Improving plant allometry by fusing forest models and remote sensing
Fabian Jörg Fischer, Isabelle Maréchaux, Jérôme Chave

To cite this version:

Fabian Jörg Fischer, Isabelle Maréchaux, Jérôme Chave. Improving plant allometry by fusing forest models and remote sensing. New Phytologist, Wiley, 2019, 223 (3), pp.1159-1165. 10.1111/nph.15810. hal-02191073

HAL Id: hal-02191073
https://hal.umontpellier.fr/hal-02191073
Submitted on 22 Jul 2021

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.
Improving plant allometry by fusing forest models and remote sensing

F. Jörg Fischer1, I. Maréchaux2 and J. Chave1
1Laboratoire Evolution et Diversité Biologique, UMR5174, CNRS–Université Paul Sabatier–IRD, Bâtiment 4R1, 118 route de Narbonne, F-31062 Toulouse Cedex 9, France; 2AMAP, INRA, IRD, CIRAD, CNRS, University of Montpellier, F-34000 Montpellier, France

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
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 belowground 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 to 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.
Tree height (m) vs DBH (m) plot. The plot shows 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°N, 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.

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 (m³), 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...
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
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.
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


