821–827.

Wright A, Schnitzer SA, Reich PB. 2014. Living close to your neighbors: The importance of both competition and facilitation in plant communities. *Ecology* **95**: 2213–2223.

Zanne AE, Lopez-Gonzalez G, Coomes DA, Ilic J, Jansen S, Lewis SL, Miller RB, Swenson NG, Wiemann MC, Chave J. 2009. Data from: Global wood density database. *Dryad Digital Repository*.

Zemp DC, Schleussner CF, Barbosa HMJ, Hirota M, Montade V, Sampaio G, Staal A, Wang-Erlandsson L, Rammig A. 2017. Self-amplified Amazon forest loss due to vegetation-atmosphere feedbacks. *Nature Communications* **8**: 14681.

Zhu Z, Piao S, Myneni RB, Huang M, Zeng Z, Canadell JG, Ciais P, Sitch S, Friedlingstein P, Arneth A, *et al.* 2016. Greening of the Earth and its drivers. *Nature Climate Change* **6**: 791–795.

Chapter 1: Improving plant allometry by fusing forest models and remote sensing

(Tansley Insight, published on 21 March, 2019 in *New Phytologist*)

This chapter is a conceptual paper on how the complexity of individual-based forest growth models and the data-richness of remote sensing inform each other to yield ecological insights. At its heart is the analysis of plant allometries, i.e. the various scaling relationships that exist between plant size and function and that have been a core component of vegetation models and research on the global carbon cycle. We review the challenges in allometric scaling, provide an example with the individual-based forest model TROLL of how they can be tackled by advances in data-model fusion, and outline how, in doing so, such models can serve as data integrators for dynamic global vegetation models.

New Phytologist

Review

Tansley insights

Improving plant allometry by fusing forest models and remote sensing

Author for correspondence:
Jérôme Chave
Tel: +33 561 556 760
Email: jerome.chave@univ-tlse3.fr

Received: *29 December 2018*
Accepted: *5 March 2019*

Fabian Jörg Fischer[1], Isabelle Maréchaux[2] and Jérôme Chave[1]

[1]Laboratoire Evolution et Diversité Biologique, UMR5174, CNRS–Université Paul Sabatier–IRD, Bâtiment 4R1, 118 route de Narbonne, F-31062 Toulouse Cedex 9, France; [2]AMAP, INRA, IRD, CIRAD, CNRS, University of Montpellier, F-34000 Montpellier, France

Contents

	Summary	1159
I.	Introduction	1159
II.	Tree allometry and transferability	1160
III.	Condensing the point cloud: allometry from space	1161
IV.	Bayesian merging of data in IBMs	1162
V.	Challenges and perspectives	1163
	Acknowledgements	1164
	References	1164

New Phytologist (2019) **223:** 1159–1165
doi: 10.1111/nph.15810

Key words: allometry, Approximate Bayesian Computation, data-model fusion, forest model, LiDAR.

Summary

Allometry determines how tree shape and function scale with each other, related through size. Allometric relationships help scale processes from the individual to the global scale and constitute a core component of vegetation models. Allometric relationships have been expected to emerge from optimisation theory, yet this does not suitably predict empirical data. Here we argue that the fusion of high-resolution data, such as those derived from airborne laser scanning, with individual-based forest modelling offers insight into how plant size contributes to large-scale biogeochemical processes. We review the challenges in allometric scaling, how they can be tackled by advances in data-model fusion, and how individual-based models can serve as data integrators for dynamic global vegetation models.

I. Introduction

Forests provide important services to societies globally, sequestering large amounts of carbon, limiting erosion, regulating the water cycle, and providing a habitat for many species. Size, shape and function relationships among plants, or allometries, play a key role in understanding these services. Such relationships encapsulate ontogenetic, ecological and evolutionary constraints (Niklas, 1994) and have been widely used in quantitative tools to aid forest management. How much carbon is stored in the world's forests, for instance, is estimated from forest inventories using allometric models and then scaled up to regional and global scales, based on Earth observation data and modelling (Pan *et al.*, 2011).

Allometries also describe how metabolic functions, such as respiration rates and net primary production, scale with each other. A theory has been developed to infer allometric scaling from evolutionary optimisation principles (Enquist & Niklas, 2002), but this theory does not account for recent advances in plant physiology (Rogers *et al.*, 2017; Scoffoni *et al.*, 2017), and its predictions do not match empirical data well (Muller-Landau *et al.*, 2006; Poorter *et al.*, 2012). Our ability to simulate the vegetation response to environmental change in Dynamic Global Vegetation Models (DGVMs) is, however, directly dependent on the robustness of these scaling relationships. Because DGVMs adopt a coarse-grained description of forests, allometries are often used to link fluxes and pools, but the results do not always correspond to empirical observations (Wolf *et al.*, 2011).

A great opportunity to bring processes and field information into a consistent modelling framework is offered by individual-based models (IBMs) of forest dynamics (DeAngelis & Grimm, 2014). In

© 2019 The Authors
New Phytologist © 2019 New Phytologist Trust

IBMs, the forest ecosystem emerges from a combination of individual tree physiological and demographic processes at a scale that is relevant for forest resource management and ecological data assimilation, as in the FORMIND model (Rödig *et al.*, 2017). This approach can be extended to larger scales, either by informing DGVMs through IBMs (for example ED2; Medvigy *et al.*, 2009; LPJ-GUESS, Smith *et al.*, 2014) or by directly scaling them up (SEIB-DGVM, Sato *et al.*, 2007; FORMIND, Fischer *et al.*, 2016). Like DGVMs, forest IBMs often rely on empirical allometric models to predict tree shape and function but, during model calibration, information can also be gained about the allometric models themselves and the processes that shape them.

Proper calibration and validation of forest IBMs should be based on a variety of independent data sources, ranging from forest inventories to eddy-flux data, as recently exemplified with the TROLL model, a physiology-based and fully spatially explicit forest IBM (Maréchaux & Chave, 2017). A promising additional data source is provided by remote sensing. With its ability to generate detailed information over unprecedented scales and at locations that are otherwise hard to access (for example upper canopy layers, remote ecosystems), remote sensing has already had a transformative effect on vegetation modelling (Shugart *et al.*, 2015).

Here, we examine how a fusion of IBMs and airborne laser scanning (ALS), a remote-sensing technology that provides structural information at landscape scale, can be used to improve allometric relationships and better understand the processes that shape them. We argue that by linking forest IBMs with ALS, we can reduce unexplained variation in allometric estimates and extend these to large spatial scales, as displayed in Fig. 1. This is an important step towards increased biological knowledge and improved predictions of ecosystem functioning. It is also a test case for the integration of future remote-sensing sources such as hyperspectral imaging or spaceborne laser scanning.

II. Tree allometry and transferability

When tree size, shape and function relate to each other across scales and environmental conditions, then the measurement of a single dimension can already provide a rough estimate of whole-tree attributes. This factor is particularly relevant when one quantity is more easily measured (for example trunk diameter) than the others (for example metabolic rate or biomass). Empirical studies provide a strong support for generalised allometric relationships. Whole-plant autotrophic respiration, for example, scales predictably with biomass across several orders of magnitude and from boreal to tropical forests (Mori *et al.*, 2010), and general patterns of allocation into aboveground vs belowgound plant organs exist at individual and stand levels globally (Poorter *et al.*, 2012; Chen *et al.*, 2019). Similarly, allometries that relate trunk diameter to tree height, as shown in Fig. 2, can be found across forest types and have been used to supplement height measurements that are error prone and time consuming without optimised protocols (Sullivan *et al.*, 2018).

The notion that a model developed at one site may be valid elsewhere is called transferability (Wenger & Olden, 2012). An important application is exemplified by the calculation of carbon stocks from forest inventories. The product of wood density, trunk cross-sectional area, and tree height turns out the be a good

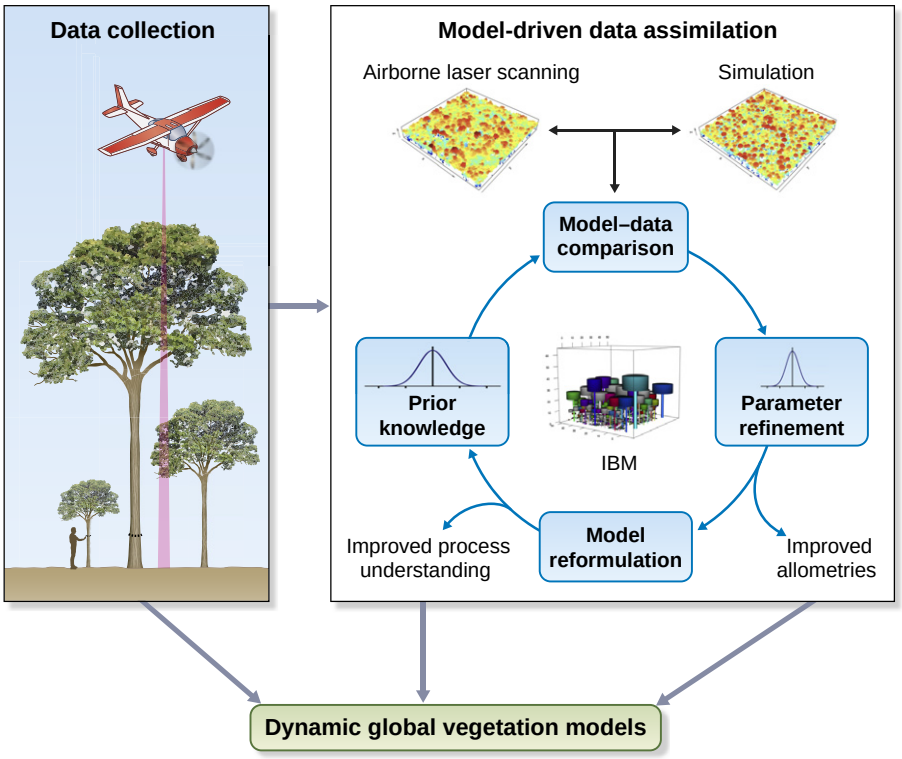


Fig. 1 Individual-based models (IBMs) as data assimilators, in interface with dynamic global vegetation models (DGVMs): application to allometric inference. Ground-based censuses and airborne laser scanning (ALS) provide complementary views on trees and forest canopies. Both techniques can be incorporated into the model–data fusion cycle, as formalised by Approximate Bayesian Computation (ABC). Increasingly diverse data can therefore be used to improve model representation and allometric parameter inference. Such improvement can be a benefit to DGVMs, whose simulations typically reach larger extents than IBMs, but which are currently run at coarse resolution, preventing them from making direct comparisons with data provided at finer spatial resolution.

© 2019 The Authors
New Phytologist © 2019 New Phytologist Trust

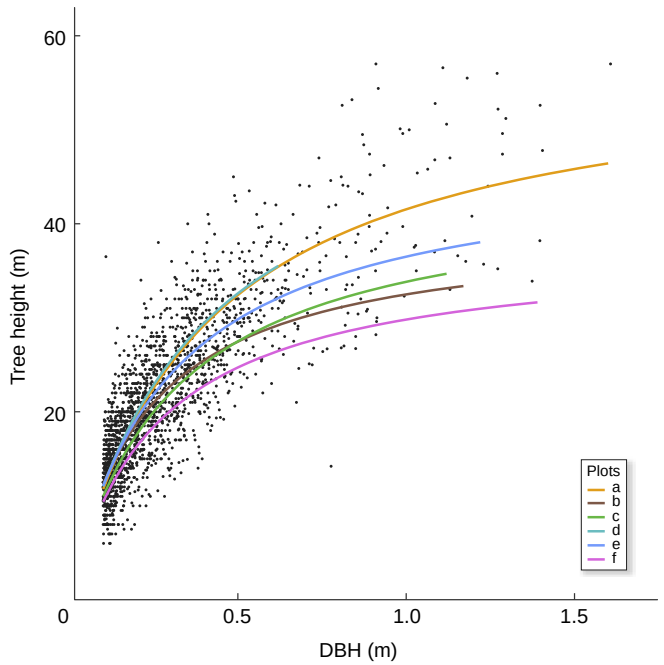


Fig. 2 Empirical allometric relations between tree height and trunk diameter (DBH). Michaelis–Menten type allometric models were fitted with nonlinear least squares and a heteroscedastic error structure at six sites, typical of tropical forests, as follows: (a) Ulu Ulu National Park, Brunei (4.54°N, 115.15°E); (b) Parque Estadual Cristalino, Mato Grosso, Brazil (9.06°S, 55.94°W); (c) Grebo National Forest, southeast Liberia (5.4°N, 7.62°W); (d) Nouragues Ecological Research Station, French Guiana (4.09°N, 52.67°W); (e) Dja Faunal Reserve, Cameroon (1.89°S, 13.22°E); (f) Tambopata National Reserve, Peru (12.84S, 69.29W). Data are from Sullivan *et al.* (2018), and metadata can be accessed on the forestplots.net data portal.

predictor of tree biomass obtained from destructive harvesting (Chave *et al.*, 2014). This holds true across a wide range of values for the predictor variables and broad bioclimatic gradients, from dry forest woodlands to tropical rainforests. Recent work based on an extensive destructive harvest experiment in African tropical forests suggests that relatively simple biomass models are transferable (Fayolle *et al.*, 2018), and could therefore be useful in biomass assessments across the tropics.

However, in most cases, allometries are influenced by environmental factors, both abiotic and biotic, and are not easily transferable. The scaling of tree height with trunk diameter, for example, depends on bioclimatic constraints (Lines *et al.*, 2012; Olson *et al.*, 2018), and tree growth is shaped by interactions with other trees (Coomes *et al.*, 2011; Jucker *et al.*, 2015). Furthermore, allometries typically have a multiplicative error structure. Residual standard deviations for predictions translate into large absolute errors for the biggest individuals and result in inflated uncertainty in the predicted variables.

To quantify variation in scaling of tree shape, remote sensing offers new perspectives. Terrestrial laser scanning (TLS), for example, provides accurate estimates of tree dimensions without requiring destructive harvesting (Momo Takoudjou *et al.*, 2017). It therefore holds great potential for exploring geometric scaling properties in forest trees and their dependence on environmental conditions (Disney, 2019).

III. Condensing the point cloud: allometry from space

Where TLS is a type of remote sensing 'from the ground', airborne LiDAR scanning extends the 3D-mapping capacity of forest and tree structure to the landscape scale. The technology and its application to forest scanning have been developed for over 3 decades (Schreier *et al.*, 1985; Nelson *et al.*, 1988), and studies now commonly cover several 1000 hectares of forest at high point densities, that is high resolution. As a result, individual tree shapes can be measured in open woodlands, allowing researchers to monitor the growth and death of individual plants (Levick & Asner, 2013; Duncanson & Dubayah, 2018). Even more impressively, clustering algorithms have been developed to segment ALS point clouds into individual tree crowns in closed-canopy forests (Ferraz *et al.*, 2016). As tree trunk diameter was recently found to be correlated with the product of tree height and crown size, the segmented crowns can then be used to estimate ground-based measurements (Jucker *et al.*, 2017); this technology is being increasingly used in routine forest monitoring programmes.

Tree-delineation from ALS is not without its problems, however. Trees often have irregular crowns, they may partly overlap, and the sharp light attenuation within dense canopy means that understorey trees are sparsely scanned, rendering the direct retrieval of tree dimensions difficult. IBMs such as TROLL (Maréchaux & Chave, 2017) offer an indirect, yet powerful alternative.

The spatially explicit rendering of treefalls and the competition for light resources introduce ecological constraints on the simulated forest structure, limiting tree density and dimensions across size classes. Instead of translating point clouds back into individual tree dimensions, we can create better fits between virtual and empirical canopies by adjusting vital rates and allometric parameters that can therefore be derived from mechanistic principles – even for trees that are difficult to observe directly from ALS. As TROLL's virtual canopies have a high spatial resolution (m^3), they compare naturally to ALS data and a few statistics are often suffice to link them. For example, Fig. 3 shows the match between top-canopy height obtained by ALS and a TROLL-based reconstruction. In the future it would be critical to extend this approach to other data sources, including TLS and spaceborne missions. Examples are the spaceborne laser scanner GEDI, a LiDAR now on board the International Space Station, and the BIOMASS synthetic aperture radar satellite, scheduled for launch in 2022, that will both provide a radically new view of the world's forests.

Because vegetation models and remote sensing have long proven mutually informative (Sellers *et al.*, 1997), the available approaches for data-model fusion have been well tested. Possibilities include the derivation of tree-level data from ALS for model parameterisation, the comparison of outputs with observed canopies for model validation (Seidl *et al.*, 2012, Fig. 3), and so-called model inversion, in which models are run with a wide range of parameter combinations and systematically compared with remotely derived metrics (Fig. 1). A hybrid between these approaches – partly inverse modelling, partly initialisation – was developed early on and has recently been applied to derive biomass maps across Amazonia using spaceborne LiDAR (Hurtt *et al.*, 2004; Rödig *et al.*, 2017). Moreover, when models

© 2019 The Authors
New Phytologist © 2019 New Phytologist Trust

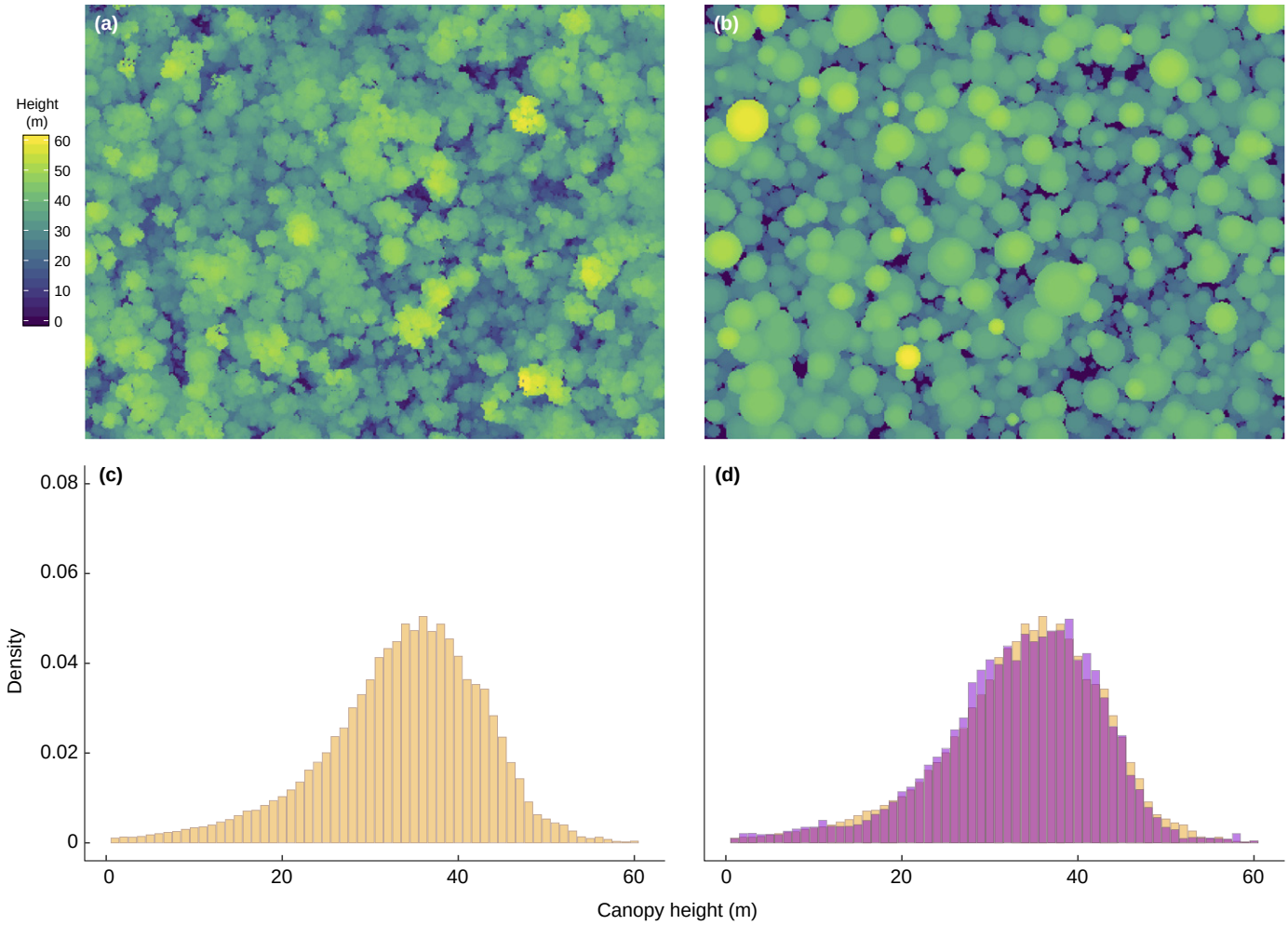


Fig. 3 Comparison of canopies sensed by airborne laser scanning (ALS) and simulated by individual-based models (IBMs). Shown are canopy-height models (height in m across 12 ha, 1-m^2 resolution; (a, b), and the corresponding height distributions (c, d). (a, c) Canopy-height model derived from an ALS campaign at the Nouragues field station, French Guiana, in 2012. (b, d) Canopy-height model as constructed by TROLL, an IBM of forest growth. Input information are a tree inventory, and allometries predicting tree height and crown dimensions from trunk diameter accounting for individual variation around the allometric trend.

provide realistic representations of forest structure, virtual ALS data can be produced and tested before using empirically observed canopies (Fassnacht *et al.*, 2018; Knapp *et al.*, 2018).

IV. Bayesian merging of data in IBMs

One efficient way to merge data and models is offered by Bayesian approaches such as Approximate Bayesian Computation (ABC), a widespread method in biological and ecological applications (Beaumont, 2010; Hartig *et al.*, 2011). Fig. 4 illustrates the inference of crown allometry parameters based on ABC. In qualitative terms, the approach is as follows: large numbers of simulations are performed with variations in crown allometry parameters (the prior in Bayesian statistics), the resulting virtual canopies are then compared with an empirically observed canopy (through statistics such as canopy height; Fig. 3) and, finally, the parameter values of the best-performing simulations are selected (the posterior). Inference on tree allometries is therefore turned into a parameter optimisation problem, and uncertainty around the parameter estimate reflects how informative is the data regarding a particular allometry. In the example given in Fig. 4, the inference is considerably improved by using ALS data in addition to ground data, providing more precise estimates for allometry parameters across diameter-size classes.

When harmonising high-dimensional data, as obtained from ALS and IBMs such as TROLL, some issues emerge. Inferences can be markedly different, depending on how virtual and empirical canopies are compared, and dimension reduction and cross-validation techniques are needed to find an appropriate set of statistics (Csilléry *et al.*, 2012; Nunes & Prangle, 2015). But even when summary statistics are well chosen, a pattern (for example a virtual canopy) can be the result of several parameter combinations or ways to represent processes (for example allometries). In this case, inference methods such as ABC are not well posed. This type of uncertainty, usually referred to as 'equifinality' (Luo *et al.*, 2009), cannot always be avoided, but it can be mitigated. Additional data sources can help to narrow down the parameter space (Fig. 4). Furthermore, it is desirable to implement mechanistic models over statistical ones, because mechanistic simulations are restricted to a generally smaller universe of possibilities. They therefore

© 2019 The Authors
New Phytologist © 2019 New Phytologist Trust

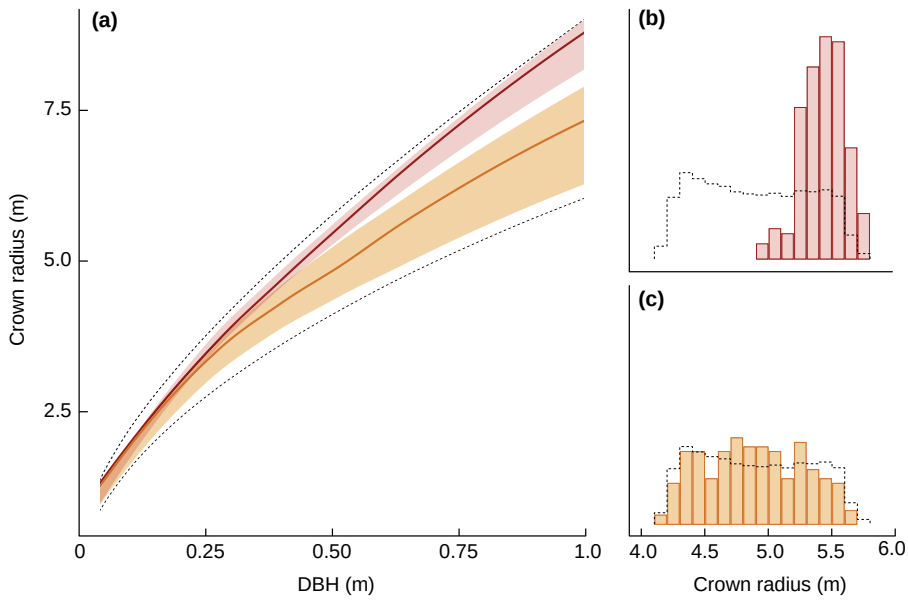


Fig. 4 Crown radius allometries inferred by Approximate Bayesian Computation (ABC) in TROLL. (a) Shape of the crown-radius allometry included in the TROLL model. Two posterior distributions of crown-radius allometry are shown, one constrained by ground data only (orange), the other also constrained by data derived from airborne laser scanning (ALS, red). The thick lines represent the posterior mode, the coloured areas the 70% highest-density intervals, and the dotted lines the extent of the prior distribution. (b, c) Cuts through the allometric distribution at 0.5 m in trunk diameter (DBH), for both simulations, with priors indicated by dotted lines. The addition of ALS data in ABC inference considerably narrows down the crown allometry parameters. The inference is based on 20 000 simulations, with a posterior composed of the best 200 simulations. Summary statistics included tree diameter-size distributions and ALS-derived canopy-height distributions. The overlap between simulated and empirical distributions was quantified and, to determine the posterior, we used the rejection scheme implemented in the R package ABC (Csilléry *et al.*, 2012) with an acceptance rate of 1%.

complement machine-learning techniques that are increasingly popular across science, including forest modelling, but are especially prone to the equifinality problem.

Another approach to narrow down the parameter space for models such as TROLL is the construction of an initial canopy state whose spatial arrangement is consistent with both the mechanistic principles of TROLL and the ALS-derived canopy structure. One method to produce such an initial state consists in sequentially assigning trees to spatial positions such that they receive enough light and that their size matches ALS observations. A space-filling rule is then iterated until all available space in the scene has been filled by trees (Taubert *et al.*, 2015). This initialisation can be useful to explore the range of validity of forest structure parameters (canopy gaps, crown exposure) and, therefore, yields both priors for the IBM and an evaluation of summary statistics. The IBM can then use this information to focus on ecological dynamics and provide distributions for tree trunks, crown dimensions and heights that represent a predictive check on ecological inferences and a new prior for the parameterisation of DGVMs (Fig. 1).

V. Challenges and perspectives

In this paper we argue that the explicit merging of plant allometry, forest observations, and individual-based modelling contributes to a unified vision of forest ecology. A fully spatially explicit IBM, when used for Bayesian data-model fusion, can inform quantities such as crown size and shape that are difficult to measure in dense canopies, but to which spatially explicit models are highly sensitive. The approach also helps to gain an understanding of ecological processes, as it captures the fine-grained structure of forest canopies. It could, therefore, better explain tree regeneration and simulate the dynamics of nontree life forms, including lianas and epiphytes or even canopy-dwelling animals. This challenge is one of the greatest in biodiversity research today (Singer *et al.*, 2016). Further ecological insights can be gained regarding submodels, such as the tree growth equations implemented in gap models (Shugart *et al.*, 2018), or the autotrophic respiration equation (Atkin *et al.*, 2015). The obtained information could then constrain the parameters of physiological models that are usually prescribed in DGVMs (but see Wang *et al.*, 2017), and offer a direct benchmark of upscaling simplifications of canopy structure, such as the perfect plasticity approximation (Purves *et al.*, 2008).

Where the focus of DGVMs has traditionally been on satellite data, forest IBMs have instead been developed for and from ground inventories, with trunk diameters and their growth the main predictors of all simulated ecological processes. The remote-sensing revolution calls, however, for a new paradigm in forest modelling, including new data sets, and new approaches to model building. This mirrors the larger change of direction in global forest research in which remotely sensed metrics are increasingly used to predict ground metrics (Jucker *et al.*, 2017). This also represents a timely challenge because spaceborne missions such as GEDI and BIOMASS will acquire global forest structure datasets, but it is likely that a correct interpretation of these datasets will require an explicit linkage with models (Fisher *et al.*, 2018). Model upscaling raises the question of spatial model transferability. It is crucial to test this by validating the model at places where it has not been calibrated.

© 2019 The Authors
New Phytologist © 2019 New Phytologist Trust

The assimilation of global remote-sensing data will be greatly helped by recent advances in computing technology that have shifted the limits of what forest extent can be simulated at tree level (Shugart *et al.*, 2015). For calibration, which requires 1000s of simulations for data-model fusion, computational cost can be reduced by classic assimilation techniques (Hurtt *et al.*, 2004). Future increases in computational power and emulators will further speed up inference (Fer *et al.*, 2018), turning IBMs into Bayesian data integrators that create a common vision of forest functioning and structure and the allometric relationships that link both across scales.

Complementary to this effort is the need to explain allometries from evolutionary optimisation arguments, for plant form (Enquist & Niklas, 2002), plant function (Wolf *et al.*, 2016), and forest structure (Farrior *et al.*, 2016). A better fundamental knowledge on allometric relationships can only improve our confidence in the parameters and simplify model calibration. This can only be achieved if theory is consistent with the known constraints of plant physiology. In return, data-model fusion, as explained here, provides a strong validation for theory.

Acknowledgements

This work has benefitted from 'Investissement d'Avenir' grants managed by Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01; ANAEE-France: ANR-11-INBS-0001, TULIP: ANR-10-LABX-0041), from CNES, and from ESA (CCI-Biomass). This work was granted access to the HPC resources of CALMIP supercomputing centre in Toulouse under allocation p17013.

ORCID

Jérôme Chave https://orcid.org/0000-0002-7766-1347
Fabian Jörg Fischer https://orcid.org/0000-0003-2325-9886
Isabelle Maréchaux https://orcid.org/0000-0002-5401-0197

References

Atkin OK, Bloomfield KJ, Reich PB, Tjoelker MG, Asner GP, Bonal D, Bönisch G, Bradford MG, Cernusak LA, Cosio EG *et al.* 2015. Global variability in leaf respiration in relation to climate, plant functional types and leaf traits. *New Phytologist* **206**: 614–636.

Beaumont MA. 2010. Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics* **41**: 379–406.

Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WBC, Duque A, Eid T, Fearnside PM, Goodman RC *et al.* 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology* **20**: 3177–3190.

Chen G, Hobbie SE, Reich PB, Yang Y, Robinson D. 2019. Allometry of fine roots in forest ecosystems. *Ecology Letters* **22**: 322–331.

Coomes DA, Lines ER, Allen RB. 2011. Moving on from Metabolic Scaling Theory: hierarchical models of tree growth and asymmetric competition for light. *Journal of Ecology* **99**: 748–756.

Csilléry K, François O, Blum MGB. 2012. Abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution* **3**: 475–479.

DeAngelis DL, Grimm V. 2014. Individual-based models in ecology after four decades. *F1000Prime Reports* **6**: 39.

Disney M. 2019. Terrestrial LiDAR: a three-dimensional revolution in how we look at trees. *New Phytologist* **222**: 1736–1741.

Duncanson L, Dubayah R. 2018. Monitoring individual tree-based change with airborne lidar. *Ecology and Evolution* **8**: 5079–5089.

Enquist BJ, Niklas KJ. 2002. Global allocation rules for patterns of biomass partitioning in seed plants. *Science* **295**: 1517–1520.

Farrior CE, Bohlman SA, Hubbell S, Pacala SW. 2016. Dominance of the suppressed: power-law size structure in tropical forests. *Science* **351**: 155–157.

Fassnacht FE, Latifi H, Hartig F. 2018. Using synthetic data to evaluate the benefits of large field plots for forest biomass estimation with LiDAR. *Remote Sensing of Environment* **213**: 115–128.

Fayolle A, Ngomanda A, Mbasi M, Barbier N, Bocko Y, Boyemba F, Couteron P, Fonton N, Kamdem N, Katembo J *et al.* 2018. A regional allometry for the Congo basin forests based on the largest ever destructive sampling. *Forest Ecology and Management* **430**: 228–240.

Fer I, Kelly R, Moorcroft PR, Richardson AD, Cowdery EM, Dietze MC. 2018. Linking big models to big data: efficient ecosystem model calibration through Bayesian model emulation. *Biogeosciences* **15**: 5801–5830.

Ferraz A, Saatchi S, Mallet C, Meyer V. 2016. Lidar detection of individual tree size in tropical forests. *Remote Sensing of Environment* **183**: 318–333.

Fisher RA, Koven CD, Anderegg WR, Christoffersen BO, Dietze MC, Farrior CE, Holm JA, Hurtt GC, Knox RG, Lawrence PJ *et al.* 2018. Vegetation demographics in Earth System Models: a review of progress and priorities. *Global Change Biology* **24**: 35–54.

Hartig F, Calabrese JM, Reineking B, Wiegand T, Huth A. 2011. Statistical inference for stochastic simulation models–theory and application. *Ecology Letters* **14**: 816–827.

Fischer R, Bohn F, Dantas de Paula M, Dislich C, Groeneveld J, Gutiérrez AG, Kazmierczak M, Knapp N, Lehmann S, Paulick S *et al.* 2016. Lessons learned from applying a forest gap model to understand ecosystem and carbon dynamics of complex tropical forests. *Ecological Modelling* **326**: 124–133.

Hurtt GC, Dubayah R, Drake J, Moorcroft PR, Pacala SW, Blair JB, Fearon MG. 2004. Beyond potential vegetation: combining lidar data and a height-structured model for carbon studies. *Ecological Applications* **14**: 873–883.

Jucker T, Bouriaud O, Coomes DA. 2015. Crown plasticity enables trees to optimize canopy packing in mixed-species forests. *Functional Ecology* **29**: 1078–1086.

Jucker T, Caspersen J, Chave J, Antin C, Barbier N, Bongers F, Dalponte M, van Ewijk KY, Forrester DI, Haeni M *et al.* 2017. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Global Change Biology* **23**: 177–190.

Knapp N, Fischer R, Huth A. 2018. Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states. *Remote Sensing of Environment* **205**: 199–209.

Levick SR, Asner GP. 2013. The rate and spatial pattern of treefall in a savanna landscape. *Biological Conservation* **157**: 121–127.

Lines ER, Zavala MA, Purves DW, Coomes DA. 2012. Predictable changes in aboveground allometry of trees along gradients of temperature, aridity and competition. *Global Ecology and Biogeography* **21**: 1017–1028.

Luo Y, Weng E, Wu X, Gao C, Zhou X, Zhang L. 2009. Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models. *Ecological Applications* **19**: 571–574.

Maréchaux I, Chave J. 2017. An individual-based forest model to jointly simulate carbon and tree diversity in Amazonia: description and applications. *Ecological Monographs* **87**: 632–664.

Medvigy D, Wofsy SC, Munger JW, Hollinger DY, Moorcroft PR. 2009. Mechanistic scaling of ecosystem function and dynamics in space and time: ecosystem Demography model version 2. *Journal of Geophysical Research: Biogeosciences* **114**: G01002.

Momo Takoudjou S, Ploton P, Sonké B, Hackenberg J, Griffon S, Coligny F, Kamdem NG, Libalah M, Mofack GII, Le Moguédec G *et al.* 2017. Using terrestrial laser scanning data to estimate large tropical trees biomass and calibrate allometric models: a comparison with traditional destructive approach. *Methods in Ecology and Evolution* **9**: 905–916.

Mori S, Yamaji K, Ishida A, Prokushkin SG, Masyagina OV, Hagihara A, Rafiqul Hoque ATM, Suwa R, Osawa A, Nishizono T *et al.* 2010. Mixed-power scaling of whole-plant respiration from seedlings to giant trees. *Proceedings of the National Academy of Sciences, USA* **107**: 1447–1451.

© 2019 The Authors
New Phytologist © 2019 New Phytologist Trust

Muller-Landau HC, Condit RS, Chave J, Thomas SC, Bohlman SA, Bunyavejchewin S, Davies S, Foster R, Gunatilleke S, Gunatilleke N *et al.* 2006. Testing metabolic ecology theory for allometric scaling of tree size, growth and mortality in tropical forests. *Ecology Letters* **9**: 575–588.

Nelson R, Krabill W, Tonelli J. 1988. Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment* **24**: 247–267.

Niklas KJ. 1994. Morphological evolution through complex domains of fitness. *Proceedings of the National Academy of Sciences, USA* **91**: 6772–6779.

Nunes MA, Prangle D. 2015. abctools: an R package for tuning approximate Bayesian computation analyses. *R Journal* **7**: 1–16.

Olson ME, Soriano D, Rosell JA, Anfodillo T, Donoghue MJ, Edwards EJ, León-Gómez C, Dawson T, Camarero Martínez JJ, Castorena M *et al.* 2018. Plant height and hydraulic vulnerability to drought and cold. *Proceedings of the National Academy of Sciences, USA* **115**: 7551 LP-7556.

Pan Y, Birdsey RA, Fang J, Houghton R, Kauppi PE, Kurz WA, Phillips OL, Shvidenko A, Lewis SL, Canadell JG *et al.* 2011. A large and persistent carbon sink in the world's forests. *Science* **333**: 988 LP-993.

Poorter H, Niklas KJ, Reich PB, Oleksyn J, Poot P, Mommer L. 2012. Biomass allocation to leaves, stems and roots: meta-analyses of interspecific variation and environmental control. *New Phytologist* **193**: 30–50.

Purves DW, Lichstein JW, Strigul N, Pacala SW. 2008. Predicting and understanding forest dynamics using a simple tractable model. *Proceedings of the National Academy of Sciences, USA* **105**: 17018–17022.

Rödig E, Cuntz M, Heinke J, Rammig A, Huth A. 2017. Spatial heterogeneity of biomass and forest structure of the Amazon rain forest: linking remote sensing, forest modelling and field inventory. *Global Ecology and Biogeography* **26**: 1292–1302.

Rogers A, Medlyn BE, Dukes JS, Bonan G, von Caemmerer S, Dietze MC, Kattge J, Leakey ADB, Mercado LM, Prentice IC. 2017. A roadmap for improving the representation of photosynthesis in Earth system models. *New Phytologist* **213**: 22–42.

Sato H, Itoh A, Kohyama T. 2007. SEIB-DGVM: a new Dynamic Global Vegetation Model using a spatially explicit individual-based approach. *Ecological Modelling* **200**: 279–307.

Schreier H, Lougheed J, Tucker C, Leckie D. 1985. Automated measurements of terrain reflection and height variations using an airborne infrared laser system. *International Journal of Remote Sensing* **6**: 101–446.

Scoffoni C, Albuquerque C, Brodersen CR, Townes SV, John GP, Cochard H, Buckley TN, McElrone AJ, Sack L. 2017. Leaf vein xylem conduit diameter influences susceptibility to embolism and hydraulic decline. *New Phytologist* **213**: 1076–1092.

Seidl R, Spies TA, Rammer W, Steel EA, Pabst RJ, Olsen K. 2012. Multi-scale drivers of spatial variation in old-growth forest carbon density disentangled with Lidar and an individual-based landscape model. *Ecosystems* **15**: 1321–1335.

Sellers PJ, Dickinson RE, Randall DA, Betts AK, Hall FG, Berry JA, Collatz GJ, Denning AS, Mooney HA, Nobre CA *et al.* 1997. Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science* **275**: 502–509.

Shugart HH, Asner GP, Fischer R, Huth A, Knapp N, Le Toan T, Shuman JK. 2015. Computer and remote-sensing infrastructure to enhance large-scale testing of individual-based forest models. *Frontiers in Ecology and the Environment* **13**: 503–511.

Shugart HH, Wang B, Fischer R, Ma J, Fang J, Yan X, Huth A, Armstrong AH. 2018. Gap models and their individual-based relatives in the assessment of the consequences of global change. *Environmental Research Letters* **13**: 033001.

Singer A, Johst K, Banitz T, Fowler MS, Groeneveld J, Gutiérrez AG, Hartig F, Krug RM, Liess M, Matlack G *et al.* 2016. Community dynamics under environmental change: how can next generation mechanistic models improve projections of species distributions? *Ecological Modelling* **326**: 63–74.

Smith B, Wärlind D, Arneth A, Hickler T, Leadley P, Siltberg J, Zaehle S. 2014. Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences* **11**: 2027–2054.

Sullivan MJP, Lewis SL, Hubau W, Qie L, Baker TR, Banin LF, Chave J, Cuni-Sanchez A, Feldpausch TR, Lopez-Gonzalez G *et al.* 2018. Field methods for sampling tree height for tropical forest biomass estimation. *Methods in Ecology and Evolution* **9**: 1179–1189.

Taubert F, Jahn MW, Dobner H-J, Wiegand T, Huth A. 2015. The structure of tropical forests and sphere packings. *Proceedings of the National Academy of Sciences, USA* **112**: 15125–15129.

Wang H, Prentice IC, Keenan TF, Davis TW, Wright IJ, Cornwell WK, Evans BJ, Peng C. 2017. Towards a universal model for carbon dioxide uptake by plants. *Nature Plants* **3**: 734–741.

Wenger SJ, Olden JD. 2012. Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods in Ecology and Evolution* **3**: 260–267.

Wolf A, Anderegg WRL, Pacala SW. 2016. Optimal stomatal behavior with competition for water and risk of hydraulic impairment. *Proceedings of the National Academy of Sciences, USA* **113**: E7222–E7230.

Wolf A, Ciais P, Bellassen V, Delbart N, Field CB, Berry JA. 2011. Forest biomass allometry in global land surface models. *Global Biogeochemical Cycles* **25**: GB3015.

About *New Phytologist*

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews and Tansley insights.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <26 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)
- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**

© 2019 The Authors
New Phytologist © 2019 New Phytologist Trust

Chapter 2: A new method to infer forest structure and tree allometry from airborne laser scanning and forest inventories

(Target Journal: Remote Sensing of the Environment, to be submitted October 30, 2019)

Chapter 2 builds on the approach outlined in Chapter 1 and, using the geometric principles of the TROLL model, develops a new method called the Canopy Constructor. The Canopy Constructor uses a combination of field inventory data and Airborne Lidar scans to create virtual 3D representations of forest stands. The approach consists of two steps: At the plot scale, the Canopy Constructor creates 3D scenes that best fit ground and airborne data and then infers the underlying forest structure (allometry, crown packing density). In a second step, the results of the first step are extrapolated over the whole lidar scene to create virtual tree inventories across thousands of hectares in a spatially explicit way. In the paper, we present results from an application to two tropical rain forests, one in French Guiana and one in Gabon, where we used the Canopy Constructor to infer forest structure, created high resolution maps of above-ground biomass and tree abundance, and validated both steps against ground data.

A new method to infer forest structure and tree allometry from airborne laser scanning and forest inventories

Fabian Jörg Fischer [1,*], Nicolas Labrière [1], Grégoire Vincent [2], Bruno Hérault[3,4], Alfonso Alonso[5], Hervé Memiaghe[6], Pulchérie Bissiengou[7], David Kenfack[8], Sassan Saatchi[9], and Jérôme Chave [1]

[1] Laboratoire Évolution et Diversité Biologique, UMR 5174 (CNRS/IRD/UPS), 31062 Toulouse Cedex 9, France

[2] Botanique et Modélisation de l'Architecture des Plantes et des Végétations (AMAP), UMR 5120 (CIRAD/CNRS/INRA/IRD/UM2), 34398 Montpellier Cedex 5, France

[3] Cirad, Univ Montpellier, UR Forests & Societies, F-34000 Montpellier, France.

[4] INPHB, Institut National Polytechnique Félix Houphouët-Boigny, Yamoussoukro, Ivory Coast

[5] Center for Conservation and Sustainability, Smithsonian Conservation Biology Institute, 1100 Jefferson Drive SW, Suite 3123, Washington DC 20560-0705, USA

[6] Institut de Recherche en Écologie Tropicale (IRET), Centre National de la Recherche Scientifique et Technologique (CENAREST), B.P. 13354, Libreville, Gabon

[7] Herbier National du Gabon, Institut de Pharmacopée et de Médecine Traditionnelle (IPHAMETRA), Centre National de la Recherche Scientifique et Technologique (CENAREST), B.P. 13354, Libreville, Gabon

[8] Center for Tropical Forest Science -Forest Global Earth Observatory, Smithsonian Tropical Research Institute, West Loading Dock, 10th and Constitution Ave NW, Washington DC 20560, USA

[9] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

* Correspondence: fabian.j.d.fischer@gmx.de

Keywords: Forest structure, Individual-based modelling, Airborne Lidar, Approximate Bayesian Computation, Biomass

---

1. Introduction

Tropical forests sequester large amounts of carbon and thus play a pivotal role in carbon mitigation strategies (Chazdon *et al.*, 2016; Grassi *et al.*, 2017). Of particular importance to biomass stocks and ecosystem functioning is forest structure, i.e. the vertical and horizontal arrangement of tree stems and crowns (Shugart *et al.*, 2010). In order to improve carbon mitigation strategies, we need methods to quantify forest structure that account for local heterogeneities and are also applicable over the large areas covered by tropical forests (Fischer, R., *et al.*, 2019). Here, we propose a new method for quantifying forest structure by constraining an individual-based model with airborne lidar data.

Field-based inventories provide detailed description of three-dimensional forest structure across time and space and form the bedrock of research in forest ecology. However, they are often limited to a few hectares in sampled area, and typically involve mapping, measuring and identifying all trees above a trunk diameter threshold (e.g. above 10cm) within the sampled area. Furthermore, reliable measurement of tree height and other crown dimensions from the ground is difficult (Feldpausch *et al.*, 2012; Sullivan *et al.*, 2018). Therefore, detailed description of the three-dimensional forest structure has long been limited to drawings illustrating the stratification of tropical rain forests (Oldeman, 1974).

Much has changed, however, with the advent of laser scanning. Aircraft-mounted laser scanning devices (aerial laser scanning, ALS), are now commonly used to survey thousands of hectares of forest, and to obtain information such as canopy height and leaf density at centimetric resolution (Riaño *et al.*, 2004; Rosette *et al.*, 2008). In some situations, individual tree dimensions – especially tree height, crown area and depth –

can be deduced from dense ALS point clouds by segmentation methods (Morsdorf *et al.*, 2004; Ferraz *et al.*, 2016). However, in multistoried forests, many trees are overtopped and often difficult to delineate, so a large part of the individual tree size information is veiled. Here we propose an alternative model-based strategy that assimilates ALS data together with tree inventories. Our method, which we call "Canopy Constructor", is related to the canopy-filling algorithms published recently (Taubert *et al.*, 2015; Bohn & Huth, 2017).

The method consists of constructing simulated forest canopies using the assumptions of a spatially explicit individual-based forest model, here the model TROLL (Maréchaux and Chave 2017), then optimizing the model structural parameters to match the observations, using a Bayesian inversion technique (see e.g., Hartig *et al.*, 2011). Our method specifically infers the static three-dimensional structure of a forest from a combination of three basic elements: (1) forest inventories collected over a few ha, (2) ALS surveys collected over a few hundred to thousand ha, (3) allometric relationships relating tree dimensions, such as tree height, trunk diameter, and crown size. Building on a few simple assumptions about space-filling and scaling relationships, the Canopy Constructor builds up canopies from below, tree by tree. Importantly, it includes interindividual variation in tree architecture, and, in an iterative process, redraws and shifts tree crown dimensions until reaching a high similarity between an empirical, ALS-derived canopy and the simulated canopy. The Canopy Constructor thus infers forest structure, as quantified through allometric equations, from a field inventory and a co-registered ALS campaign. In addition, it is also used to extrapolate forest structure over the whole ALS-observed area.

We explored the performance of Canopy Constructor on two tropical rain forest sites with extensive ground inventory data (> 20ha) and ALS surveys of several thousand hectares in total – one at the Nouragues field station in French Guiana, the other at the Rabi site in Gabon (Labriere *et al.*, 2018). For both sites, we tested the predictive quality of the algorithm, compared results to empirical measurements, and analyzed biomass predictions with regard to previous estimates. Specifically, we asked the following questions: (1) how well can the allometric relationships between trunk diameter and tree dimensions be inferred based on the combined knowledge of tree inventories and airborne lidar scanning; (2) how well can we predict tree diameters and biomass from purely ALS-based metrics and how does local forest heterogeneity affect these predictions; (3) what are the predictions of the individual-based tree reconstruction approach regarding aboveground biomass stocks at landscape scale and how do they compare to simpler regression models?

2. Materials and Methods

2.1 Data

Two sites were chosen based on availability of large (≥ 10 ha) field-based measurements and co-registered ALS campaigns (Labriere *et al.*, 2018).

Part of the study was conducted at the Nouragues Ecological Research Station in French Guiana (4.06°N, 52.68°W). The site is characterised by a lowland tropical rainforest (except for a granitic outcrop at 430m asl), ca. 2900 mm rain per year and one 3-month dry season, in September-November and a shorter one in March. Field inventories have been carried out on a regular basis since the early 1990s (Chave *et al.*, 2008; Labriere *et al.*, 2018), and several ALS surveys have been conducted since 2008 (Réjou-Méchain *et*

al., 2015). We here use a ground inventory at two plots (a 10ha plot called "Grand Plateau" and a 12ha plot called "Petit Plateau") together with an ALS campaign. The field inventory was conducted at the end of 2012, with trees mapped on both plots, measured at a height of 1.30m dbh (diameter at breast height) and identified at the species level. ALS acquisition was done with a Riegl laser rangefinder (LMS-Q560) earlier in March of 2012 and covers 2,400 ha (Réjou-Méchain *et al.*, 2015) at an average pulse density of ~12 per m^2 (based on density of last returns) and an overall point density of ~18 per m^2 (all returns).

The second site, Rabi, is in Gabon (1.92°S, 9.88°E) and is part of the AfriSAR campaign (Fatoyinbo *et al.*, 2017; Labriere *et al.*, 2018). It is characterised by a lowland tropical rain forest – partly disturbed by oil operation. The plot is located in southwestern Gabon's Gamba Complex, and is representative of the Guineo-Congolian rainforest that contains a diverse mix of upland and wet-forest habitats. Annual rainfall is of ca. 2300 mm per year on average. A forest inventory, covering 25ha and including all trees ≥ 1 cm dbh, was conducted between 2010-2012. An airborne lidar campaign was carried out three years later, using a helicopter-based RIEGL VQ-480i, with point densities of 2.5 per m^2. For validation purposes (cf. 2.3 below), we split the 25-ha plot into two rectangular strips of 10ha and 15ha respectively, corresponding roughly to the 10ha and 12ha sizes at Nouragues.

ALS observations were converted into canopy height models (CHMs) to minimize site-specific biases. CHMs are defined as the top-of-canopy height above ground for a given grid cell, here at a $1m^2$ resolution. CHMs provide a robust baseline, since they are not too

sensitive to the technical specifications of lidar instruments. To create the CHMs, the lidar data were classified via TerraScan and then post-processed with LAStools to obtain pit-free CHMs (Khosravipour *et al.*, 2014; Isenburg, 2018).

2.2 Model description

We now describe the core forest reconstruction and fitting algorithm implemented in the Canopy Constructor. We proceed from a forest inventory, a co-registered ALS canopy-height model, and a set of allometries, and convert them into spatially explicit tree reconstructions. For each tree, variation around allometric means is assigned following a prescribed distribution, and the Canopy Constructor then seeks to optimize tree dimensions spatially by moving or redrawing tree crowns, until virtual and empirical canopy are sufficiently similar. Via model inversion, the final reconstructions can then be used for allometric inference. Optionally, the results can serve as a calibration step to predict a wall-to-wall forest inventory over the full area of the ALS survey. This virtual inventory is inferred from an ALS-only model and space-filling principles (cf. Section 3, "Extrapolating forest surveys across landscapes").

All simulations were developed in C++. Statistical analysis and visualization were carried out in R (R Development Core Team, 2019), including the packages *data.table*, *raster*, *ggplot2*, and *viridis* (Wickham, 2011; Hijmans, 2016; Dowle & Srinivasan, 2018; Garnier, 2018).

Allometric relationships

In the initial step, the Canopy Constructor inputs tree diameters and locations from a forest inventory. Tree heights and crown shape and size are simulated through allometry, as explained below. Using these crown shapes, as well as the allometrically predicted tree dimensions, we then fill up the 3D-canopy (resolution of $1m^3$).

First we assign each tree to a grid with $1m^2$ cell size. If several trees co-occur on the same $1m^2$ grid cell, a single effective tree is retained, with an effective stem diameter at breast height (dbh) equal to $dbh_{eff} = \sqrt{\sum_i dbh_i{}^2}$. For simplicity, we refer to dbh_{eff} as dbh. If the field inventory has a cutoff value above 1cm (e.g., 10cm), the non-measured trees are filled up. In this study, the Rabi plots had all trees ≥ 1 cm measured, but the Nouragues plots had to be gap-filled and we parameterized the dbh-size distribution as follows $P(dbh) = \exp(-dbh/4.2)$. While power laws or Weibull distributions generally provide a better fit for small trees (Muller-Landau *et al.*, 2006), this simple parameterization of an exponential function yielded overall tree densities (>= 1cm in dbh) upwards of 4,500 ha^{-1} consistent with observations at the site and sufficient for the purposes of the Canopy Constructor (data not shown). The gap-filled trees were then placed on the grid randomly.

The Canopy Constructor then predicts the tree dimensions through the following allometric models:

$$h = \frac{hmax \times dbh}{(a_h + dbh)} \times \exp(\varepsilon_h) \qquad (1)$$

$$cr = \exp(a_{cr} + \varepsilon_{cr}) \times dbh^{b_{cr}} \quad (2)$$

Here, h is tree height, dbh diameter at breast height, $hmax$ and a_h Michaelis Menten coefficients. Similarly, cr is the tree's crown radius, and a_{cr} and b_{cr} are the intercept and slope of a log-log regression, i.e. a power law model. For tree height, Equation (1) is chosen instead of a power model to better capture the saturating relationships typically found in tropical rain forests (Cano *et al.*, 2019). The ε_h and ε_{cr} are the respective error terms – i.e. the natural variation in allometry –, given by:

$$\varepsilon_h \sim N(0, \sigma_h) \quad (3)$$

and

$$\varepsilon_{cr} \sim N(0, \sigma_{cr}) \quad (4)$$

In both cases, they are exponentiated to yield a multiplicative error structure that is more relevant biologically and accounts for the heteroscedasticity in crown and height allometries (Molto *et al.*, 2014). Here, we assume that allometric variation does not depend on species identity, that there is no covariance between tree height and crown radius, and that crown depth can be simply calculated as a proportion of h, as in the TROLL model (Maréchaux & Chave, 2017).

We also modelled variation in crown shape. We defined the ratio γ between the radius at the top of the crown and its base, with a linear slope linking both layers. γ varies between 0 and 1: if $\gamma = 0$, the tree crown is a cone, while if $\gamma = 1$, it is a cylinder, as assumed in the TROLL model. For the purposes of this study, we set γ to 0.8. We have chosen this parameterization to account for the less clear-cut edges found empirically and the fact that tree crown volume is always smaller than a cylindric envelope.

Optimization algorithm

Once the trees are reconstructed to create an initial forest mockup, we improve its overlap with the ALS-derived canopy by reshuffling the tree crowns in space. We loop through all trees on the grid, in random order, and adjust their crown dimensions on a tree-by-tree basis.

Because this optimization step can be time-consuming, we here choose to perform the following algorithm. For the majority of trees, we pick pairs of trees and swap their respective values of ε_h (deviation in height) and ε_{cr} (deviation in crown radius). We then recalculate the new dimensions of both trees and keep the change if it results in an increase in the overall goodness of fit when compared with the ALS-derived CHM. To keep the overall variance structure, trees are binned in logarithmic dbh classes and we only pick pairs of trees within the same dbh class. This procedure rapidly redistributes deviations from the allometric means across the population of trees so as to create better spatial fits, all the while preserving the initial, randomly drawn, allometric structure.

A special case are trunk diameter classes that contain only a small number of trees (here set to < 10 trees for the whole plot). Unless plots have been heavily disturbed, this is typically the case for the largest diameter classes (e.g. > 1m). Since low tree numbers mean that there are limited opportunities to swap tree dimensions, we redraw altogether new values ε_h and ε_{cr}. If the new draw creates a better fit to empirical data, it is retained. Each new draw must preserve the allometric mean within the *dbh* class. If trees within a class are, for example, smaller on average than their expected allometric height, the newly drawn deviations are accepted only if they exceed the mean height and

thus compensate for the negative bias. Another special case are trees that have been randomly placed, such as gap-filled trees < 10cm *dbh*. Since their initial positions have been chosen at random, we do not change the trees' dimensions, but only their position: a tree is moved at random within a radius equal to its height. If the new location increases the goodness of fit, the change is accepted.

This loop across all trees is iterated several times, until improvements in canopy structure become marginal, i.e. low acceptance rates are reached (< 1%). This part of the algorithm is similar to the one described in Taubert *et al.* (2015). In practice, we have observed that a small number of iterations (~100-200) are sufficient.

We considered the issue of boundary conditions within tree inventories (see also Mascaro *et al.*, 2011). For each tree *i*, we calculated the crown area outside the plot CA_i^{out} and the total crown area CA_i, summed across the n inventoried trees to compute the whole plot ratio $R = \frac{\sum_{i=1}^{n} CA_i^{out}}{\sum_{i=1}^{n} CA_i}$. During optimization procedure, we forced R to remain constant, close to the starting value. If during the fitting process, *R* exceeds its initial value, then the next fitting at the edge for tree *i* is only accepted if it decreases *R*, and vice versa. We also ruled out cases of trees with small crown radius, but large height, growing through trees with large crown radius, but small height. To do so, for every newly fitted crown, we circled through all trees within a distance $dist = CR_{tree} + CR_{treemax}$, where CR_{tree} is the current tree's crown radius and $CR_{treemax}$ is the maximum crown radius of trees allowed. This gives the maximum distance within which the current crown could theoretically overlap with another crown. We then determined the

2D-overlap of the crown areas, and reject crown fittings that create a forest patch where the crown of a taller tree is fully encompassed by the crown of a smaller tree.

Goodness-of-fit metrics

In the algorithm, each time a tree crown is updated, we test whether this change increases the match with empirical values. To assess the goodness of the fit between virtual and empirical CHMs, we use two metrics. The first one is the mean of the absolute errors:

$$MAE = \frac{1}{sites} \sum_{s=1}^{sites} \left| chm_{emp}(s) - chm_{sim}(s) \right| \tag{5}$$

where each s represents a 1m^2 grid cell of forest, chm_{emp} and chm_{sim} the empirical and simulated canopy heights of that grid cell and *sites* the total number of grid cells, respectively. This metric adjusts trees locally to reproduce canopy height patterns. We opt for the mean absolute error instead of a mean squared error, because it is more robust with regard to outliers (Hill & Holland, 1977). Such outliers are frequent in our procedure, since real tree crowns tend to have gaps and crown irregularities that can create large deviations from the idealized crown shapes (cylinders, cones) presupposed here.

Since large trees are important to forest structure and biomass estimates, but would be underestimated by shrinkage towards the mean from the optimization of *MAE*, we also constrain by the dissimilarity index of the canopy height distributions:

$$D = \frac{1}{2} \sum_{h=0}^{h=hmax} \left| \mathrm{d}_{\mathrm{emp}}(h) - \mathrm{d}_{\mathrm{sim}}(h) \right| \tag{6}$$

where h is a discrete height index (in m), and d_{emp} and d_{sim} are the densities of the empirical and simulated height histogram across the surveyed area, i.e. total number of height occurrences, normalized by the number of 1m² grid cells. This index can be interpreted as a measure of distribution overlap, i.e. the lower the dissimilarity, the higher the overlap. In the limit of $D = 0$, both distributions are identical. Formally, if OVL is the distribution overlap, then $D = 1 - OVL$, with $OVL = \sum_{h=0}^{hmax} \min\left(d_{emp}(h), d_{sim}(h)\right)$ (Inman & Bradley, 1989).

To combine the metrics, we first fit the tree crowns using each metric separately, until a low acceptance rate is achieved for each metric (< 1%, typically reached within 50 iterations for the *MAE* metric, and within 5 iterations for the dissimilarity metric). This gives us a maximum (initial) and minimum (fitted) value for both metrics, and we use the difference of these values to normalize each metric. The normalized values are combined to an overall error as follows:

$$\varepsilon = \sqrt{MAE_{norm}^2 + D_{norm}^2} \qquad (7)$$

We then run a final number of iterations to minimize ε. In using the combined metric, we ensure that crowns do not only fit spatially at local scales, encapsulated by a low *MAE*, but also preserve the overall canopy height model distribution and prevent shrinkage towards the mean, encapsulated by a low dissimilarity D.

Forest structure characterization

Once the canopy has been reconstructed, we calculate the aboveground biomass (in kg) for each tree as in Chave *et al.* (2014): $AGB = 0.0673 \times (\rho \times dbh^2 \times h)^{0.976}$. Here ρ

represents wood density and is either directly assigned from the field census or drawn from a local distribution of wood densities. We also calculate crown packing density, an important descriptor of forest structure that summarizes the proportion of space occupied by tree crowns within the canopy (Taubert *et al.*, 2015). Our algorithm does not restrict crown overlap, so a useful definition of packing density is the ratio of unit crown volume to unit canopy volume (m^3 per m^3). This value can be locally larger than 1.0, if two or more tree crowns overlap, but is equivalent to the Taubert et al. (2015) definition in the limiting case of no crown overlap. Particularly useful for the characterization of forest structure is the crown packing density at height h, with $0 \leq h \leq h_{max}$, and with h_{max} top-of-canopy height. The result is best described by a matrix, where columns represent top-of-canopy height and rows represent within-canopy height layers (cf. Figure 1, left panel). We call this quantity the packing density matrix.

2.3 Model calibration and prediction

Inferring Allometric Parameters by Approximate Bayesian Computation

The core routine of the Canopy Constructor finds the best canopy reconstruction, given a certain set of allometric parameters. Here, we used this routine to solve the inverse problem: which combination of allometric parameters is the most likely to match the observed data?

To provide an answer to the question, we used an Approximate Bayesian Computation rejection scheme (Csilléry *et al.*, 2010; Hartig *et al.*, 2014; Fischer, FJ, *et al.*, 2019). We drew 10,000 combinations for a set of six allometric parameters: ($hmax$, a_h, a_{cr}, b_{cr}) and the two variance terms (σ_h, σ_{cr}). We used these to approximate the prior probability

distribution over all parameters. We then applied the Canopy Constructor to all 10,000 allometric parameter combinations, reconstructed a best fit canopy for each and retained only the results close enough to the empirical ones. The retained parameter values provided a posterior probability distribution over credible allometric parameterizations given the data.

We chose flat priors by drawing from uniform distributions within ranges of tree allometry observed globally (Jucker *et al.*, 2017). Parameters were drawn on the logscales they were described at, except for the crown allometry intercept a_{cr}, drawn from a uniform distribution on the back-transformed scale. We applied a Latin Hypercube scheme, and accounted for correlation between allometric coefficients as described in the R package 'pse' (Chalom *et al.*, 2013), but rewritten in C++ for speed. Covariance coefficients were also taken from the data set in Jucker *et al.*, (2017). Since crown depth does not influence canopy height – and thus does not directly affect the fitting procedure –, it was fixed to 20% of tree height.

As summary statistics to assess each simulation's fit with the empirical CHM, we used the same metrics as for the Canopy Constructor fitting procedure, i.e. mean absolute error (*MAE*) and the dissimilarity *D*. We used the difference between maximum and minimum values across all simulations to normalize them (instead of within-simulation minimum and maximum) and combined them as before to $\varepsilon_{ABC} = \sqrt{MAE_{normABC}^2 + D_{normABC}^2}$. Only the best 1% of reconstructions (i.e. 100 parameter sets) were retained (Csilléry *et al.*, 2010).

ALS-based extrapolation of forest surveys across landscapes

Once the allometries were inferred, we used the Canopy Constructor to extend the tree-by-tree reconstructions over the whole ALS-covered area. To do so, we essentially created a second model of forest structure, based only on ALS-data and a few assumptions concerning forest structure: (1) Stem diameter distributions across the whole area are similar to the field inventory; (2) The allometric parameters inferred locally can be extrapolated the whole area; (3) The local crown packing densities are representative of the whole lidar-covered area.

We implemented the following routine: For each posterior simulation from the calibration step, we extracted the packing density matrix (cf. 2.2.4), extrapolated it across the whole ALS-covered area (assumption 3) and could thus infer the average crown packing densities at any height underneath any top-of-canopy height. To speed up the calculation, we divided the ALS scans into 400m x 400m patches and ran the full suite of 100 posterior draws on each patch's CHM. Areas at the edges of the ALS-scans (with an area < 400m x 400m) were discarded. Based on the measured CHM distribution, we calculated an average crown packing density for each canopy voxel of the chosen patch and then summed across the whole canopy height distribution to obtain the total crown volume per height layer. This procedure can be implemented efficiently as a matrix multiplication:

$$\bar{v} = \boldsymbol{P}\bar{c}$$

Where $\boldsymbol{P}$ is the n x n crown packing density matrix, $\bar{c}$ the patch's CHM distribution, formalized as a row vector of top-of-canopy height frequencies (from 0 to n, in m), and $\bar{v}$ a column vector, summarizing total volume per height layer (also from 0 to n, in m) across the whole patch (for a visual demonstration cf. Figure 1). The dimension n is the

number of height bins into which the canopy is discretized, from 0m to maximum canopy height. We here chose a discretization step of 1m, so that n corresponds to the maximum canopy height. If the maximum height of the calibration plot was lower than the maximum height of the patch where we predicted (i.e. $n_{matrix} < n_{vector}$), the crown packing matrix had to be extrapolated by rescaling and averaging across top-of-canopy heights immediately below the missing top-of-canopy height value.

The distribution of total crown volume per height layer was then used as a reference for space-filling within the forest patch under consideration. To fill the forest with trees, we drew random stem diameters from the local distribution (assumption 1), applied the locally calibrated allometries (assumption 2), and randomly placed trees on the grid until the distribution over crown volume per height layer corresponded to the reference distribution. As a stopping rule, we determined by how much the newly added tree improved the overlap for every height layer (i.e. filling volume underneath its reference value) and by how much it reduced the overlap (i.e. filling volume beyond its reference value). If the reduction in overlap was greater than the improvements, we rejected the tree. If after one full cycle through the stem diameter distribution, the rejection rate reached 100%, we stopped the procedure.

This procedure results in a virtual forest inventory with random allometric variation, as required as starting point for the Canopy Constructor algorithm. Given that all trees were placed randomly, we could simply move the trees until an optimal spatial fit was achieved, as described above (cf. "Optimization Algorithm"). This process was repeated for all of the posterior draws, i.e. for a 100 distinct sets of allometric parameters and

crown packing densities, thus turning the posterior from step 1 into a prior for step 2 and propagating uncertainties across the whole procedure.

Inference and validation at the two sample sites

In this paper, we explored the procedure at two sites, at Nouragues in French Guiana, and Rabi in Gabon. To assess within-site heterogeneity and test the accuracy of our predictions, we split the field inventories into two parts, corresponding to roughly half of the data at each site. At Nouragues we used the already geographically separated Petit Plateau (12 ha) and Grand Plateau (10 ha) plots, at Rabi the contiguous 10-ha and 15-ha subplots of the 25-ha plot. We used large subplots rather than representative samples, because non-random splitting of data is often better-suited for transferability assessments (Wenger & Olden, 2012). Furthermore, plot sizes of > 10ha allowed us to minimize edge effects and keep a balance between the computational burden of the procedure and the sample sizes needed to swap variance between crowns.

On all of the four plots, we separately inferred tree allometries and forest structure properties (crown packing densities). To validate the allometric inference, we used data and estimates that had not been used in the fitting procedure ("predictive check"). These included allometric relationships derived from field measurements of tree height and diameter (Labriere *et al.*, 2018; Sullivan *et al.*, 2018) as well as plot biomass estimates for both sites, reported in Labriere et. al. (2018). For Nouragues, two separate allometric relationships were available for Petit Plateau and Grand Plateau. For Rabi, we compared both plots to the overall site allometry.

The ground-reconstructed forest was then used to validate the extrapolation step. First we compared the ALS-based predictions for all four plots against their own ground-based data, i.e. we treated the plot from which we had obtained packing densities and allometric parameters as if we needed to extrapolate to it ("local validation"). This allowed us to estimate how well inferred allometries and packing densities summarized forest structure and the error introduced by inference "from above". We then assessed the effects of between-plot heterogeneity in forest structure and the transferability of the extrapolation model by using plots as training and validation data for each other ("crossvalidation"). As metrics we chose stem diameter distributions and above-ground biomass estimates. We computed aboveground biomass estimates on 1ha and 0.25ha grid cells, and quantified predictive accuracy through the overall RMSE (given in t/ha) across all four plots.

Finally, we used the larger calibration plots at both sites (i.e. Petit Plateau and the 15ha Rabi plot) to predict tree positions and biomass across the landscape. We then reported how above-ground biomass scaled up to the whole area and compared our results to biomass estimates from Labriere *et al.*, 2018, in terms of precision and accuracy.

3. Results

The overall approach of the Canopy Constructor is summarized in Figure 2, applied on the Nouragues site (Petit Plateau plot). Figure 2 shows that the initial draw already depicts the mean canopy structure, but not the spatial location of features. Swapping the trees ('final fit') greatly improved the spatial structure. Although the match was not perfect at metric resolution, the final mean error was typically < 0.5m and the final mean

absolute error < 3m, or 10% of the mean canopy height, mainly due to crown gaps and large deviations from idealized geometric shapes at the crown edges.

In terms of allometric inference, height allometries were better constrained than crown radius allometries (cf. Figure 3 for an example at Petit Plateau, compare top and bottom panels, also Table 1). In both cases, we found substantial covariation between the allometric parameters (cf. Table 2, and Figure 3, left panels). High within-site similarity was found for height allometries at both Nouragues and Rabi (Figure 4). Crown allometries, on the other hand, showed a divergence at Nouragues, with larger crown radii predicted at Petit Plateau than at Grand Plateau. Between sites, the Rabi and Nouragues site were clearly separated by their height allometries, but not by their crown allometries (cf. Figure 4, righthand panels).

Compared to empirical results, the Canopy Constructor produced parameter estimates very close to previously obtained allometries at both sites (cf. Figure 4, top row), also mirroring qualitative patterns at Nouragues, i.e. the slightly larger heights predicted at Petit Plateau compared to Grand Plateau (cf. Figure 4, lefthand column). Similarly, biomass estimates of 355.4 and 438.7 t ha^{-1} for Grand Plateau and Petit Plateau, or a combined 400.8 t ha^{-1}, matched very closely previous estimates of 404.6 t ha^{-1} (Labriere *et al.*, 2018). The combined 302.2 t ha^{-1} at Rabi, on the other hand, was lower by ~12 t ha^{-1} than the AfriSAR reference estimates (314.6 t ha^{-1}), in keeping with slightly lower height allometry estimates (Figure 4, upper middle panel).

When validating the extrapolation step, i.e. the inference of tree diameters and positions from ALS-data, against ground-informed reconstructions, the model reproduced stem

diameter distributions very well, both at the plots where packing densities and allometries have been inferred (local validation from "above",Figure 5, blue bars) and when the model was transferred from one plot to another (crossvalidation, Figure 5, green bars). In terms of biomass inference, the model exhibited good overall predictive quality as well. We obtained an RMSE of 53.2 t ha^{-1} at the one-hectare scale, and 87.3 t ha^{-1} at 0.25 ha scale, with a mean bias error (MBE) of -16.1 t ha^{-1}, or roughly 5% of the total biomass (cf. Figure 6, and also Table 3). Between-plot heterogeneity did not greatly affect the quality of ALS-based inference, as can be seen from model transfer, where we obtained nearly identical predictions, with highly similar bias (MBE of -17.1 t ha^{-1}) and RMSEs (53.73 t ha^{1} at the one-hectare scale, and 87.59 at the 0.25 ha scale).

At landscape level, the model predicted overall stem densities of 443.4 trees ha^{-1} and aboveground biomass of 299.8 t ha^{-1} at Nouragues, and 418.8 trees ha^{-1} and 251.8 t ha^{-1} at Rabi, both lower than at the calibration plots, due to heterogeneity in vegetation features (cf. Figure 7, top panels, for maps at the 0.25 ha scale). Posterior uncertainty, as quantified by the coefficient of variation, was highest at vegetation edges and low biomass areas, and generally higher at Rabi (median CV of ~0.24) than at Nouragues (~0.16, cf. also Figure 7, middle panels). At both Nouragues and Rabi, biomass reached similar extreme values, of over 1100 t ha^{-1} at the 0.25-ha scale, but with differently shaped distributions. Compared to regression-based estimates, our approach resulted in a much larger variation in aboveground biomass density (Figure 7, lower panels).

4. Discussion

In this study we have developed a method, called the Canopy Constructor, to assimilate forest inventory information and airborne lidar scanning data, so as to (i) infer the

allometric relationships among tree dimensiosn, (ii) recreate 3D canopies from simple assumptions about tree geometry, and (iii) provide large-scale inference of tree dimensions for further uses, such as aboveground biomass mapping. The initial tests with the Canopy Constructor presented here are promising.

First, we were able to successfully infer the allometric parameters. Such parameters are difficult to obtain in the field (Sullivan *et al.*, 2018), but are crucial for biomass estimates (Feldpausch *et al.*, 2012). Conceptually, our method differs from individual tree crown segmentation using ALS datasets (Dalponte & Coomes 2016), where tree crown dimensions are individually isolated from an ALS point cloud and then used for the construction of tree allometries. Here we rather assumed that tree shapes and distributions emerge from the space-filling rules of the canopy, and that the empirical forest can be gradually approximated by a virtual reconstruction. By assuming predefined functional forms of the allometric equations, we were able to reduce the inference of tree dimensions to a parameter optimization problem, which we solved through an approximate Bayesian computing (ABC) approach. We applied our method to two study sites where tree allometric data were available and where we were therefore able to validate the approach.

For the sake of simplicity and to reduce computational efforts, we assumed fixed functional forms for the allometries; we also assumed fixed parameters for crown shape and crown depth. An extension of the ABC routine to include these parameters should be considered in future applications. The description of crown shape could also be made more complex if needed. Furthermore, we did not impose any restrictions on crown overlap, which is at odds with observations (Goudie *et al.*, 2009). Also, crown

overlapping may be responsible for the uncertainty in crown radius allometries, since crowns can be more easily hidden within other tree crowns.

Second, our approach showed good predictive quality and spatial transferability for forest structure at the study plots. Crucially, the metrics we used to quantify and extrapolate forest structure, i.e. stem diameter distributions, and canopy packing densities, appeared to remain relatively constant within sites. This was even true for the Nouragues plots that are characterized by very different canopy height models. While some of our assumptions might not hold across all latitudes and biomes (Spriggs *et al.*, 2019), we successfully decomposed forest structure into horizontal aspects that were highly variable across the landscape, for example different disturbance regimes encapsulated by larger and more frequent gaps, and vertical aspects of forest structure that were more homogeneous, most likely due to physiological constraints or evolutionary history (Niklas, 1994). This suggests that our approach, once calibrated with local forest inventories, can generally be extrapolated across entire landscapes.

As seen from the striking similarity between ALS-only predictions at plots used for allometric inference and at crossvalidation plots (Figure 6, left panels vs. right panels), the major source of uncertainty and bias is not forest structure heterogeneity, but the loss of ground-based information on exact tree locations and stem diameter distributions. While there is a natural limit on predictive accuracy of canopy height models, this suggests that there is still room for improvement regarding the prediction of stem diameter distributions from CHMs. Especially in the largest diameter classes we found that the algorithm was the least accurate due to low sampling intensity. For these largest diameter classes, individual tree crown segmentation methods could be used

during a pre-processing step to isolate individuals that can be clearly delineated. This is particularly promising, because large canopy trees could also provide natural constraints on possible allometric laws and the most appropriate crown shapes. We will return to this possible improvement in the future. Furthermore, so far, we have used mainly summary statistics of static quantities at the 1-ha and 0.25-ha scale, but other applications could include the inference of more complex variables, such as species-specific allometric equations with intraspecific variation, or the inclusion of repeated ALS acquisitions to yield estimates of tree mortality and growth for top-of-canopy layers.

Third, a major insight of this analysis were the benefits of the individual-based approach for biomass mapping. While overall, the Canopy Constructor showed similar patterns of high- and low-biomass areas as in regression-based methods of biomass inference, we here were able to detect a much larger variation. Our new method better accounts for natural variation in biomass, since it does not suffer from regression towards the mean. Indeed, in statistical models such as the $AGB = f(MCH)$ it is assumed that variation in biomass is a random error term and thus locally high or low biomass values are replaced by a mean value. Our method thus has the potential to be more widely applicable across biomes and environmental conditions than area-based regression-models that need to be locally calibrated (Coomes *et al.*, 2017). In the future, our approach would provide an efficient model-based approach to assimilate forest inventories and ALS surveys into high-resolution aboveground biomass maps that could be used in the validation of remote-sensing biomass missions (Le Toan *et al.*, 2011; Chave *et al.*, 2019; Duncanson *et al.*, 2019).

A general challenge concerns the typical ground plot sizes found across the tropics. For this study, we have selected two sites with inventories in large continuous forest areas and subdivided these inventories into simple rectangular plots. This allowed us to largely ignore edge effects, i.e. trees reaching into the considered plot from outside, and trees within the plot having parts of their crown outside of the plot (Mascaro *et al.*, 2011). In the tropics, field inventories are typically of smaller area (0.25 ha or 1 ha) with a large number of plots to better sample environmental variation (Chave *et al.*, 2019). It would be useful to improve the Canopy Constructor to allow the inclusion of a collection of small, non-contiguous plots – and with non-rectangular shapes. This would imply that we would have to account for edge effects, but would help make the Canopy Constructor operational across data sets, environmental gradients and biomes.

To conclude, the Canopy Constructor is a model-based approach which simulates assemblies of individual trees. Further data sources such as topography or other remote-sensing products could be integrated. In particular, due to its high spatial resolution, modelling every individual tree down to 1cm *dbh*, the Canopy Constructor can be used to initialize individual-based models of forests, such as the TROLL model (Maréchaux & Chave, 2017), or other similar models. Once these leaf-on canopies are constructed, the Canopy Constructor can provide forest structure descriptions to process-based vegetation models, infer initial canopy constructions for individual-based models of forest dynamics or mockups for radiative transfer models.

Acknowledgements

This work has benefitted from "Investissement d'Avenir" grants managed by Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01; ANAEE-France: ANR-11-INBS-0001, TULIP: ANR-10-LABX-0041), from CNES, and from ESA (CCI-Biomass). This work was granted access to the HPC resources of CALMIP supercomputing centre in Toulouse under allocation p17012. In Gabon, CENAREST granted permission an provided personnel to establish and study the Rabi plot. Shell Gabon, the society Compagnie des Bois du Gabon provided financial and logistical support. The Center for Tropical Forest Science provided technical advice and funding for fieldwork and data analysis. This is contribution #193 of the Gabon Biodiversity Program

Literature

Anderson-Teixeira K., S. Davis, A. Bennett, E. Gonzalez-Akre, H. Muller-Landau, J. Wright, K. Abu, A. Zambrano, A. Alonso, J. Baltzer, Y. Basset, et al. 2015. CTFS-ForestGEO: A worldwide network monitoring forests in an era of global change. Global Change Biology **21**:528-549.

Bohn FJ, Huth A. 2017. The importance of forest structure to biodiversity-productivity relationships. *Royal Society Open Science* **4**: 160521.

Cano IM, Muller-Landau HC, Joseph Wright S, Bohlman SA, Pacala SW. 2019. Tropical tree height and crown allometries for the Barro Colorado Nature Monument, Panama: A comparison of alternative hierarchical models incorporating interspecific variation in relation to life history traits. *Biogeosciences* **16**: 847–862.

Chalom A, Mandai C, Prado P. 2013. Sensitivity analyses: a brief tutorial with Rpackage pse. *Cran.Rstudio.Com*: 1–14.

Chave J, Davies SJ, Phillips OL, Lewis SL, Sist P, Schepaschenko D, Armston J, Baker TR, Coomes D, Disney M, *et al.* 2019. Ground Data are Essential for Biomass Remote Sensing Missions. *Surveys in Geophysics* **40**: 863–880.

Chave J, Olivier J, Bongers F, Châtelet P, Forget PM, Van Der Meer P, Norden N, Riéra B, Charles-Dominique P. 2008. Above-ground biomass and productivity in a rain forest of eastern South America. *Journal of Tropical Ecology* **24**: 355–366.

Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WBC, Duque A, Eid T, Fearnside PM, Goodman RC, *et al.* 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology* **20**: 3177–3190.

Chazdon RL, Broadbent EN, Rozendaal DMA, Bongers F, Zambrano AMA, Aide TM,

Balvanera P, Becknell JM, Boukili V, Brancalion PHS, *et al.* 2016. Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. *Science Advances* **2**: e1501639.

Coomes DA, Asner GP, Lewis SL, Dalponte M, Phillips OL, Qie L, Nilus R, Phua M-H, Banin LF, Burslem DFRP, *et al.* 2017. Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asian forests from airborne laser scanning data. *Remote Sensing of Environment* **194**: 77–88.

Csilléry K, Blum MGB, Gaggiotti OE, François O. 2010. Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology and Evolution* **25**: 410–418.

Dowle M, Srinivasan A. 2018. data.table: Extension of 'data.frame'. R package version 1.11.2. https://cran.r-project.org/package=data.table.

Duncanson L, Armston J, Disney M, Avitabile V, Barbier N, Calders K, Carter S, Chave J, Herold M, Crowther TW, *et al.* 2019. The Importance of Consistent Global Forest Aboveground Biomass Product Validation. *Surveys in Geophysics* **40**: 979–999.

Fatoyinbo L, Pinto N, Hofton M, Simard M, Blair B, Saatchi S, Lou Y, Dubayah R, Hensley S, Armston J, *et al.* 2017. The 2016 NASA AfriSAR campaign: Airborne SAR and Lidar measurements of tropical forest structure and biomass in support of future satellite missions. In: International Geoscience and Remote Sensing Symposium (IGARSS). 4286–4287.

Feldpausch TR, Lloyd J, Lewis SL, Brienen RJW, Gloor M, Monteagudo Mendoza A, Lopez-Gonzalez G, Banin L, Abu Salim K, Affum-Baffoe K, *et al.* 2012. Tree height integrated into pantropical forest biomass estimates. *Biogeosciences* **9**: 3381–3403.

Ferraz A, Saatchi S, Mallet C, Meyer V. 2016. Lidar detection of individual tree size in tropical forests. *Remote Sensing of Environment* **183**: 318–333.

Fischer R, Knapp N, Bohn F, Shugart HH, Huth A. 2019a. The Relevance of Forest

Structure for Biomass and Productivity in Temperate Forests: New Perspectives for Remote Sensing. *Surveys in Geophysics*: 1–26.

Fischer FJ, Maréchaux I, Chave J. 2019b. Improving plant allometry by fusing forest models and remote sensing. *New Phytologist* **223**: 1159–1165.

Garnier S. 2018. viridis: Default Color Maps from 'matplotlib'. *R package version 0.5.1*.

Grassi G, House J, Dentener F, Federici S, Den Elzen M, Penman J. 2017. The key role of forests in meeting climate targets requires science for credible mitigation. *Nature Climate Change* **7**: 220–226.

Hartig F, Calabrese JM, Reineking B, Wiegand T, Huth A. 2011. Statistical inference for stochastic simulation models - theory and application. *Ecology Letters* **14**: 816–827.

Hartig F, Dislich C, Wiegand T, Huth A. 2014. Technical note: Approximate bayesian parameterization of a process-based tropical forest model. *Biogeosciences* **11**: 1261–1272.

Hijmans RJ. 2016. raster: Geographic Data Analysis and Modeling. R package version 2.5-8. *R package*.

Hill RW, Holland PW. 1977. Two Robust Alternatives to Least-Squares Regression. *Journal of the American Statistical Association* **72**: 828–833.

Inman HF, Bradley EL. 1989. The Overlapping Coefficient as a Measure of Agreement Between Probability Distributions and Point Estimation of the Overlap of two Normal Densities. *Communications in Statistics - Theory and Methods* **18**: 3851–3874.

Isenburg M. 2018. LAStools - efficient LiDAR processing software. : obtained from http://rapidlasso.com/LAStools.

Jucker T, Caspersen J, Chave J, Antin C, Barbier N, Bongers F, Dalponte M, van Ewijk KY, Forrester DI, Haeni M, *et al.* 2017. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Global Change Biology* **23**: 177–

190.

Khosravipour A, Skidmore AK, Isenburg M, Wang T, Hussin YA. **2014**. Generating Pit-free Canopy Height Models from Airborne Lidar. *Photogrammetric Engineering & Remote Sensing* **80**: 863–872.

Labriere N, Tao S, Chave J, Scipal K, Toan T Le, Abernethy K, Alonso A, Barbier N, Bissiengou P, Casal T, *et al.* 2018. In Situ Reference Datasets from the TropiSAR and AfriSAR Campaigns in Support of Upcoming Spaceborne Biomass Missions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **11**: 3617–3627.

Maréchaux I, Chave J. **2017**. An individual-based forest model to jointly simulate carbon and tree diversity in Amazonia: description and applications. *Ecological Monographs* **87**: 632–664.

Mascaro J, Detto M, Asner GP, Muller-Landau HC. **2011**. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sensing of Environment* **115**: 3770–3774.

Molto Q, Hérault B, Boreux JJ, Daullet M, Rousteau A, Rossi V. **2014**. Predicting tree heights for biomass estimates in tropical forests -A test from French Guiana. *Biogeosciences* **11**: 3121–3130.

Morsdorf F, Meier E, Kötz B, Itten KI, Dobbertin M, Allgöwer B. **2004**. LIDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. In: Remote Sensing of Environment. 353–362.

Muller-Landau HC, Condit RS, Harms KE, Marks CO, Thomas SC, Bunyavejchewin S, Chuyong G, Co L, Davies S, Foster R, *et al.* 2006. Comparing tropical forest tree size distributions with the predictions of metabolic ecology and equilibrium models. *Ecology Letters* **9**: 589–602.

Niklas KJ. **1994**. Morphological evolution through complex domains of fitness.

Proceedings of the National Academy of Sciences **91**: 6772–6779.

Oldeman RAA. **1974**. *The architecture of the forest of French Guiana. L'architecture de la foret Guyanaise*.

R Development Core Team. **2019**. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*.

Réjou-Méchain M, Tymen B, Blanc L, Fauset S, Feldpausch TR, Monteagudo A, Phillips OL, Richard H, Chave J. **2015**. Using repeated small-footprint LiDAR acquisitions to infer spatial and temporal variations of a high-biomass Neotropical forest. *Remote Sensing of Environment* **169**: 93–101.

Riaño D, Valladares F, Condés S, Chuvieco E. **2004**. Estimation of leaf area index and covered ground from airborne laser scanner (Lidar) in two contrasting forests. *Agricultural and Forest Meteorology* **124**: 269–275.

Rosette JAB, North PRJ, Suárez JC. **2008**. Vegetation height estimates for a mixed temperate forest using satellite laser altimetry. In: International Journal of Remote Sensing. 1475–1493.

Shugart HH, Saatchi S, Hall FG. **2010**. Importance of structure and its measurement in quantifying function of forest ecosystems. *Journal of Geophysical Research: Biogeosciences* **115**: G2.

Spriggs RA, Vanderwel MC, Jones TA, Caspersen JP, Coomes DA. **2019**. A critique of general allometry-inspired models for estimating forest carbon density from airborne LiDAR. *PLoS ONE* **14**: e0215238.

Sullivan MJP, Lewis SL, Hubau W, Qie L, Baker TR, Banin LF, Chave J, Cuni-Sanchez A, Feldpausch TR, Lopez-Gonzalez G, *et al.* 2018. Field methods for sampling tree height for tropical forest biomass estimation. *Methods in Ecology and Evolution* **9**: 1179–1189.

Taubert F, Jahn MW, Dobner H-J, Wiegand T, Huth A. **2015**. The structure of tropical forests and sphere packings. *Proceedings of the National Academy of Sciences* **112**: 15125–15129.

Le Toan T, Quegan S, Davidson MWJ, Balzter H, Paillou P, Papathanassiou K, Plummer S, Rocca F, Saatchi S, Shugart H, *et al.* 2011. The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. *Remote Sensing of Environment* **115**: 2850–2860.

Wenger SJ, Olden JD. **2012**. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods in Ecology and Evolution* **3**: 260–267.

Wickham H. **2011**. ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*.

	a_h	h_{max}	σ_h	a_{CR}	b_{CR}	σ_{CR}
GP	0.41	56.88	0.39	2.19	0.55	0.24
PP	0.39	58.38	0.23	2.29	0.56	0.22
Rabi10	0.32	47.52	0.37	2.2	0.53	0.25
Rabi15	0.28	43.67	0.35	2.23	0.55	0.27

Table 1: Inferred parameters: Mean of posterior distributions for the tested allometric parameters at the two sites. Plots are Grand Plateau (GP) and Petit Plateau (PP) at Nouragues, as well as the 10-ha and 15-ha rectangular strips at Rabi (Rabi10 and Rabi15, respectively).

	a_h	h_{max}	σ_h	a_{CR}	b_{CR}	σ_{CR}
a_h	1					
h_{max}	0.95	1				
σ_h	-0.31	-0.52	1			
a_{CR}	0.42	0.19	0.27	1		
b_{CR}	0.15	0.06	0.02	0.65	1	
σ_{CR}	0.13	0.22	-0.30	-0.34	-0.30	1

Table 2: Correlation structure of the allometric parameters after inference: This shows the correlation between the allometric parameters after inference for the Petit Plateau plot.

	GP	***PP***	***Rabi*10**	***Rabi15***
$trees_{ground}$	478.3	512.1	452.5	461.7
$trees_{above}$	468.1	499.5	445.2	454.1
$trees_{crossval}$	471.5	477.1	457.2	441.7
AGB_{ground}	355.4	438.7	311.9	295.8
AGB_{above}	337.8	419.4	299.9	280.7
$AGB_{crossval}$	324.8	430.0	301.1	276.3

Table 3: Prediction of above-ground biomass across the four (sub-)plots. Shown are overall mean AGB values in t ha^{-1}, calculated by summing over all trees > 10cm in *dbh*. Trees below that threshold are not included to ensure comparability with previous estimates. Given are values for ground-based inference ($trees_{ground}$, AGB_{ground}), for inference from ALS data alone ($trees_{above}$, AGB_{above}), and for crossvalidation with the respective other plot from the same site ($trees_{crossval}$, $AGB_{crossval}$).

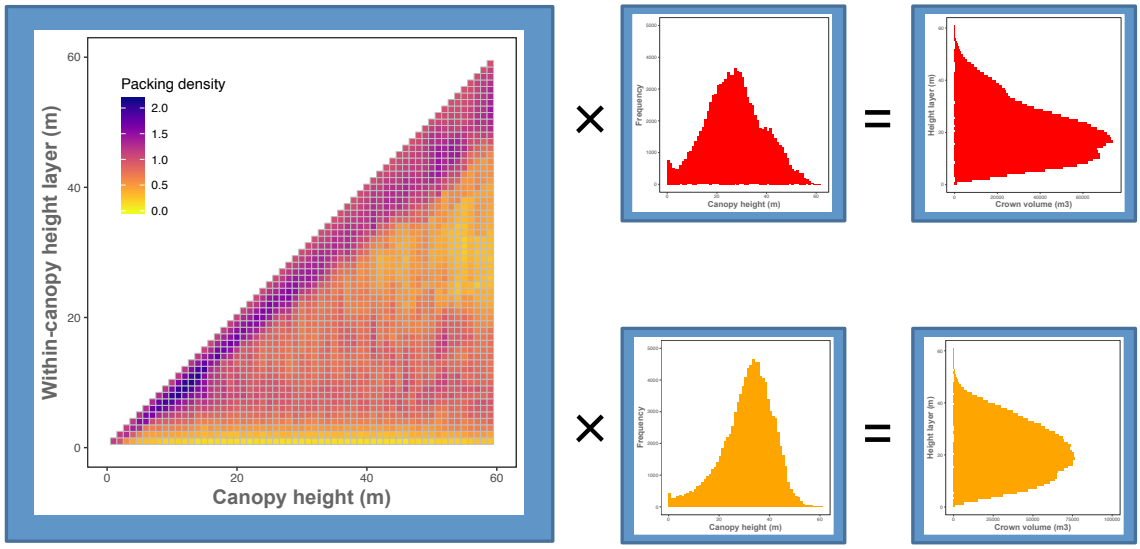


Figure 1: Converting packing densities into crown volume distributions: The figure displays the crown packing density matrix for a sample run at Nouragues (Petit Plateau) as well as two 10ha CHM distributions (middle panels, derived from the Grand Plateau and Petit Plateau plots, respectively), and the result of a multiplication with the packing density matrix. Packing densities are given as unit crown volume per unit canopy volume (m^3 m^{-3}). When a particular canopy height (i.e. column) occurs less than 1000 times within the sampled canopy, values represent an average over neighboring canopy heights, obtained from rescaling of all heights involved to percentage of top-of-canopy height and then converting back to absolute height values. For both plots, we obtain a distribution of total crown volume (in cubic meter) per height layer (in m), which can then be filled up with random draws from a stem diameter distribution and allometric predictions.

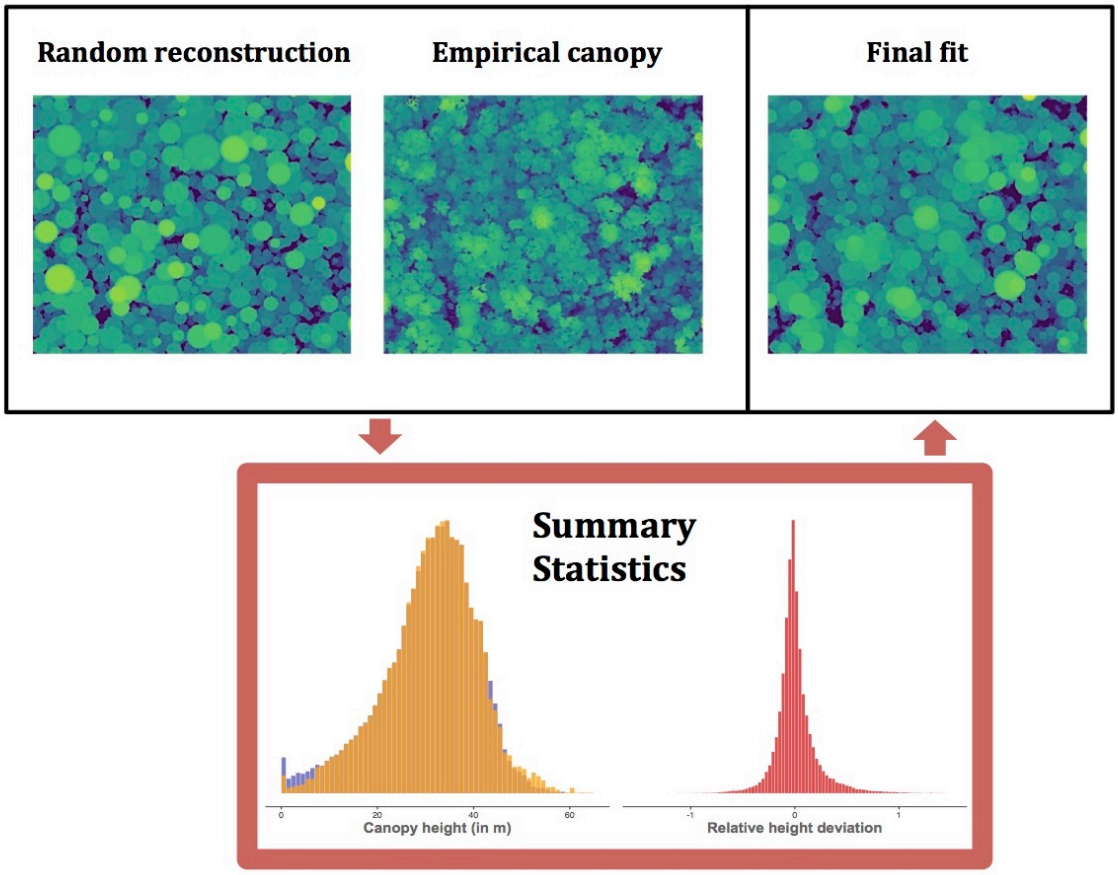


Figure 2. Results of canopy fitting at the Petit Plateau plot at Nouragues Field Station, based on both field inventory and canopy height model: Upper panels: The left image shows the initial canopy height model (CHM) for Petit Plateau where tree dimensions are randomly drawn from site-specific allometries and an empirical diameter distribution ("initial/random fit"). The middle image shows the corresponding empirically derived canopy, and the righthand image shows the final reconstruction ("spatial fit") of the Canopy Constructor. Similarities between the initial fit and the empirical canopy height model, particularly in gappy areas, are due to known tree diameters and positions. Divergences are due to random variation around allometric means. The Canopy Constructor improves upon this by swapping deviations between trees or drawing alternative tree dimensions from the allometric models until a better fit is created. Lower panels: Shown are the two summary statistics used to create a better fit. The left panel shows the canopy height distribution of the Petit Plateau field plot, overlaid by a fitted canopy height distribution (in orange). The dissimilarity D between the two normalized distributions (where D = 1 – the overlapping area) is used to quantify goodness of fit. The righthand panel shows the distribution of per-pixel deviations, here rescaled with the mean empirical canopy height. In the fitting algorithm, we convert the devations to absolute values and take their mean (mean absolute error, *MAE*).

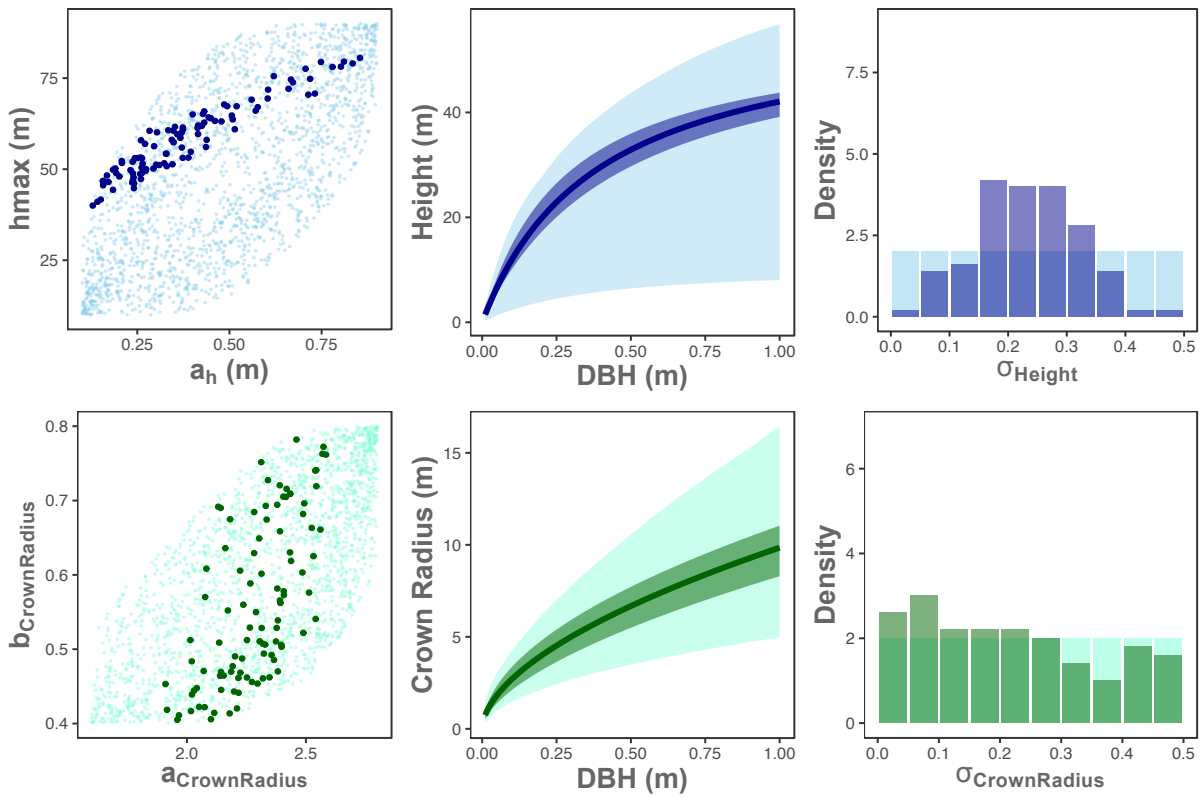


Figure 3: Parameter inference at Petit Plateau, Nouragues, with a 1% cutoff. Results of the ABC rejection scheme, for the Petit Plateau plot retaining only the best 1% of simulations. The top row shows the parameter space for the Michaelis Menten parameters of the height allometry (left panel, prior in light blue, with 2,000 out of 10,000 reconstructions displayed, posterior, i.e. the best 100 reconstructions, in dark blue), the prior and posterior allometries (middle panel) and the prior and posterior distribution of the variance term (right panel). The bottom row shows the same information for the crown radius intercept and slope (a_{CR} and b_{CR}), i.e. the parameter space, the corresponding prior and posterior distributions (prior in light green, posterior in dark green) and the variance term. The best simulation (mean parameter combination) is given as dark blue/green line in the middle panels, the uncertainty interval is derived from the 75% highest density intervals of the joint posterior distribution, with best-fit allometric equations smoothing the upper and lower limits of the interval.

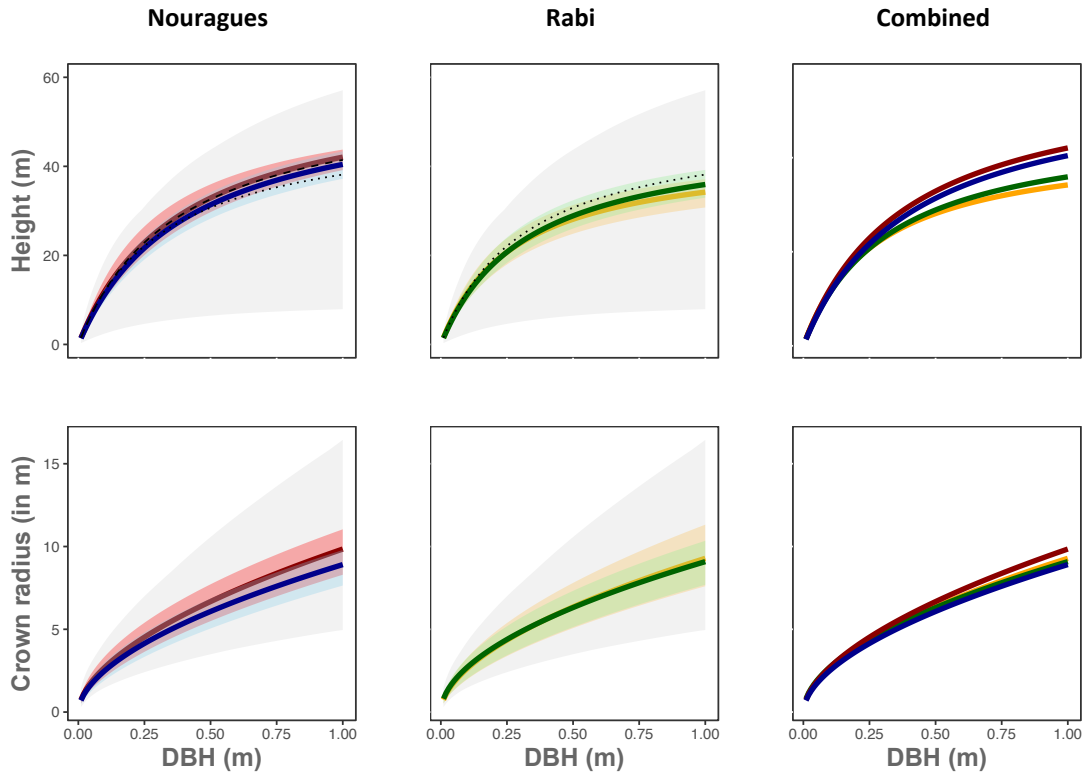


Figure 4: Inferred allometries at Nouragues and Rabi. The panels show height allometries (top row) and crown allometries (bottom row), as inferred by the Canopy Constructor, for Nouragues (lefthand side), Rabi (middle panels) and both sites combined (righthand side). The grey background indicates the prior range. Mean and 75% highest density intervals are given for each plot separately, i.e. for Grand Plateau (blue) and Petit Plateau (red) at Nouragues, and for the 10ha (green) and 15ha (orange) plot at Rabi. As comparison, we have plotted ground-inferred height allometries for both Grand Plateau (dotted) and Petit Plateau (dashed) in the top panels, as well as a single ground-inferred allometry at Rabi. The grey background indicates the prior range.

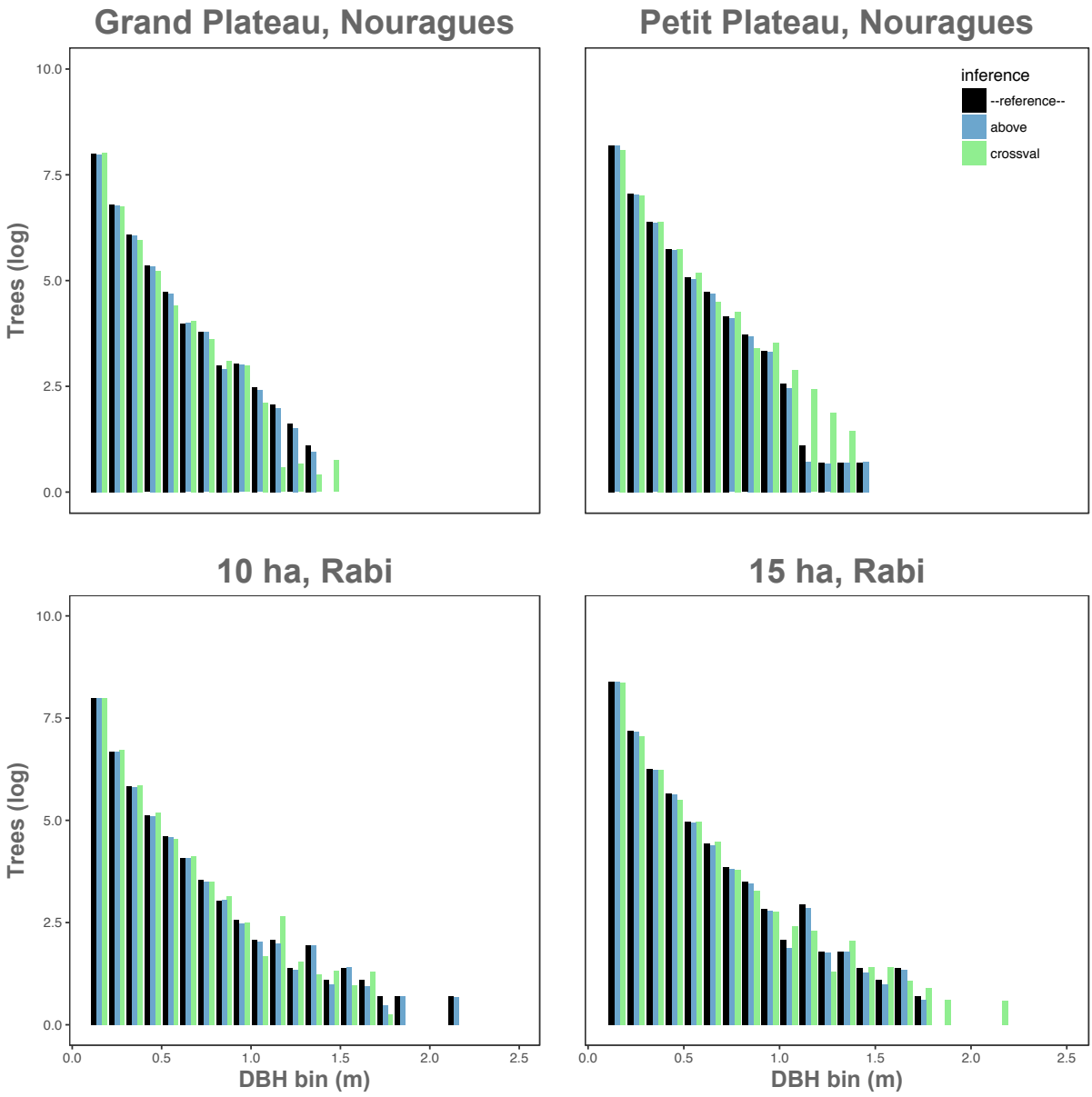


Figure 5: Stem diameter distributions (trees > 10cm in diameter), as inferred "from above": Shown are log-transformed stem diameter distributions starting at a diameter at breast height of 10cm, compared across inference procedures. The black bars are reference values (from field inventories), the blue bar represents ALS-based inference of the stem diameter distribution at the local plot where allometries and packing densities were inferred (local validation from "above"), and the green bar represents inference when the model is transferred between plots at the same site (crossvalidation). Crossvalidation is visible in how patterns from the ground distribution in the lefthand panels translate into patterns of the "crossval" distribution on the righthand site, and vice versa.

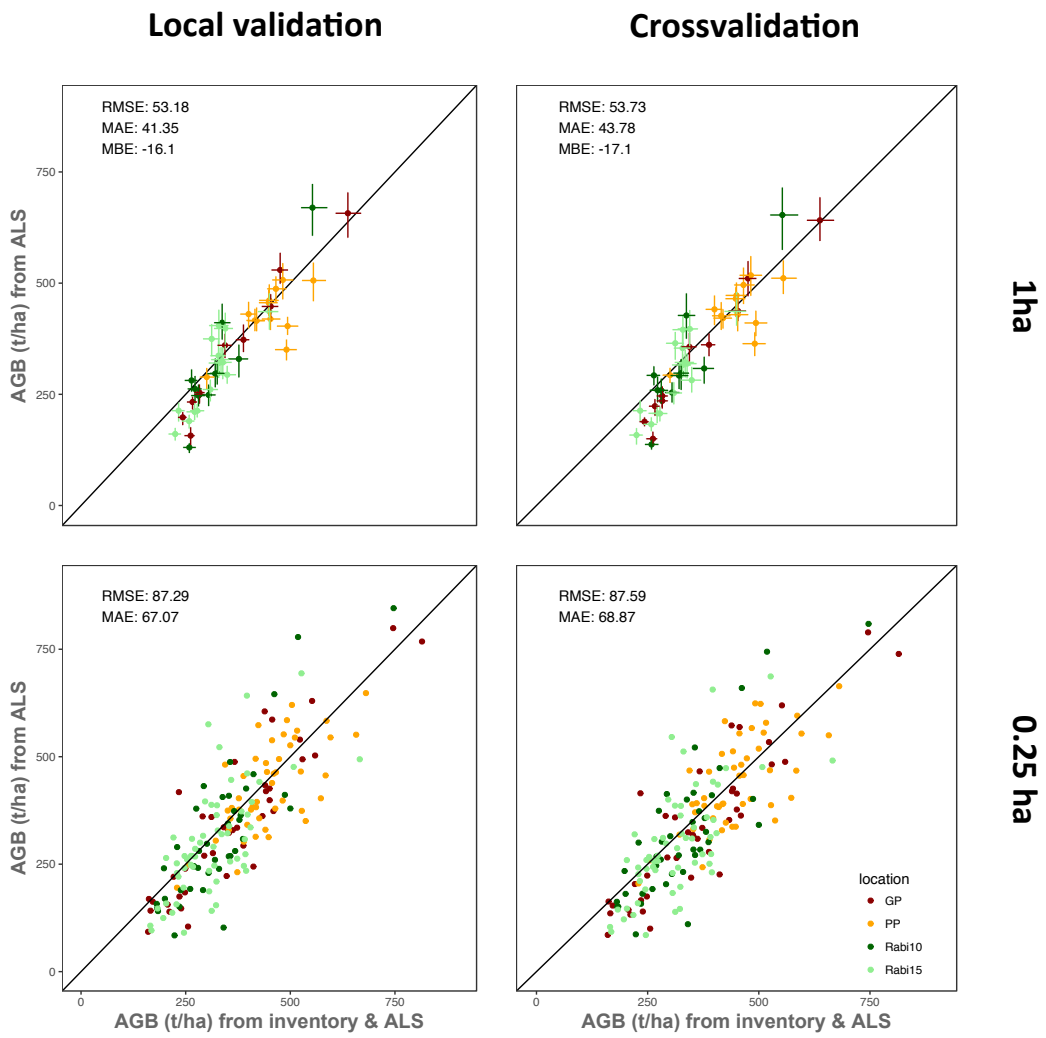


Figure 6: Aboveground biomass predictions, as inferred "from above": Shown are the predictions of aboveground biomass (median of 100 posterior simulations, given in t ha^{-1}) from ALS-derived CHMs only, both at the hectare scale (top panels) and quarterhectare scale (bottom panels). The left column describes the results when the Canopy Constructor inversion is applied to the local plot where allometries and packing densities were calibrated, the right column the results from crossvalidation. We see the two Nouragues plots in red and orange, and the two Rabi plots in dark and light green. RMSE (root mean squared error), MAE (mean absolute error) and MBE (mean bias of the error) are given in the top-left corner of the panels. MBE does not change between hectare and quarterhectare scales and is thus only given in the top panels. For visualization purposes, we only plot error bars at the hectare scale, representing the interquartile ranges of estimates from 100 posterior simulations.

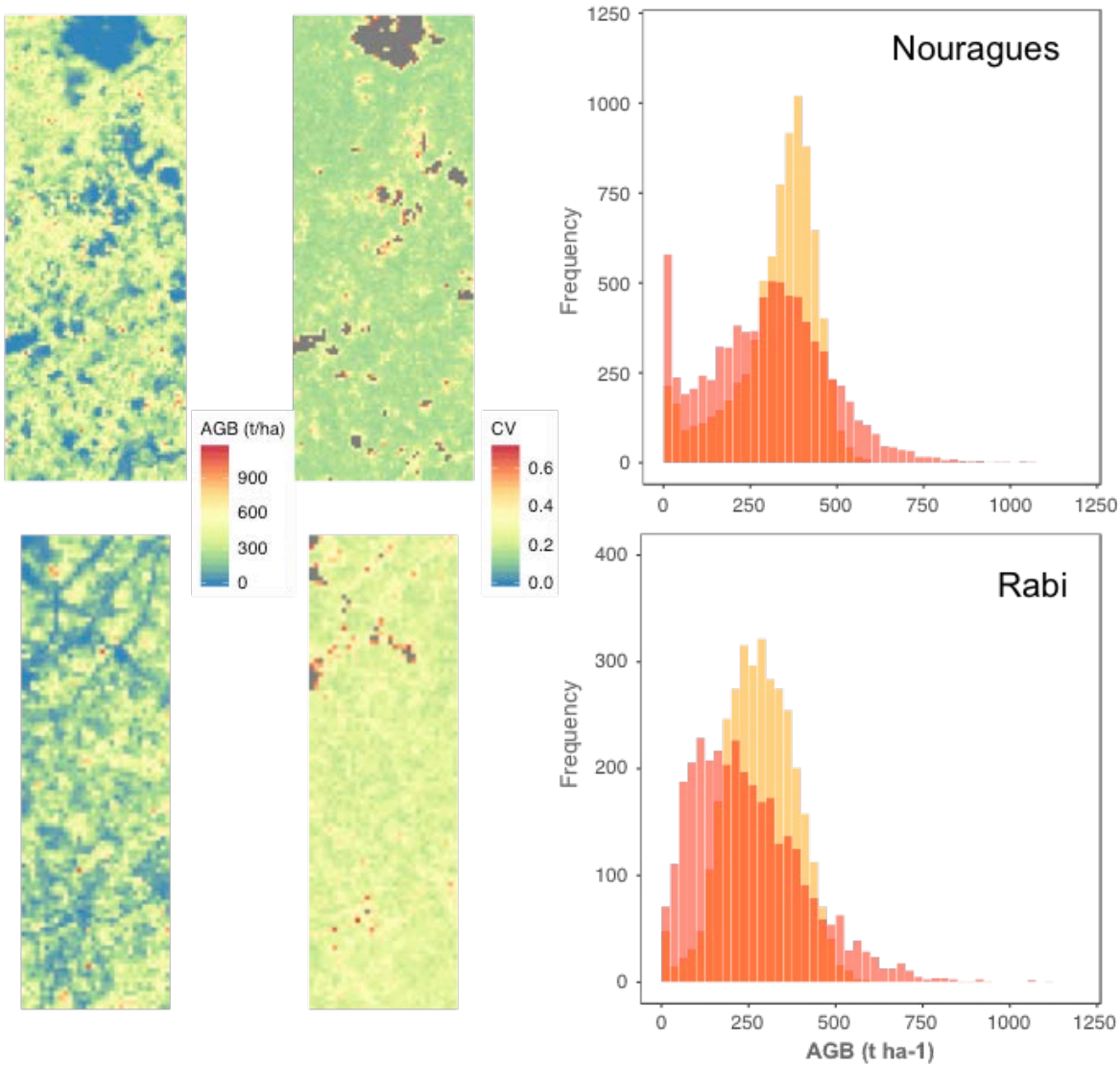


Figure 7: Aboveground biomass predictions for ALS campaign at Nouragues (2,016 ha) and Rabi (832 ha). Maps show the mean predicted aboveground biomass values (t ha^{-1}) across the ALS covered areas (left panels, Nouragues upper panel, Rabi lower panel), the respective coefficient of variation across 100 simulations (middle panels, dimensionless), and the overall distributions of aboveground biomass (right panels, red distributions, in t ha^{-1}). Also given are the reference estimates (in yellow), as derived from a regression-based approach (Labrière et al. 2018). Clearly evident is the shrinkage towards the mean in the regression-based approach, as opposed to much stronger variation in the Canopy Constructor approach. Please note that the geographic extent of the maps has been rescaled for visualization purposes.

Chapter 3: Calibrating the short-term dynamics of the TROLL individual-based model in an old-growth tropical forest

(Target Journal: Ecological Modelling)

Chapter 3 extends the new method developed in Chapter 2 to create not only geometric tree representations, but also leaf-filled canopies, and then couples it with the individual-based forest model TROLL. Based on a new, improved version of TROLL, TROLL v.2.5, and an adapted version of the Canopy Constructor we reconstructed a physiologically realistic old-growth forest on a plot in French Guiana, simulated its dynamics over a 400 year period, and calibrated parameters related to mortality and carbon allocation. We studied the stability of the inferred ecosystem structure in time, inferred mortality rates, qualitatively assessed trunk diameter growth and canopy dynamics with regard to empirically derived values and compared overall stand development to forest regrowth from bare ground.

Calibrating the short-term dynamics of the TROLL individual-based model in an old-growth tropical forest

Fabian Jörg Fischer[1,*], Isabelle Maréchaux[2], Grégoire Vincent[3] and Jérôme Chave[1]

[1] Laboratoire Évolution et Diversité Biologique, UMR 5174 (CNRS/IRD/UPS), 31062 Toulouse Cedex 9, France

[2] AMAP, INRA, IRD, CIRAD, CNRS, University of Montpellier, F-34000Montpellier, France

[3] Botanique et Modélisation de l'Architecture des Plantes et des Végétations (AMAP), UMR 5120 (CIRAD/CNRS/INRA/IRD/UM2), 34398 Montpellier Cedex 5, France

* Correspondence: fabian.j.d.fischer@gmx.de

Keywords: Forest structure, Individual-based modelling, Airborne Lidar, Approximate Bayesian Computation, Vital Rates

Target Journal: Ecological Modelling

1. Introduction

Tropical forests are a crucial component of the global carbon cycle (Pan *et al.*, 2011; Malhi, 2012), but are changing rapidly under the influence of climate change and anthropogenic pressures such as deforestation (Hansen *et al.*, 2013). As a result, their future as well as the feedback on the global carbon cycle remain highly uncertain.

Vegetation models are essential to explore the impact of global change on the terrestrial carbon cycle by simulating matter and energy exchange from the vegetation and the atmosphere. Such large-scale models have long been parameterized for the tropical biome (Cramer et al. 2000, Cox et al. 2000). However, a detailed account of carbon allocation into woody components of forests requires to simulate demographic processes (Fisher *et al.*, 2018). This challenge has, for example, been addressed by cohort-based plant demography models that describe coarse-grained dynamics on average forest patches (Moorcroft *et al.*, 2001; Medvigy *et al.*, 2009).

In recent years, there has, however, been a renewed interest for the development of individual-based models (IBMs), because these models represent forest dynamics tree by tree and thus allow for a more explicit simulation of demographic processes (Pacala *et al.*, 1996). In particular, since every individual is explicitly accounted for, field data such as stem diameter or functional trait measurements are more easily assimilated for calibration or validation purposes. At the same time, innovative strategies are required to deal with high model complexity, spatial extent and computational costs (Grimm & Railsback, 2012; Fischer *et al.*, 2019). They are particularly important for the inference of mortality processes that are often coarsely represented, but have a strong impact on simulated dynamics (Bugmann *et al.*, 2019).

Previously, we have shown how forest plot inventories and airborne lidar can be used together to reconstruct detailed 3D-forest scenes (the Canopy Constructor

algorithm; Fischer *et al.*, in preparation). The approach provides a key connection between field data and fine resolution forest dynamic models, and a move towards generating locally parameterized tropical forest simulators. Here, we describe how these static reconstructions of tropical rainforests can be translated into dynamically evolving forests, using the spatially-explicit and individual-based forest simulator model TROLL (Maréchaux & Chave, 2017).

To do so, we first revisit some crucial features of the TROLL model and improve model stability and transferability (sensu Wenger & Olden, 2012). Notable improvements of this version of the TROLL model include the consideration of intra-specific variability in traits, a more detailed description of within-crown variation of photosynthetic assimilation, and the development of a method to account for the plastic response to light gradients.

We then extend the Canopy Constructor algorithm to provide biologically viable representations of an old-growth forest in French Guiana and initialise TROLL directly from these best-fit reconstructions. We vary its dynamic parameters and calibrate them by imposing the condition that the initial forest structure be largely preserved. Finally, we assess the simulated old-growh forest and compare our predictions both to simulations from bare ground and to empirical data from repeated inventories and airborne lidar acquisitions.

We ask: (1) How well can we reconstruct viable, leaf-filled canopies and how does the inclusion of leaf physiology impact on allometric inference? (2) Is the parameter space well-constrained? How do estimates of mortality and growth compare to empirical data? (3) How stable is the old-growth forest ecosystem and how does it compare to forest regrowth from bare ground?

2. Methods

2.1 The TROLL model v.2.5

Our study uses the spatially explicit and individual-based forest growth simulator TROLL (Chave, 1999, 2001; Maréchaux & Chave, 2017). Here we provide the context of the model, and improvements included in the latest version 2.5, compared with previously released version 2.3.1 (Maréchaux & Chave, 2017).

The TROLL model simulates individual trees ≥ 1 cm in trunk diameter within a voxel space of 1 m^3 spatial resolution. Tree crowns occupy the voxel space, assimilate carbon and shade other plants. Each individual tree is assigned a species, its mean plant functional traits (leaf nutrients, leaf mass per area, wood density), and allometric relationships. When maturity is reached, each tree has the potential to disperse propagules in the neighborhood and to recruit seedlings into the community.

Every month, tree growth and mortality are calculated. Photosynthesis or gross primary production is based on the FvCB-model (Farquhar *et al.*, 1980), with the main parameters ($Jmax$, $Vcmax$, C_i) estimated from species traits (Domingues *et al.*, 2010; Medlyn *et al.*, 2011), and dark respiration from an equation for broadleaf trees (Atkin *et al.*, 2015). To calculate the vertical change in the environmental variables – photosynthetic photon flux density (*PPFD*), temperature (*T*), and water vapour pressure deficit (*VPD*) –, the Beer-Lambert extinction law is applied (Maréchaux & Chave, 2017). Once carbon losses from leaf and stem respiration are deducted, primary productivity is translated into biomass gain, and allocated to various plant organs according to preset ratios. Tree mortality is simulated through a baseline mortality that declines linearly with wood density (Kraft *et al.*, 2010), carbon starvation when respiration exceeds photosynthesis for prolonged periods, and treefall – typically simulated with a simple height threshold, although more complex formulations are available (Chave, 1999).

In version 2.5 of TROLL, traits are allowed to vary among individuals within species. For every trait i, we assume a lognormal distribution, i.e. a multiplicative factor e^{ε_i}, where $\varepsilon_i \sim N(0, \sigma_i)$. For wood specific gravity, we assume normal variation around the mean, or an additive error term $\varepsilon_{wsg} \sim N(0, \sigma_{wsg})$, as observed in empirical data sets (Kattge *et al.*, 2011). Traits are then drawn from a multivariate Gaussian distribution, preserving intraspecific covariance (cf. covariance matrix in Supplementary Material 1 and a previous study, Fischer *et al.*, in preparation).

In TROLL v.2.5, the light flux is not computed at the top of a each layer in the 3D voxel space as in the previous version. Rather, the absorbed photons per m^2 of leaf area are calculated. In a layer of thickness v at canopy height z, the absorbed photosynthetic photon flux density (*PPFD*, in μmol m^{-2} s^{-1}) is: $PPFD(z)_{abs} = [PPFD(z+v) - PPFD(z)]/dens(v)$, where *dens* is the average leaf area density (m^2/m^2) in layer v. We define $PPFD(z) = \exp(-k^*_{PAR} \times LAI(z))$ where *LAI* is the leaf area index at height z (in m^2/m^2) and k^*_{PAR} the Beer-Lambert extinction factor multiplied by ε that represents leaf absorptance: $k^*_{PAR} = \varepsilon \times k$. With this definition, the quantum yield parameter θ does not need to be converted to absorptance-based values anymore (Medlyn *et al.*, 2002). In the new equation, an increase in transmitted radiation at low k is balanced out by a decrease in intercepted radiation from leaves. Temperature T and water vapour pressure deficit *VPD* are now averaged across each layer instead of being taken from the top of each layer (details on the equations cf. Supplementary Material S2).

TROLL v.2.5 also proposes a new concept to model leaf lifespan based on leaf turnover optimization (Kikuzawa, 1991). The Kikuzawa model assumes that leaf area is limited, construction cost is incurred once in a leaf's lifetime and photosynthesis declines with leaf age. Leaf lifespan is computed as the condition at which a leaf represents the optimal investment (Kikuzawa, 1991), a relationship validated against

empirical data (Kitajima *et al.*, 1997, 2002; Kikuzawa & Lechowicz, 2006). The major uncertainty in this model is *b*, the maximum potential lifespan. Recently, Xu et al. (2017) have compiled *b* values from the literature and found a good correlation with *Vcmax* values. Given that *Vcmax* is also strongly related to *LMA*, we derived a direct relationship between *LMA* and *b*. from values in Xu et al. (2017). Our formula is: $b = e^{1.071+1.211 \times \log(LMA)}$, and the modified Kikuzawa formula reads:

$$LLS = 1.0 + \frac{1}{30} \times \sqrt{\frac{2.0 \times CC \times LMA \times b}{NPPmax}}$$

Where NPP_{max} is the net primary production at full leaf expansion, *CC* the construction costs of the leaf (typically assumed to be 1.5 g/g to account for growth respiration, cf. Kikuzawa, 1991) and the factor 1/30 the conversion factor into monthly *LLS*. One month is added to account for leaf lifespan before full expansion.

Another improvement of TROLL v.2.5 is the representation of plasticity to light, a crucial feature of plant growth (Bloor & Grubb, 2004; Curt *et al.*, 2005; Niinemets, 2010) with considerable influences on ecosystem functioning (Williams *et al.*, 2017). We hypothesize that leaf allocation balances leaf litterfall, and that leaves are not allocated beyond their light compensation point *LCP*, i.e. the incident light at which carbon gains from leaf photosynthesis equals carbon losses through leaf respiration (Kitajima *et al.*, 2005). To calculate the *LCP*, we inverse the FvCB-model (Farquhar *et al.*, 1980) and calculate the maximum amount of leaves trees can support under mean climatic conditions at the study site. Excess carbon – i.e. carbon that cannot be allocated to leaves – is stored in a pool of non-structural carbohydrates (set to 10% of total carbon, half of which is accessible, cf. Martínez-Vilalta *et al.*, 2016) and is allocated to stem growth only once the storage pool is saturated. At periods when net primary productivity is negative, the storage pool is used for maintenance.

Finally, in TROLL v.2.5, a crown shape parameter γ describes the ratio between radius at the top and the bottom of the crown. This was motivated by the development of the Canopy Constructor and the need to create more realistic crown surfaces and volumes (cf. Fischer *et al.*, in preparation). We also define a new gap fraction, i.e. a fraction of each tree crown's pixels that will not be or only partially filled with leaves. This simulates physiological constraints on the trees' ability to fill up the canopy space and has important consequences on ecological dynamics, as the light penetrating through tree crown gaps is crucial to recruitment and regeneration (Way & Pearcy, 2012). Because empirical canopies display high variation in tree leaf area index, sometimes well below their maximal capacity, we set this gap fraction to 0.4, with some inter-individual variation due to variation in crown radius (Further details cf. Supplementary Material S4).

2.2 Initial conditions and calibration of the TROLL model

We used the Canopy Constructor algorithm to create an initial forest state and to infer the allometric relationships that underlie the TROLL model (Fischer *et al.*, in preparation). We adapted the Canopy Constructor to jointly optimize the spatial tree configurations that reflect forest structure and to ensure tree viability. To determine biological viability, we kept track of the overall carbon balance of trees > 10cm in diameter and optimized the 3D assembly of tree so that only a minimal fraction of trees experienced a negative carbon balance (cf. Supplementary Material S3). As before, we used a simple Approximate Bayesian Computation approach and selected the best reconstructions to infer parameters (Fischer *et al.*, in preparation). Including tree viability as an additional constraint was predicted to constrain the relationship between trunk diameter and crown radius because trees must have a large enough crown to

ensure positive carbon balance, but small enough to not interfere to much with the crowns of other trees, particularly in the understorey.

The Canopy Constructor uses ALS data only through a canopy height model (CHM). Lidar scans contain, however, a substantial amount of information on plant densities (Vincent *et al.*, 2017). Provided that the effects of the previous CHM fitting are separated out, this information can serve as a source of validation for the reconstructed leaf-on canopies. Since CHMs essentially describe the ratio of voxels within the canopy to voxels outside of the canopy at a particular height level (i.e. the number of voxels above which is vegetation vs. those above which is no vegetation), we separated this information out by considering only the properties of within-canopy voxels. Specifically, we compared the number of leaf-filled voxels inside the canopy as well as the mean transmittance of these voxels. To account for effects introduced by the ALS acquisition procedure – i.e. lower sampling densities and energy fraction in the understorey, resulting in higher uncertainties and potential bias –, the comparison was based on a virtual lidar scan (details for the simplified lidar scan cf. S5).

To translate the static inference of forest structure into dynamic forest growth, we picked the ten best Canopy Constructor reconstructions and initialised the TROLL model from them, relying on the allometric relationships of each reconstruction. We parameterized all species that had been taxonomically identified within the study plot, including morphospecies (622 species overall). Unidentified trees (ca. 5% in our empirical dataset) were assigned to one of the identified species proportionally to the species' relative abundances. To infer the most likely trait values at species and genus level, we used a local trait collection (Baraloto *et al.*, 2010) and hierarchical Bayesian modelling with the package *rstan* (Stan Development Team, 2019), assuming lognormal distributions for leaf-level traits and normal distributions for wood density. Where no

data was available for a particular species, we assumed traits to be phylogenetically conserved, and used genus means. Where no genus-level data was available, we assigned community-weighted plot means.

To infer stand dynamics, we varied three parameters relating to the trees' vital rates and to whom TROLL's dynamics are particularly sensitive (Maréchaux & Chave, 2017). These include the baseline mortality rate as well as allocation rules for newly produced biomass to either canopy or trunk biomass. Since we hypothesized treefall to have a strong impact on dynamics, we further included the treefall threshold parameter for calibration. For each of the best 10 reconstructions, we ran 100 simulations of old-growth dynamics for 400 years, with random combinations of the parameters within realistic prior ranges (cf. Table 1).

Out of the resulting 1000 simulations, we again selected the best simulations based on a simple rejection scheme. We assumed that a realistic rendering of dynamics would require the old-growth forest structure to be largely preserved over 400 years of growth. To quantify stability, we recorded stand characteristics every ten years, including stand-level aboveground biomass, trunk diameter distributions and canopy height distributions. At each timestep t, we assessed the overlap between the distributions and their initial shape and calculated dissimilarity metrics $D_{dbh}(t)$ and $D_{chm}(t)$, with dissimilarity between distributions defined as in our previous study (Fischer *et al.*, in preparation). Furthermore, dissimilarity between aboveground biomass $AGB(t)$ at timestep t and initial AGB_i was calculated as $D_{agb}(t) = 1 - AGB(t)/AGB_i$ when $AGB(t) < AGB_i$ and $D_{agb}(t) = 1 - AGB_i/AGB(t)$ when $AGB(t) > AGB_i$. All dissimilarity metrics thus ranged between 0 and 1 and could be merged to form a combined dissimilarity index $D_{comb}(t) = \sqrt{D_{dbh}(t)^2 + D_{chm}(t)^2 + D_{agb}(t)^2}$. To assess the stability of the inferred forest, we calculated the coefficient of variation of D_{comb} over

the entire 400 year span (40 samples in total, excluding the initial state) and selected the 10% of simulations that had the lowest coefficient of variation (i.e. did not vary strongly in their similarity to the initial state). We hypothesized that this would exclude both highly unstable simulations and simulations that remained stable in the long term, but were very dissimilar from the initial configuration, since they would show large deviations in the initial decades until reaching equilibrium. To further test this assumption, we compared our results to a second calibration, based on mean dissimilarity over 400 years.

Based on the best simulations, we then assessed overall tree mortality and qualitatively compared the dynamics simulated by the TROLL model to changes in canopy height from successive ALS campaigns and trunk diameter growth from successive forest inventories. Since measurement error in trunk diameter considerably alters the distribution of empirical diameter growth rates (resulting in spurious decreases in diameter growth) and thus renders comparisons between empirical and virtual data difficult, we applied a measurement error model to the true diameter values in TROLL (Chave *et al.*, 2004; Réjou-Méchain *et al.*, 2017).

Finally, we assessed the overall stability of the simulated old-growth dynamics by comparing initial and final trunk diameter and canopy height distributions and reran the 10 best simulations over 400 years from bare ground to compare old-growth dynamics and forest regeneration.

2.3 Data

All data used in this study were obtained at the Nouragues Ecological Research Station in French Guiana (4.06°N, 52.68°W), a site with a lowland tropical rainforest, rainfall of ca. 2900 mm per year and with a 2-mo dry season from September to November and a

shorter one in March. Forest inventories have been conducted since the early 1990s (Chave *et al.*, 2008b; Labrière *et al.*, 2018), with several ALS surveys conducted since 2007 (Réjou-Méchain *et al.*, 2015). We here use two successive ground inventories at the 12ha "Petit Plateau" plot, one carried out in november 2012 and a partial reinventory in october 2015. All trees were tagged, mapped, and their dbh was measured when above 10 cm. Trunk dbh was measured 130 cm above ground, or 50 cm above buttresses or deformities. They were also identified to the species level for about 95% of the stems. We also used data from two corresponding ALS campaigns, one with a Riegl laser rangefinder (LMS-Q560) earlier in March of 2012 (Réjou-Méchain *et al.*, 2015) at an average pulse density of ~12 per m^2 (based on density of last returns) and an overall point density of ~18 per m^2 (all returns), and one in October 2015, using a Riegl laser rangefinder (LMS-Q780) at an average pulse density of 23 per m^2 and an overall point density of 37 per m^2. For both lidar campaigns, we derived spike-free canopy height models, based on the LAStools software (Isenburg, 2018) at m^2 resolution. For validation of model outputs, we further relied on local field and fluxtower data (Chave *et al.*, 2008; Aguilos *et al.*, 2018).

Statistical analysis and visual rendering were conducted in the R software (R Development Core Team, 2019), including the packages *data.table* (Dowle & Srinivasan, 2018), ggplot2 (Wickham, 2011), *viridis* (Garnier, 2018), *hdi* (Meredith & Kruschke, 2018), *rstan* (Stan Development Team, 2019) and *coda* (Plummer *et al.*, 2006).

3. Results

In general, the modified Canopy Constructor was highly successful in creating viable canopies. Gross primary productivity of the best leaf-on canopies was 46.7 MgC ha^{-1} yr^{-1} [45.0-47.8 MgC ha^{-1} yr^{-1}]. Mean net primary productivity amounted to 17.0 MgC ha^{-1} yr^{-1} [16.2-17.7 MgC ha^{-1} yr^{-1}]. Furthermore, a comparison between empirical and simulated leaf densities and transmittances shows that the inferred forest reconstructions represented well observations, both quantitatively and qualitatively (Figure 1).

As hypothesized, the viability constraint in allometric inference affected the inferred crown allometry parameters, compared to previous results from the Canopy Constructor algorithm (Fischer *et al.*, in preparation). Crown diameters were generally inferred to be smaller and showed less dispersion than when inferred merely from geometric principles. Variance around mean crown diameter, in particular, showed a clear peak, compared to an uninformative posterior for the purely geometric fitting. Height allometries, on the other hand, were nearly identical, with 75% highest density intervals of both methods overlapping almost completely over the whole range (cf. Supplementary Material, Figure S1).

The dynamic constraint imposed by the dissimilarity index was efficient for mortality rate and the treefall threshold *vC*, with strong correlation between both parameters, but had no clear effect on the allocation parameters (Figure 2, cf. also Supplementary Material, Figure S2 for correlation matrices). Inferred tree mortality was 0.015 yr^{-1} [0.011-0.02 yr^{-1}]. Yearly treefall estimates were more variable, with a mean of 0.005 yr^{-1} [0.001-0.011 yr^{-1}], responsible for 10-50% of annual mortality. These patterns were very stable, irrespective of whether the model was constrained by the coefficient

of variation of the dissimilarity index or its mean (cf. Supplementary Material, Figures S2 and S3).

A qualitative comparison between the best model calibration and empirical data showed that changes in the canopy height model due to treefall were consistent with obervations (Figure 3) and diameter growth patterns were also similar to observed ones (Figure 4).

Finally, overall canopy structure was well-preserved in the old-growth forest, as can be seen from comparisons between initial and final trunk diameter distributions and canopy height distributions (Figure 5). Stand metrics such as above-ground biomass or tree numbers were stable, suggesting that the fundamental dynamics were rendered accurately (Figure 6). Forest regeneration from bare ground quickly reached similar tree numbers, but converged much more slowly towards a stable aboveground biomass configuration, sometimes not even reaching it within the 400 years of simulation.

4. Discussion

4.1 Model calibration and validation

Here we have shown how to infer and predict ecosystem functioning from a combination of successive forest inventories, airborne lidar data and individual-based modelling. A step-wise inference procedure, constraining forest structure first, then the dynamics of forest growth, enabled us to simulate the dynamics of an old-growth forest in French Guiana. This represents a significant advance in the predictive modelling of vegetation dynamics.

We have demonstrated that initial forest reconstructions by the Canopy Constructor rendered adequately the leaf distribution of canopies, with gross primary productivity

and its ratio to net primary productivity (~1/3) close to empirical values (Malhi *et al.*, 2011; Aguilos *et al.*, 2018). The simulated short-term forest dynamics mirrored empirial dynamics quantitatively, with estimates of mortality rates of 1-2% per year and treefall rates of around 0.5% close to empirical estimates (Chave *et al.*, 2008) and good qualitative agreement with treefall patterns and diameter growth. Importantly, we found that the simulations were stable over the whole 400 year period, with good preservation of the underlying stand metrics, indicating that the model does not need an extended spin-up phase. Growth parameters were not well-constrained, but there was substantial covariation between mortality parameters, indicating that mortality rates can be efficiently narrowed down by the model.

4.2 Towards large-scale predictions of tropical forest dynamics

The 'divide-and-conquer' approach presented here is a powerful tool for prediction of ecosystem dynamics. We applied the method at a site where both repeated ground inventories and ALS-derived CHMs are available, i.e. a "supersite" (Fischer *et al.*, 2011; Chave *et al.*, 2019). However, it can be adapted to less ideal circumstances. In cases where only ground data is available, the Canopy Constructor could, for example, still use some basic assumptions about allometric scaling to reconstruct a viable old-growth forest for TROLL initialisation. Conversely, a CHM-model together with limited information about diameter distributions and leaf densities (LAI, crown packing densities, GPP) would also allow to infer a 3D-reconstruction, even in the absence of ground data.

The fact that TROLL seems to be able to translate such reconstructions almost seemlessly into forest growth dynamics – i.e. with little to no spin-up time –, not only

suggests that the dynamics are rendered well, but also opens up the possibility of creating a standardized approach where remotely sensed forest canopies are converted into individual tree assemblies, which in turn can be projected into the future. Here, we have used successive inventories and ALS-observed canopies only for broad, qualitative validation. In the future they could, however, be incorporated in a more stringent way, either for quantitative assessments of transferability (Wenger & Olden, 2012) or for the calibration procedure itself and thus further inform the modelled dynamics.

Most importantly, however, we have kept environmental conditions stable in this study. TROLL is a mechanistic model where community dynamics and diameter growth emerge directly from the underlying physiology of individual trees and their interactions in space. This makes it well-suited for predictive purposes under changing conditions (Railsback, 2001). With increasing airborne lidar coverage, new remote sensing missions and wider availability of trait data (Kattge *et al.*, 2011), fine-scale individual-based predictions of environmental change could thus be extended to large geographic areas and either be incorporated or serve as validations for global vegetation models.

4.3 Improvements for the future.

This study offers a number of opportunities for improving this approach further. First, the initial forest reconstructions from the Canopy Constructor currently assume that all tree species follow broadly the same allometry (with individual variation around the mean). This translates into similar growth- and maturation trajectories. However, we could also assume that tree species' allometries align with their ecological roles, as suggested empirically (King, 1996; Bohlman & O'Brien, 2006; Thomas *et al.*, 2015).

Much of the species-specific information contained in our forest reconstructions, particularly on the relation between diameter and height, is thus not yet transformed into ecological knowledge and could be used for future TROLL modelling efforts – as long as care is taken to avoid overparameterization.

Furthermore, TROLL v.2.5 proposes a more flexible representation of tree crown geometry, including crown heterogeneity and plasticity to its environment. Such crown plasticity is a well-documented feature of natural forests (Purves *et al.*, 2007; Jucker *et al.*, 2015). In the future, a more natural concept of modelling tree crowns would be a fully plastic tree growth into empty space, that is to say a "light-foraging" model. Trees would be allowed to dynamically expand toward voxel cells – if connected to the trunk and if not too costly energetically. Observations at the Nouragues field station, for example, suggests rapid lateral growth in tree crowns that TROLL does not capture adequately. If modular growth was to be combined with the assumption of crown shyness (Franco, 1986), this new approach would also likely result in a reduction of the computational burden of TROLL.

Finally, when simulating the dynamics of an old-growth forest directly from the initial condition, we kept track only of stand-scale patterns and did not analyze the underlying community patterns or patchy aggregation of trees in space. This touches directly on important ecological dynamics such as seed dispersal (Price *et al.*, 2001) and disturbances (Bugmann *et al.*, 2019). Some of these issues, such as the mortality of canopy trees could, for example, be investigated by applying the Canopy Constructor to two successive lidar scans, directly estimating treefall rates and assess how canopy dynamics differ compared to the TROLL modules.

All in all, the study at hand has laid the foundations to turn TROLL into a model that can be calibrated with minimal data requirements, project the dynamics of old-growth forests into the future and thus contribute to a predictive ecology necessary in the face of a changing environment.

Acknowledgements

This work has benefitted from "Investissement d'Avenir" grants managed by Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01; ANAEE-France: ANR-11-INBS-0001, TULIP: ANR-10-LABX-0041), from CNES, and from ESA (CCI-Biomass). This work was granted access to the HPC resources of CALMIP supercomputing centre in Toulouse under allocation p17012.

Literature

Aguilos M, Hérault B, Burban B, Wagner F, Bonal D. 2018. What drives long-term variations in carbon flux and balance in a tropical rainforest in French Guiana? *Agricultural and Forest Meteorology* **253–254**: 114–123.

Atkin OK, Bloomfield KJ, Reich PB, Tjoelker MG, Asner GP, Bonal D, Bönisch G, Bradford MG, Cernusak LA, Cosio EG, *et al.* 2015. Global variability in leaf respiration in relation to climate, plant functional types and leaf traits. *New Phytologist* **206**: 614–636.

Baraloto C, Paine CET, Poorter L, Beauchene J, Bonal D, Domenach AM, Hérault B, Patiño S, Roggy JC, Chave J. 2010. Decoupled leaf and stem economics in rain forest trees. *Ecology Letters* **13**: 1338–1347.

Bloor JMG, Grubb PJ. 2004. Morphological plasticity of shade-tolerant tropical rainforest tree seedlings exposed to light changes. *Functional Ecology* **18**: 337–348.

Bohlman S, O'Brien S. 2006. Allometry, adult stature and regeneration requirement of 65 tree species on Barro Colorado Island, Panama. *Journal of Tropical Ecology* **22**: 123–136.

Bugmann H, Seidl R, Hartig F, Bohn F, Brůna J, Cailleret M, François L, Heinke J, Henrot AJ, Hickler T, *et al.* 2019. Tree mortality submodels drive simulated long-term forest dynamics: assessing 15 models from the stand to global scale. *Ecosphere* **10**: 2616.

Chave J. 1999. Study of structural, successional and spatial patterns in tropical rain forests using TROLL, a spatially explicit forest model. *Ecological Modelling* **124**: 233–254.

Chave J. 2001. Spatial Patterns and Persistence of Woody Plant Species in Ecological Communities. *The American Naturalist* **157**: 51–65.

Chave J, Condit R, Aguilar S, Hernandez A, Lao S, Perez R. 2004. Error propagation

and sealing for tropical forest biomass estimates. In: Philosophical Transactions of the Royal Society B: Biological Sciences. 409–420.

Chave J, Davies SJ, Phillips OL, Lewis SL, Sist P, Schepaschenko D, Armston J, Baker TR, Coomes D, Disney M, *et al.* 2019. Ground Data are Essential for Biomass Remote Sensing Missions. *Surveys in Geophysics* **40**: 863–880.

Chave J, Olivier J, Bongers F, Châtelet P, Forget PM, Van Der Meer P, Norden N, Riéra B, Charles-Dominique P. **2008**. Above-ground biomass and productivity in a rain forest of eastern South America. *Journal of Tropical Ecology* **24**: 355–366.

Curt T, Coll L, Prévosto B, Balandier P, Kunstler G. **2005**. Plasticity in growth, biomass allocation and root morphology in beech seedlings as induced by irradiance and herbaceous competition. *Annals of Forest Science* **62**: 51–60.

Domingues TF, Meir P, Feldpausch TR, Saiz G, Veenendaal EM, Schrodt F, Bird M, Djagbletey G, Hien F, Compaore H, *et al.* 2010. Co-limitation of photosynthetic capacity by nitrogen and phosphorus in West Africa woodlands. *Plant, Cell and Environment* **33**: 959–980.

Dowle M, Srinivasan A. **2018**. data.table: Extension of 'data.frame'. R package version 1.11.2. https://cran.r-project.org/package=data.table.

Farquhar GD, von Caemmerer S, Berry JA. **1980**. A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. *Planta* **149**: 78–90.

Fischer R, Aas W, de Vries W, Clarke N, Cudlin P, Leaver D, Lundin L, Matteucci G, Matyssek R, Mikkelsen TN, *et al.* 2011. Towards a transnational system of supersites for forest monitoring and research in Europe - An overview on present state and future recommendations. *IForest* **4**: 167–171.

Fischer FJ, Maréchaux I, Chave J. **2019**. Improving plant allometry by fusing forest models and remote sensing. *New Phytologist* **223**: 1159–1165.

Fisher RA, Koven CD, Anderegg WRL, Christoffersen BO, Dietze MC, Farrior CE, Holm JA, Hurtt GC, Knox RG, Lawrence PJ, *et al.* 2018. Vegetation demographics in Earth System Models: A review of progress and priorities. *Global Change Biology* **24**: 35–54.

Franco M. 1986. The influence of neighbours on the growth of modular organisms with an example from trees. *Philosophical Transactions - Royal Society of London, Series B* **313**: 209–225.

Garnier S. 2018. viridis: Default Color Maps from 'matplotlib'. *R package version 0.5.1*.

Grimm V, Railsback SF. 2012. Pattern-oriented modelling: a 'multi-scope' for predictive systems ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**: 298–310.

Hansen MC, Potapov P V., Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman S V., Goetz SJ, Loveland TR, *et al.* 2013. High-resolution global maps of 21st-century forest cover change. *Science* **342**: 850–853.

Isenburg M. 2018. LAStools - efficient LiDAR processing software. : obtained from http://rapidlasso.com/LAStools.

Jucker T, Bouriaud O, Coomes DA. 2015. Crown plasticity enables trees to optimize canopy packing in mixed-species forests. *Functional Ecology* **29**: 1078–1086.

Kattge J, Díaz S, Lavorel S, Prentice IC, Leadley P, Bönisch G, Garnier E, Westoby M, Reich PB, Wright IJ, *et al.* 2011. TRY - a global database of plant traits. *Global Change Biology* **17**: 2905–2935.

Kikuzawa K. 1991. A cost-benefit analysis of leaf habit and leaf longevity of trees and their geographical pattern. *American Naturalist* **138**: 1250–1263.

Kikuzawa K, Lechowicz MJ. 2006. Toward synthesis of relationships among leaf longevity, instantaneous photosynthetic rate, lifetime leaf carbon gain, and the gross

primary production of forests. *American Naturalist* **168**: 373–383.

King DA. **1996**. Allometry and life history of tropical trees. *Journal of Tropical Ecology* **12**: 25–44.

Kitajima K, Mulkey SS, Samaniego M, Wright SJ. **2002**. Decline of photosynthetic capacity with leaf age and position in two tropical pioneer tree species. *American Journal of Botany* **89**: 1925–1932.

Kitajima K, Mulkey SS, Wright SJ. **1997**. Decline of photosynthetic capacity with leaf age in relation to leaf longevities for five tropical canopy tree species. *American Journal of Botany* **84**: 702–708.

Kitajima K, Mulkey SS, Wright SJ. **2005**. Variation in crown light utilization characteristics among tropical canopy trees. *Annals of Botany* **95**: 535–547.

Kraft NJB, Metz MR, Condit RS, Chave J. **2010**. The relationship between wood density and mortality in a global tropical forest data set. *New Phytologist* **188**: 1124–1136.

Labriere N, Tao S, Chave J, Scipal K, Toan T Le, Abernethy K, Alonso A, Barbier N, Bissiengou P, Casal T, *et al.* 2018. In Situ Reference Datasets from the TropiSAR and AfriSAR Campaigns in Support of Upcoming Spaceborne Biomass Missions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **11**: 3617–3627.

Malhi Y. **2012**. The productivity, metabolism and carbon cycle of tropical forest vegetation. *Journal of Ecology* **100**: 65–75.

Malhi Y, Doughty C, Galbraith D. **2011**. The allocation of ecosystem net primary productivity in tropical forests. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**: 3225–3245.

Maréchaux I, Chave J. **2017**. An individual-based forest model to jointly simulate carbon and tree diversity in Amazonia: description and applications. *Ecological Monographs* **87**: 632–664.

Martínez-Vilalta J, Sala A, Asensio D, Galiano L, Hoch G, Palacio S, Piper FI, Lloret F. 2016. Dynamics of non-structural carbohydrates in terrestrial plants: A global synthesis. *Ecological Monographs* **86**: 495–516.

Medlyn BE, Dreyer E, Ellsworth D, Forstreuter M, Harley PC, Kirschbaum MUF, Le Roux X, Montpied P, Strassemeyer J, Walcroft A, *et al.* 2002. Temperature response of parameters of a biochemically based model of photosynthesis. II. A review of experimental data. *Plant, Cell and Environment* **25**: 1167–1179.

Medlyn BE, Duursma RA, Eamus D, Ellsworth DS, Prentice IC, Barton CVM, Crous KY, De Angelis P, Freeman M, Wingate L. 2011. Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology* **17**: 2134–2144.

Medvigy D, Wofsy SC, Munger JW, Hollinger DY, Moorcroft PR. 2009. Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2. *Journal of Geophysical Research: Biogeosciences* **114**: G01002.

Meredith M, Kruschke J. 2018. HDInterval: Highest (Posterior) Density Intervals.

Moorcroft PR, Hurtt GC, Pacala SW. 2001. A method for scaling vegetation dynamics: The ecosystem demography model (ED). *Ecological Monographs* **71**: 557–586.

Niinemets Ü. 2010. A review of light interception in plant stands from leaf to canopy in different plant functional types and in species with varying shade tolerance. *Ecological Research* **25**: 693–714.

Pacala SW, Canham CD, Saponara J, Silander JA, Kobe RK, Ribbens E. 1996. Forest models defined by field measurements: estimation, error analysis and dynamics. *Ecological Monographs* **66**: 1–43.

Pan Y, Birdsey RA, Fang J, Houghton R, Kauppi PE, Kurz WA, Phillips OL, Shvidenko A, Lewis SL, Canadell JG, *et al.* 2011. A Large and Persistent Carbon Sink in the World's Forests. *Science* **333**: 988 LP – 993.

Plummer M, Best N, Cowles K, Vines K. 2006. CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News* **6**: 7–11.

Price DT, Zimmermann NE, Van Der Meer PJ, Lexer MJ, Leadley P, Jorritsma ITM, Schaber J, Clark DF, Lasch P, Mcnulty S, *et al.* 2001. Regeneration in gap models: Priority issues for studying forest responses to climate change. *Climatic Change* **51**: 475–508.

Purves DW, Lichstein JW, Pacala SW. 2007. Crown plasticity and competition for canopy space: A new spatially implicit model parameterized for 250 North American tree species. *PLoS ONE* **2**: e870.

R Development Core Team. 2019. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*.

Railsback SF. 2001. Concepts from complex adaptive systems as a framework for individual-based modelling. *Ecological Modelling* **139**: 47–62.

Réjou-Méchain M, Tanguy A, Piponiot C, Chave J, Hérault B. 2017. Biomass: an R Package for Estimating Above-Ground Biomass and Its Uncertainty in Tropical Forests. *Methods in Ecology and Evolution* **8**: 1163–1167.

Réjou-Méchain M, Tymen B, Blanc L, Fauset S, Feldpausch TR, Monteagudo A, Phillips OL, Richard H, Chave J. 2015. Using repeated small-footprint LiDAR acquisitions to infer spatial and temporal variations of a high-biomass Neotropical forest. *Remote Sensing of Environment* **169**: 93–101.

Stan Development Team. 2019. RStan: the R interface to Stan. R package version 2.19.2. http://mc-stan.org/.

Thomas SC, Martin AR, Mycroft EE. 2015. Tropical trees in a wind-exposed island ecosystem: Height-diameter allometry and size at onset of maturity. *Journal of Ecology* **103**: 594–605.

Vincent G, Antin C, Laurans M, Heurtebize J, Durrieu S, Lavalley C, Dauzat J. **2017**. Mapping plant area index of tropical evergreen forest by airborne laser scanning. A cross-validation study using LAI2200 optical sensor. *Remote Sensing of Environment* **198**: 254–266.

Way DA, Pearcy RW. **2012**. Sunflecks in trees and forests: From photosynthetic physiology to global change biology. *Tree Physiology* **32**: 1066–1081.

Wenger SJ, Olden JD. **2012**. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods in Ecology and Evolution* **3**: 260–267.

Wickham H. **2011**. ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*.

Williams LJ, Paquette A, Cavender-Bares J, Messier C, Reich PB. **2017**. Spatial complementarity in tree crowns explains overyielding in species mixtures. *Nature Ecology and Evolution* **1**.

Xu X, Medvigy D, Joseph Wright S, Kitajima K, Wu J, Albert LP, Martins GA, Saleska SR, Pacala SW. **2017**. Variations of leaf longevity in tropical moist forests predicted by a trait-driven carbon optimality model. *Ecology Letters* **20**: 1097–1106.

	m0	*vC*	$f_{allocwood}$	$f_{alloccanopy}$
prior range	0.0 – 0.05	0.0 – 0.15	0.2 – 0.4	0.2 – 0.4
posterior range	0.0001 – 0.0231	0.0438 – 0.1498	0.2163 – 0.3948	0.2489 – 0.3977
best simulation	0.0161	0.0567	0.2209	0.3295

Table 1: Prior and posterior distributions: Indicated are the prior (1,000 simulations) and posterior ranges (best 10% simulations). Since the priors were chosen as uniform distributions, we show the minimum and maximum values. The marginal posterior distributions were not informative (hence we show only the ranges), but there was substantial covariation between mortality parameters (cf. Figure 2).

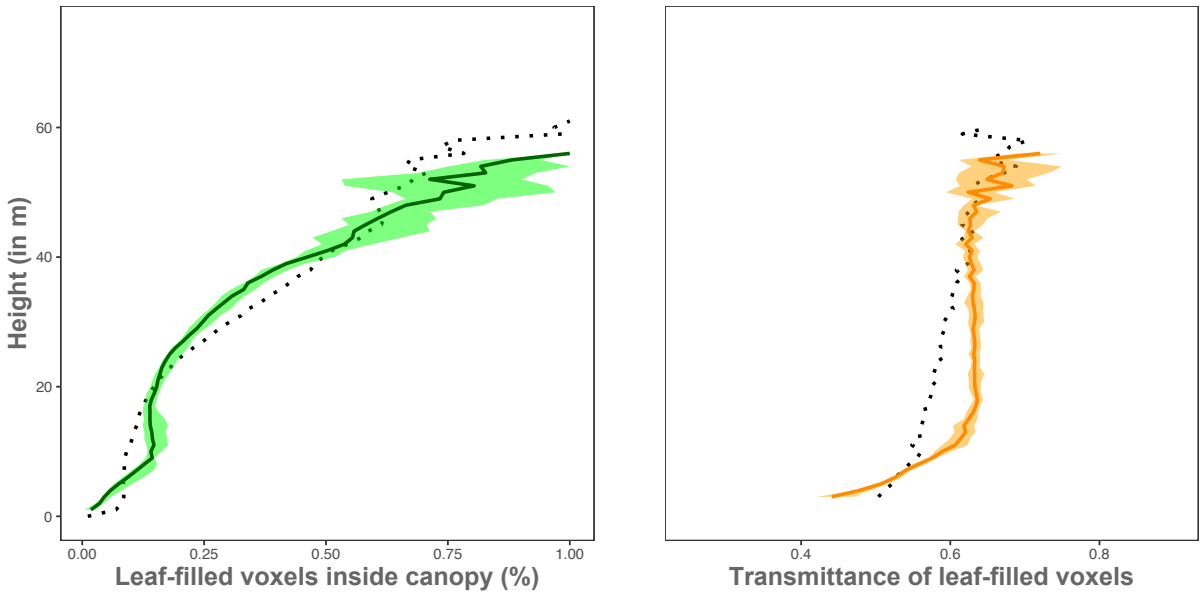


Figure 1: Comparison of modelled and observed leaf densities inside the forest canopy at Petit Plateau, Nouragues. This figure shows two metrics that quantify how well the physiological version of the Canopy Constructor represents leaf densities and crown packing at Nouragues field station. Plotted is the ratio of leaf-filled voxels inside the canopy to the total number of voxels inside the canopy (i.e. how densely space is filled by crowns), as well as the average transmittance of these leaf-filled voxels (i.e. how densely crowns are filled with leaves). The coloured lines represent mean estimates from a synthetic lidar run on the ten best forest canopy reconstructions, surrounded by 75% credibility intervals. The dotted lines are respective empirical estimates from a 2012 airborne lidar campaign. Both metrics show good correspondence, both qualitatively and quantitatively, with a tendency for the modelled forest to have higher densities in the upper layers than empirically observed and lower densities than observed in the lower layers.

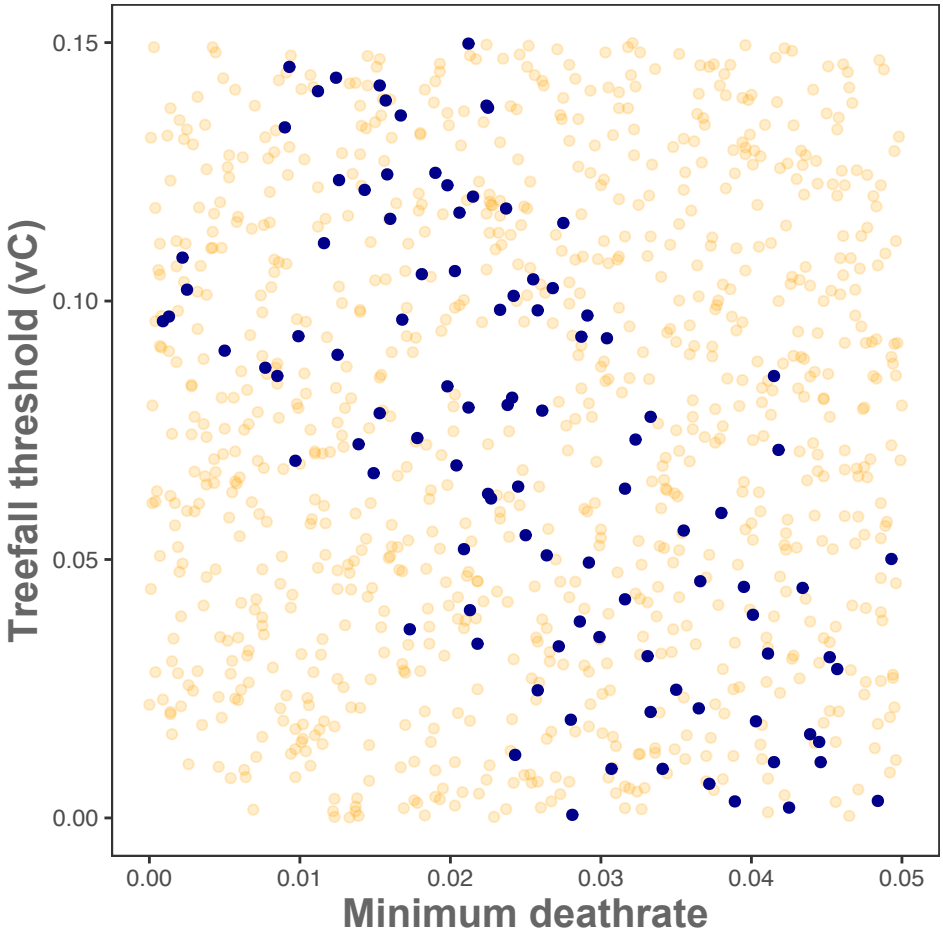


Figure 2: Mortality-parameter calibration of the TROLL model. Shown are the prior parameter distribution (light orange dots) for two mortality related parameters (treefall threshold parameter *vC*, and minimum deathrate), as well the 10% best posterior simulations (dark blue dots). While the parameters are not well-constrained individually, they are inversely correlated, indicating lower and upper limits on tree mortality imposed by the Approximate Bayesian Calibration procedure.

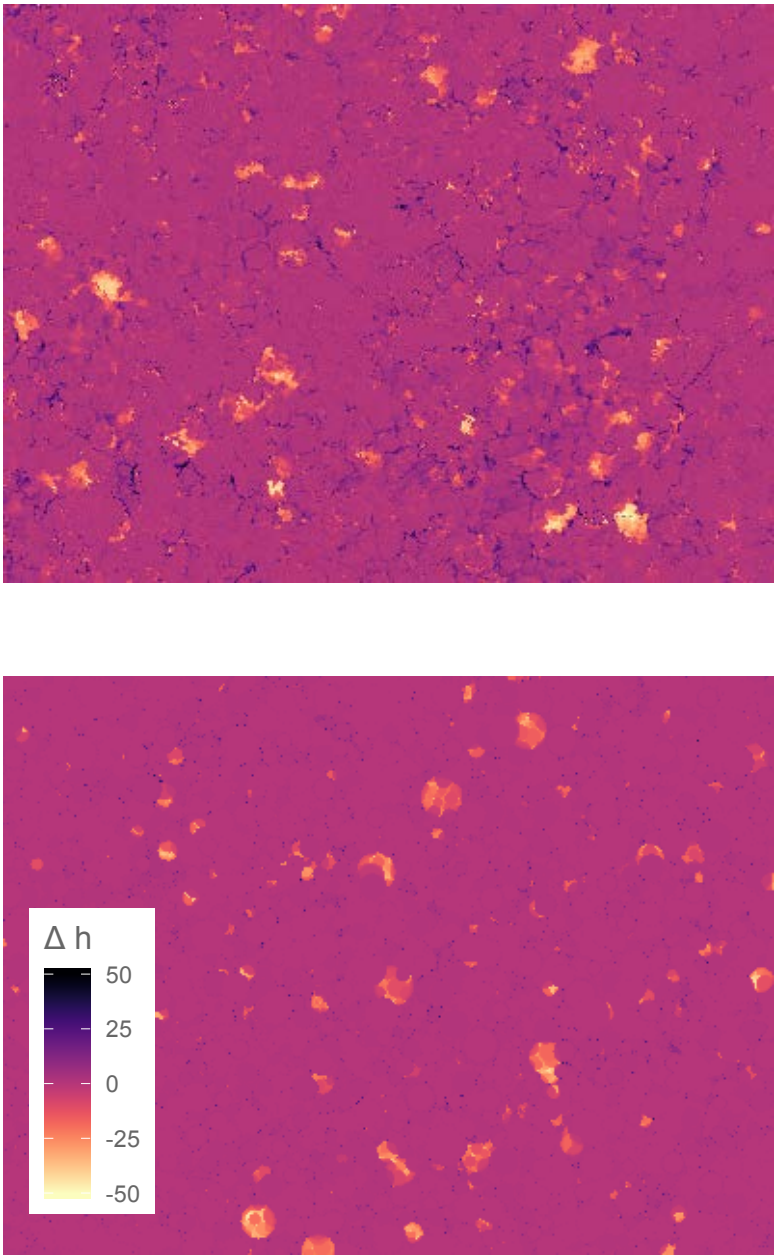


Figure 3: Canopy dynamics at Nouragues, Petit Plateau, observed and simulated. Comparison between ALS-observed canopy height changes between March 2012 and October 2015 (upper panel) and TROLL-simulated height changes over a time period of the same length (~42 months, lower panel). The TROLL-run is based on a representative simulation from the posterior after calibration. Treefall gaps are light specks, whereas dark spots indicate large height growth. The latter is typically due to crowns growing

sideways into gaps. Particularly noticeable is how TROLL replicates well the patchy treefall dynamics observed empirically. Lateral crown growth, on the other hand is less well modelled, as can be seen from a lack of large height increases in the lower panel.

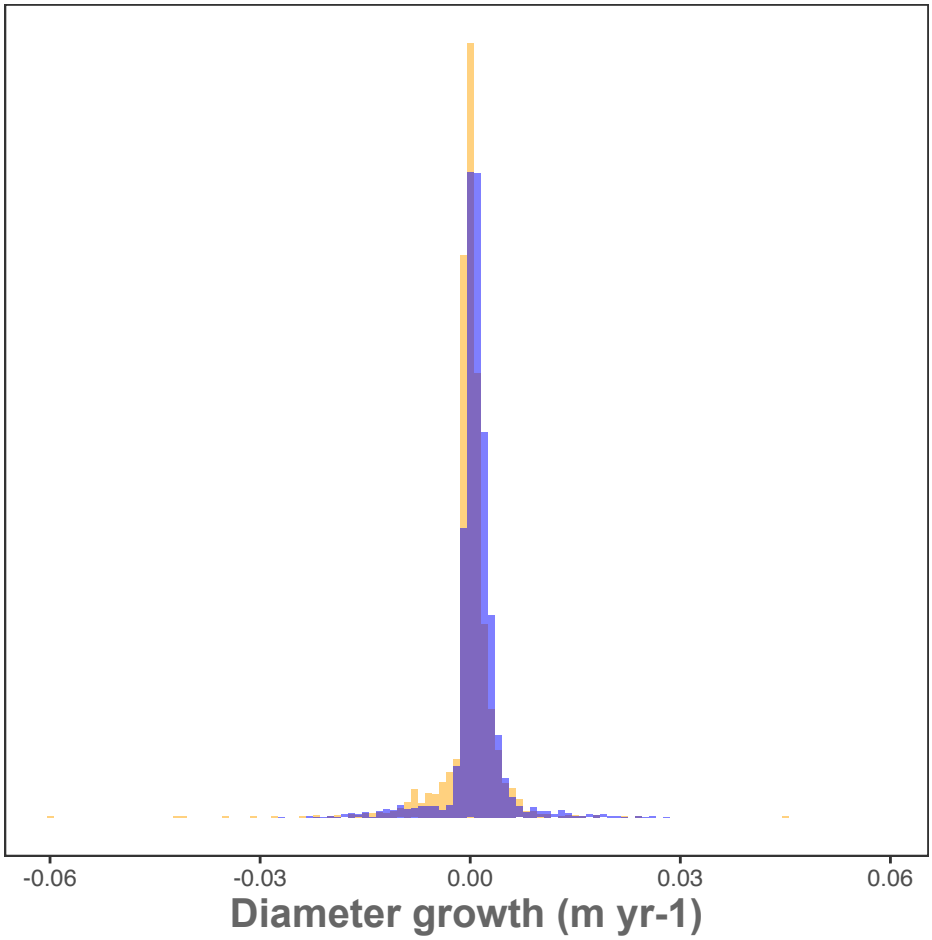


Figure 4: Stem diameter growth at Nouragues, Petit Plateau, observed and simulated. Comparison between observed (orange distribution, background) and TROLL diameter growth rates (light blue distribution) for trees > 10cm in stem diameter at Petit Plateau between 2012 and 2015. The TROLL simulation is based on the previous calibration. A measurement error model (Chave et al. 2004) has been applied to the simulated stem diameters to account for similar sources of error as in empirical inventories.

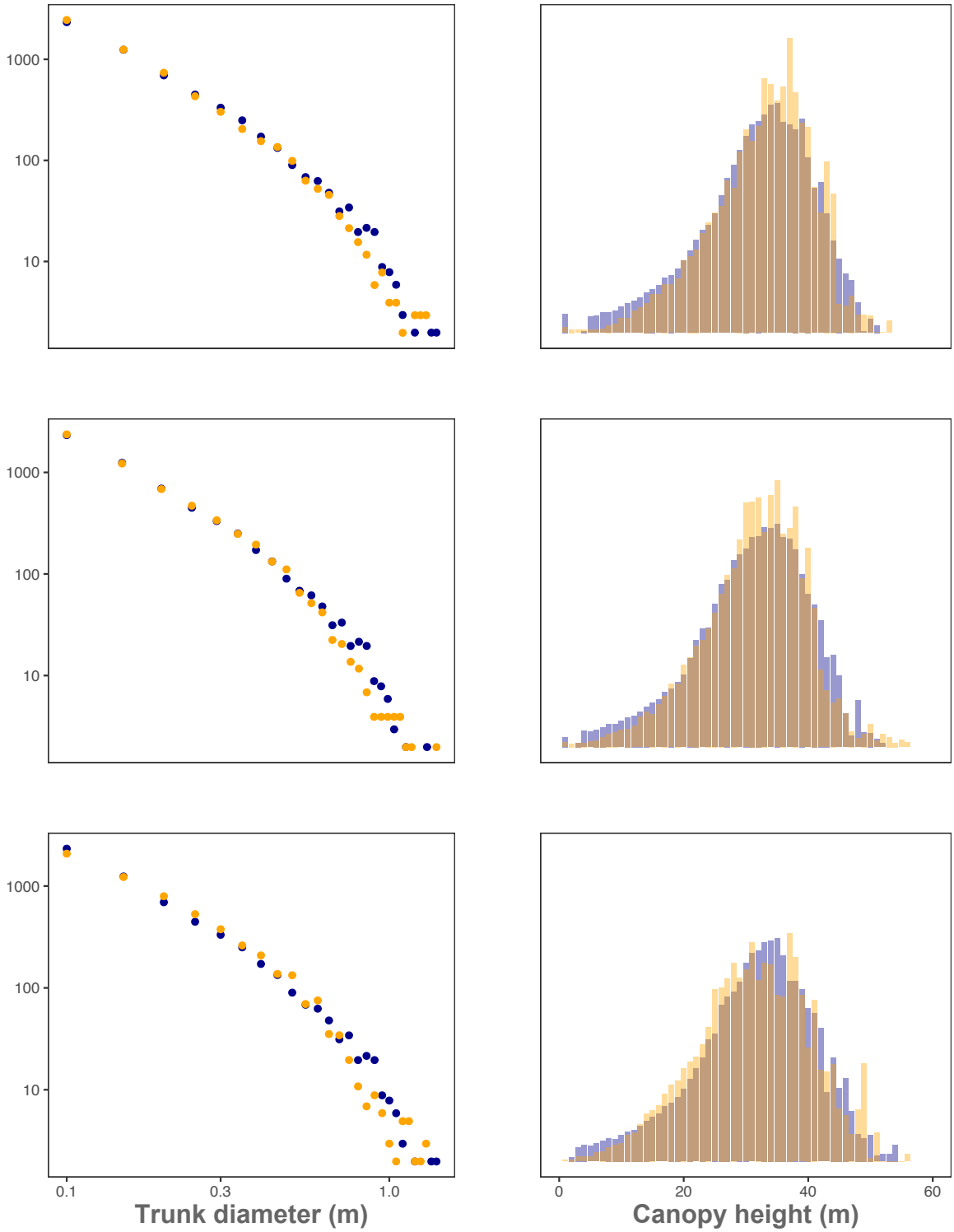


Figure 5: Overall preservation of canopy structure after 400 years of old-growth forest dynamics, as simulated by TROLL. The lefthand panels show a comparison between initial trunk diameter distribution (blue) and final trunk diameter distribution (orange), with both diameter bins and frequencies plotted logarithmicallly, for three

sample simulations from the posterior parameter distribution. The righthand panels show a comparison between the initially inferred canopy height distribution (blue) and the final canopy height distribution (orange), plotted on the original scales, for the same three simulations.

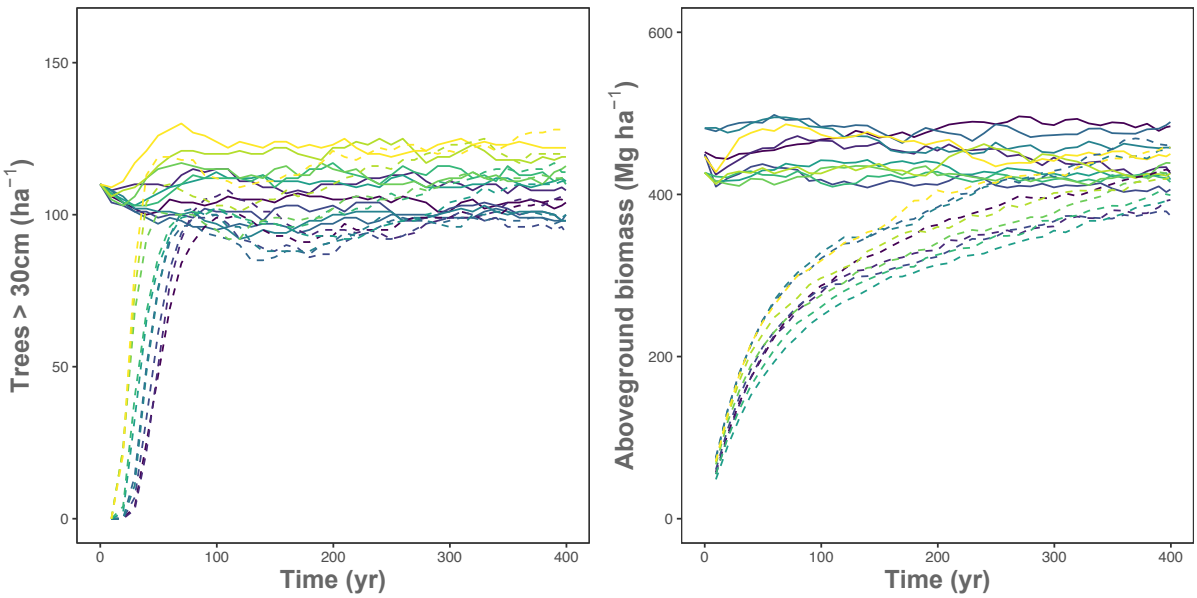


Figure 6: Comparison between 400 years of old-growth forest dynamics to 400 years of regeneration from bare soil. Shown are the number of trees > 30cm in trunk diameter (left) and aboveground biomass per hectare (right) for the 10 best posterior simulations with the forest growth model TROLL. Simulations are either initialised from an inferred old-growth forest (solid lines) or from bare ground (dashed lines). From the picture, it can be seen that tree numbers quickly reach similar levels, irrespective of initial state, but that aboveground biomass growth takes much longer to recuperate, with only some of the simulations reaching the equilibrium state after 400 years.

Supplementary Material to

Calibrating the short-term dynamics of the TROLL individual-based model in an old-growth tropical forest

S1: Covariance matrix for leaf traits

	N	P	LMA
N	0.0144		
P	0.01872	0.0576	
LMA	-0.012384	-0.022464	0.0576

S2: Calculations for VPD and Temperature

We use the equations for VPD and temperature

$$VPD(v)_{avg} = VPD_{top} \times \left\{0.25 + \sqrt{\max(0; 0.08035714 \times (7 - LAI(z))}\right\}$$

$$T(v)_{avg} = T_{top} - 0.4285714 \times \min(7;\ LAI(z))$$

, integrate over a layer and obtain the following equations:

$$VPD(v)_{avg} = VPD_{top} \times \left\{0.25 + \frac{0.188982}{dens} \times \left[(7 - LAI(z))^{3/2} - (7 - LAI(z) - dens(z)^{3/2}\right]\right\}$$

$$T(v)_{avg} = T_{top} - 0.4285714 \times \{LAI(z) + 0.5 \times dens(a)\}$$

Special cases are empty voxels where the average values across the layer are just the values at the top of the layer – i.e. the original equations:

$$VPD(v)_{avg} = VPD_{top} \times \left\{0.25 + \sqrt{\max(0.0; 0.08035714 \times (7.0 - LAI(a))}\right\}$$

$$T(v)_{avg} = T_{top} - 0.4285714 \times \min(7.0, LAI(a))$$

Where LAI exceeds the critical value of 7, our equations reduce to:

$$VPD(v)_{avg} = VPD_{top} \times 0.25$$

$$T(v)_{avg} = T_{top} - 3.0$$

The latter equations will also be used for any voxel where the LAI reaches 7 m^2/m^3 within the voxel, i.e. the cutoff values introduced in an earlier version of TROLL (Maréchaux and Chave 2017). The discontinuity in the first derivative would theoretically necessitate a more complex integral, but the errors introduced by our simplifications are negligible.

S3: Crown heterogeneities

While the crown gap fraction is given irrespective of species identity, we modify the gap fraction for trees with large variation in crown radius so that, for a given *dbh*, all trees have the same leafarea. This means that trees with larger crowns than typical for their diameter class will increase their gap fraction and have more porous crowns, while

smaller crowns will have lower gap fractions, be denser and intercept more light. Our reasoning is based on theories of plant functioning (Shinozaki *et al.* 1964) that hold that leafarea should scale with by stem diameter and thus be largely independent of variations of crown radius at a given *dbh*. This is largely confirmed by empirical data (Falster *et al.*, 2015; Sirri *et al.*, 2019, unpublished data from Piste Sainte-Elie, French Guiana). While the inclusion of crown extent and crown depth as predictors typically results in improved models, the improvement is often small and most of the variance is usually already explained by the *dbh* alone.

S4 Canopy Constructor modification

The basic fitting procedure is explained in a separate paper (Fischer *et al.*, 2019). The updated fitting procedure works as follows: All trees in the Canopy Constructor are assigned the same functional traits as trees in TROLL, either based on measurements, species labels or through a trait distribution. These values are then converted into estimates of the light compensation point *LCP*, as in TROLL v.2.5. After the initial assignment of height and crown dimensions, we determine the maximum amount of leaf area that the tree can allocate based on its light environment, again as in TROLL v.2.5, and calculate photosynthesis and respiration (Maréchaux and Chave 2017). This serves as an initial estimate of all trees' carbon balances. Whenever a new tree crown is fit, we update the maximum leaf area as well as gross and net primary productivity of all the trees affected by the change and calculate the changes to the trees' carbon balances. Since we do not have data on trees < 10cm in stem diameter, we only calculate changes in photosynthesis and respiration for trees above the 10cm threshold. This also reduces the computational burden.

A tree crown is accepted as a better fit when it simultaneously better fits the canopy height model geometrically and reduces the number of trees with a negative carbon balance. This is based on the assumption that only a small fraction of trees > 10cm in stem diameter would experience shading so extreme as to be under prolonged negative carbon balance. Accordingly, we calculate the new goodness-of-fit metric in the following way:

$$\varepsilon = \sqrt{MAE_{norm}^2 + D_{norm}^2 + \rho_{cs,norm}}$$

where MAE and D are the previously used mean absolute error and the dissimilarity index of the canopy height model distributions and ρ_{cs} the newly added fraction of trees under negative carbon balance (cs for "carbon starvation"). All three metrics are normalized by the difference between maximum and minimum values obtained through first fitting them separately (cf. Fischer *et al.*, in preparation). As before, we then run this algorithm a large number of times (10,000 typically) and select the best 100 forest reconstructions.

S5: Simulated lidar

When validating models, there are often additional sources of error or variation in empirical data sets that make difficult comparisons to model output, where the true value is given. To allow for better comparison between TROLL and airborne lidar data, we include a simulation of a simplified synthetic laser scan on the reconstructed forest scene. As input the lidar model takes a pulse density (mean and standard deviation), draws a number of beams n_{beam}^0, and, consistent with TROLL's light model, simulates vertical beam extinction with the Beer-Lambert law. The probability of a hit/return is calculated as: $p_{hit} = n_{beam} * (1.0 - \exp(-k \times LAI(x)))$. Where k is the same k as used

before and n_{beam} the number of laser beams reaching the voxel. A binomial model $\mathcal{B}(n_{beam}, p_{hit})$ then calculates the number of intercepted beams. Given a high transmittance of leaves in the near-infrared spectrum, a probability of getting secondary returns $p_{transm} = 0.4$ is given. This results in a lower effective k^*_{NIR} than for visible light. Assuming, for example, $k = 0.5$, we obtain $k^*_{NIR} \cong 0.3$ as opposed to $k^*_{PAR} \cong 0.45$. These factors are in agreement with observations and have been used in previous simulation studies (Knapp *et al.* 2018). The ratioes for second and third returns (~0.4 and ~0.16, respectively) are also close to empirical lidar surveys at the Nouragues field station (~0.45 and ~0.1, respectively, data not shown).

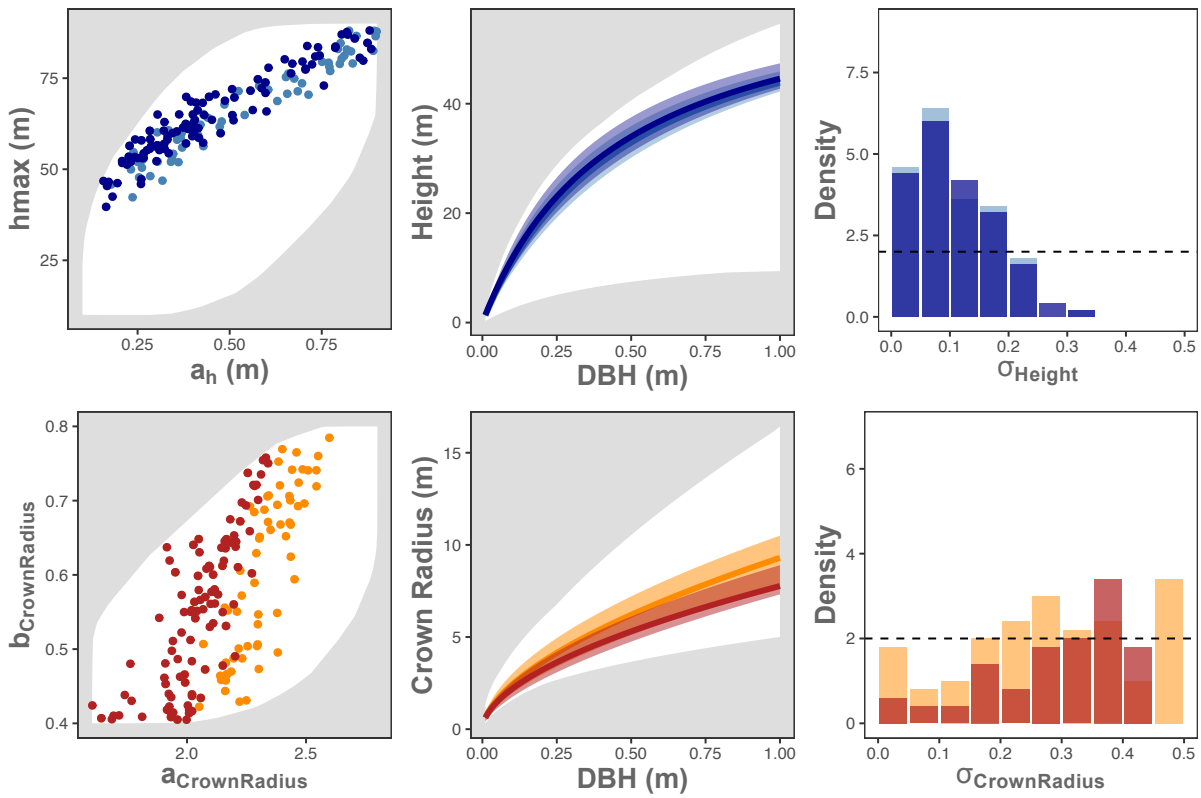


Figure S1: Inference of allometric parameters with the Canopy Constructor – geometric-only vs. physiological reconstruction. This figure compares inference of allometric scaling with two versions of the Canopy Constructor algorithm. Light colors (skyblue in top panels and orange in bottom panels) depict inferences based on a purely geometric fitting with empirically derived Canopy Height Models (CHMs). Dark colors (darkblue and darkred) depict the results of extending the approach to physiological principles, i.e. including leaf densities. The top row shows the relation between the two posterior distribution of height allometric parameters (left panel), the prior and posterior allometries (middle panel) and the prior and posterior distribution of the variance term (right panel). The bottom row shows the same information for the crown radius intercept and slope, i.e. the posterior distributions, allometries and the variance term. The best simulation (mean parameter combination) is given dark line in the middle panels, the uncertainty interval is derived from the 75% highest density

intervals of the joint posterior distribution, with best-fit allometric equations smoothing the upper and lower limits of the interval. Prior ranges are indicated as white background in the left and middle panels and through a dashed line in the righthand side panel. As can be seen from the figure, the addition of physiological information does not change inference about tree height, but clearly impacts and narrows down crown allometric parameters.

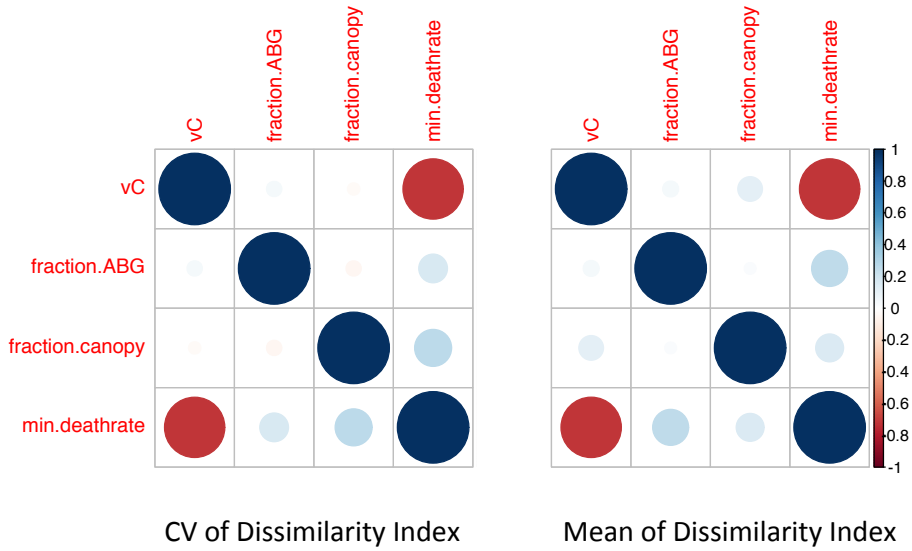


Figure S2: Correlation between parameters, as inferred through Approximate Bayesian Computation, based on two measures of canopy stability. The two panels show the inferred covariation between process-related parameters of the TROLL model, when successive old-growth forest states are compared to a reconstructed initial state. The left-hand panel shows results when the coefficient of variation of the dissimilarity index is applied, the righthand panel the results when the mean of the dissimilarity index is applied.

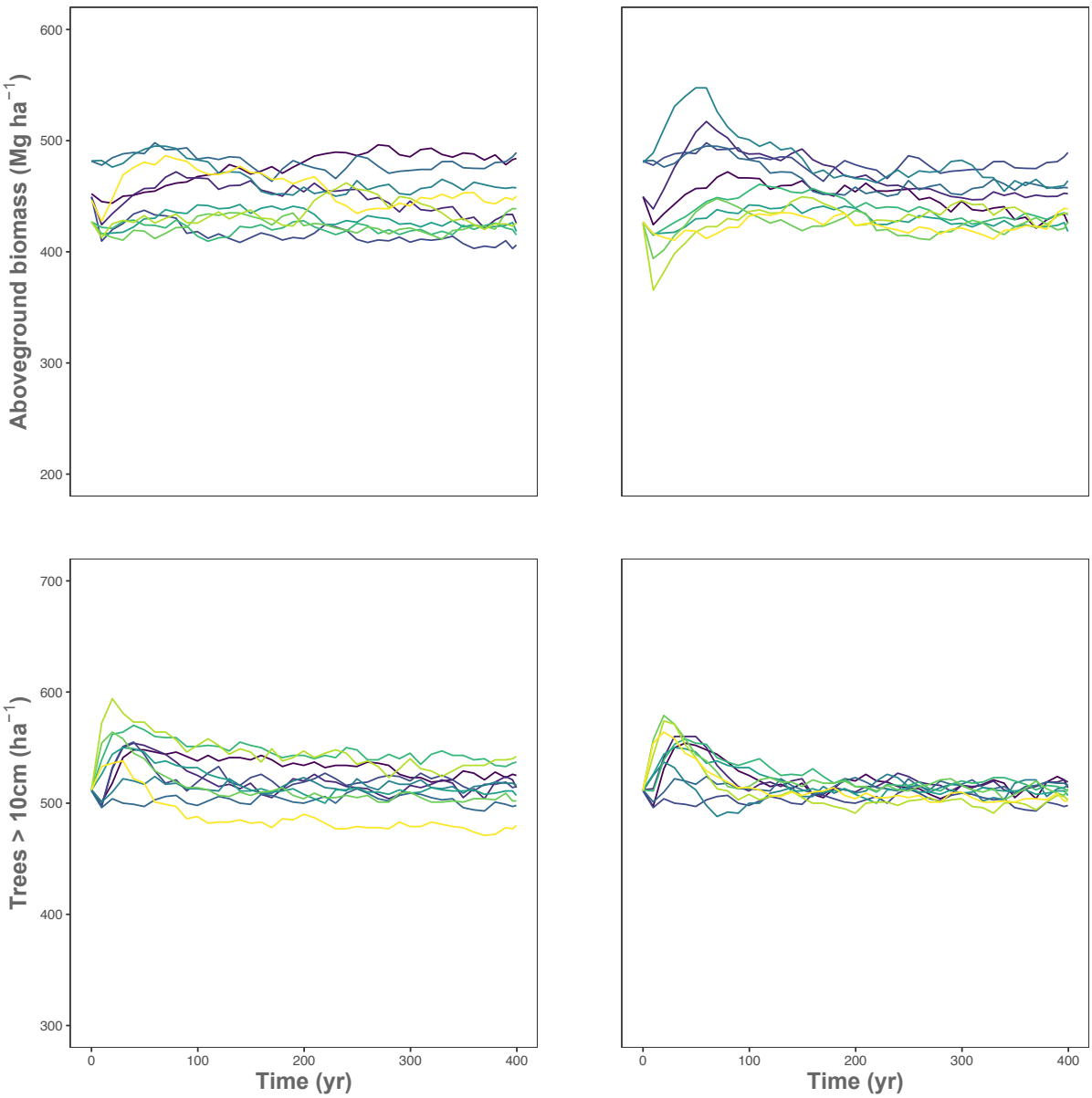


Figure S3: Overall canopy statistics, based on inference with two different measures of canopy stability. The panels show the results of dynamic inference (10 best simulations) in terms of aboveground biomass (upper panels) and overall tree numbers (lower panels). The left column shows results from inference based on the coefficient of variation of the dissimilarity index, the right column the same inference based on the mean of the dissimilarity index. It can be seen that both types of inference lead to very similar results. Constraining the variation improves slightly overall stability in terms of aboveground biomass, but leads to a larger spread in overall tree numbers.

Chapter 4: The importance of considering inter-individual variation in allometric and functional traits in an individual-based forest model

(Target Journal: TBD)

Chapter 4 builds on the previous results from Chapter 3 and uses the newly developed and calibrated TROLL v.2.5 model to investigate the effect of crown plasticity and inter-individual variation in traits on overall ecosystem functioning and structure in a tropical rainforest in French Guiana. By gradually reducing inter-individual variation from empirically calibrated values and testing the resulting simulations with and without plasticity in crown leaf area, we were able to demonstrate that crown plasticity and concomitant inter-individual variation in crown radii had a strong effect on overall ecosystem and structure, augmenting stand biomass by more than 10% and improving the efficiency of carbon uptake by trees. Inter-individual variation in leaf and woody traits, on the other hand, did not substantially change ecosystem functioning, and ontogenetic changes in trait distributions were not substantially affected.

The importance of considering inter-individual variation in allometric and functional traits in an individual-based forest model

Fabian Jörg Fischer[1,*], Isabelle Maréchaux[2], and Jérôme Chave[1]

[1] Laboratoire Évolution et Diversité Biologique, UMR 5174 (CNRS/IRD/UPS), 31062 Toulouse Cedex 9, France

[2] Botanique et Modélisation de l'Architecture des Plantes et des Végétations (AMAP), UMR 5120 (CIRAD/CNRS/INRA/IRD/UM2), 34398 Montpellier Cedex 5, France

* Correspondence: fabian.j.d.fischer@gmx.de

Keywords: Intraspecific variation, Plasticity, Individual-based modelling

1. Introduction

Ecological communities are complex adaptive systems (Levin, 1998), and as such, system dynamics emerge from the interaction of individual organisms. In forest canopies, individual plants display widely varying shapes and physiological properties. Some of the variation in tree architecture and traits is found among species and thus reflects differences related to their evolutionary history and ecological roles within the community (Paine *et al.*, 2011). Another part of the variation is found at the inter-individual or within-species level (Messier *et al.*, 2010; Vieilledent *et al.*, 2010; Baraloto *et al.*, 2010b) and can be genetic, reflect ontogenetic changes or be due to plastic responses to the environment (Sterck & Bongers, 2001; Bolnick *et al.*, 2003; Rozendaal *et al.*, 2006).

Despite its importance in forest ecosystems, vegetation models have rarely incorporated individual-level variation. While Dynamic Global Vegetation Models (DGVMs), designed to reflect the interactions between atmosphere and biosphere, increasingly take into account the demography of forest communities and incorporate individual-based approaches (Moorcroft *et al.*, 2001; Sato *et al.*, 2007; Medvigy *et al.*, 2009; Smith *et al.*, 2014), most representations of canopy structure take the approach of a mean representative plant per size class (Fisher *et al.*, 2018). A few studies have explored some forms of variability in tree dimensions or traits (Purves *et al.*, 2007; Vincent & Harja, 2008; Fyllas *et al.*, 2014), however even highly detailed individual-based models of forest growth that have emerged from the gap-modelling tradition (Fischer *et al.*, 2016) ignore variation in traits or architecture beyond that encapsulated in plant functional type concepts, i.e. groupings of plants that share the same physiology and ecological strategy (but see Pacala *et al.*, 1996).

The problem when individual variability and the organism's unique position in space and time is ignored from community dynamic models is that the dynamics of ecological systems can be seriously misrepresented (Chesson, 1986; Clark, 2003; Des Roches *et al.*, 2018). If ecological processes, for example, are not related to each other linearly, then they cannot be reliably predicted from average individuals (Bolnick *et al.*, 2011). In forests, in particular, crown plasticity results in an optimization of crown packing and thus influences stand productivity (Strigul, 2012; Pretzsch, 2014; Jucker *et al.*, 2015; Williams *et al.*, 2017). It is thus important to explore the impact of individual-level variation on ecological processes.

Here, we study the effect of individual variation in functional traits, crown extent and plasticity with a spatially explicit, individual- and trait-based forest simulator, called TROLL (Maréchaux & Chave, 2017). We build on the recent model version TROLL v.2.5, calibrated for 622 species at the Nouragues field station (Fischer *et al*., in preparation), and focus on a suite of leaf- and stem-level traits, as well as variation in crown radius allometry and crown plasticity, simulated through dynamically regulated leaf densities. In our study, we considered various representations of variability in traits from the field-inferred values to mean species values, then simulated growth both with and without crown plasticity, and linked the variability to changes in ecosystem structure and functioning. To illucidate the underlying mechanism, we examined changes in demographic rates and trait distributions for the most extreme scenarios.

2. Methods

2.1 The TROLL model

We here rely on TROLL v.2.5, parameterized for an old-growth tropical forest in French Guiana (Fischer *et al.*, in preparation). The TROLL model is an individual-based forest growth simulator that simulates tree growth in three-dimensional space (m^3 resolution). Species are represented through combinations of traits (leaf nutrients and leaf mass per area, wood density). Trees > 1cm in stem diameter are grown from seed equivalents (Maréchaux & Chave, 2017), assimilate carbon through photosynthesis (Farquhar *et al.*, 1980), and if their gross primary production exceeds respiration, they allocate biomass to tissues to grow and extend their leaf area. Above a given trunk diameter threshold, tree are considered mature and they produce new seeds. Mortality is modelled through a baseline mortality, with higher risks for low-wood density species, and additional mortality risks through carbon starvation or treefall (details cf. Maréchaux and Chave 2017).

TROLL v.2.5. includes intraspecific variation. Every trait *i*, except wood density, is assumed to follow a lognormal distribution, i.e. a multiplicative factor e^{ε_i}, where $\varepsilon_i \sim N(0, \sigma_i)$. Intraspecific variation in wood density is assumed to have an additive error term $\varepsilon_{wsg} \sim N(0, \sigma_{wsg})$. Scaling laws that govern crown and height dimensions are also modified with multiplicative factors.

In TROLL v.2.5, an important assumption is that trees adjust their leaf area so that they remain viable under typical climatic conditions (Fischer *et al.*, in preparation). This crown plasticity is achieved by calculating the average leaf density above a tree's crown at every timestep, and adjusting the total leaf area of a tree so that no leaves are allocated beyond their light compensation point. The latter is obtained by inversing the photosynthesis equations for an average day at the study site (Farquhar *et al.*, 1980) and

determines the light intensity level below which leaves are unable to balance their carbon expenditure. Trees will thus only allocate resources into leaf production when this improves their net carbon balance.

A detailed field calibration based on physiological traits, permanent inventories and airborne lidar is described elsewhere (Fischer *et al.*, in preparation).

2.2 Levels of individual variation and plasticity

To explore the effect of individual variation and crown plasticity on the dynamics in TROLL, we carried out the following test. Our baseline simulation was a fully calibrated model with intraspecific variation in leaf and wood traits and full variation in crown radius allometry (values cf. Table 1). We then ran several simulations where we jointly lowered the variation in all leaf and wood traits from 100% to 0% in steps of 10%. To explore the influence of variation on tree geometry, we applied the same scheme to variation in crown radius, running simulations with variation ranging from 100% to 0% of the empirically calibrated variation.

We then combined these simulations with two levels of crown plasticity. In one suite of simulations ("no plasticity"), crowns allocate leaves up to their theoretical maximum leaf area, irrespective of the amount of light they receive. In the other suite of simulations ("plasticity"), they use the crown plasticity module as implemented in TROLL v.2.5 and only allocate leaves up to their light compensation point.

This simulation strategy resulted in 11 x 11 x 2 = 242 parameter combinations. For each of the parameter combinations, we carried out 5 simulations for a total 400 simulated years of forest growth to account for stochasticity. Simulations were started from bare ground to provide the same initial conditions for all parameter combinations.

For lognormally distributed traits, a reduction in variance changes the mean value of the variable. If a variable v follows a lognormal distribution then $log(v) \sim N(m, \sigma)$, i.e. the logarithm of v follows a normal distribution with mean m and a standard deviation σ . It can be shown that the mean of variable v is given by the formula $e^{m+\sigma^2/2}$ (Baskerville, 1972). If we reduce variation to a new σ_i which is only a fraction i of the original σ, then the mean of v would be lowered as well. To avoid this, a compensation factor C can be calculated as follows: $e^{\sigma^2/2} = e^{\sigma_i^2/2} \times C$. It follows that $C = e^{(\sigma^2 - \sigma_i^2)/2}$. We applied this compensation factor to all varied traits except wood density, since the latter is normally distributed.

2.3 Tree growth and trait distributions

Finally, since we expected changes in trait variation and crown plasticity to not only affect ecosystem functioning, but also community dynamics, we analyzed ontogenetic shifts in traits as well as individual-level resource acquisition across simulations. We picked the most extreme scenarios, i.e. varying crown extent vs. non-varying crown extent, varying traits vs. non-varying traits and crown plasticity vs. no plasticity (8 simulations in total, 10 replicates), and simulated another 400-yr period of forest growth. In order to separate out successional effects, we did not start from bare ground, but from the already simulated old-growth forests. In year one, we selected the cohort of newly recruited saplings and followed their development for the following 400 years. In particular, we recorded their productivity, the light environment they experienced and their relative growth rates, defined as $\frac{\log(AGB_2) - \log(AGB_1)}{t_2 - t_1}$ where t_1 and t_2 are the times of recording and AGB_1 and AGB_2 the respective above-ground biomass values. We then

compared these quantities, their relation to individual traits and the community-level shifts in trait distributions from the sapling to adult stage across scenarios.

2.4 Study site

The simulations were parameterized for a lowland tropical rainforest at the Nouragues Ecological Research Station in French Guiana (4.06°N, 52.68°W). This forest experiences 2900 mm rainfall per year as well as a two month long dry season from September to November and a shorter one in March. Forest inventories have been conducted since the early 1990s (Chave *et al.*, 2008b; Labrière *et al.*, 2018), and several ALS surveys have been conducted since 2007 (Réjou-Méchain *et al.*, 2015). The TROLL model was calibrated for the Petit Plateau site so that it simulates all 622 species that have been measured there. Traits were assigned based on a large local trait collection (Baraloto *et al.*, 2010a). If no traits were available for a particular species, we assigned values either based on genera, and if that was not possible, we assigned mean community traits. More details can be found in a previous study (Fischer *et al.*, in preparation).

Statistical analysis and visual rendering were conducted in the R software (R Development Core Team, 2019), including the packages *data.table* (Dowle & Srinivasan, 2018), ggplot2 (Wickham, 2011), *viridis* (Garnier, 2018).

3. Results

Crown plasticity had a strong effect on overall forest structure after 400 years of simulation (Figure 1). Mean abundance of trees > 10cm dbh increased across all simulations with crown plasticity from 418 ha^{-1} [383, 462] to 465 ha^{-1} [425,518]. Likewise mean aboveground biomass (AGB) increased from 396 Mg ha^{-1} [381.8, 410.2] to 436 Mg ha^{-1} [420, 451.6] under crown plasticity, while there was a decrease in both litterfall from 6.1 Mg ha^{-1} yr^{-1} [5.9,6.3] to 5.4 Mg ha^{-1} yr^{-1} [5.3,5.47] and gross primary productivity (GPP) from 49.6 MgC ha^{-1} yr^{-1} [48.9, 50.3] to 47.7 MgC ha^{-1} yr^{-1} [47.2, 48.2]. Distributions of stand-scale statistics exhibited a clear shift with the inclusion of crown plasticity and were well separated, except for tree numbers. Both GPP and litterfall had narrower distributions when crown plasticity was simulated.

The effects of inter-individual variation in traits and crown extent on forest structure and functioning were less uniform than those of crown plasticity. Increasing variation in crown extent resulted in higher numbers of trees per hectare (cf. Figure 2, upper panels) and higher GPP (cf. Figure 2, lower panels), both in the case of simulations without crown plasticity and in simulations with crown plasticity. The effect on AGB was more subtle, with no clear pattern discernible in simulations without crown plasticity, but a tendency towards higher AGB in simulations with plasticity (Figure 2, middle panels). Increasing intra-specific variation in leaf and wood traits led to a slight decrese in GPP, but no effect on overall tree density or AGB (cf. S1 and Figure S1 therein for a more detailed regression analysis).

In the mature forest, out of an average of 3,805 [3,725 - 3,922] simulated saplings ha^{-1} yr^{-1}, only 2.8 [1.75 - 4.25] ever reached maturity, indicating a massive mortality rate due to competition for light. We selected individual trees when they reached 10cm in trunk diameter and analyzed their relative growth rates (RGR) from sapling stage up to

that point. Overall, growth rates uniformly increased with GPP and decreased with shading. Higher wood density values also led to a decrease in growth, but explained much less variance (Figure 3). Among the scenarios, crown plasticity led to consistently stronger effects – higher efficiency in transforming GPP into growth, lower susceptibility to shading and a ratio between net and gross primary productivity that increased from 0.299 [0.266 - 0.361] to 0.321 [0.280 – 0.384] (Figure 3, upper panels). Consistent with ecosystem-wide patterns, variation in crown extent or traits had less or no effect (cf. Figure 3, where relations are mostly overlapping).

Overall, there were also few differences between the trait distributions of saplings and trees that reached maturity, indicating little selection across tree ontogeny. However we did notice a few shifts, most notably towards lower leaf mass per area (LMA) and lower wood density in mature trees, as well as subtle shifts towards higher phosphorus (cf. Figure 4, as well as Figure S2 in the Supplementary Material). These patterns were generally found both with and without intra-specific variation and crown plasticity.

4. Discussion

This study explored the role of inter-individual variation in functional traits on the emergent properties of a forest, especially with respect to stand productivity and biomass. We modelled crown plasticity through adaptive leaf densities, and showed that the ibnclusion of this model had a strong effect on overall forest characteristics, increasing the stand biomass by roughly 10%. This effect was even more pronounced when inter-individual variation in crown diameter around allometric means was allowed. Overall carbon ecosystem turnover decreased with increased trait plasticity, with both lower primary productivity and lower leaf litterfall. While the reduction in

litterfall is linked to both lower primary productivity and more plastic plant growth, the decrease in gross primary productivity with plasticity is likely due to trees overregulating when in the shade of other trees. Our assumption in TROLL v.2.5 is that trees adjust their leaf area so that they remain viable under average climatic conditions. Since photosynthesis is, however, non-linearly dependent on environmental conditions (Farquhar *et al.*, 1980), monthly changes can cause significant upwards departures in photosynthetic production and thus trees could potentially allocate more leaves than they currently do.

The effect of individual variation in crown extent also led to notable changes in ecosystem functioning, such as higher tree numbers and higher gross primary productivity, as would be expected from improved crown packing due to heterogeneity in crown shapes (Pretzsch, 2014; Jucker *et al.*, 2015). There also was influence on overall biomass, but only when trees had plastic crowns. Otherwise, there was no discernible effect – presumably, because improved crown packing decreases competition mainly for small trees with low biomass. On the other hand, intra-specific variation in leaf traits and wood density, did not show any substantial effect on forest structure. There was a decrease in overall productivity with increasing variation, but this did not translate into similar reductions in biomass. That effects were marginal might be partly due to the fact that important mechanisms such as hydrology and nutrient cycling are not yet represented in TROLL, meaning that high-dimensional trade-offs are likely not adequately represented (Clark *et al.*, 2010).

Increases in aboveground biomass with plasticity were also confirmed at the individual tree level, where crown plasticity, but not trait variation, led to increased efficiency in the conversion of photosynthates into biomass growth, lower respiration loads on plastic crowns and generally higher growth rates. General patterns such as trait

shifts towards lower leaf mass per area from the sapling stage to adult stage were, however, not changed, irrespective of underlying variation, and similar, although much subtler changes were observed towards lower wood densities and higher leaf nutrient contents. Although we restricted our analysis to saplings regrowing in a mature forests, this would indicate that faster-growing strategies had a slight competitive advantage over shade tolerant strategies in our simulations.

Together with the fact that the light environment experienced by trees was a more important determinant of relative growth rates than traits such as wood density (cf. Figure 3), the lack of strong shifts in trait patterns suggests that the randomness of an individual tree's life history, i.e. its spatial location and the temporal patterns creating its micro-environment, plays a crucial role in determining tree growth. This offers an explanation for why functional traits often leave a lot of variation in tree vital rates unexplained (Poorter *et al.*, 2008; Paine *et al.*, 2015; Visser *et al.*, 2016).

Overall, we have shown that individual variability, particularly crown plasticity, plays an important role in predicting forest structure and functioning. We also related individual variability to the underlying tree-level growth rates. This suggests that individual variation, at least in architectural traits, is necessary to adequately assess the effects of changing environments on tropical forests.

Global climate change does not only affect the ecological dynamics of plant communities, but also has evolutionary consequences (Aitken *et al.*, 2008). Since selection is mediated through individual variation in traits, its incorporation into vegetation models opens up the possibility of propagating traits from parent trees to their offspring – allowing for the modelling of microgeographic adaptations (Richardson *et al.*, 2014) and an integrated modelling framework for ecological and evolutionary

dynamics. In the future, it would thus be possible to use the TROLL model to explore the micro-evolutionary implications of environmental changes.

Acknowledgements

This work has benefitted from "Investissement d'Avenir" grants managed by Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01; ANAEE-France: ANR-11-INBS-0001, TULIP: ANR-10-LABX-0041), from CNES, and from ESA (CCI-Biomass). This work was granted access to the HPC resources of CALMIP supercomputing centre in Toulouse under allocation p17012.

Literature

Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S. 2008. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* **1**: 95–111.

Baraloto C, Paine CET, Poorter L, Beauchene J, Bonal D, Domenach AM, Hérault B, Patiño S, Roggy JC, Chave J. 2010a. Decoupled leaf and stem economics in rain forest trees. *Ecology Letters* **13**: 1338–1347.

Baraloto C, Timothy Paine CE, Patiño S, Bonal D, Hérault B, Chave J. 2010b. Functional trait variation and sampling strategies in species-rich plant communities. *Functional Ecology* **24**: 208–216.

Baskerville G. 1972. Use of logarithmic regression in the estimation of plant biomass. *Can. J. For. Res.* **2**: 49–53.

Bolnick DI, Amarasekare P, Araújo MS, Bürger R, Levine JM, Novak M, Rudolf VHW, Schreiber SJ, Urban MC, Vasseur DA. 2011. Why intraspecific trait variation matters in community ecology. *Trends in Ecology and Evolution* **26**: 183–192.

Bolnick DI, Svanbäck R, Fordyce JA, Yang LH, Davis JM, Hulsey CD, Forister ML. 2003. The ecology of individuals: Incidence and implications of individual specialization. *American Naturalist* **161**: 1–28.

Chave J, Olivier J, Bongers F, Châtelet P, Forget PM, Van Der Meer P, Norden N, Riéra B, Charles-Dominique P. 2008. Above-ground biomass and productivity in a rain forest of eastern South America. *Journal of Tropical Ecology* **24**: 355–366.

Chesson PL. 1986. Environmental variation and the coexistence of species. *Community Ecology*: 240–256.

Clark JS. 2003. Uncertainty and variability in demography and population growth: A hierarchical approach. *Ecology* **84**: 1370–1381.

Clark JS, Bell D, Chu C, Courbaud B, Dietze M, Hersh M, Hillerislambers J, Ibáñez I, Ladeau S, McMahon S, *et al.* 2010. High-dimensional coexistence based on individual variation: A synthesis of evidence. *Ecological Monographs* **80**: 569–608.

Dowle M, Srinivasan A. **2018**. data.table: Extension of 'data.frame'. R package version 1.11.2. https://cran.r-project.org/package=data.table.

Farquhar GD, von Caemmerer S, Berry JA. **1980**. A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. *Planta* **149**: 78–90.

Fischer R, Bohn F, Dantas de Paula M, Dislich C, Groeneveld J, Gutiérrez AG, Kazmierczak M, Knapp N, Lehmann S, Paulick S, *et al.* 2016. Lessons learned from applying a forest gap model to understand ecosystem and carbon dynamics of complex tropical forests. *Ecological Modelling* **326**: 124–133.

Fisher RA, Koven CD, Anderegg WRL, Christoffersen BO, Dietze MC, Farrior CE, Holm JA, Hurtt GC, Knox RG, Lawrence PJ, *et al.* 2018. Vegetation demographics in Earth System Models: A review of progress and priorities. *Global Change Biology* **24**: 35–54.

Fyllas NM, Gloor E, Mercado LM, Sitch S, Quesada CA, Domingues TF, Galbraith DR, Torre-Lezama A, Vilanova E, Ramírez-Angulo H, *et al.* 2014. Analysing Amazonian forest productivity using a new individual and trait-based model (TFS v.1). *Geoscientific Model Development* **7**: 1251–1269.

Garnier S. **2018**. viridis: Default Color Maps from 'matplotlib'. *R package version 0.5.1*.

Jucker T, Bouriaud O, Coomes DA. **2015**. Crown plasticity enables trees to optimize canopy packing in mixed-species forests. *Functional Ecology* **29**: 1078–1086.

Labriere N, Tao S, Chave J, Scipal K, Toan T Le, Abernethy K, Alonso A, Barbier N, Bissiengou P, Casal T, *et al.* 2018. In Situ Reference Datasets from the TropiSAR and AfriSAR Campaigns in Support of Upcoming Spaceborne Biomass Missions. *IEEE Journal*

of Selected Topics in Applied Earth Observations and Remote Sensing **11**: 3617–3627.

Levin SA. **1998**. Ecosystems and the biosphere as complex adaptive systems. *Ecosystems* **1**: 431–436.

Maréchaux I, Chave J. **2017**. An individual-based forest model to jointly simulate carbon and tree diversity in Amazonia: description and applications. *Ecological Monographs* **87**: 632–664.

Medvigy D, Wofsy SC, Munger JW, Hollinger DY, Moorcroft PR. **2009**. Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2. *Journal of Geophysical Research: Biogeosciences* **114**: G01002.

Messier J, McGill BJ, Lechowicz MJ. **2010**. How do traits vary across ecological scales? A case for trait-based ecology. *Ecology Letters* **13**: 838–848.

Moorcroft PR, Hurtt GC, Pacala SW. **2001**. A method for scaling vegetation dynamics: The ecosystem demography model (ED). *Ecological Monographs* **71**: 557–586.

Pacala SW, Canham CD, Saponara J, Silander JA, Kobe RK, Ribbens E. **1996**. Forest models defined by field measurements: estimation, error analysis and dynamics. *Ecological Monographs* **66**: 1–43.

Paine CET, Amissah L, Auge H, Baraloto C, Baruffol M, Bourland N, Bruelheide H, Daïnou K, de Gouvenain RC, Doucet JL, *et al.* 2015. Globally, functional traits are weak predictors of juvenile tree growth, and we do not know why. *Journal of Ecology* **103**: 978–989.

Paine CET, Baraloto C, Chave J, Hérault B. **2011**. Functional traits of individual trees reveal ecological constraints on community assembly in tropical rain forests. *Oikos* **120**: 720–727.

Poorter L, Wright SJ, Paz H, Ackerly DD, Condit R, Ibarra-Manríquez G, Harms KE, Licona JC, Martínez-Ramos M, Mazer SJ, *et al.* 2008. Are functional traits good

predictors of demographic rates? Evidence from five neotropical forests. *Ecology* **89**: 1908–1920.

Pretzsch H. **2014**. Canopy space filling and tree crown morphology in mixed-species stands compared with monocultures. *Forest Ecology and Management* **327**: 251–264.

Purves DW, Lichstein JW, Pacala SW. **2007**. Crown plasticity and competition for canopy space: A new spatially implicit model parameterized for 250 North American tree species. *PLoS ONE* **2**: e870.

R Development Core Team. **2019**. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*.

Réjou-Méchain M, Tymen B, Blanc L, Fauset S, Feldpausch TR, Monteagudo A, Phillips OL, Richard H, Chave J. **2015**. Using repeated small-footprint LiDAR acquisitions to infer spatial and temporal variations of a high-biomass Neotropical forest. *Remote Sensing of Environment* **169**: 93–101.

Richardson JL, Urban MC, Bolnick DI, Skelly DK. **2014**. Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology and Evolution* **29**: 165–176.

Des Roches S, Post DM, Turley NE, Bailey JK, Hendry AP, Kinnison MT, Schweitzer JA, Palkovacs EP. **2018**. The ecological importance of intraspecific variation. *Nature Ecology and Evolution* **2**: 57–64.

Rozendaal DMA, Hurtado VH, Poorter L. **2006**. Plasticity in leaf traits of 38 tropical tree species in response to light; relationships with light demand and adult stature. *Functional Ecology* **20**: 207–216.

Sato H, Itoh A, Kohyama T. **2007**. SEIB-DGVM: A new Dynamic Global Vegetation Model using a spatially explicit individual-based approach. *Ecological Modelling* **200**: 279–307.

Smith B, Wärlind D, Arneth A, Hickler T, Leadley P, Siltberg J, Zaehle S. **2014**.

Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences* **11**: 2027–2054.

Sterck FJ, Bongers F. **2001**. Crown development in tropical rain forest trees: Patterns with tree height and light availability. *Journal of Ecology* **89**: 1–13.

Strigul N. **2012**. Individual-Based Models and Scaling Methods for Ecological Forestry: Implications of Tree Phenotypic Plasticity. In: Sustainable Forest Management - Current Research. 359–384.

Vieilledent G, Courbaud B, Kunstler G, Dhôte JF, Clark JS. **2010**. Individual variability in tree allometry determines light resource allocation in forest ecosystems: A hierarchical Bayesian approach. *Oecologia* **163**: 759–773.

Vincent G, Harja D. **2008**. Exploring ecological significance of tree crown plasticity through three-dimensional modelling. *Annals of Botany* **101**: 1221–1231.

Visser MD, Bruijning M, Wright SJ, Muller-Landau HC, Jongejans E, Comita LS, de Kroon H. **2016**. Functional traits as predictors of vital rates across the life cycle of tropical trees. *Functional Ecology* **30**: 168–180.

Wickham H. **2011**. ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*.

Williams LJ, Paquette A, Cavender-Bares J, Messier C, Reich PB. **2017**. Spatial complementarity in tree crowns explains overyielding in species mixtures. *Nature Ecology and Evolution* **1**.

Nitrogen SD (log scale)	**Phosphorus SD** (log scale)	**Leaf mass per area SD** (log scale)	**Wood specific gravity SD** (orig. scale, g/g)	**Crown radius SD** (log scale)
0.12	0.24	0.24	0.06	0.28

Table 1: Variation around traits in calibrated TROLL v.2.5. The standard deviation is given on logscales for all traits except wood density (additive error). Leaf and wood trait variation reflects variation around species means, whereas crown radius variation reflects variation around a global mean allometry.

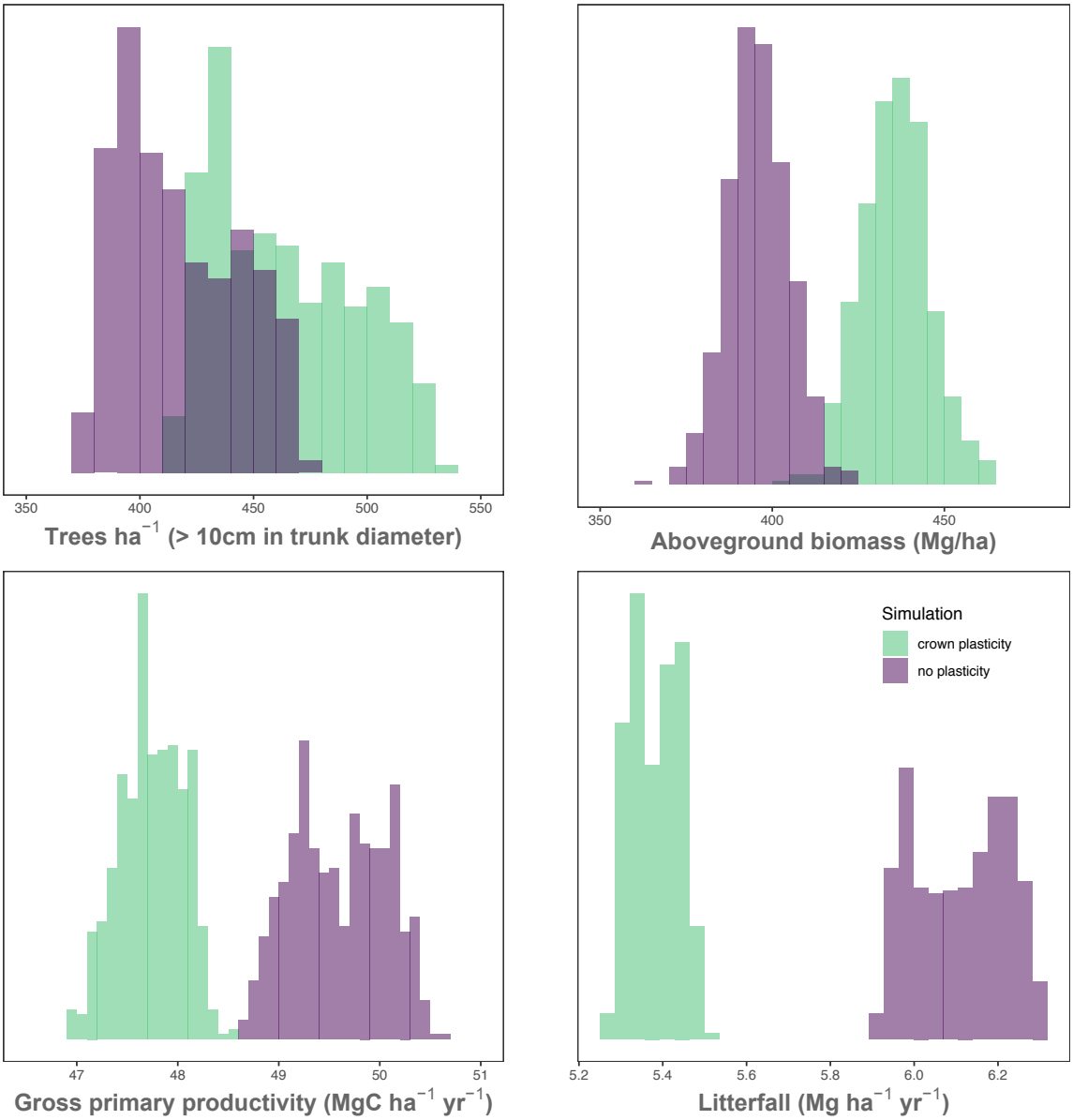


Figure 1: Comparisons of simulated ecosystem structure and functioning with and without crown plasticity: Here we show the frequency distribution of all simulations regarding tree numbers per hectare (> 10cm in trunk diameter), aboveground biomass (in Mg ha^{-1}), gross primary productivity (MgC ha^{1} yr^{-1}) and leaf litterfall (Mg ha^{1} yr^{-1}) after 400 years of forest growth. Simulations are divided only according to whether plasticity in leaf dynamics was considered or not. Simulations with crown plasticity are clearly separated from simulations without plasticity across all four metrics, with generally increased biomass (~ 40 Mg ha^{-1}) and tree numbers (~ 75 ha^{-1}), accompanied

by lower turnover in biomass, as evidenced by reduced productivity and sharply reduced leaf litterfall.

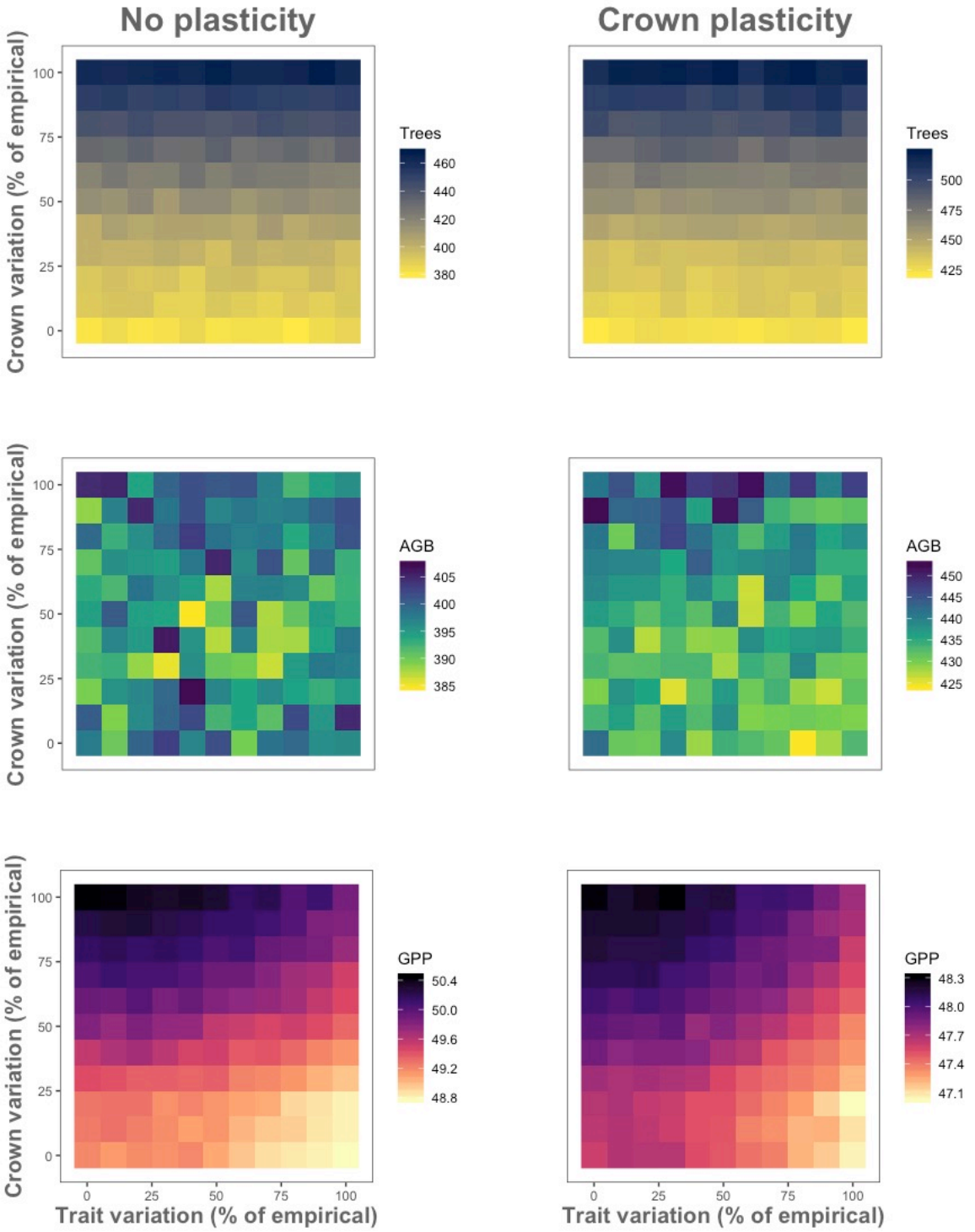


Figure 2: The effect of increasing levels of trait variation on forest structure and functioning. Shown are 5-simulation averages of forest structure and functioning metrics across increasing levels of trait variation (in percentage of the empirically parameterized values) after 400 years of forest growth. These include mean tree numbers per hectare (upper panels), mean aboveground biomass (in Mg ha^{-1}) and mean

GPP (MgC ha^{-1} yr^{-1}). To separate out crown plasticity effects, panels are subdivided (no plasticity left, crown plasticity right). Note that color scales differ etween left and righthand panels to better visualize the effects of gradual differences in trait variation.

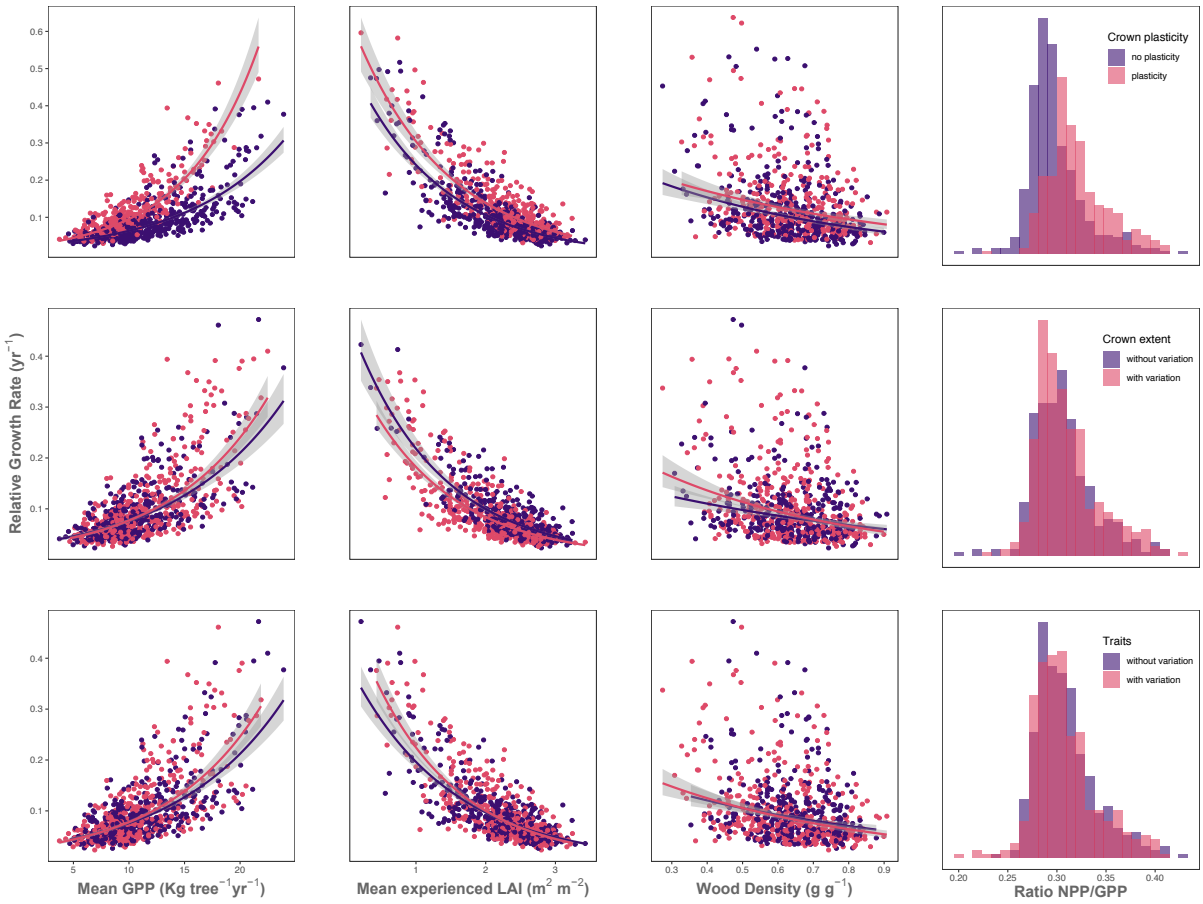
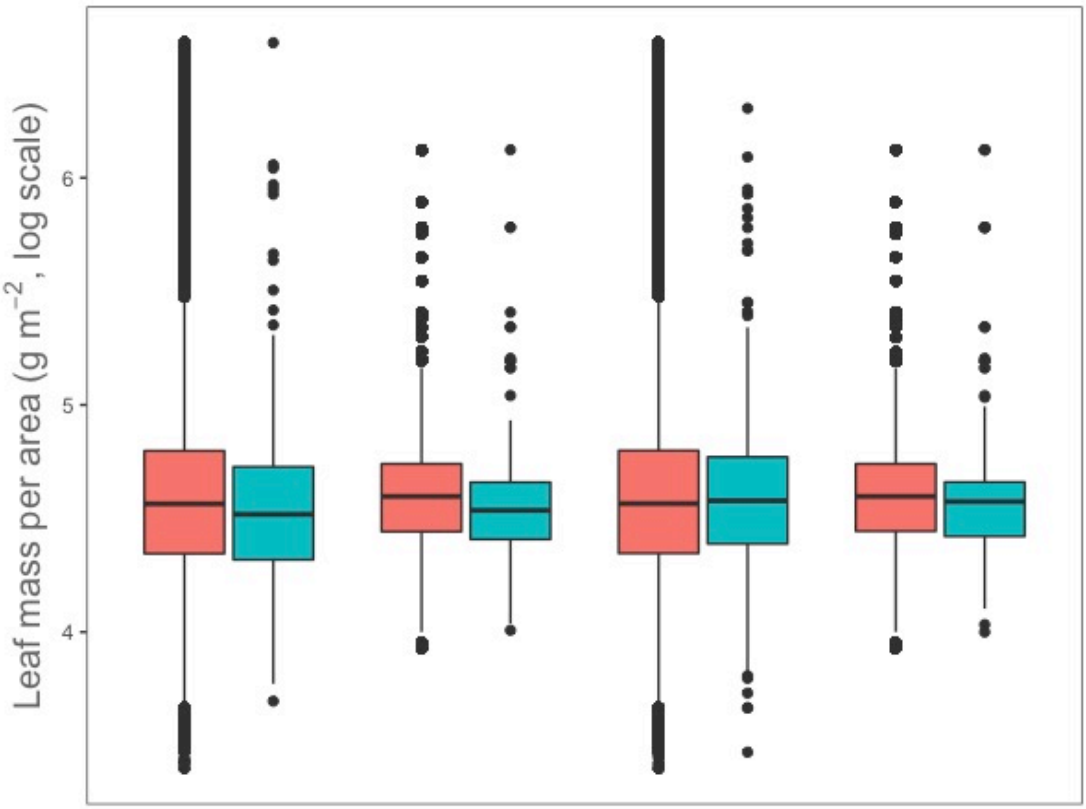


Figure 3: Relationship between relative growth rates and individual life history. Shown are, from left to right, how individual trees' relative growth rates (yr^{-1}) relate to gross primary productivity (kg $tree^{-1}$ yr^{-1}), mean local leaf area index (m^2 m^{-2}) and wood specific gravity (g g^{-1}), separated according to whether individual variability is considered (red) or not (purple). From top to bottom, we separate according to crown plasticity, variability in crown extent and variability in leaf and woody traits. The histograms to the right show the ratio of net to gross primary productivity, a measure of how efficiently trees convert photosynthetic assimilates into growth of leaves and stem biomass. Coloured lines are linear regression lines fit on logscales and backtransformed to the original scales. Only in the case of crown plasticity (upper panels), a clear separation of the fitted relationships is visible. Every point represents one tree, all trees

that are plotted have just surpassed 10cm in trunk diameter, all quantities are calculated for the trees' whole lifetime.


Leaf mass per area (g m^{-2}, log scale)
6
5
4

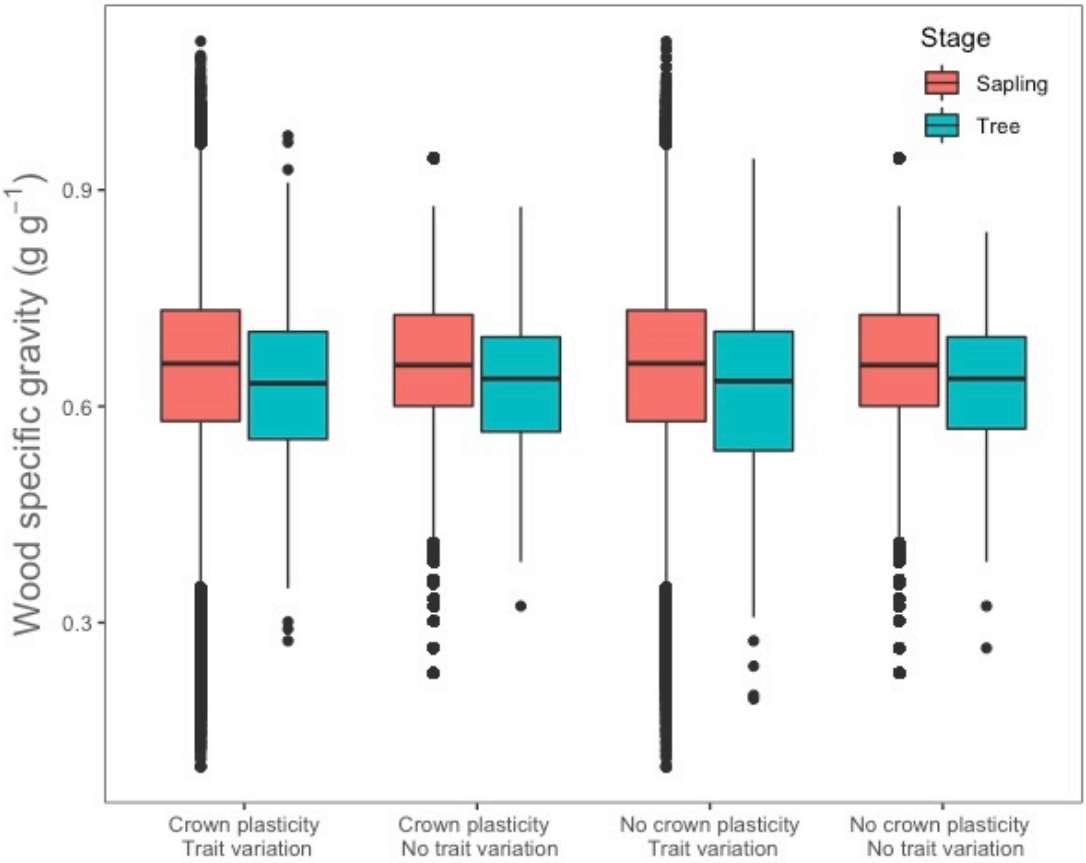

Wood specific gravity (g g^{-1})
0.9
0.6
0.3
Stage
Sapling
Tree
Crown plasticity
Trait variation
Crown plasticity
No trait variation
No crown plasticity
Trait variation
No crown plasticity
No trait variation

Figure 4: Ontogenetic changes in trait distributions for leaf mass per area and wood specific gravity. Shown is ontogenetic trait variation between the sapling and tree stages across a range of scenarios (simulating crown plasticity or not, simulating intra-specific trait variation or not). Clearly visible are minor, but generally consistent shifts towards lower leaf mass per area and lower wood densities between sapling and adult stages, suggesting that trees with low leaf mass per area and low wood density have a slight competitive advantage.

Supplementary material to:

The importance of considering inter-individual variation in allometric and functional traits in an individual-based forest model

S1: Trends in biomass with increasing intra-specific variation

To separate out the effect of crown extent on aboveground biomass from the effect of increasing variation in leaf and wood traits, we normalized aboveground biomass by the mean biomass across all simulations with the same variation in crown extent and then plotted it against increasing levels of trait variation. There was no clear effect without crown plasticity and a small decrease in biomass when crown plasticity was included in the model. The regression slopes were -9.3×10^{-7} without crown plasticity and -8.9×10^{-5}. The latter corresponds to a reduction of not even 1% from 0 to 100% of intraspecific trait variation, or, given that mean aboveground biomass in case of plasticity is 436 ha^{-1} Mg, of 3.9 Mg ha^{-1}.

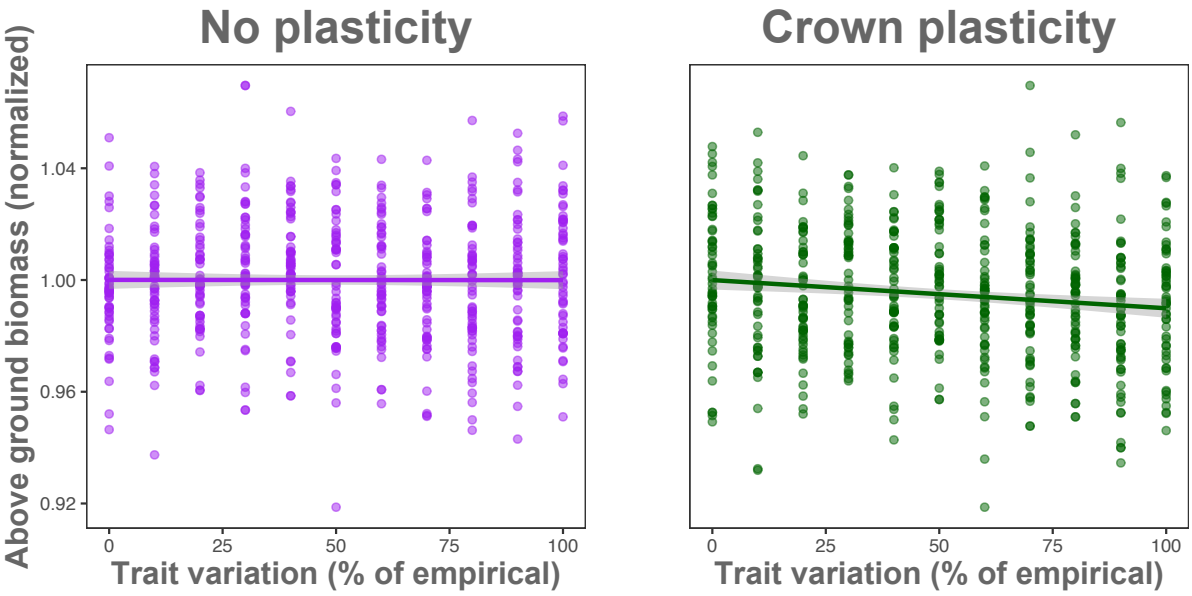


Figure S1: Trends of increasing intra-specific variation on aboveground biomass.

This figure shows the regression lines described in Supplementary Material S1.

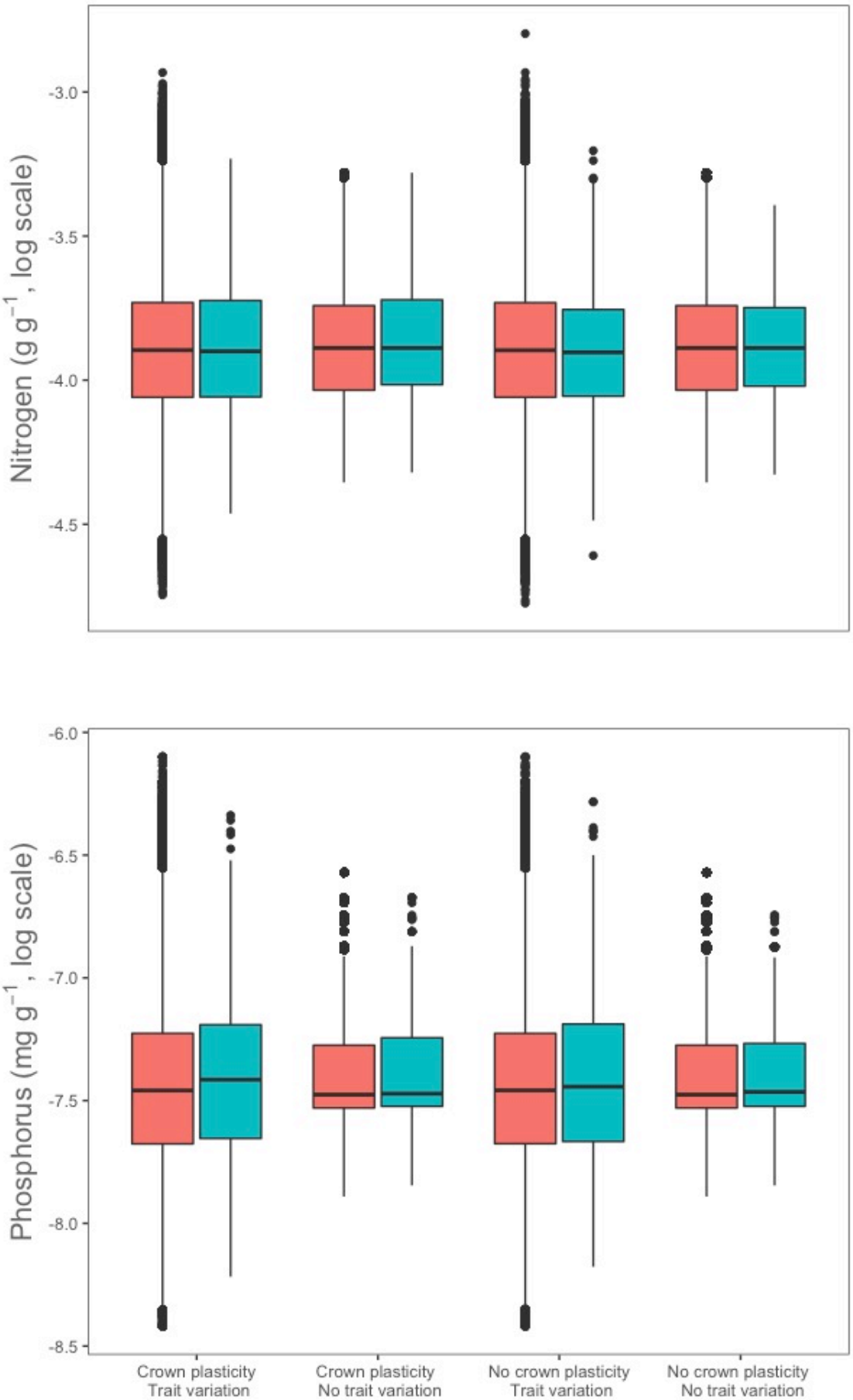

Nitrogen (g g^{-1}, log scale)
-3.0
-3.5
-4.0
-4.5
Phosphorus (mg g^{-1}, log scale)
-6.0
-6.5
-7.0
-7.5
-8.0
-8.5
Crown plasticity
Trait variation
Crown plasticity
No trait variation
No crown plasticity
Trait variation
No crown plasticity
No trait variation

Figure S2: Ontogenetic changes in trait distributions for leaf nitrogen and phosphorus. This figure is equivalent to Figure 4 from the main text, only for leaf phosphorus and leaf nitrogen. Shown is ontogenetic trait variation between the sapling and tree stages across a range of scenarios (simulating crown plasticity or not, simulating intra-specific trait variation or not).

Chapter 5: Global patterns and evolutionary trends in wood density

(Target Journal: Nature Plants)

Functional traits play an important part in the assessment of ecosystem structure and function, such as biomass estimations in Chapter 2, and in predictive ecoystem modelling, as carried out in Chapter 3. In Chapter 5, we depart from the strictly model-based approaches of the previous chapters, focus on a single trait, wood density, and update a large trait data base, the Global Wood Density Database. Assembling a wealth of new data and using improved conversion factors between various woody quantities, we increase the number of records from ca. 14,000 to over 70,000, and the number of species covered from ~8,000 to ~15,000. We then use the assemblied data base to examine within-species variation of wood density, merge it with recently published phylogenies of seed plants to estimate which evolutionary lineages contributed most to the functional distinctiveness in current woody diversity, and finally, match records to GBIF tree occurrences world wide to derive a global map of wood density.

Global patterns and evolutionary trends in wood density

Fabian Jörg Fischer[1] *, Amy Zanne[2], Shengli Tao[1], Ghislain Vieilledent[3 65 80], Renato A. Ferreira de Lima[4], Tom De Mil[5], Hans Beeckman[5], S. Joseph Wright[6], Steven Jansen[7], Brendan Choat[8], Dmitry Schepaschenko[9], Luisa Casas[10], Patrick Langbour[11 66], Sébastian Paradis[11 66], Bernard Thibaut[11 66], Julieta Rosell[12], Mark Olson[13], Nadia S. Santini[14 67], André Luiz Alves de Lima[15 68], James Lawson[16], Jürgen Homeier[17], Jean-Francois Bastin[18], Vincent Hervé[19 69], Enrique G. De la Riva[20 70], Sarah Richardson[21], Marco Njana[22], Thomas Ibanez[23 71 81], Rob Salguero-Gomez[24], Tahiana Ramananantoandro[25], Alex Fajardo[26], Cate Cate Macinnis-Ng[27], Onja Razafindratsima[28], Tommaso Jucker[29], Roberto Cazzolla Gatti[30 72], Teresa Rosas[31 73], J. Julio Camarero[32], Euler Nogueira[33], David Yue Phin Tng[34], Deborah Mattos Guimaraes Apgaua[35 34], Keryn Paul[36], Arne Sellin[37], James McCarthy[38 74], Peter Hietz[39], Arcanjo Fátima[40], Oris Rodríguez-Reyes[41 75], Rosa Goodman[42], Vanessa Boukili[43], Kasia Zieminska[44], Rosana Lopez[45 76], Luciana Alves[46], Tanaka Kenzo[47], Javid Ahmad Dar[48], Adam Martin[49], Somaiah Sundarapandian[50], Joli Borah[51], David Edwards[51], Luiz Fernando Magnago[52], Luitgard Schwendenmann[53], Daniel Falster[54], Olivier Flores[55], Bernhard Schuldt[56], Yusuke Onoda[57], Martyna Kotowska[58 77], Ashley Matheny[59], Masha Van der Sande[60 78 82], Timothy Staples[61], Rafael Oliveira[62], Esteban Alvarez[63 79], Rinku Moni Kalita[64], Jerome Chave[1]

1 Laboratoire Évolution et Diversité Biologique, UMR 5174 (CNRS/IRD/UPS), 31062 Toulouse Cedex 9, France

2 Department of Biological Sciences, George Washington University, Washington, DC, 20052, USA

3 CIRAD, UPR Forêts et Sociétés, F-34398 Montpellier, France

4 Universidade de São Paulo, Instituto de Biociências, Rua do Matão, nº 321, 05508-090, São Paulo, SP, Brasil

5 Royal Museum for Central Africa, Wood Biology Service, Leuvensesteenweg 13, 3080 Tervuren, Belgium

6 Smithsonian Tropical Research Institute, Balboa, Republic of Panama

7 Institute of Systematic Botany and Ecology, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany

8 Hawkesbury Institute for the Environment, Western Sydney University, Locked Bag 1797, Penrith, 2751 NSW, Australia

9 International Institute for Applied Systems Analysis, Schlossplatz 1, 2361 Laxenburg, Austria

10 Departamento de Ciencias Biológicas, Laboratorio de Ecología de Bosques Tropicales y Primatología, Universidad de los Andes, Bogota, Colombia

11 CIRAD, UPR BioWooEB, F-34398 Montpellier, France

12 Laboratorio Nacional de Ciencias de la Sostenibilidad, Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad de México, Mexico

13 Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad de México, Mexico

14 Cátedra Consejo Nacional de Ciencia y Tecnología, Av. Insurgentes Sur 1582, Crédito Constructor, Benito Juárez, 03940 Ciudad de México, Mexico

15 Instituto Federal de Educacao Ciencia e Tecnologia do CearaQuixada - CEBrazil

16 Department of Biological Sciences, Macquarie University, North Ryde, New South Wales, 2109 Australia

17 Plant Ecology, Albrecht von Haller Institute for Plant Sciences, University of Göttingen, 37037 Göttingen, Germany

18 Crowther Lab, Department of Environmental Systems Science, Institute of Integrative Biology, ETH Zürich, Zürich, Switzerland

19 Laboratory of Microbiology, Institute of BiologyUniversity of NeuchâtelNeuchâtelSwitzerland

20 Área de Ecología, Facultad de CienciasUniversidad de CordobaCordobaSpain

21 Landcare Research, Lincoln, New Zealand

22 National Carbon Monitoring Centre, Box 3013, Morogoro, Tanzania

23 Institut Agronomique néo-Calédonien (IAC)NouméaFrance

24 Department of Zoology, University of Oxford, OX2 6GG Oxford, United Kingdom

25 Département des Eaux et Forêts, Ecole Supérieure des Sciences AgronomiquesUniversité d'Antananarivo, Antananarivo, Madagascar

[26] Centro de Investigación en Ecosistemas de la Patagonia (CIEP), Camino Baguales s/n, Coyhaique, Chile

[27] School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

[28] Department of Biology, College of Charleston, Charleston, SC, USA

[29] CSIRO Land and Water, 147 Underwood Avenue, Floreat, Washington, 6014 Australia

[30] Department of Forestry and Natural Resources, Purdue University, West Lafayette, United States

[31] CREAF, Bellaterra, 08193 Barcelona, Spain

[32] Instituto Pirenaico de Ecología (IPE-CSIC), Avda. Montañana 1005, Zaragoza, Spain

[33] Centro Universitário UniFG, Guanambi, BA, Brasil

[34] Centre for Tropical Environmental and Sustainability Sciences, College of Marine and Environmental Sciences, James Cook University, 14–88 McGregor Road, Smithfield, Qld 4878, Cairns, Qld, Australia

[35] Department of Forest Sciences, Federal University of Lavras, C.P. 3037, CEP 37200-000, Lavras, Minas Gerais, Brazil

[36] CSIRO Land and Water, Canberra, Australia

[37] Institute of Ecology and Earth Sciences, University of Tartu, Lai 40, 51005 Tartu, Estonia

[38] School of Biological Sciences, The Univ. of Queensland, St Lucia, Brisbane, QLD, Australia

[39] Institute of Botany, University of Natural Resources and Life Sciences, Vienna, Austria

[40] Universidade Estadual de Londrina, Centro de Ciências Biológicas, Programa de Pós-Graduação em Ciências Biológicas

[41] Universidad de Panamá, Facultad de Ciencias Naturales, Exactas y Tecnología, Departamento de Botánica. Apartado 000 17, Panamá 0824, Panamá

[42] Department of Forest Ecology and Management, Swedish University of Agricultural Sciences, Umeå, Sweden

[43] Somerville City Hall, Somerville, MA, USA

[44] Arnold Arboretum of Harvard University, 1300 Centre St, Boston, MA, 02130 USA

[45] Hawkesbury Institute for the Environment, Western Sydney University, Locked Bag 1797, Penrith, NSW, Australia

[46] Center for Tropical Research, Institute of the Environment and Sustainability, UCLA, Los Angeles, CA, 90095 USA

[47] Forestry and Forest Products Research Institute, Tsukuba 305-8687, Japan

[48] Biodiversity Conservation Laboratory, Department of Botany, Dr. Harisingh Gour Vishwavidyalaya (A Central University), Sagar, M.P., 470003, India

[49] Department of Physical and Environmental Sciences and Centre for Critical Development Studies, University of Toronto Scarborough, Toronto, Canada

[50] Department of Ecology and Environmental Sciences, Pondicherry University, Puducherry-605014, India

[51] Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, S10 2TN United Kingdom

[52] Centro de Formação em Ciências e Tecnologias Agroflorestais, Universidade Federal do Sul da Bahia, Ilhéus, BA, Brazil

[53] School of Environment, University of Auckland, Auckland, New Zealand

[54] Evolution & Ecology Research Centre, and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney NSW 2052, Australia

[55] Université de La Réunion, UMR PVBMT, Saint-Pierre, La Réunion, France

[56] Plant Ecology, Albrecht von Haller Institute for Plant Sciences, University of Goettingen, Untere Karspüle 2, Goettingen 37073, Germany

[57] Graduate School of Agriculture, Kyoto University, Kyoto, 606-8502 Japan

[58] University of Goettingen, Plant Ecology and Ecosystems Research, Germany

[59] Department of Geological Sciences, Jackson School of Geosciences, 2305 Speedway Stop, C1160 Austin, TX, USA

[60] Institute for Global Ecology, Florida Institute of Technology, Melbourne, FL, USA

[61] School of Biological Sciences, The University of Queensland, Brisbane, Qld, Australia

[62] Department of Plant Biology, Institute of Biology, P.O. Box 6109, University of Campinas – UNICAMP, 13083-970 Campinas, SP, Brazil

[63] Escuela ECAPMA, UNAD, Calle 14 Sur No. 14-23, Bogotá, Colombia

[64] Department of Ecology and Environmental Science, Assam University, Silchar 788 011, India

[65] Forêts et Sociétés, Univ Montpellier, CIRAD, Montpellier, France

[66] BioWooEB, Univ Montpellier, CIRAD, Montpellier, France

[67] Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04500 Ciudad de México, Mexico

[68] Departamento de BiologiaUniversidade Federal Rural de PernambucoRecife - PEBrazil

[69] Laboratory of Biogeosciences, Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland

[70] Estación Biológica de Doñana (EBD-CSIC)SevilleSpain
[71] AMAP, CIRAD, CNRS, INRA, IRD, Univ MontpellierMontpellierFrance
[72] Biological Institute, Tomsk State University, Tomsk, Russian Federation
[73] Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
[74] CSIRO, Black Mountain, Canberra, ACT, Australia
[75] Smithsonian Tropical Research Institute, Box 0843-03092, Balboa, Ancón Republic of Panamá
[76] Université Clermont-Auvergne, INRA, PIAF, 63000 Clermont-Ferrand, France
[77] Macquarie University, Department of Biological Sciences, North Ryde,NSW, Australia
[78] Institute for Biodiversity & Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands
[79] Fundación Con Vida, Avenida del Río # 20-114, Medellín, Colombia
[80] Joint Research Centre of the European Commission, Bio-economy unit, I-21027 Ispra, Italy
[81] Department of BiologyUniversity of Hawai´i at HiloHiloUSA
[82] Forest Ecology and Forest Management Group, Wageningen University and Research, P.O. Box 47, 6700 AA Wageningen, The Netherlands

* Correspondence: fabian.j.d.fischer@gmx.de

Keywords: Wood density, Plant traits, Functional Distinctiveness

Main

Wood is a global store of carbon[1] and understanding how plants build and maintain woody organs is essential biogeochemistry, with great significance for ecology[2]. Coordination of woody traits, arising from selective pressures, has played a major role in the evolution of trees[3,4]. One of the major axis of woody variation is specific gravity, weight of an anhydrous sample divided by its water-saturated volume. Trees with denser woods tend to be more resistant to compression, bending, breaking, and to shear stress, so wood specific gravity is a summary trait for wood mechanical properties[5]. Wood density has also been attributed to an increased resistance to pathogens and xylophagous insects[6]. Finally, while the relation between wood ultrastructure and hydraulic properties is complex, there is some evidence for increased hydraulic safety at high wood densities[7–9]. Thus, woody plants with higher wood density tend to have higher survival rates[10]. Increased allocation to to the construction of woody organs comes, however, at higher costs, resulting in a trade-off of ecological strategies[11,12].

We here explore global patterns in wood density. In particular, we evaluate how wood density as a plant trait has evolved along the history of seed plants. We do so by testing which plant clades contribute the most to the extant variation in wood density. Furthermore, we assess global coverage of wood density measurements and explore how wood density varies across continents and across plant organs. These advances are based on the building and careful verification of a global wood density database, the GWDD v.2, which superseeds a previously published compilation of trait values[5,13].

The Global Wood Density Database v.2

The new wood density database contains 71,028 records, of which 66,685 are identified at the species level, comprising 13,124 accepted species and 2,933 genera, and a further 1,575 species that have not been taxonomically resolved. This new effort has resulted in an increase of 56% of the species coverage and of 74% of the genus coverage. The GWDD v.2 includes 20% of the existing tree species[14], and 59% of the tree genera. For 8,006 species, more than one record is available, including records from several sources in 5,878 species. Globally, our database has a mean wood density of 0.589 and a standard deviation of 0.163. From 546 species where we had more than 15 individual plant records, we observed twice as much between-species variation as intra-specific variation, with standard deviations of 0.145 and 0.073, respectively.

Species coverage of the GWDD v.2 varied geographically, reaching up 75-85% for European countries. In Amazonia, the database matched 30% of the tree species and 75% of the genera in a recent taxonomically verified database[15]. In tropical Africa, it matched 40% of the tree species and 80% of the genera[16]. Three regions had low coverage: the South Pacific, the Caribbean, and the Arabic Peninsula (cf. also Supplementary Material S3 and Figures S1-S3).

Evolutionary and geographic patterns

In terms of evolutionary history, the GWDD v.2 matches 13,185 species from a recent phylogenetic tree including all land plants[17]. Based on a test of functional distinctiveness[18], Eudicotyledons were identified as the most distinctive clade, followed – in descending order – by Myrtaceae, the genus *Ficus*, and Fabales and Ericales (cf. Figure 1, upper panel). All distinctive clades identified showed clear shifts towards

higher wood densities, except the genus *Ficus*, which diverged strongly towards low wood densities (cf. Figure 1, lower panel).

We further combined the GWDD v.2 with occurrence data from the Global Biodiversity Information Facility[19], and inferred a global map of wood density by linking these occurrence records with several layers of remote sensing, climate and topography through random forest modelling. Our analysis revealed broad patterns of wood density variation across the globe (Figure 2). While low wood density wood dominates at high latitude, high wood densities were found in Amazonia, across the African continent and particularly high values in Australia. While we found a gradient of increasing wood density from Western to Eastern South America, we did not find particularly strong increases within Amazonia[20].

Major axis regression furthermore indicated that branch and trunk wood densities were well correlated, with branch wood density slightly lower than trunk wood density across the data set, and a correlation coefficient of 0.77 (n=749 species, Figure 3). Furthermore, wood density is well-preserved within species that occur across contintents, with correlation coefficients of ca. 0.85 (cf. Figure S4).

Discussion

Our study provides an integrative picture of wood density across time and space, revealing broad patterns in variation in woodiness, adaptations to the environment and how they have shaped their evolutionary history.

First, our study indicates an early shift towards high wood densities in angiosperms that far predated the Cenozoic[21]. A strong shift towards high wood density was also found due to the speciation of Myrtaceae and a similar, albeit lesser, shift in Fabales. Both Myrtaceae and Fabales thought to have diversified in dry environments[22–

[25] and the high wood density in the lineages may be attributable to drought tolerance. This stands in stark contrast to the genus *Ficus* that has much lower wood density. Part of the strong downwards shift here may be driven by functional specialisation due to a tight mutualistic relationship with fig wasps and resulting physiological constraints on transpiration and water storage[26,27]. For Ericales, also among the five most distinctive clades, previous studies on evolution of vessel characteristics have indicated two major types of wood structure – one "primitive" type derived from cornalean-ericalean ancestors found in temperate and montane tropical regions, and a derived wood structure type due to shifts into lowland tropical rainforests[28]. Our results mirror these patterns, with a close match in low wood densities between Cornales and the so-called primitive type, and high wood densities in tropical families such as Sapotaceae (cf. Figure S5). This indicates that a large part of their functional distinctiveness has evolved as a result of migration into and adaptation to the environments of lowland tropical forests.

The underlying causes for adaptations in wood density are changes in wood composition that reflect mechanical resistance[5,29] or varying demands on transpiration and embolism resistance between climates[30]. This is reflected in changes of wood density with environmental conditions, such as a decrease with elevation and decreasing temperatures[31,32]. Several studies have also found a negative correlation between precipitation and wood density[33–37], but it might not be as universal as temperature dependence[32].

Our global wood density map provides evidence for environmental determinants of wood density, with notable decreases of wood density in high-altitude regions across continents and towards Northern latitudes, high wood densities in lowland tropical forests and particularly high wood densities in the arid or semi-arid climates of

Australia. Previously found patterns such as notable shifts in wood density across Amazonia[20] could, however, not be confirmed. Part of this might be due to our estimates being derived from mean species values and an aggregation procedure that relies on taxonomic records, but does not reflect the local abundance or biomass of species.

While we thus expect our map to reveal broad patterns, some of the local variation might be lost (cf. Figure S6 for a correlation between different types of wood density aggregates). Future studies might substantially improve on our wood density map by also considering the respective abundances and sizes of species, but attaining vast geographic coverage of field sampling and precise estimates of local abundances at the same time will remain a considerable challenge. Recently, several procedures have, for example, been proposed to create global trait maps, relying on broad functional types to scale local measurements up, but the delimitation of plant functional types is likely introducing considerable error of its own[38,39].

Finally, intra-specific variation in wood density is an important factor to consider. Estimates were generally lower in branch wood than in trunk wood, but there was considerable variation, presumably due to whole-tree morphology and changes in wood anatomy between branch and trunk wood[40–43]. A limiting factor in our study is the fact that we relied on measurements of branch and trunk wood from different individuals of the same species, thus not taking into account individual-level variation and coordination between traits. In species with amphi-Atlantic or global distributions, we also observed strong correlations of wood density across continents, indicating an important phylogenetic component in variation.

Methods

Data Assembly

To update the data base, we merged records from the previous compilation[5,13] with wood density values assembled from a wide variety of sources, including published and unpublished data, and added additional columns to reflect a new variety of measurements (such as intra-specific and intra-individual variation). Entries from the GWDD v.1. were transformed into the new format, reviewed and, when approved, transferred into the new database. Species names for the whole data base were spell-checked and taxonomically resolved with *The Plant List v.1.1*[44]. For more details on the data assembly process, please refer to S1.

Data harmonization

Wood density is defined as basic density, i.e. weight of the anhydrous sample (P_0) divided by the volume of the water-saturated sample (V_s).[45]. Many wood density measurements have been published in agroforestry studies, and wood density at a reference water content (w) is generally reported (D_w), where conventionally w = 12% or w = 15%. To harmonize the large range of wood density measurements collated from the literature, we converted air-dried values of wood density (D_w) into basic density (D_b). Recent research, based on a standardized set of measurements, indicates that simple conversion factors can be used across seed plants, such that $D_b = 0.828 \times D_{12}$[46], where D_{12} is the wood density measured at a reference 12% moisture, the most frequent reference in the forestry industry. This choice has serious implications for quantitative assessments and we were therefore led to correct some of the values in the original database. More details, including conversion factors for other moisture values, are available in S2.

Identifying functionally distinctive lineages

A recent study has used a phylogenetic tree together with plant trait data to estimate the most influential lineages in the distribution of traits[18]. Influential lineages are lineages without whom present day trait distributions would be very different. We here extended this approach to wood density values. We used a recently published phylogenetic tree of all seedplants [17], resolved it taxonomically with *The Plant List v.1.1*[44] and computed the Kolmogorov–Smirnov Importance index (KSI) as implemented in the package ksi[47] to find the five most functionally distinctive lineages. To match nodes to taxonomic categories, we used the package *taxonlookup*[48].

Mapping wood density based on taxonomic records

Finally, we analyzed how wood density is distributed across the world's forests. To do so, we extracted all records for gymnosperms and angiosperms from the *Global Biodiversity Information Facility (GBIF)*. We downsampled the data set to include only one species occurrence per squarekilometer, which avoids oversampling of particular species due to local trait collections, and applied the *CoordinateCleaner* package[49]. We then aggregated the resulting data set to a resolution of 25km, selected several layers of biophysical, climatic variables and soil and topographic variables and used random forest modelling[50] to predict grid cells where no wood density information was available (more details are provided in S4 and Table S2).

Analyzing intraspecific variation in traits

To analyze the relationship of densities across plant organs, we compared the species mean values of branch and trunk densities. We used Major Axis regression, as

implemented in the *lmodel2* package[51]. Furthermore, since our data set also includes species with amphi-Atlantic distributions (*Ceiba pentandra*, *Symphonia globulifera*) or species planted globally (e.g. *Eucalyptus globulus*, *Casuarina equisetifolia*, *Albizia lebbeck*, *Melia azedarach*), we grouped species with three or more measurements per continent and analyzed how their values varied across continents.

Literature

1. Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B. C., Duque, A., Eid, T., Fearnside, P. M., Goodman, R. C., Henry, M., Martínez-Yrízar, A., Mugasha, W. A., Muller-Landau, H. C., Mencuccini, M., Nelson, B. W., Ngomanda, A., Nogueira, E. M., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C. M., Saldarriaga, J. G. & Vieilledent, G. Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Chang. Biol.* **20,** 3177–3190 (2014).
2. Niklas, K. J. & Spatz, H. C. Worldwide correlations of mechanical properties and green wood density. *Am. J. Bot.* **97,** 1587–1594 (2010).
3. Zanne, A. E., Westoby, M., Falster, D. S., Ackerly, D. D., Loarie, S. R., Arnold, S. E. J. & Coomes, D. A. Angiosperm wood structure: Global patterns in vessel anatomy and their relation to wood density and potential conductivity. *Am. J. Bot.* **97,** 207–215 (2010).
4. Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., Fitzjohn, R. G., McGlinn, D. J., O'Meara, B. C., Moles, A. T., Reich, P. B., Royer, D. L., Soltis, D. E., Stevens, P. F., Westoby, M., Wright, I. J., Aarssen, L., Bertin, R. I., Calaminus, A., Govaerts, R., Hemmings, F., Leishman, M. R., Oleksyn, J., Soltis, P. S., Swenson, N. G., Warman, L. & Beaulieu, J. M. Three keys to the radiation of angiosperms into freezing environments. *Nature* **506,** 89–92 (2014).
5. Chave, J., Coomes, D., Jansen, S., Lewis, S. L., Swenson, N. G. & Zanne, A. E. Towards a worldwide wood economics spectrum. *Ecol. Lett.* **12,** 351–366 (2009).
6. Augspurger, C. K. & Kelly, C. K. Pathogen mortality of tropical tree seedlings: experimental studies of the effects of dispersal distance, seedling density, and light conditions. *Oecologia* **61,** 211–217 (1984).
7. Markesteijn, L., Poorter, L., Paz, H., Sack, L. & Bongers, F. Ecological differentiation

in xylem cavitation resistance is associated with stem and leaf structural traits. *Plant, Cell Environ.* **34,** 137–148 (2011).

8. Hacke, U. G., Sperry, J. S., Pockman, W. T., Davis, S. D. & McCulloh, K. A. Trends in wood density and structure are linked to prevention of xylem implosion by negative pressure. *Oecologia* **126,** 457–461 (2001).

9. Meinzer, F. C., Johnson, D. M., Lachenbruch, B., McCulloh, K. A. & Woodruff, D. R. Xylem hydraulic safety margins in woody plants: Coordination of stomatal control of xylem tension with hydraulic capacitance. *Funct. Ecol.* **23,** 922–930 (2009).

10. Kraft, N. J. B., Metz, M. R., Condit, R. S. & Chave, J. The relationship between wood density and mortality in a global tropical forest data set. *New Phytol.* **188,** 1124–1136 (2010).

11. Reich, P. B. The world-wide 'fast-slow' plant economics spectrum: A traits manifesto. *J. Ecol.* **102,** 275–301 (2014).

12. Poorter, L., Rozendaal, D. M. A., Bongers, F., de Almeida-Cortez, J. S., Almeyda Zambrano, A. M., Álvarez, F. S., Andrade, J. L., Villa, L. F. A., Balvanera, P., Becknell, J. M., Bentos, T. V., Bhaskar, R., Boukili, V., Brancalion, P. H. S., Broadbent, E. N., César, R. G., Chave, J., Chazdon, R. L., Colletta, G. D., Craven, D., de Jong, B. H. J., Denslow, J. S., Dent, D. H., DeWalt, S. J., García, E. D., Dupuy, J. M., Durán, S. M., Espírito Santo, M. M., Fandiño, M. C., Fernandes, G. W., Finegan, B., Moser, V. G., Hall, J. S., Hernández-Stefanoni, J. L., Jakovac, C. C., Junqueira, A. B., Kennard, D., Lebrija-Trejos, E., Letcher, S. G., Lohbeck, M., Lopez, O. R., Marín-Spiotta, E., Martínez-Ramos, M., Martins, S. V., Massoca, P. E. S., Meave, J. A., Mesquita, R., Mora, F., de Souza Moreno, V., Müller, S. C., Muñoz, R., Muscarella, R., de Oliveira Neto, S. N., Nunes, Y. R. F., Ochoa-Gaona, S., Paz, H., Peña-Claros, M., Piotto, D., Ruíz, J., Sanaphre-Villanueva, L., Sanchez-Azofeifa, A., Schwartz, N. B., Steininger, M. K.,

Thomas, W. W., Toledo, M., Uriarte, M., Utrera, L. P., van Breugel, M., van der Sande, M. T., van der Wal, H., Veloso, M. D. M., Vester, H. F. M., Vieira, I. C. G., Villa, P. M., Williamson, G. B., Wright, S. J., Zanini, K. J., Zimmerman, J. K. & Westoby, M. Wet and dry tropical forests show opposite successional pathways in wood density but converge over time. *Nat. Ecol. Evol.* **3,** 928–934 (2019).

13. Zanne, A. E., Lopez-Gonzalez, G., Coomes, D. A., Ilic, J., Jansen, S., Lewis, S. L., Miller, R. B., Swenson, N. G., Wiemann, M. C. & Chave, J. Data from: Global wood density database. *Dryad Digit. Repos.* (2009). doi:10.5061/dryad.234

14. Beech, E., Rivers, M., Oldfield, S. & Smith, P. P. GlobalTreeSearch: The first complete global database of tree species and country distributions. *J. Sustain. For.* **36,** 454–489 (2017).

15. Cardoso, D., Särkinen, T., Alexander, S., Amorim, A. M., Bittrich, V., Celis, M., Daly, D. C., Fiaschi, P., Funk, V. A., Giacomin, L. L., Goldenberg, R., Heiden, G., Iganci, J., Kelloff, C. L., Knapp, S., Cavalcante de Lima, H., Machado, A. F. P., dos Santos, R. M., Mello-Silva, R., Michelangeli, F. A., Mitchell, J., Moonlight, P., de Moraes, P. L. R., Mori, S. A., Nunes, T. S., Pennington, T. D., Pirani, J. R., Prance, G. T., de Queiroz, L. P., Rapini, A., Riina, R., Rincon, C. A. V., Roque, N., Shimizu, G., Sobral, M., Stehmann, J. R., Stevens, W. D., Taylor, C. M., Trovó, M., van den Berg, C., van der Werff, H., Viana, P. L., Zartman, C. E. & Forzza, R. C. Amazon plant diversity revealed by a taxonomically verified species list. *Proc. Natl. Acad. Sci.* **114,** 10695–10700 (2017).

16. Gilles, D., Zaiss, R., Blach-Overgaard, A., Catarino, L., Damen, T., Deblauwe, V., Dessein, S., Dransfield, J., Droissart, V., Duarte, M. C., Engledow, H., Fadeur, G., Figueira, R., Gereau, R. E., Hardy, O. J., Harris, D. J., de Heij, J., Janssens, S., Klomberg, Y., Ley, A. C., MacKinder, B. A., Meerts, P., van de Poel, J. L., Sonké, B.,

Sosef, M. S. M., Stévart, T., Stoffelen, P., Svenning, J.-C., Sepulchre, P., van der Burgt, X., Wieringa, J. J. & Couvreur, T. L. P. RAINBIO: a mega-database of tropical African vascular plants distributions. *PhytoKeys* **74,** 1–18 (2016).

17. Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.* **105,** 302–314 (2018).

18. Cornwell, W. K., Westoby, M., Falster, D. S., Fitzjohn, R. G., O'Meara, B. C., Pennell, M. W., McGlinn, D. J., Eastman, J. M., Moles, A. T., Reich, P. B., Tank, D. C., Wright, I. J., Aarssen, L., Beaulieu, J. M., Kooyman, R. M., Leishman, M. R., Miller, E. T., Niinemets, Ü., Oleksyn, J., Ordonez, A., Royer, D. L., Smith, S. A., Stevens, P. F., Warman, L., Wilf, P. & Zanne, A. E. Functional distinctiveness of major plant lineages. *J. Ecol.* **102,** 345–356 (2014).

19. GBIF Occurrence Download. at <GBIF.org>

20. Ter Steege, H., Pitman, N. C. A., Phillips, O. L., Chave, J., Sabatier, D., Duque, A., Molino, J. F., Prévost, M. F., Spichiger, R., Castellanos, H., Von Hildebrand, P. & Vásquez, R. Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature* **443,** 444–447 (2006).

21. Li, H. T., Yi, T. S., Gao, L. M., Ma, P. F., Zhang, T., Yang, J. B., Gitzendanner, M. A., Fritsch, P. W., Cai, J., Luo, Y., Wang, H., van der Bank, M., Zhang, S. D., Wang, Q. F., Wang, J., Zhang, Z. R., Fu, C. N., Yang, J., Hollingsworth, P. M., Chase, M. W., Soltis, D. E., Soltis, P. S. & Li, D. Z. Origin of angiosperms and the puzzle of the Jurassic gap. *Nat. Plants* **5,** 461–470 (2019).

22. Pennington, R. T., Lavin, M., Prado, D. E., Pendry, C. A., Pell, S. K. & Butterworth, C. A. Historical climate change and speciation: Neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. in *Philos. Trans. R. Soc. B Biol. Sci.* **359,** 515–537 (2004).

23. Crisp, M., Cook, L. & Steane, D. Radiation of the Australian flora: What can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities? in *Philos. Trans. R. Soc. B Biol. Sci.* **359,** 1551–1571 (2004).

24. Biffin, E., Lucas, E. J., Craven, L. A., Da Costa, I. R., Harrington, M. G. & Crisp, M. D. Evolution of exceptional species richness among lineages of fleshy-fruited Myrtaceae. *Ann. Bot.* **106,** 79–93 (2010).

25. Pennington, R. T., Lavin, M. & Oliveira-Filho, A. Woody Plant Diversity, Evolution, and Ecology in the Tropics: Perspectives from Seasonally Dry Tropical Forests. *Annu. Rev. Ecol. Evol. Syst.* **40,** 437–457 (2009).

26. Patiño, S., Herre, E. A. & Tyree, M. T. Physiological determinants of Ficus fruit temperature and implications for survival of pollinator wasp species: comparative physiology through an energy budget approach. *Oecologia* **100,** 13–20 (1994).

27. Cook, J. M. & Rasplus, J. Y. Mutualists with attitude: Coevolving fig wasps and figs. *Trends Ecol. Evol.* **18,** 241–248 (2003).

28. Lens, F., Schönenberger, J., Baas, P., Jansen, S. & Smets, E. The role of wood anatomy in phylogeny reconstruction of Ericales. *Cladistics* **23,** 229–254 (2007).

29. Read, J. & Stokes, A. Plant biomechanics in an ecological context. *Am. J. Bot.* **93,** 1546–1565 (2006).

30. Baas, P. & Schweingruber, F. H. Ecological Trends in the Wood Anatomy of Trees, Shrubs and Climbers from Europe. *IAWA J.* **8,** 245–274 (1987).

31. Chave, J., Muller-Landau, H. C., Baker, T. R., Easdale, T. A., Hans Steege, T. E. R. & Webb, C. O. Regional and phylogenetic variation of wood density across 2456 neotropical tree species. *Ecol. Appl.* **16,** 2356–2367 (2006).

32. Swenson, N. G. & Enquist, B. J. Ecological and evolutionary determinants of a key

plant functional trait: Wood density and its community-wide variation across latitude and elevation. *Am. J. Bot.* (2007). doi:10.3732/ajb.94.3.451

33. Barajas-Morales, J. Wood specific gravity in species from two tropical forests in Mexico. *IAWA J.* **8,** 143–148 (1987).

34. Wiemann, M. C. & Williamson, G. B. Geographic variation in wood specific gravity: Effects of latitude, temperature, and precipitation. *Wood Fiber Sci.* **34,** 96–107 (2002).

35. Martínez-Cabrera, H. I., Jones, C. S., Espino, S. & Jochen Schenk, H. Wood anatomy and wood density in shrubs: Responses to varying aridity along transcontinental transects. *Am. J. Bot.* **96,** 1388–1398 (2009).

36. Fortunel, C., Ruelle, J., Beauchêne, J., Fine, P. V. A. & Baraloto, C. Wood specific gravity and anatomy of branches and roots in 113 Amazonian rainforest tree species across environmental gradients. *New Phytol.* **202,** 79–94 (2014).

37. Ibanez, T., Chave, J., Barrabé, L., Elodie, B., Boutreux, T., Trueba, S., Vandrot, H. & Birnbaum, P. Community variation in wood density along a bioclimatic gradient on a hyper-diverse tropical island. *J. Veg. Sci.* **28,** 19–33 (2017).

38. Moreno-Martínez, Á., Camps-Valls, G., Kattge, J., Robinson, N., Reichstein, M., van Bodegom, P., Kramer, K., Cornelissen, J. H. C., Reich, P., Bahn, M., Niinemets, Ü., Peñuelas, J., Craine, J. M., Cerabolini, B. E. L., Minden, V., Laughlin, D. C., Sack, L., Allred, B., Baraloto, C., Byun, C., Soudzilovskaia, N. A. & Running, S. W. A methodology to derive global maps of leaf traits using remote sensing and climate data. *Remote Sens. Environ.* **218,** 69–88 (2018).

39. Butler, E. E., Datta, A., Flores-Moreno, H., Chen, M., Wythers, K. R., Fazayeli, F., Banerjee, A., Atkin, O. K., Kattge, J., Amiaud, B., Blonder, B., Boenisch, G., Bond-Lamberty, B., Brown, K. A., Byun, C., Campetella, G., Cerabolini, B. E. L., Cornelissen,

J. H. C., Craine, J. M., Craven, D., De Vries, F. T., Díaz, S., Domingues, T. F., Forey, E., González-Melo, A., Gross, N., Han, W., Hattingh, W. N., Hickler, T., Jansen, S., Kramer, K., Kraft, N. J. B., Kurokawa, H., Laughlin, D. C., Meir, P., Minden, V., Niinemets, Ü., Onoda, Y., Peñuelas, J., Read, Q., Sack, L., Schamp, B., Soudzilovskaia, N. A., Spasojevic, M. J., Sosinski, E., Thornton, P. E., Valladares, F., Van Bodegom, P. M., Williams, M., Wirth, C., Reich, P. B. & Schlesinger, W. H. Mapping local and global variability in plant trait distributions. *Proc. Natl. Acad. Sci. U. S. A.* **114,** E10937–E10946 (2017).

40. Schuldt, B., Leuschner, C., Brock, N. & Horna, V. Changes in wood density, wood anatomy and hydraulic properties of the xylem along the root-to-shoot flow path in tropical rainforest trees. *Tree Physiol.* **33,** 161–174 (2013).

41. Fajardo, A. Insights into intraspecific wood density variation and its relationship to growth, height and elevation in a treeline species. *Plant Biol.* **20,** 456–464 (2018).

42. Sarmiento, C., Patiño, S., Timothy Paine, C. E., Beauchêne, J., Thibaut, A. & Baraloto, C. Within-individual variation of trunk and branch xylem density in tropical trees. *Am. J. Bot.* **98,** 140–149 (2011).

43. Swenson, N. G. & Enquist, B. J. The relationship between stem and branch wood specific gravity and the ability of each measure to predict leaf area. *Am. J. Bot.* **95,** 516–519 (2008).

44. The Plant List. The Plant List. Version 1.1. *Internet* (2013).

45. Williamson, G. B. & Wiemann, M. C. Measuring wood specific gravity...correctly. *Am. J. Bot.* **97,** 519–524 (2010).

46. Vieilledent, G., Fischer, F. J., Chave, J., Guibal, D., Langbour, P. & Gérard, J. New formula and conversion factor to compute basic wood density of tree species

using a global wood technology database. *Am. J. Bot.* **105,** 1653–1661 (2018).

47. Tempo and Mode in Plant Trait Evolution Working Group. ksi: Phylogeny-based Komogorov-Smirnov Importance Statistic. R package version 0.1-3. (2019).

48. Cornwell, W., FitzJohn, R. & Pennell, M. taxonlookup: A dynamically-updating versioned taxonomic resource for vascular plants. R package version 1.1.5. (2016).

49. Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D., Farooq, H., Herdean, A., Ariza, M., Scharn, R., Svantesson, S., Wengström, N., Zizka, V. & Antonelli, A. CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods Ecol. Evol.* **10,** 744–751 (2019).

50. Liaw, A. & Wiener, M. Classification and Regression by randomForest. *R News* **2,** 18–22 (2002).

51. Legendre, P. lmodel2: Model II Regression. R package version 1.7-3. (2018).

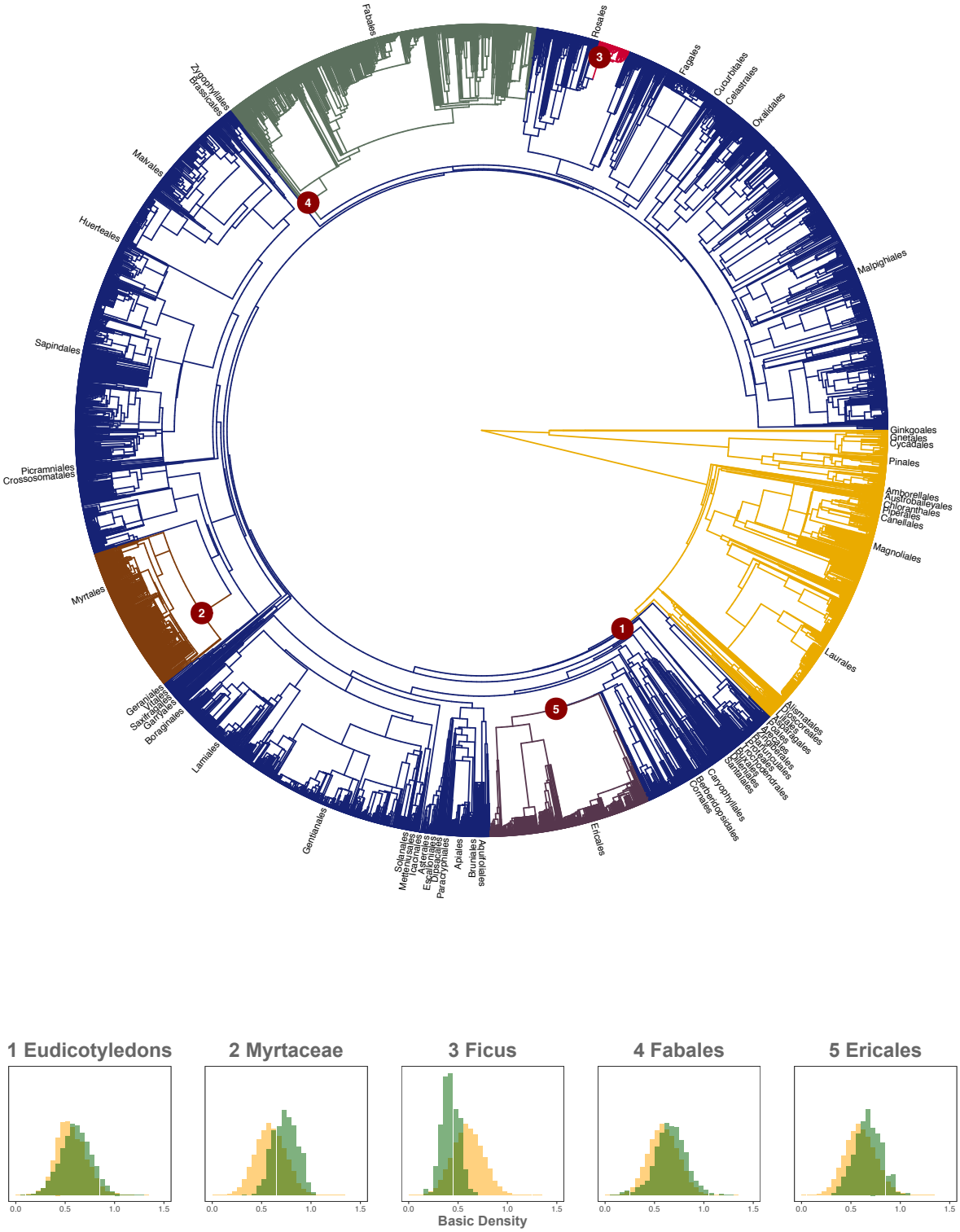


Figure 1: The five functionally most distinctive clades based on wood density from the GWDD v.2. In the upper panel, we show the Smith & Brown (2018) phylogenetic tree for seedplants, matched to the GWDD v.2 on 13,184 extant species. The most functionally distinctive clades given in the following order: Eudicotyledons (1),

Myrtaceae (2), Ficus (3), Fabales (4), Ericales (5). Taxonomic orders are provided, matched to species and jittered for better visualization. In the lower panel, the corresponding shifts in wood density are shown, ranked from left to right in functional distinctiveness. Note that histogram densities are normalized to 1.

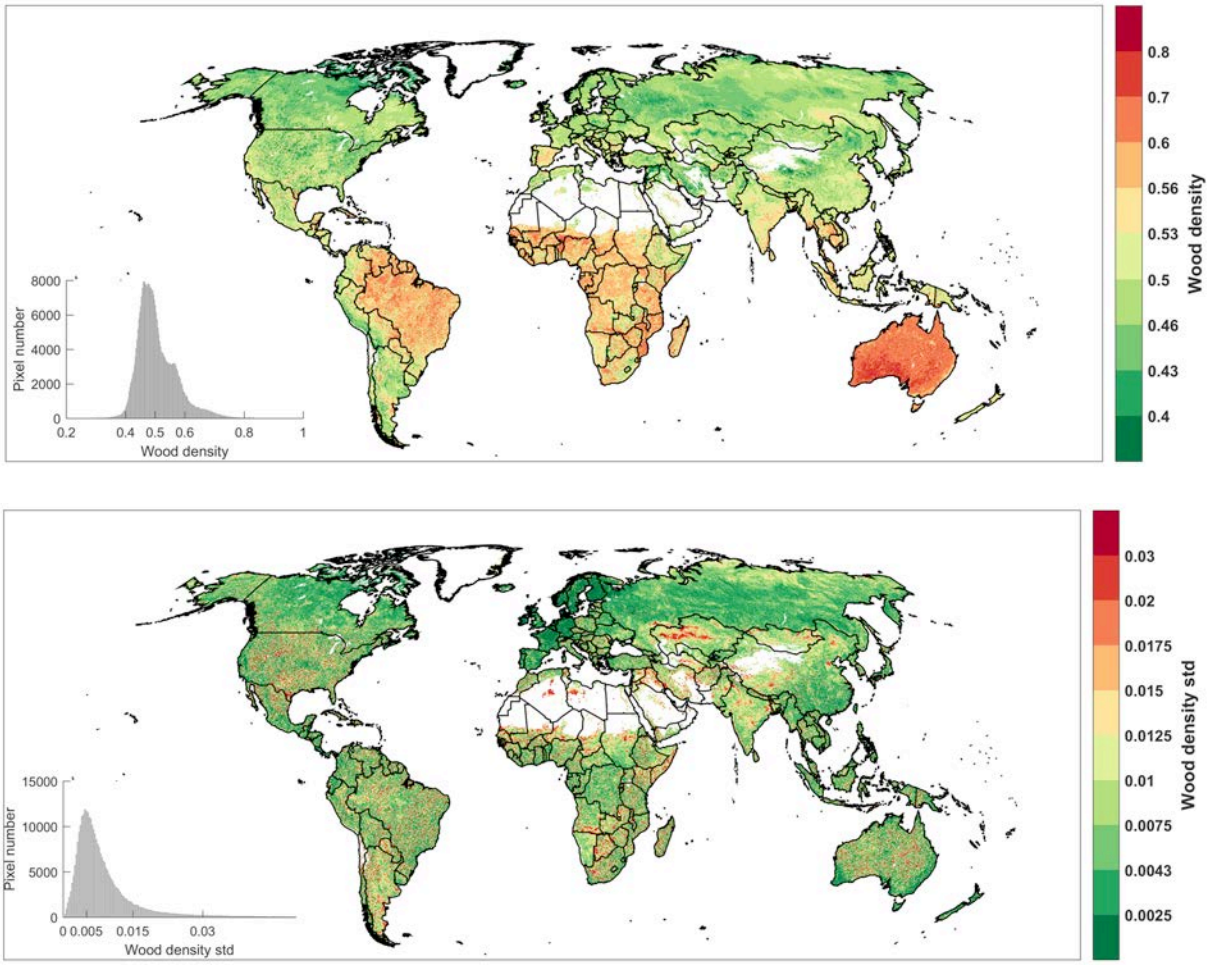


Figure 2: A global map of wood density based on taxonomic records. Shown is the predicted distribution of wood density across the globe (upper map) as well as standard deviation around the predicted means (lower map). Wood density information is based on the GWDD and occurrence records derive from the *Global Biodiversity Information Facility (GBIF)*. Values have been averaged across 25 sqkm grid cells. Where no values were available, they were predicted from a random forest model relying on multiple environmental layers (biophysical, climate, soil and topography, cf. S4 and Table S2).

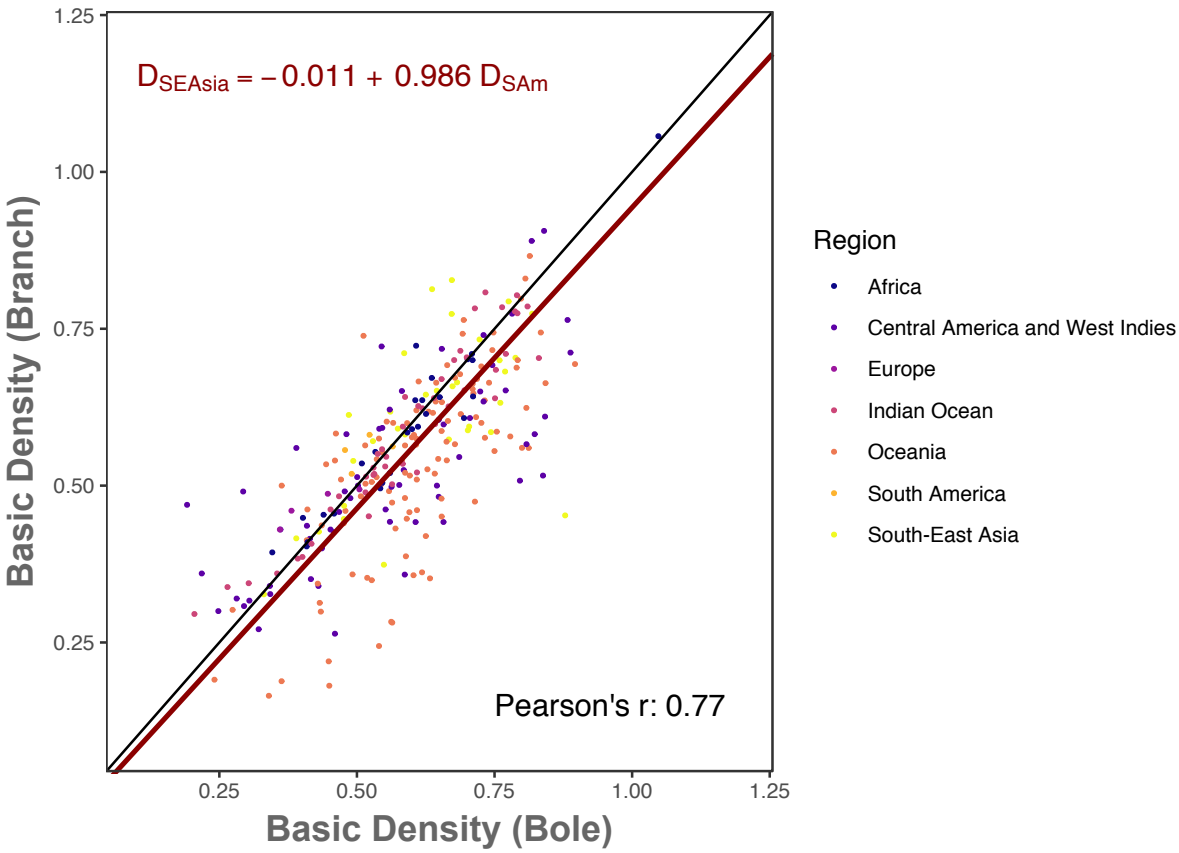


Figure 3: Intraspecific variability in wood density – branch and trunk wood. Shown are mean branch wood densities plotted against mean trunk densities on a per-species basis. The geographic origin of the species is indicated through the colors of each point. The red line shows the result of a Major Axis regression, with the coefficients given in the upper left of the panel and Pearson's correlation coefficient provided in the lower right of the panel. The black line is the reference line (intercept of zero, slope of 1).

Supplementary Information to
Global patterns and evolutionary trends in wood density

S1: Assembling the GWDD v.2

New compilation

A key part in assembling the new data base was the inclusion of new records. To do so, we contacted a large number of colleagues who were involved in plant trait or carbon assessment research, and invited them to collaborate on improving the coverage of the GWDD. We thus assembled wood density data from a wide range of novel sources, including published records, but also unpublished data. Recent publications provided a greater diversity of data. To reflect this high diversity, we documented additional fields (Table S1).

Species names were harmonized by correcting spelling-errors via online resources such as *Tropicos*, based on the *taxize* package[1], and taxonomic name resolution via *The Plant List v.1.1*[2], as embedded in the package *taxonstand*[3]. Where names could not be found with *The Plant List*, we carried out an additional manual search and correction. Finally, we updated the taxonomic classification, indicating the status of each record via *Accepted* and *Unresolved*.

Airdry and ovendry densities were then converted via the conversion factors at the respective water content. When the water content of airdry densities was not reported, we assumed w = 12%. While this might introduce biases in the case of higher or lower

water contents, these biases will likely have little impact on the overall estimates, because differences are typically less than 5%.

Each record was assigned a weight that corresponds to the number of trees that have been sampled. Where the wood density value represents an aggregate sample, but the exact sampling numbers are not known (such as values provided by wood technology compendia), we assume an average weight of 4. This will underestimate the weight of some well-tested tree species and overestimate the weight for species where only one or two specimen have been tested, but provides a useful rule of thumb for aggregating species records.

Correction of GWDD v.1

We corrected the values of GWDD v.1 when necessary. We reported the original D_w and the corresponding w values. Wherever other conversion factors, based on local estimates, had been used, we backconverted with these factors. To improve the reproducibility of conversions, the GWDD v.2 provides the new and old and values.

We also thoroughly checked the reference list and removed minor inconsistencies. Some records in GWDD v.1 were obtained via the ICRAF data base (http://db.worldagroforestry.org/wd, last accessed on 29 July 2019), itself a collection of primary resources. More recent versions of the ICRAF data base have incorporated and overwritten records using the GWDD v.1. Since this introduces a circularity, we remove records where the quality of the original data could not be ascertained. We also replaced in full the Sallenave data base and related data[4,5], included in GWDD v.1, by the more comprehensive CIRAD wood technology data base[6].

S2: Conversion factors

In a nutshell, the explanation of the conversion factors is as follows [6]: Wood is composed of a variable fraction of water, some of which is free to move in the vessels, and the rest associated to wood fibers and cells. During drying, free water is fully lost at the *fiber saturation point*. Beyond this point, volume shrinks. The moisture at fiber saturation is quite variable across species from ca 10% to 50%, with a typical value of ~30%[6,7]. If V_s is the volume of the sample at fiber saturation S, and V_0 is the volume when the sample has lost all of its water, the volumetric shrinkage, or retractability, is the percent loss in volume $R_T = (V_s - V_0)/V_s \times 100$, which varies from 5% to 25% across species (Vieilledent *et al.*, 2018). This analysis results in the following conversion formula:

$$D_b = \frac{1 - (R/100) \times (S - w)}{1 + w/100} \times D_w$$

The verified data base contains measurements of S, R, and D_{12} for 3,832 individual trees, based on >10 samples per individual and measurements at four different moisture contents w (from 18% to 0%). Using these measurements and the formula, it was possible to obtain D_b and D_{12} for each sampled tree, and also to derive D_w for any w. Each D_w was regressed against D_b, and, by forcing the regression through zero, obtained simple conversion factors for any w. In particular, we obtained the following conversion factors: 0.819 for airdry densities w = 15%, 0.828 for airdry densities at w = 12%, 0.840 at airdry densities w = 8%, and 0.868 for ovendry densities (or w = 0%). These various conditions span the range of the published values of wood density

S3: Geographic coverage (details)

The GWDD v.2 contains 71,028 records, of which 66,685 are identified at the species level, comprising 14,698 taxa (Figure S1). Of these, 13,124 of the species names are accepted as of *The Plant List 1.1*, a further 1,326 species are unresolved, and 249 species missed a reference in *The Plant List*. For 8,006 species, more than one record is available, including records from several sources in 5,878 species. The GWDD v.2 comprises 2,933 genera. Accounting for a further 4,343 records identified only at the genus level, the data base gains an additional 38 genus-level records and thus has records on 2,951 genera. The majority of wood density measurements are either directly D_b (> 45% of records), or airdry densities D_w with a water content w of 8-15% (also >45% of records). Ovendry densities amounted to 7% of the records, with green densities and dry fraction accounting for < 1% of the records.

The GWDD v.2. matches 11,838 species out of the 60,011 tree species recorded in the GlobalTreeSearch database[8], or ~20%, and 2,509 out of the 4,277 genera, or ~60% (Figure S3). Sampling coverage varied geographically reaching 75-85% for most European countries, from Scandinavia down to the Balkan states. Sample coverage decreased at low latitudes. Three regions had coverage ~20%: 1) the South Pacific (Fiji, Papua New Guinea, New Caledonia, French Polynesia), 2) the Caribbean (Cuba, Haiti, Dominican Republic), and 3) the Arab Peninsula (Yemen and Oman). Sampling coverage also varied regionally. In Amazonia, the Guiana Shield had higher coverage (~50%) than Brazil (~29%). Similarly, in tropical Africa, West African species are generally better sampled (50-60%) than East African species (~30%).

These tendencies were confirmed by local data bases. In Amazonia, a plant checklist was recently published[9], the GWDD v.2 matched 2,011 out of 6,727 tree species records, or ~30% (or 75% at genus level). Genera such as *Platycarpus* and *Kutchubaea*, with 11 species each, are the largest missing genera. The highest sample coverage for Amazonian trees was in the Guiana Shield (with ~60% of species covered) and the lowest coverage in Brazil (with ~36% of species covered).

For tropical Africa, we also used a recent checklist[10]. Our database matched 1,479 out of 3,662 tree species records (80% of the genera). The largest missing genera are *Raphia* and *Uvariodendron*. West African forests were better explored in terms of wood density (80% coverage) than East African ones (coverage varying between 20-50%). GWDD v.2 represents a particularly high improvement in African countries where sampling coverage went up by >40%, and Madagascar, from only 3% of species to 26%.

S4: Mapping wood density with random forest algorithm

Environmental layers

To predict wood density from environmental variables, we assembled 54 environmental layers, including 7 biophysical, 26 climatic, and 21 soil and topographic layers (Table S2). All covariate layers were aggregated to 25km resolution using GRASS GIS.

We used the ClustOfVar R package to select the most relevant layers for each of the three layer groups[11,12]. The ClustOfVar identifies the environmental layers which maximize the variations in the GBIF observed wood density values in the environmental space. 3 biophysical layers (Modis PET, GPP, and PALSAR2 HH polarization signals), 4 climatic layers (bioclimatic variable 5, 6, 16 and 17), and 5 topographic and soil layers

(probability of occurrence of R horizon, Soil organic carbon density, soil pH, silt content, and slope) were selected.

Modelling

We used the random forest modelling technique[13] to predict wood density at a global scale. The 10-fold cross-valiation approach was used to test the goodness of performance of the model. To this end, the GBIF observed dataset was randomly divided into 10 subsets. 10 RF models were built, each time using nine subsets for model building and one for validation. The performance of the RF approach was validated by regressing all predicted and observed values (S7).

Field	Description	Typical values
species	the species binomial, after spell-checking and taxonomic name resolution	
genus	the genus, after spell-checking and taxonomic name resolution	
epithet	the species epithet, after spell-checking and taxonomic name resolution	
family	the family, after spell-checking and taxonomic name resolution	
authority	the taxonomic authority, after spell-checking and taxonomic name resolution	
wsg	the basic wood density value, as obtained from conversion or direct measurement (unitless)	
species_reference	the species binomial supplied to the GWDD v.2	
value_reference	the wood density value supplied to the GWDD v.2	
backtransformed	is the value backconverted from the GWDD v.1	0 or 1
quantity_reference	the type of wood density that was measured	"Airdry","Ovendry","Basic"
moisture_airdry	the moisture at which wood was considered airdry (%)	
wsg_conversion	the conversion factor used to convert airdry or ovendry densities	
tree_agg	is the wood density measurement aggregated across several trees or not?	0 or 1
trees_sampled	the number of trees sampled for the measurement	
weight_value	the weight for averaging density values across measurements	equal to *trees_sampled*, if *trees_sampled* is NA, then set to 4 to reflect an average sampling value of 3-5 trees
location_sample	the within-tree location of the sample	"root","bole","branch","twig", or combinations thereof
type_tissue	the type of woody tissue sampled	"sapwood","heartwood","bark", "total (bark to pith)", or combinations thereof
type_sample	the type of sample	"core" or "disk"
instrument	the instrument used to obtain wood density	
temperature_drying	the temperature at which samples have been dried	
source_short	the short name of the source	
source_long	the full name of the source	
site	site of measurement, at various levels of precision	
latitude/longitude	coordinates of the site of measurement	
country	country of measurement	
region	one of nine regions	"South America", "Central America and West Indies", "North America", "Africa","Indian Ocean", "Europe", "Asia", "South-East Asia", "Oceania"
type_forest	type of forest	local ecological descriptors
id_dboriginal	the id of the sample in the original database	
id_plant/age/dbh	individual-plant level information on IDs in the source data, plant age and plant diamaeter	
experiment	if the data have been collected during an experiment 1, else 0, details in the next field	0 or 1
experiment design	the design of the experiment	

Table S1. Fields of the GWDD v.2. Given are the names of the fields, as they appear in the GWDD v.2, a short description of the field, and – if applicable –, the number of typical and possible values that the field can contain.

ID	Layer name	Explanation	Source
1	MOD ET 2000-2013 mean	MODIS evapotranspiration, 2000-2013 yearly mean	
2	MOD LE 2000-2013 mean	MODIS evaporation, 2000-2013 yearly mean	
3	**MOD PET 2000-2013 mean**	**MODIS potential evapotranspiration, 2000-2013 yearly mean**	
4	**MOD GPP 2000-2015 mean**	**MODIS GPP, 2000-2015 yearly mean**	
5	MOD NPP 2000-2015 mean	MODIS NPP, 2000-2015 yearly mean	
6	**PALSAR2 HH 20172018 man**	**PALSAR2 radar data, HH polarization signal, 2017-2018 yearly mean**	
7	PALSAR2 HV 20172018 man	PALSAR2 radar data, HV polarization signal, 2017-2018 yearly mean	
1	CWD	Climatic water deficit	Chave et al. 2015. GCB
2	WC2.0, srad12mean	world climate2, solar radiance, 12 months mean	
3	WC2.0, srad12std	world climate2, solar radiance, 12 months std	
4	WC2.0, vapr12mean	world climate2, vapor pressure, 12 months mean	
5	WC2.0, vapr12std	world climate2, vapor pressure, 12 months std	
6	WC2.0, wind12mean	world climate2, wind speed, 12 months mean	
7	WC2.0, wind12std	world climate2, wind speed, 12 months std	
8	WC2.0, bio1	world climate2, bioclimatic 1	
9	WC2.0, bio2	world climate2, bioclimatic 2	
10	WC2.0, bio3	world climate2, bioclimatic 3	
11	WC2.0, bio4	world climate2, bioclimatic 4	
12	**WC2.0, bio5**	**world climate2, bioclimatic 5**	
13	**WC2.0, bio6**	**world climate2, bioclimatic 6**	
14	WC2.0, bio7	world climate2, bioclimatic 7	
15	WC2.0, bio8	world climate2, bioclimatic 8	
16	WC2.0, bio9	world climate2, bioclimatic 9	
17	WC2.0, bio10	world climate2, bioclimatic 10	
18	WC2.0, bio11	world climate2, bioclimatic 11	
19	WC2.0, bio12	world climate2, bioclimatic 12	
20	WC2.0, bio13	world climate2, bioclimatic 13	
21	WC2.0, bio14	world climate2, bioclimatic 14	
22	WC2.0, bio15	world climate2, bioclimatic 15	
23	**WC2.0, bio16**	**world climate2, bioclimatic 16**	
24	**WC2.0, bio17**	**world climate2, bioclimatic 17**	
25	WC2.0, bio18	world climate2, bioclimatic 18	
26	WC2.0, bio19	world climate2, bioclimatic 19	
1	BDRICM_M_250m_ll.tif	Depth to bedrock (R horizon) up to 200 cm	https://soilgrids.org
2	**BDRLOG_M_250m_ll.tif**	**Probability of occurrence of R horizon**	**https://soilgrids.org**
3	BDTICM_M_250m_ll.tif	Absolute depth to bedrock (in cm)	https://soilgrids.org
4	BLDFIE_M_sl1_250m_ll.tif	Bulk density (fine earth, oven dry) in kg / cubic-meter at depth 0-2 m mean	https://soilgrids.org
5	CECSOL_M_sl1_250m_ll.tif	Cation exchange capacity of soil in cmolc/kg at depth 0-2m mean	https://soilgrids.org
6	CLYPPT_M_sl1_250m_ll.tif	Clay content (0-2 micro meter) mass fraction in % at depth 0-2m mean	https://soilgrids.org
7	CRFVOL_M_sl1_250m_ll.tif	Coarse fragments volumetric in % at depth 0-2m mean	https://soilgrids.org
8	**OCDENS_M_sl1_250m_ll.tif**	**Soil organic carbon density in kg per cubic-m at depth 0-2m mean**	**https://soilgrids.org**
9	OCSTHA_M_100cm_250m_ll.tif	Soil organic carbon stock in tons per ha for depth interval 0-1m	https://soilgrids.org
10	**PHIHOX_M_sl1_250m_ll.tif**	**Soil pH x 10 in H2O at depth 0-2m mean**	**https://soilgrids.org**
11	PHIKCL_M_sl1_250m_ll.tif	Soil pH x 10 in KCl at depth 0-2m mean	https://soilgrids.org
12	**SLTPPT_M_sl1_250m_ll.tif**	**Silt content (2-50 micro meter) mass fraction in % at depth 0-2m mean**	**https://soilgrids.org**
13	SNDPPT_M_sl1_250m_ll.tif	Sand content (50-2000 micro meter) mass fraction in	https://soilgrids.org

		% at depth 0-2m mean	
14	TAXNWRB_250m_ll.tif	Predicted WRB 2006 subgroup classes (as integers)	https://soilgrids.org
15	TAXOUSDA_250m_ll.tif	Predicted USDA 2014 suborder classes (as integers)	https://soilgrids.org
16	aspect_cosin	Topographic aspect, calculated using cosin method	https://www.earthenv.org/texture
17	aspect_eastness	Topographic aspect, eastness	https://www.earthenv.org/texture
18	aspect_northness	Topographic aspect, northness	https://www.earthenv.org/texture
19	elevation	Elevation	https://www.earthenv.org/texture
20	roughness	Roughness	https://www.earthenv.org/texture
21	**slope**	**Slope**	**https://www.earthenv.org/texture**

Table S2: Layers used for creation of global wood density map. Shown are biophysical layers (blue), climate layers (orange) and soil and topographic layers (green). Layers that have been retained after a preliminary analysis are marked in darker colors and in bold.

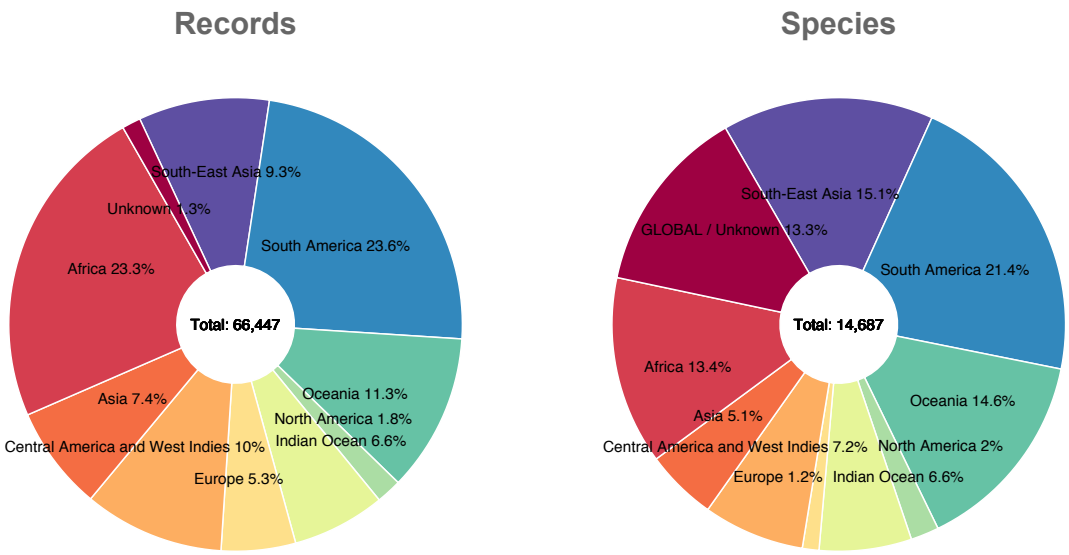


Figure S1. Geographic distribution of records and species: Shown is the geographic distribution of measured or converted basic density values in the GWDD v.2. Excluded are records identified only at the genus level. The lefthand panel describes the repartition of all records across continents and subcontinental regions, irrespective of sampling size, the righthand panel the same distribution in terms of species. A small number of records could not be clearly attributed to a region ("Unknown"). In the righthand panel, these records are taken together with species that occur across several of the displayed regions ("GLOBAL") and reflect the continuity between regions such as South America and Central America or South-East Asia and the rest of Asia.

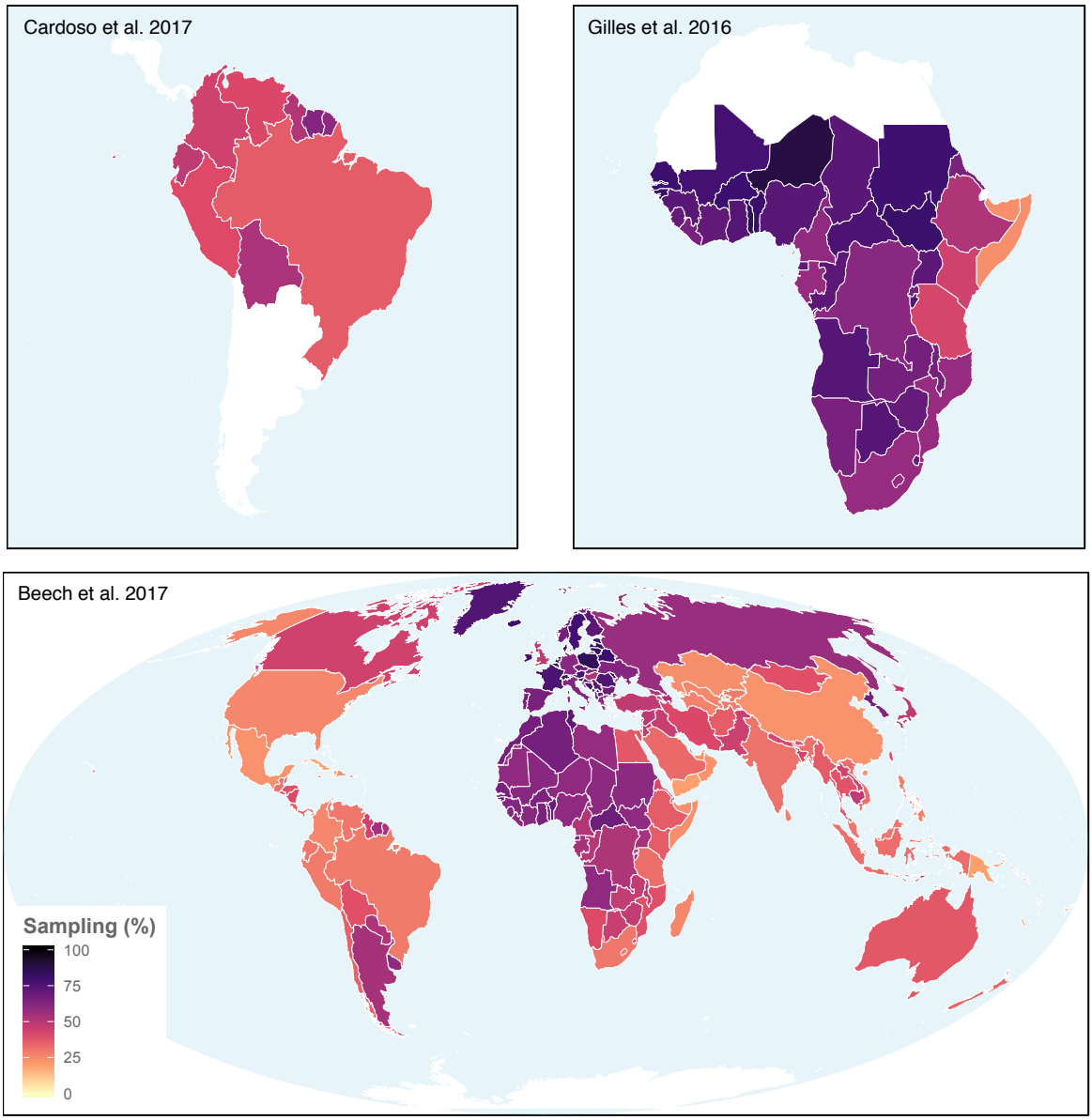


Figure S2: Estimated wood density sampling coverage of tree species around the world in the GWDD v.2. This figure shows estimates of wood density sampling coverage of tree species on a per-country level. Estimates are based on a comparison of the GWDD v.2 with three recently published lists of tree species, one for Amazonia (Cardoso et al. 2017), one covering mainly tropical Africa (Gilles et al. 2016), and a global compilation that uses a wider definition of what constitutes a tree (Beech et al. 2017). In order to avoid artificially high or low sampling numbers, countries with less than 10 recorded tree species have been excluded from the figure in the upper right

panel (African database). While there is a clear difference in absolute numbers between regional databases and the global tree list, with the GWDD v.2. covering the regional lists and their more narrow tree density definitions better, similar qualitative pictures emerge across databases. We find, for example, higher sampling coverage in the Guyanas, compared to the rest of Amazonia, and higher coverage in West Africa than in East Africa.

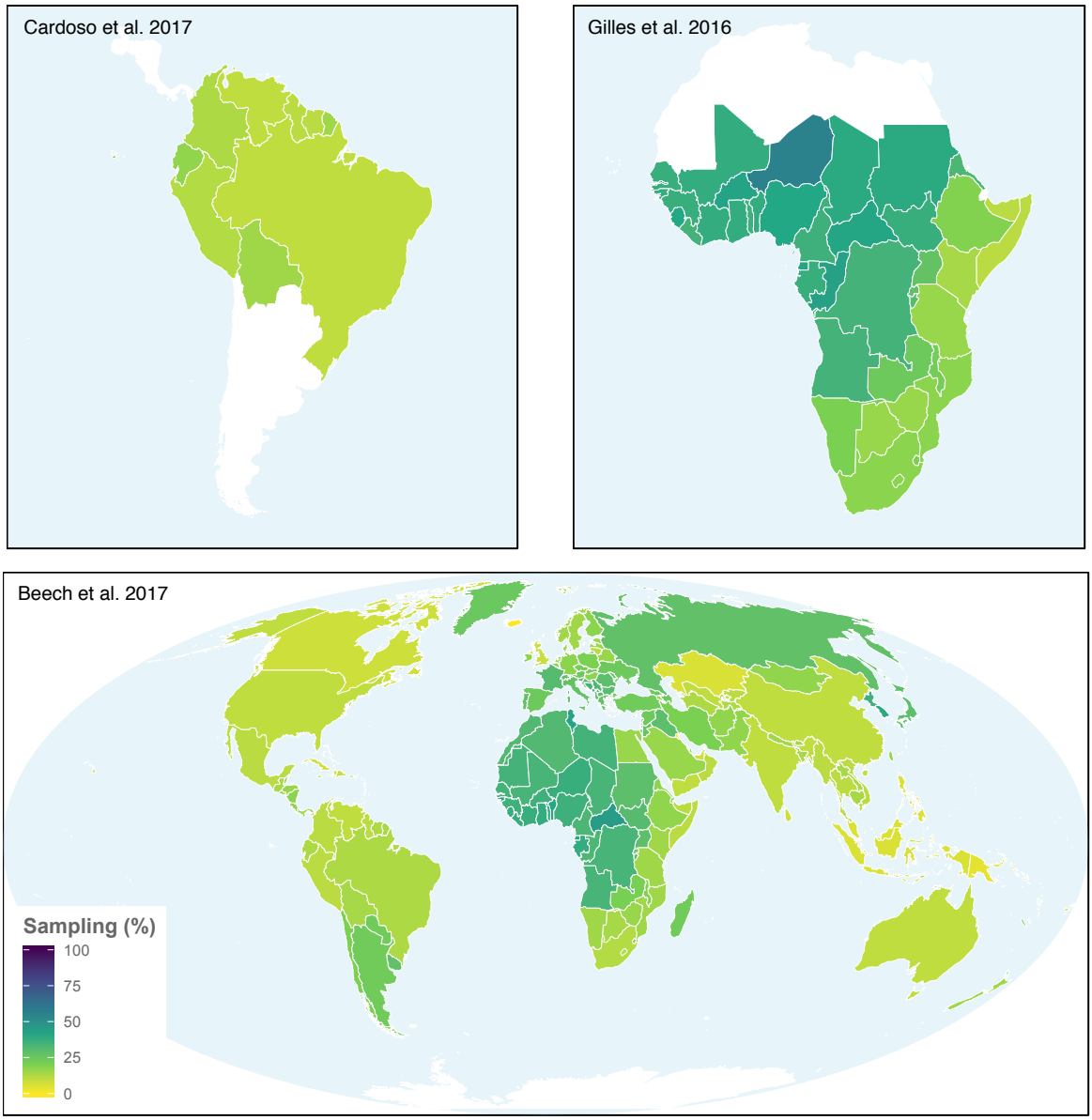


Figure S3: Improvement of wood density sampling coverage of tree species around the world from GWDD v.1 to GWDD v.2. This figure shows estimates of how wood density sampling coverage of tree species has improved between the original GWDD and the newest update, based on a comparison with three recently published lists of tree species, one for Amazonia (Cardoso et al. 2017), one covering mainly tropical Africa (Gilles et al. 2016), and a global compilation that uses a wider definition of what constitutes a tree (Beech et al. 2017). In order to avoid artificially high or low sampling numbers, countries with less than 10 recorded tree species have been excluded from the figure in the upper right panel (African database). Improvement is

given in percentage with respect to the total number of species recorded in each of the three data bases.

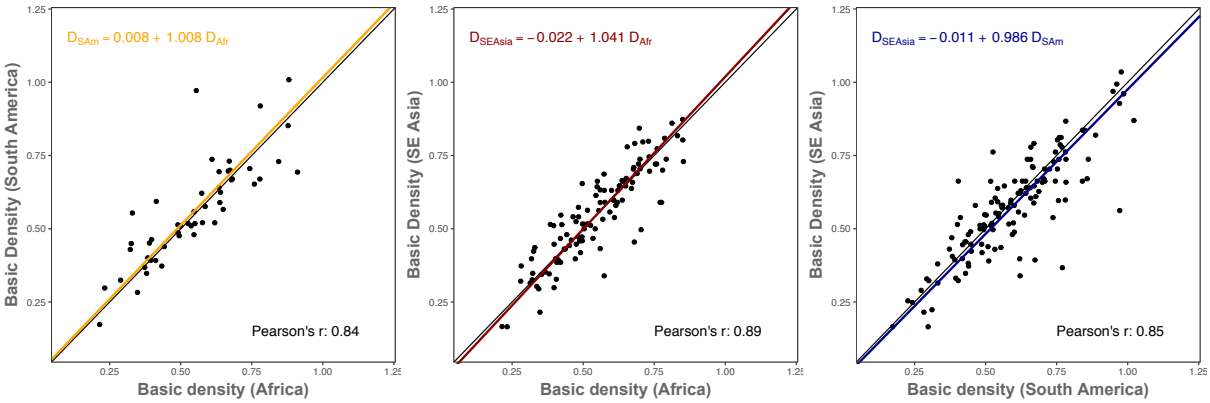


Figure S4: Pairwise comparison of mean species densities between three continents. This figure shows a pairwise plot of wood densities averaged for species that occur across at least two of the following three (sub-)continents: Africa, South America, South-East Asia and have at least 3 measurements recorded on each of them. For the purposes of this comparison, South America is considered to include Central America and the Carribbean, and South-East Asia is supplemented with species from Oceania. Each dot represents one species. The colored lines are major axis regression fits, the obtained formula is provided in the upper left of each panel, the Pearson correlation coefficient in the lower right of each panel. Also shown is the reference line in black (intercept of zero, slope of one).

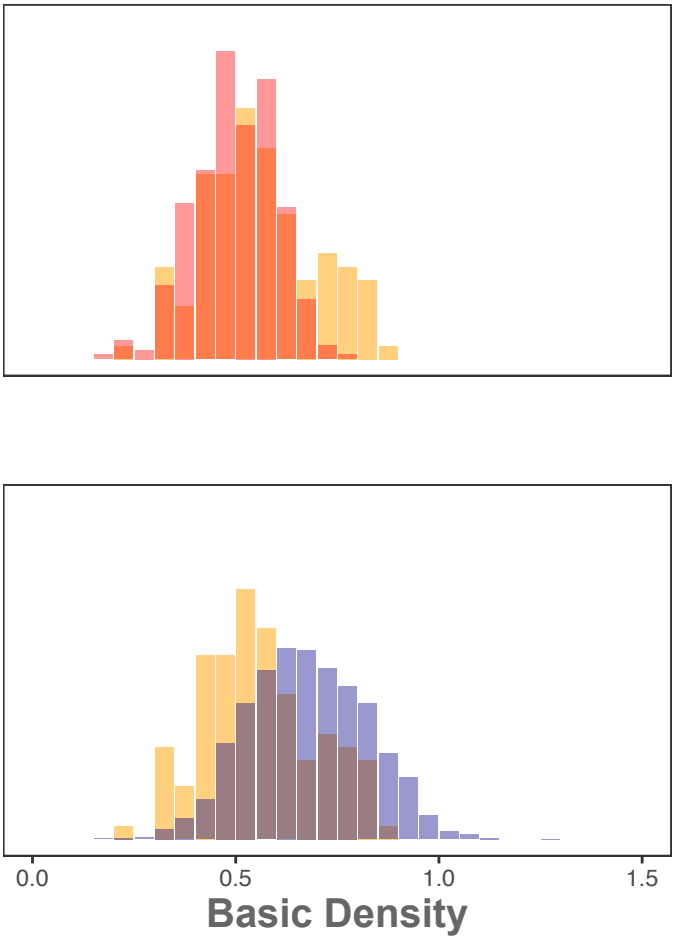


Figure S5: Wood density variation within the order Ericales, compared to variation within Cornales. Shown are wood density values of different clades within Ericales (red in the upper panel and blue in the lower panel) compared to the order Cornales (in yellow). The upper panel shows genera belonging to the so-called primitive wood structure type that has developed in temperate and tropical montane forests and largely conserves woody properties compared to the common ancestor of Ericales and Cornales [14], and the lower panel the remaining clades that have spread to lowland tropical forests. A clear shift towards higher wood density values is visible in the lower panel.

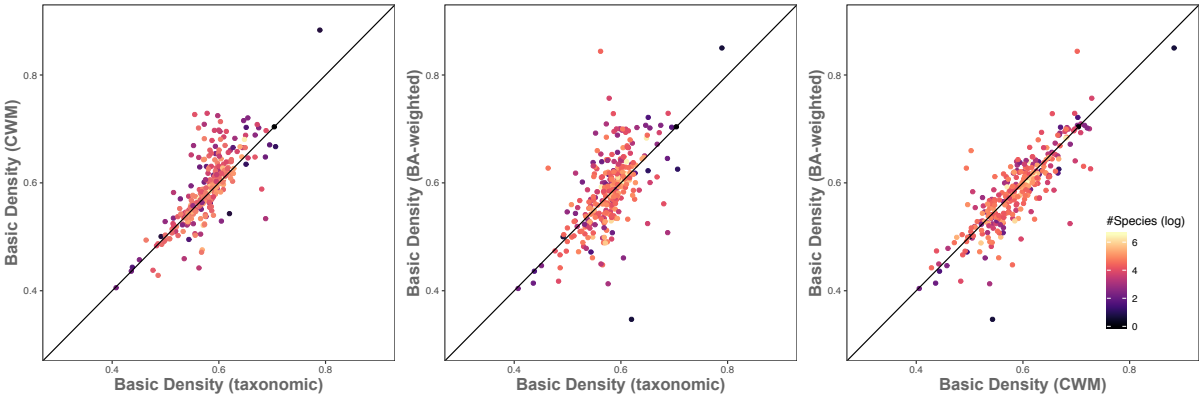


Figure S6: Relation between taxonomic, community weighted and basal-area weighted wood density. Shown is the relationship between taxonomically averaged wood density, community-weighted wood density and basal-area weighted wood density for trees > 10cm in trunk diameter, as inferred from a large collection of plot- and transect-based wood density measurements. Colouring shows the number of species per plot or transect (on log scales). The plots and transects data have been accumulated from our own field data and various openly available sources[15,16,25–30,17–24].

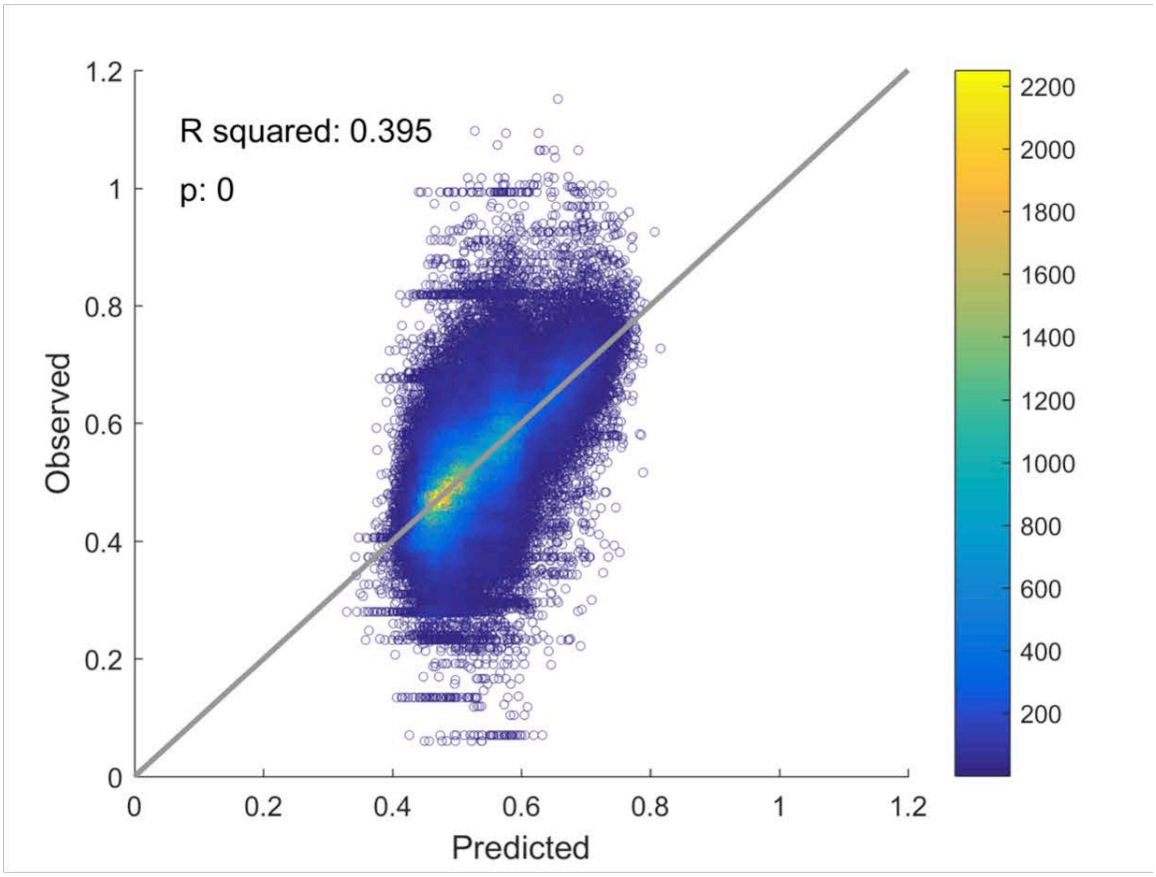


Figure S7: 10-fold cross-validation of random forest modelling. Shown are observed vs. predicted wood density values from the ten folds, as well as R^2.

Bibliography

1. Chamberlain, S. A. & Szöcs, E. taxize: taxonomic search and retrieval in R. *F1000Research* **2,** 191 (2013).
2. The Plant List. The Plant List. Version 1.1. *Internet* (2013).
3. Cayuela, L., Granzow-de la Cerda, Í., Albuquerque, F. S. & Golicher, D. J. Taxonstand: An r package for species names standardisation in vegetation databases. *Methods Ecol. Evol.* **3,** 1078–1083 (2012).
4. Détienne, P., Jacquet, P. & Mariaux, A. *Manuel d'Identification de Bois Tropicaux 3. Guyane française.* (Centre Technique Forestier Tropical, 1982).
5. Détienne, P. & Jacquet, P. *Atlas d'Identification des Bois de l'Amazonie et des Régions Voisines*. (Centre Technique Forestier Tropical, 1983).
6. Vieilledent, G., Fischer, F. J., Chave, J., Guibal, D., Langbour, P. & Gérard, J. New formula and conversion factor to compute basic wood density of tree species using a global wood technology database. *Am. J. Bot.* **105,** 1653–1661 (2018).
7. Berry, S. L. & Roderick, M. L. Plant-water relations and the fibre saturation point. *New Phytol.* **168,** 25–37 (2005).
8. Beech, E., Rivers, M., Oldfield, S. & Smith, P. P. GlobalTreeSearch: The first complete global database of tree species and country distributions. *J. Sustain. For.* **36,** 454–489 (2017).
9. Cardoso, D., Särkinen, T., Alexander, S., Amorim, A. M., Bittrich, V., Celis, M., Daly, D. C., Fiaschi, P., Funk, V. A., Giacomin, L. L., Goldenberg, R., Heiden, G., Iganci, J., Kelloff, C. L., Knapp, S., Cavalcante de Lima, H., Machado, A. F. P., dos Santos, R. M., Mello-Silva, R., Michelangeli, F. A., Mitchell, J., Moonlight, P., de Moraes, P. L. R., Mori, S. A., Nunes, T. S., Pennington, T. D., Pirani, J. R., Prance, G. T., de Queiroz, L. P., Rapini, A., Riina, R., Rincon, C. A. V., Roque, N., Shimizu, G., Sobral, M., Stehmann,

J. R., Stevens, W. D., Taylor, C. M., Trovó, M., van den Berg, C., van der Werff, H., Viana, P. L., Zartman, C. E. & Forzza, R. C. Amazon plant diversity revealed by a taxonomically verified species list. *Proc. Natl. Acad. Sci.* **114,** 10695–10700 (2017).

10. Gilles, D., Zaiss, R., Blach-Overgaard, A., Catarino, L., Damen, T., Deblauwe, V., Dessein, S., Dransfield, J., Droissart, V., Duarte, M. C., Engledow, H., Fadeur, G., Figueira, R., Gereau, R. E., Hardy, O. J., Harris, D. J., de Heij, J., Janssens, S., Klomberg, Y., Ley, A. C., MacKinder, B. A., Meerts, P., van de Poel, J. L., Sonké, B., Sosef, M. S. M., Stévart, T., Stoffelen, P., Svenning, J.-C., Sepulchre, P., van der Burgt, X., Wieringa, J. J. & Couvreur, T. L. P. RAINBIO: a mega-database of tropical African vascular plants distributions. *PhytoKeys* **74,** 1–18 (2016).
11. Chavent, M., Kuentz-Simonet, V., Liquet, B. & Saracco, J. ClustOfVar: An R package for the clustering of variables. *J. Stat. Softw.* **50,** 1–16 (2012).
12. Bastin, J. F., Finegold, Y., Garcia, C., Mollicone, D., Rezende, M., Routh, D., Zohner, C. M. & Crowther, T. W. The global tree restoration potential. *Science (80-.).* (2019). doi:10.1126/science.aax0848
13. Liaw, A. & Wiener, M. Classification and Regression by randomForest. *R News* **2,** 18–22 (2002).
14. Lens, F., Schönenberger, J., Baas, P., Jansen, S. & Smets, E. The role of wood anatomy in phylogeny reconstruction of Ericales. *Cladistics* **23,** 229–254 (2007).
15. Wang, H. F. & Xu, M. Individual size variation reduces spatial variation in abundance of tree community assemblage, not of tree populations. *Ecol. Evol.* **7,** 10815–10828 (2017).
16. Craven, D., Knight, T. M., Barton, K. E., Bialic-Murphy, L., Cordell, S., Giardina, C. P., Gillespie, T. W., Ostertag, R., Sack, L. & Chase, J. M. OpenNahele: The open Hawaiian

forest plot database. *Biodivers. Data J.* **6,** (2018).

17. Pélissier, R., Pascal, J.-P., Ayyappan, N., Ramesh, B. R., Aravajy, S. & Ramalingam, S. R. Tree demography in an undisturbed Dipterocarp permanent sample plot at Uppangala, Western Ghats of India. *Ecology* **92,** 1376–1376 (2011).
18. Ramesh, B. R., Swaminath, M. H., Patil, S. V., Dasappa, Pélissier, R., Venugopal, P. D., Aravajy, S., Elouard, C. & Ramalingam, S. Forest stand structure and composition in 96 sites along environmental gradients in the central Western Ghats of India. *Ecology* **91,** 3118–3118 (2010).
19. Clark, D. B., Clark, D. A. & Oberbauer, S. F. Annual wood production in a tropical rain forest in NE Costa Rica linked to climatic variation but not to increasing CO2. *Glob. Chang. Biol.* **16,** 747–759 (2010).
20. Condit, R., Lao, S., Pérez, R., Dolins, S. B., Foster, R. B. & Hubbell, S. P. Barro Colorado Forest Census Plot Data, 2012 Version. *Cent. Trop. For. Sci. Databases.* (2012). doi:10.5479/data.bci.20130603
21. Bradford, M. G., Murphy, H. T., Ford, A. J., Hogan, D. L. & Metcalfe, D. J. Long-term stem inventory data from tropical rain forest plots in Australia. *Ecology* **95,** 2362 (2014).
22. Goncalves, F. G., Treuhaft, R. N., Dos santos, J. R., Graca, P., Almeida, A. & Law, B. E. Tree Inventory and Biometry Measurements, Tapajos National Forest, Para, Brazil, 2010. (2018).
23. Numata, I., da Silva, S. S. & Cochrane, M. A. Forest Inventories and DBH at Burned and Unburned Forest Sites, Acre, Brazil, 2017. (2018).
24. Franklin, J., Ripplinger, J., Freid, E. H., Marcano-Vega, H. & Steadman, D. W. Regional variation in Caribbean dry forest tree species composition. *Plant Ecol.* **216,** 873–886 (2015).

25. Nyirambangutse, B., Zibera, E., Uwizeye, F. K., Nsabimana, D., Bizuru, E., Pleijel, H., Uddling, J. & Wallin, G. Carbon stocks and dynamics at different successional stages in an Afromontane tropical forest. *Biogeosciences* **14,** 1285–1303 (2017).

26. Werden, L. K., Becknell, J. M. & Powers, J. S. Edaphic factors, successional status and functional traits drive habitat associations of trees in naturally regenerating tropical dry forests. *Funct. Ecol.* **32,** 2766–2776 (2018).

27. Young, N. E., Romme, W. H., Evangelista, P. H., Mengistu, T. & Worede, A. Variation in population structure and dynamics of montane forest tree species in Ethiopia guide priorities for conservation and research. *Biotropica* **49,** 309–317 (2017).

28. Stride, G., Thomas, C. D., Benedick, S., Hodgson, J. A., Jelling, A., Senior, M. J. M. & Hill, J. K. Contrasting patterns of local richness of seedlings, saplings, and trees may have implications for regeneration in rainforest remnants. *Biotropica* **50,** 889–897 (2018).

29. Gentry, A., Phillips, O. L. & Miller, J. S. *Global patterns of plant diversity: Alwyn H. Gentry's forest transect data set*. (Missouri Botanical Press, 2002).

30. Potts, M. D., Ashton, P. S., Kaufman, L. S. & Plotkin, J. B. Habitat patterns in tropical rain forests: A comparison of 105 plots in northwest Borneo. *Ecology* **83,** 2782–2797 (2002).

DISCUSSION

A. Individual-level variation and a predictive ecology at global scale

As has been shown in the thesis at hand, individual-based approaches can contribute greatly to a more predictive forest ecology. In particular, they provide a means to translate remotely sensed data into highly detailed representations of forest structure across thousands of hectares of forest canopy (Chapter 1 and 2). These virtual mockups of forests, in turn, can serve as the initial conditions for the inference of vital rates and the prediction of old-growth forest dynamics at fine spatial scales (Chapter 3). Finally, individual-based models of forest growth can investigate ecological questions at the level of the individual that would elude coarser-grained vegetation models, such as the role of plasticity and inter-individual variation on forest structure and function (Chapter 4). All of this would, however, not be possible without the detailed, individual- and species-based trait measurements that have been and are currently collected across the globe. Mapping these traits and understanding their evolutionary history will thus be an important pillar in the further development of predictive modelling (Chapter 5).

To achieve the overall goal, however, of predicting the fate of the world's vegetation, including the highly diverse tropical forests, a large synthesis is needed that is still in the making. Trait-based ecology, for example, has opened up one promising avenue. Based on the idea that every trait comes at a cost and thus involves trade-offs, particularly between construction costs and growth rate (Chave *et al.*, 2009; Reich, 2014), it offers a promising framework for the quantitative assessment of ecological strategies. Plant or animal traits (such as leaf nutrients) can be directly related to ecological strategies (i.e. faster or slower growth) and environmental conditions (i.e. either responding to environmental variation or influencing it, cf. Lavorel & Garnier,

2002) and can ultimately be scaled up to the functioning of whole ecosystems (i.e. higher or lower primary productivity). In particular, it has been shown that there are trade-offs and relationships with demography and plant competition that hold at global scales (Díaz *et al.*, 2016; Kunstler *et al.*, 2016).

At the same time, some of trait-based ecology's underlying assumptions will likely require renewed examination (Shipley *et al.*, 2016; Worthy & Swenson, 2019). Among them is the transfer of globally observed trait dimensions to local scales (Messier *et al.*, 2017), and the predictive power of traits for demographic rates. While the latter is critical to functional ecology (Salguero-Gómez *et al.*, 2018), plant communities have not always shown strong relationships between traits and demography (Paine *et al.*, 2015; Poorter *et al.*, 2018; Yang *et al.*, 2018) and the relative importance of traits may vary across the ontogenetic trajectory (Falster *et al.*, 2018). Furthermore, even well-established relationships such as the relation between wood density and mortality (Kraft *et al.*, 2010) come with large variation around mean values and are not easily transformed into predictive tools (Visser *et al.*, 2016). Since most trait-relationships are correlative and leave much variance unexplained, robust tests against large data sets are needed to assess transferability (Wenger & Olden, 2012).

Remote sensing, on the other hand, offers exciting new possibilities for describing trees and forest quantitatively and non-destructively (Disney, 2019), providing ecologists with a wealth of data that can help with testing predictions or derive new models of ecosystem functioning. The many different tools, ranging from lidar scanning (terrestrial, airborne, spaceborne) over hyperspectral imaging to radar technology come together to create a three-dimensional, if not higher-dimensional, picture of the Earth across all scales. They are routinely used to quantify forest structure and functioning from the local to global scales (Le Toan *et al.*, 1992; Frankenberg *et al.*, 2011; Simard *et*

al., 2011) and have become "flux towers in the sky" (Schimel & Schneider, 2019). As a result, ecology has now an unprecedented amount of information at its disposal, and increasingly so in large, collaborative, open-access projects.

To transform remotely sensed information, however, into ecological knowledge, integrative approaches are needed. The basic units of ecology have traditionally been individual organisms, from which populations and communities emerge (Railsback, 2001; Begon *et al.*, 2005), but this is not the case for remote sensing products. While at fine to medium scales, the isolation of individual organisms, such as trees is entirely possible and can provide quantitative estimates that are more precise than those obtained with traditional ecological methods (Ferraz *et al.*, 2016; Disney *et al.*, 2018), the translation of waveforms or point clouds into individual-level structure or dynamics is not always possible. It poses a particular challenge in the dense, multistoried and hyperdiverse tropical forests.

It is here that the integration with individual-based models is a highly promising field (Shugart *et al.*, 2015). One such approach, for example, allows for the translation of remotely sensed metrics into individual organisms via model inversion, as shown in Chapters 1 to 3 of this PhD. The TROLL-based simulation approach, laid out in this work, naturally incorporates understorey layers that are otherwise difficult to penetrate, and, when coupled with physiological principles, can even rely on additional sources of information – such as trait distributions and light extinction imposed by overtopping – to produce narrower estimates.

While these are important steps towards a global synthesis, the ultimate aim of individual-based models and their relatives (Moorcroft *et al.*, 2001) will be prediction. Synthesizing trait-based approaches with remotely sensed data and field inventories, they provide a unique opportunity to assess how well we actually understand forests.

B. Future improvements for predictive modelling

The keystone for any predictive science ist the testing of models against new data (Houlahan *et al.*, 2017) and benchmarking models against each other (Fisher *et al.*, 2018). Improved protocols and pattern-oriented modelling, for example, provide important routes to improved understanding (Grimm *et al.*, 2010; Grimm & Railsback, 2012). If models reproduce and predict a large number of patterns well across scales, then we should be able to be confident about their forecasting abilities. What exactly, however constitutes a good prediction, is not well-defined – apart from a consensus, that, in general, the prediction of patterns should follow the purposes and spatial scales of the model – i.e. highly detailed local models should reproduce well local patterns, Dynamic Global Vegetation Models should reproduce well global patterns.

Due to the complexity of the modelled systems and the selective availability of data, many models, including individual-based and gap models, often use different statistics to demonstrate their ability to predict patterns, rendering comparisons between models difficult. One desirable is therefore a benchmark of fundamental patterns that vegetation models should reproduce (Kelley *et al.*, 2013; Rammig *et al.*, 2014). A good integration of remotely sensed data and field surveys would provide the natural counterpart to this benchmarking system (Chave *et al.*, 2019).

A number of challenges stand out. While some processes such as competition for light, have been simulated for years and with relatively high confidence (Shugart, 1984; Purves *et al.*, 2007), a number of other processes are unsatisfyingly represented in most models. These include tree mortality (Bugmann *et al.*, 2019), often included either empirically or with semi-mechanistic approaches (Seidl *et al.*, 2014), but rarely tested against empirical patterns, seed dispersal (Price *et al.*, 2001) and nutrient cycling

(Prentice *et al.*, 2007). All of these areas, together with hydrology, are also areas where the forest growth simulator TROLL could be massively improved in the future. Some of these areas have been underexplored, because data are rarely available at large scales or empirical relationships between variables are not well known across environmental gradients. In TROLL, in particular, the amount of sapwood fraction, a crucial component of respiration (Ryan *et al.*, 1994), plays an important role in determining how much of gross primary productivity is ultimately made available for growth (i.e. net primary productivity). While its relationship with leaf area has been well-documented and is often used in forest growth models of the gap model tradition (Fyllas *et al.*, 2014), a better quantification at the species and within-species level would be highly desirable.

A much more general challenge is, however, posed by community dynamics. Plants, when regarded in their local environments, often have integrated phenotypes (Messier *et al.*, 2017), i.e. their ecological strategies are reflected by a network of individual traits that do not align along trait axes, even if these axes are well-constrained globally (Díaz *et al.*, 2016). Furthermore, they are highly modular and plastic organisms, with often important consequences for ecosystem functioning, as shown in the PhD at hand. Acclimation and small-scale variation, for example, even affects processes as thoroughly represented in vegetation models as photosynthesis (Dietze, 2014). Taken together, plasticity and the integration of phenotypic characteristics poses a range of challenges to modellers, including the question of how to accurately model mechanistic relationships when underlying traits are subject to variation.

In this context, the relationship between what constitutes data for initialisation and what constitutes simulated vegetation dynamics warrants re-examination. Many models, such as TROLL, use plant traits as parameterization of species specific strategies (Maréchaux & Chave, 2017). Since these traits are, however, not only measured on

seedlings or sapling, but also on adult individuals – wood density in particular –, then the trait distributions we find in forest ecosystems are already the result of competition, facilitation and plastic responses, and thus an emergent property of the system (cf. Chapter 4 in this manuscript). While there are ways to tackle this problem, i.e. either by focussing more strictly on sapling data or by constraining not against trait patterns at initialisation, but within the already assembled community (i.e. through inverse modelling), this poses a number of fascinating challenges to the understanding of community dynamics via model-based approaches.

C. The integration of predictive science with society

Although historical vagaries such as the printing press cannot be dissociated from the rise to dominance of science since the 17th century, it is likely that the predictive component of scientific activities – already seen in lunar calendars and the computation of the length of daylight in early Mesopotamian astronomy (Rochberg, 2011) – has played an important role in making scientific research such a prominent component of the modern world (Dear, 2005). In gradually merging pre-modern natural philosophy with problem-oriented mathematical calculations, scientific inquiry has become an important part of modern societies, providing cognitive values such as 'objectivity' or 'impartiality' for human activity (Gaukroger, 2007)[1] and transforming human lives through technology, medical discoveries and the restructuring of production processes and public policy (Krige & Pestre, 1997).

[1] Science's dominant status also seems to put pressure on other forms of knowledge. This can not only be seen in academic disciplines such as sociology and political studies that have been under pressure to become "more scientific" since their inception (Oren, 2006), but also reaches into the pop-cultural realm where comic books or films recast old myths and the supernatural in scientific terms – turning the Nordic realm of the gods Asgard into a planetoid and superhuman abilities into genetic mutations.

Planetary challenges such as the future of tropical and non-tropical forests under climate change have given new importance and responsibility to scientific disciplines, in particular to ecology. They thus also resuscitate important questions about science's role with regard to society.

This PhD, for example, has largely focussed on the improvement of a predictive approach to understand and, ultimately, forecast tropical forest dynamics into the future. This is crucial for informed political decision-making and has important repercussions for societal change (Dietze *et al.*, 2018). At the same time, science's principal mode of reasoning – the problem-solving, instrumental approach–, though attractive and powerful, does not automatically align with other forms of human reasoning, such as ethical and collective decision making (Habermas, 1971). Questions about societal structures – What is a just distribution of goods? How do we want to live as a society? Which future do we want for our children? – may be informed by, but can rarely be transferred into quantitative reasoning – By how much will average lifespan increase or decrease? Does income increase?

On the one hand, this means that scientific insights, no matter how clear and powerful, do not necessarily translate into public appreciation. On the other hand, scientific approaches that turn political challenges into technological problems can be counterproductive, insofar as some of the global challenges we face today have been created by technological progress in the first place[2].

[2] While science is a very complex and diverse institution that defies easy categorization, the relationship between scientific theory and practice and the natural world that it investigates has not always been characterized by appreciation and curiosity alone. A trend towards technological dominance can be traced throughout the past two centuries, starting from Kant's dictum that a scientist needs to treat nature as if he or she was "an appointed judge who compels witnesses to answer the questions he puts to them"(Kant, 1787), over the "divide-and-conquer" metaphoric that this PhD itself has made use of, and is still echoed in today's "ecosystem services" that, while highlighting important facets of the relation between humans and their environments, frame the relation as one between a client and a service-provider. This link is not confined to language alone, but mirrored in the close relationship between science and technological progress, such as the rise of thermodynamics or nuclear energy. A particularly prominent example are

For climate change and ecological research, this means that, even if we develop models that reliably predict future dynamcs and help us identify technological and non-technological solutions, societal questions need to be incorporated into these solutions. Large-scale reforestation projects, for example, are an exciting prospect to create a larger-scale carbon stock and generate negative climate feedback (Griscom *et al.*, 2017). But should we use such solutions, for example, on previously undisturbed grassland ecosystems, especially when the intactness of forests in the future cannot be guaranteed? Similarly, Carbon Capture and Storage (CCS), i.e. the conversion of gaseous carbon dioxide into storeable carbon compounds (a technological pendant to photosynthesis), could one day provide a powerful engineering solution to reducing carbon emissions or even producing "negative emissions", i.e. recuperating past emissions (Keith, 2009). At the same time, a substantial part of current research is spent on creating hydrocarbon fuels (Keith *et al.*, 2018), thus neither storing nor recuperating carbon, but re-releasing it. While such technologies would still be carbon-neutral, it is not far-fetched to imagine scenarios in which global demand for fossil fuels would increase and such a technology, financed to slow down climate change by storing carbon, would ultimately only be used for the purpose of re-emission.

Despite these caveats, it is also abundantly clear that political decision-making about climate change and the fate of our forests will not be possible without a strong predictive science and a good understanding of the ecological systems that are affected by it. Recent decades have seen enormous progress in this direction and offer a lot of promise for future endeavours. We have now an increasing number of metastudies and large-scale data bases at our disposal that showcase global patterns of plant functioning and ecology. We have also access to a wide range of remotely sensed products across all

---

modern statistical methods, based around Monte Carlo methods, that were made possible by military-industrial research on the atomic bomb (Robert & Casella, 2011).

scales that allow for the discovery and prediction of patterns globally. And finally, a large number of vegetation models are developed, with increasingly detailed representations of ecosystem functioning that can integrate these data, and provide an important, if not the only, way forward to understand and predict climate change.

While it is unclear whether we will ever be able to predict the exact patterns of trees on the "Indian ruins" that Darwin observed, we might at least develop the tools to make it possible for future generations to still wonder how the trees grow on the ruins that the 21st century has left behind, how they disperse and interact with each other, how they evolve over centuries and milennia, and how they come together with all the other organisms and human beings to form the biosphere.

LITERATURE

Begon M, Townsend CR, Harper JL. 2005. *Ecology: From Individuals to Ecosystems, 4th Edition*. Wiley-Blackwell.

Bugmann H, Seidl R, Hartig F, Bohn F, Brůna J, Cailleret M, François L, Heinke J, Henrot AJ, Hickler T, *et al.* 2019. Tree mortality submodels drive simulated long-term forest dynamics: assessing 15 models from the stand to global scale. *Ecosphere* **10**: 2616.

Chave J, Coomes D, Jansen S, Lewis SL, Swenson NG, Zanne AE. 2009. Towards a worldwide wood economics spectrum. *Ecology Letters* **12**: 351–366.

Chave J, Davies SJ, Phillips OL, Lewis SL, Sist P, Schepaschenko D, Armston J, Baker TR, Coomes D, Disney M, *et al.* 2019. Ground Data are Essential for Biomass Remote Sensing Missions. *Surveys in Geophysics* **40**: 863–880.

Dear P. 2005. What Is the History of Science the History Of ? . *Isis* **96**: 390–406.

Díaz S, Kattge J, Cornelissen JHC, Wright IJ, Lavorel S, Dray S, Reu B, Kleyer M, Wirth C, Colin Prentice I, *et al.* 2016. The global spectrum of plant form and function. *Nature* **529**: 167–171.

Dietze MC. 2014. Gaps in knowledge and data driving uncertainty in models of photosynthesis. *Photosynthesis Research* **119**: 3–14.

Dietze MC, Fox A, Beck-Johnson LM, Betancourt JL, Hooten MB, Jarnevich CS, Keitt TH, Kenney MA, Laney CM, Larsen LG, *et al.* 2018. Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences of the United States of America* **115**: 1424–1432.

Disney M. 2019. Terrestrial LiDAR: a three-dimensional revolution in how we look at trees. *New Phytologist* **222**: 1736–1741.

Disney MI, Boni Vicari M, Burt A, Calders K, Lewis SL, Raumonen P, Wilkes P. 2018.

Weighing trees with lasers: Advances, challenges and opportunities. *Interface Focus* **8**: 20170048.

Falster DS, Duursma RA, FitzJohn RG. **2018**. How functional traits influence plant growth and shade tolerance across the life cycle. *Proceedings of the National Academy of Sciences of the United States of America* **115**: E6789–E6798.

Ferraz A, Saatchi S, Mallet C, Meyer V. **2016**. Lidar detection of individual tree size in tropical forests. *Remote Sensing of Environment* **183**: 318–333.

Fisher RA, Koven CD, Anderegg WRL, Christoffersen BO, Dietze MC, Farrior CE, Holm JA, Hurtt GC, Knox RG, Lawrence PJ, *et al.* 2018. Vegetation demographics in Earth System Models: A review of progress and priorities. *Global Change Biology* **24**: 35–54.

Frankenberg C, Fisher JB, Worden J, Badgley G, Saatchi SS, Lee JE, Toon GC, Butz A, Jung M, Kuze A, *et al.* 2011. New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters* **38**: L17706.

Fyllas NM, Gloor E, Mercado LM, Sitch S, Quesada CA, Domingues TF, Galbraith DR, Torre-Lezama A, Vilanova E, Ramírez-Angulo H, *et al.* 2014. Analysing Amazonian forest productivity using a new individual and trait-based model (TFS v.1). *Geoscientific Model Development* **7**: 1251–1269.

Gaukroger S. **2007**. *The Emergence of a Scientific Culture: Science and the Shaping of Modernity 1210-1685*. Oxford: Clarendon Press.

Grimm V, Berger U, DeAngelis DL, Polhill JG, Giske J, Railsback SF. **2010**. The ODD protocol: A review and first update. *Ecological Modelling* **221**: 2760–2768.

Grimm V, Railsback SF. **2012**. Pattern-oriented modelling: a 'multi-scope' for predictive systems ecology. *Philosophical Transactions of the Royal Society B: Biological*

Sciences **367**: 298–310.

Griscom BW, Adams J, Ellis PW, Houghton RA, Lomax G, Miteva DA, Schlesinger WH, Shoch D, Siikamäki J V., Smith P, *et al.* 2017. Natural climate solutions. *Proceedings of the National Academy of Sciences* **114**: 11645–11650.

Habermas J. **1971**. *Toward a Rational Society*. Heinemann Educational Books.

Houlahan JE, McKinney ST, Anderson TM, McGill BJ. **2017**. The priority of prediction in ecological understanding. *Oikos* **126**: 1–7.

Kant I. **1787**. *Critique of Pure Reason* (P Guyer and AW Wood, Eds.). Cambridge University Press.

Keith DW. **2009**. Why capture CO2 from the atmosphere? *Science* **325**: 1654–1655.

Keith DW, Holmes G, St. Angelo D, Heidel K. **2018**. A Process for Capturing CO2 from the Atmosphere. *Joule*.

Kelley DI, Prentice IC, Harrison SP, Wang H, Simard M, Fisher JB, Willis KO. **2013**. A comprehensive benchmarking system for evaluating global vegetation models. *Biogeosciences* **10**: 3313–3340.

Kraft NJB, Metz MR, Condit RS, Chave J. **2010**. The relationship between wood density and mortality in a global tropical forest data set. *New Phytologist* **188**: 1124–1136.

Krige J, Pestre D. **1997**. *Science in the twentieth century* (J Krige and D Pestre, Eds.). Amsterdam: Harwood Academic Publishers.

Kunstler G, Falster D, Coomes DA, Hui F, Kooyman RM, Laughlin DC, Poorter L, Vanderwel M, Vieilledent G, Wright SJ, *et al.* 2016. Plant functional traits have globally consistent effects on competition. *Nature* **529**: 204–207.

Lavorel S, Garnier E. **2002**. Predicting changes in community composition and ecosystem functioning from plant traits: Revisiting the Holy Grail. *Functional Ecology* **16**: 545–556.

Maréchaux I, Chave J. 2017. An individual-based forest model to jointly simulate carbon and tree diversity in Amazonia: description and applications. *Ecological Monographs* **87**: 632–664.

Messier J, Lechowicz MJ, McGill BJ, Violle C, Enquist BJ. 2017. Interspecific integration of trait dimensions at local scales: the plant phenotype as an integrated network. *Journal of Ecology* **105**: 1775–1790.

Moorcroft PR, Hurtt GC, Pacala SW. 2001. A method for scaling vegetation dynamics: The ecosystem demography model (ED). *Ecological Monographs* **71**: 557–586.

Oren I. 2006. Can political science emulate the natural sciences? the problem of self-disconfirming analysis. *Polity* **38**: 72–100.

Paine CET, Amissah L, Auge H, Baraloto C, Baruffol M, Bourland N, Bruelheide H, Daïnou K, de Gouvenain RC, Doucet JL, *et al.* 2015. Globally, functional traits are weak predictors of juvenile tree growth, and we do not know why. *Journal of Ecology* **103**: 978–989.

Poorter L, Castilho C V., Schietti J, Oliveira RS, Costa FRC. 2018. Can traits predict individual growth performance? A test in a hyperdiverse tropical forest. *New Phytologist* **219**: 109–121.

Prentice IC, Bondeau A, Cramer W, Harrison SP, Hickler T, Lucht W, Sitch S, Smith B, Sykes MT. 2007. Dynamic Global Vegetation Modeling: Quantifying Terrestrial Ecosystem Responses to Large-Scale Environmental Change. In: Canadell JG, Pataki DE, Pitelka LF, eds. Terrestrial Ecosystems in a Changing World. Berlin, Heidelberg: Springer Berlin Heidelberg, 175–192.

Price DT, Zimmermann NE, Van Der Meer PJ, Lexer MJ, Leadley P, Jorritsma ITM, Schaber J, Clark DF, Lasch P, Mcnulty S, *et al.* 2001. Regeneration in gap models: Priority issues for studying forest responses to climate change. *Climatic Change* **51**: 475–

508.

Purves DW, Lichstein JW, Pacala SW. 2007. Crown plasticity and competition for canopy space: A new spatially implicit model parameterized for 250 North American tree species. *PLoS ONE* **2**: e870.

Railsback SF. 2001. Concepts from complex adaptive systems as a framework for individual-based modelling. *Ecological Modelling* **139**: 47–62.

Rammig A, Wiedermann M, Donges JF, Babst F, von Bloh W, Frank D, Thonicke K, Mahecha MD. 2014. Tree-ring responses to extreme climate events as benchmarks for terrestrial dynamic vegetation models. *Biogeosciences Discussions* **11**: 2537–2568.

Reich PB. 2014. The world-wide 'fast-slow' plant economics spectrum: A traits manifesto. *Journal of Ecology* **102**: 275–301.

Robert C, Casella G. 2011. A short history of Markov Chain Monte Carlo: Subjective recollections from incomplete data. *Statistical Science* **26**: 102–115.

Rochberg F. 2011. Natural Knowledge in Ancient Mesopotamia. In: Harrison P, Numbers RL, Shank MH, eds. Wrestling with Nature. From Omens to Science. 9–36.

Ryan MG, Hubbard RM, Clark DA, Sanford RL. 1994. Woody-tissue respiration for Simarouba amara and Minquartia guianensis, two tropical wet forest trees with different growth habits. *Oecologia* **100**: 213–220.

Salguero-Gómez R, Violle C, Gimenez O, Childs D. 2018. Delivering the promises of trait-based approaches to the needs of demographic approaches, and vice versa. *Functional Ecology* **32**: 1424–1435.

Schimel D, Schneider FD. 2019. Flux towers in the sky: global ecology from space. *New Phytologist* **224**: 570–584.

Seidl R, Rammer W, Blennow K. 2014. Simulating wind disturbance impacts on forest landscapes: Tree-level heterogeneity matters. *Environmental Modelling and Software* **51**:

1–11.

Shipley B, De Bello F, Cornelissen JHC, Laliberté E, Laughlin DC, Reich PB. 2016. Reinforcing loose foundation stones in trait-based plant ecology. *Oecologia* **180**: 923–931.

Shugart, HH. 1984. *A theory of forest dynamics: the ecological implications of forest succession models*. Springer-Verlag.

Shugart HH, Asner GP, Fischer R, Huth A, Knapp N, Le Toan T, Shuman JK. 2015. Computer and remote-sensing infrastructure to enhance large-scale testing of individual-based forest models. *Frontiers in Ecology and the Environment* **13**: 503–511.

Simard M, Pinto N, Fisher JB, Baccini A. 2011. Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences* **116**: G04021.

Le Toan T, Beaudoin A, Riom J, Guyon D. 1992. Relating Forest Biomass to SAR Data. *IEEE Transactions on Geoscience and Remote Sensing* **30**: 403–411.

Visser MD, Bruijning M, Wright SJ, Muller-Landau HC, Jongejans E, Comita LS, de Kroon H. 2016. Functional traits as predictors of vital rates across the life cycle of tropical trees. *Functional Ecology* **30**: 168–180.

Wenger SJ, Olden JD. 2012. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods in Ecology and Evolution* **3**: 260–267.

Worthy SJ, Swenson NG. 2019. Functional perspectives on tropical tree demography and forest dynamics. *Ecological Processes* **8**.

Yang J, Cao M, Swenson NG. 2018. Why Functional Traits Do Not Predict Tree Demographic Rates. *Trends in Ecology and Evolution* **33**: 326–336.

Inferring the structure and dynamics of tropical rain forests with individual-based forest growth models

Climate change presents society and science with a challenge that goes beyond the temporal and spatial scales of most practical problems. It therefore requires approaches that reflect the complexity of the Earth's system. This holds particularly true for the biosphere and forest ecosystems, one of the most important sources of uncertainty in climate projections. Concerted data collection efforts, such as forest inventories, trait data bases, and new technologies, such as remote sensing, have considerably increased our ability to observe and analyze the current state of the Earth's vegetation. However, to extrapolate findings into the future and understand the feedbacks between vegetation and climate change, models are needed that assimilate these data and translate them into ecosystem dynamics. Mechanistic and individual-based forest models are a particular promising approach, since they simulate dynamics bottom-up, reconstruct forests tree by tree, and are thus able to predict patterns across scales. This PhD further develops the trait- and individual-based forest growth simulator TROLL, including intraspecific variation and plasticity in tree growth, derives a new method to translate Airborne Lidar data into virtual forest inventories and uses it to infer forest structure and ecosystem dynamics in tropical rain forests. Finally, in line with TROLL's trait-based approach, an update to a global trait base, the Global Wood Density Database is presented, exploring the contribution of evolutionary lineages to wood density variation and mapping wood density across the globe.

Key words: individual-based modelling, biomass, wood density, remote sensing, ecosystem functioning

Inférence de la structure et dynamique des forêts tropicales humides avec un modèle individu-centré

Le changement climatique constitue un défi qui dépasse les échelles temporelles et spatiales de la plupart des problèmes. Il nécessite donc des approches qui reflètent la complexité du système terrestre. Cela est particulièrement vrai pour la biosphère et les écosystèmes forestiers, l'une des principales sources d'incertitude dans les projections climatiques. Les efforts concertés de collecte de données, tels que les inventaires forestiers, les bases de données des traits et les nouvelles technologies, telles que la télédétection, ont considérablement accru notre capacité à observer et à analyser l'état actuel de la végétation de la Terre. Cependant, pour estimer les développements futurs et comprendre les feedbacks entre la végétation et le changement climatique, des modèles sont nécessaires pour assimiler ces données et les traduire en dynamique des écosystèmes. Les modèles forestiers mécanistes et individu-centrés sont une approche particulièrement prometteuse, car ils simulent la dynamique forestière "bottom-up", reconstruisent les forêts arbre par arbre, et sont donc capables de prédire des patrons à différentes échelles. Cette thèse continue le développement du simulateur de dynamique forestière TROLL, rajoute la variation intraspécifique et la plasticité de la croissance des arbres, dérive une nouvelle méthode pour traduire les données de télédétection en inventaires forestiers virtuels et l'utilise pour inférer la structure forestière et la dynamique des écosystèmes dans les régions tropicales. Enfin, conformément à l'approche de TROLL, basée sur les traits, une mise à jour d'une base mondiale de traits, la base de données mondiale de la densité du bois est présentée, explorant la contribution des changements évolutives et cartographiant la densité du bois à travers le monde.

Mots-clés: modèle individu-centré, biomasse, densité du bois, télédétection, fonctionnement des ecosystèmes