

1 Tackling unresolved questions in forest 2 ecology: the past and future role of 3 simulation models

4
5 Maréchaux I.*¹, Langerwisch F.*/**^{2,3}, Huth A^{4,5,6}., Bugmann H.⁷, Morin X.⁸, Reyer C.P.O.⁹,
6 Seidl R.¹⁰, Collalti A.^{11,12}, Dantas de Paula, M.¹³, Fischer R.⁴, Gutsch M.⁹, Lexer M.J.¹⁴,
7 Lischke H.¹⁵, Rammig A.¹⁶, Rödig E.⁴, Sakschewski B.⁹, Taubert F.⁴, Thonicke K.⁹,
8 Vacchiano G.¹⁷ & F. J. Bohn*⁴
9

10 Appendix A: Data availability

11 Forest models are data-demanding across the different steps of model development and
12 application. A robust parameterization of the multiple processes related to plant life cycle and
13 physiology for diverse plant types, species or individuals requires various data across scales,
14 from plant organ to population, including environmental factors. For many processes, such
15 data are often not available in the required quality and resolution, e.g. for tolerance of trees to
16 resource limitations (Craine, Engelbrecht, Lusk, McDowell, & Poorter 2012; De Kauwe et al.,
17 2015; McMahon et al., 2011) or soil characteristics (Marthews et al., 2014). Additionally, a
18 thorough initialization and validation of forest simulations over large spatial and temporal
19 scales requires observation data encompassing both fine resolution and large coverage over
20 long time spans, which can still be a challenge (Estes et al., 2018, Table 1).

21 Fortunately, data availability (see Table 2) fosters a systematic model trait-based
22 parameterization for a range of plant species and individuals. For example, Scheiter et al.

23 (2013) and Sakschewski et al. (2015) used reported trait coordination to constrain individual
24 trait combinations in simulations of forest dynamics with DGVMs. In doing so, they improved
25 model representation of functional diversity from a few discrete plant functional types to a
26 continuum of traits, while excluding unrealistic trait combinations (Van Bodegom et al., 2012).
27 Similarly, by taking advantage of comprehensive trait databases, but also of long term
28 inventories and of the detailed information they provide on tree life-histories, forest IBMs have
29 been allowed to simulate hundreds of species within diverse forest communities (Maréchaux
30 & Chave, 2017; Rüger et al., 2019).

31 Beside networks of forest plot inventories and remote sensing tools citizen science
32 programs have also been developed to create new opportunities of forest data sampling over
33 large areas (Affouard, Goëau, Bonnet, Lombardo, & Joly 2017; Delbart, Beaubien, Kergoat, &
34 Le Toan 2015; Giraud, Calenge, Coron, & Julliard, 2016; Wäldchen, Rzanny, Seeland, &
35 Mäder, 2018).

36 The development of machine learning techniques allowed Rammer & Seidl, (2019), to
37 use deep neural networks to estimate vegetation transitions across large spatial scales.
38 Additionally, an example for Bayesian modelling approaches is Van Oijen et al. (2013), who
39 found a strong reduction of uncertainty in most forest models after a Bayesian calibration.

40

41 **References**

- 42 Affouard, A., Goëau, H., Bonnet, P., Lombardo, J.-C. & Joly, A. (2017). PI@ntNet app in the
43 era of deep learning. In: *ICLR 2017 - Workshop Track - 5th International Conference*
44 *on Learning Representations*. Toulon, France, pp. 1–6.
- 45 Craine, J.M., Engelbrecht, B.M.J., Lusk, C.H., McDowell, N.G. & Poorter, H. (2012).
46 Resource limitation, tolerance, and the future of ecological plant classification.
47 *Frontiers in Plant Science*, 3.
- 48 De Kauwe, M.G., Zhou, S.-X., Medlyn, B.E., Pitman, A.J., Wang, Y.-P., Duursma, R.A., ...
- 49 Norby, R. J. (2015). Do land surface models need to include differential plant species

- 50 responses to drought? Examining model predictions across a mesic-xeric gradient in
51 Europe. *Biogeosciences*, 12, 7503–7518.
- 52 Delbart, N., Beaubien, E., Kergoat, L. & Le Toan, T. (2015). Comparing land surface
53 phenology with leafing and flowering observations from the PlantWatch citizen
54 network. *Remote Sensing of Environment*, 160, 273–280.
- 55 Giraud, C., Calenge, C., Coron, C. & Julliard, R. (2016). Capitalizing on opportunistic data for
56 monitoring relative abundances of species. *Biometrics*, 72, 649–658.
- 57 Estes, L., Elsen, P.R., Treuer, T., Ahmed, L., Caylor, K., Chang, J., ... Ellis, E.C. (2018).
58 The spatial and temporal domains of modern ecology. *Nature Ecology & Evolution*, 1.
- 59 Maréchaux, I. & Chave, J. (2017). An individual-based forest model to jointly simulate carbon
60 and tree diversity in Amazonia: description and applications. *Ecol Monogr* 87(4), 632–
61 664. doi: 10.1002/ecm.1271
- 62 Marthews, T.R., Quesada, C.A., Galbraith, D.R., Malhi, Y., Mullins, C.E., Hodnett, M.G., ...
63 Dharssi, I. (2014). High-resolution hydraulic parameter maps for surface soils in
64 tropical South America. *Geoscientific Model Development*, 7, 711.
- 65 McMahon, S. M., Harrison, S. P., Armbruster, W. S., Bartlein, P. J., Beale, C. M., Edwards,
66 M. E., ... Prentice, I. C. (2011). Improving assessment and modelling of climate
67 change impacts on global terrestrial biodiversity. *Trends in Ecology & Evolution*,
68 26(5), 249–259. doi: 10.1016/j.tree.2011.02.012
- 69 Rammer, W., & Seidl, R. (2019). A scalable model of vegetation transitions using deep
70 neural networks. *Methods in Ecology and Evolution*, 10(6), 879–890. doi:
71 10.1111/2041-210X.13171
- 72 Rüger, N., Condit, R., Dent, D. H., DeWalt, S. J., Hubbell, S. P., Lichstein, J. W., ... Farrior,
73 C. E. (2019). Demographic tradeoffs predict tropical forest dynamics. *BioRxiv*,
74 808865. doi: 10.1101/808865
- 75 Sakschewski, B., von Bloh, W., Boit, A., Rammig, A., Kattge, J., Poorter, L., ... Thonicke, K.
76 (2015). Leaf and stem economics spectra drive diversity of functional plant traits in a
77 dynamic global vegetation model. *Glob Change Biol*, 21, 2711–2725.

78 Scheiter, S., Langan, L. & Higgins, S.I. (2013). Next-generation dynamic global vegetation
79 models: learning from community ecology. *New Phytol*
80 Van Bodegom, P.M., Douma, J.C., Witte, J.P.M., Ordoñez, J.C., Bartholomeus, R.P. & Aerts,
81 R. (2012). Going beyond limitations of plant functional types when predicting global
82 ecosystem–atmosphere fluxes: exploring the merits of traits-based approaches.
83 *Global Ecology and Biogeography*, 21, 625–636.
84 Van Oijen, M., Reyer, C., Bohn, F.J., Cameron, D.R., Deckmyn, G., Flechsig, M., ...
85 Rammer, W.(2013). Bayesian calibration, comparison and averaging of six forest
86 models, using data from Scots pine stands across Europe. *Forest Ecology and*
87 *Management*, 289, 255–268.
88 Wäldchen, J., Rzanny, M., Seeland, M. & Mäder, P. (2018). Automated plant species
89 identification—Trends and future directions. *PLOS Computational Biology*, 14,
90 e1005993.
91
92

93 Appendix B Technical challenges

94 Several technical obstacles constrain model developments and runtime. Expanding
95 model development and applications relies on code and data sharing within and among larger
96 communities of model developers and users, which is also accompanied by technical
97 challenges. Several modeller teams make the model code (partly) freely available.
98 Additionally, version control systems allow to track changes and collaborate on model code in
99 an efficient way (e.g. Git, Ram, 2013; e.g. Collalti, et al., 2016). Besides code sharing,
100 simulation data are increasingly available following data open access requirements, allowing
101 subsequent analyses or model comparisons. In many modelling studies, the preparation of
102 data (e.g. for input/initialization, calibration or validation) and the analyses of model outputs
103 are very work- and time-intensive. Sharing scripts for analysing forest simulations, e.g. through

104 dedicated platforms (e.g. LeBauer, Wang, Richter, Davidson, & Dietze, 2013) or R (R Core
105 Team, 2018) package (e.g. Duursma, & Medlyn 2012), is also of great help.

106 Another example of technical challenges is the development of the simulation
107 framework within model intercomparisons and the standardisation of both model inputs and
108 outputs. Moreover, when complex process-based models are involved, whose uncertainties
109 can not simply be attributed to individual processes, a major challenge is to interpret the
110 ensemble runs and to understand which model processes actually explain the differences
111 between models. To address all these issues, transparent model documentations and
112 intensive exchange between modellers is needed accompanied by systematic tests of models
113 and their components (Reyer et al., 2020).

114 Model coupling is likewise challenging, since, in most models, some processes are
115 hidden in parameters or strongly simplified functions and the model is usually balanced by
116 fitting these parameters. If the simplified process or the parameter is replaced by a more
117 complex sub-model for the process, often the balance can be lost. Additionally, error
118 propagation among models can also prove difficult (Dunford, Harrison, & Rounsevell, 2015).
119 Several model systems and software frameworks have been developed to facilitate multi-
120 model coupling in a systematic way, and they even allow for switching between different
121 models during a simulation (Haas et al., 2013).

122

123 **References**

- 124 Collalti, A., Marconi, S., Ibrom, A., Trotta, C., Anav, A., D'Andrea, E., ... Santini, M. (2016)
125 Validation of 3D-CMCC Forest Ecosystem Model (v.5.1) against eddy covariance
126 data for 10 European forest sites. *Geoscientific Model Development* 9:479–504.
127 Dunford, R., Harrison, P.A. & Rounsevell, M.D.A. (2015). Exploring scenario and model
128 uncertainty in cross-sectoral integrated assessment approaches to climate change
129 impacts. *Climatic Change*, 132, 417–432.

- 130 Duursma, R.A. & Medlyn, B.E. (2012). MAESPA: a model to study interactions between
131 water limitation, environmental drivers and vegetation function at tree and stand
132 levels, with an example application to [CO₂]x drought interactions.
- 133 Haas, E., Klatt, S., Fröhlich, A., Kraft, P., Werner, C., Kiese, R., ... Butterbach-Bahl, K.
134 (2013). LandscapeDNDC: a process model for simulation of biosphere–atmosphere–
135 hydrosphere exchange processes at site and regional scale. *Landscape Ecol*, 28,
136 615–636.
- 137
- 138 LeBauer, D.S., Wang, D., Richter, K.T., Davidson, C.C. & Dietze, M.C. (2013). Facilitating
139 feedbacks between field measurements and ecosystem models. *Ecological
140 Monographs*, 83, 133–154.
- 141 R Core Team. (2018). *R: a language and environment for statistical computing*. R
142 Foundation for Statistical Computing, Vienna, Austria.
- 143 Ram, K. (2013). Git can facilitate greater reproducibility and increased transparency in
144 science. *Source Code for Biology and Medicine*, 8, 7.
- 145 Reyer, C., Silveyra Gonzalez, R., & Dolos, K. (2020). The PROFOUND database for
146 evaluating vegetation models and simulating climate impacts on forests. (*in press*)
- 147
- 148
- 149
- 150