On the use of individual-based models in predictive plant ecology

Christian Damgaard

Department of Ecoscience, Aarhus University, C.F. Møllers Allé 4, building 1120, 322, 8000 Aarhus C, Denmark. Email: cfd@ecos.au.dk; https://orcid.org/0000-0003-3932-4312

Author for correspondence:

Christian Damgaard

Tel: +45-30183153

Email: cfd@ecos.au.dk

Abstract

The use of IBM may be the only possible method for getting quantitative insights into some dynamical aspects of animal populations with complicated behavior, but this is surely not the case in plant ecology, where relevant plant population models have been formulated and used for decades. In my opinion, credible predictive plant ecological models is most effectively performed using population models where standard statistical methods can be used to estimate the different sources of uncertainty.

Keywords: Agent based models, plant ecology, Forecasting, predictions

Individual-based models (IBM), or agent-based models, are mathematical descriptions of the biological and ecological features and processes that characterize individuals and their interactions with other individuals in an ecosystem (Grimm 2019). All ecological models use a combination of data and hypotheses on the ecological processes to construct mathematical descriptions that are simplifications of the real world. However, when constructing IBM, the role of data is downplayed relatively to the role of hypotheses compared to empirical models (also called statistical - or phenomenological models). Typically, IBM are constructed from first principles using current ecological process knowledge, whereas empirical models mainly are constructed to capture the main features entailed in the sampled data.

Potentially, IBM are great heuristic tools for understanding the implications and consequences of our current ecological knowledge. However, some IBM researchers go further than this and implicitly assume that the ecological mechanisms are described in sufficient detail to make ecological predictions by simulating the future of an ecosystem. The terms “prediction” or “forecast” were mentioned in the title, abstract, or as a keyword in 14% of plant IBM publications (Fig. 1). In my opinion, the use of IBM as predictive tools in plant ecology is compromised by serious and unsurmountable epistemological problems.

Already Kant (1781) criticized the use of pure reasoning when making inferences about the real would, and showed how the objective input from our senses (data) together with our belief system (e.g. prior notions of space, time and causality) transcends into our subjective understanding of the world. Current ecological knowledge is inaccurate to some unknown degree and is expected to be continuously updated and refined. For example, the important notion by Liebig that plant growth is determined by the limiting resource, which has been instrumental in the construction of many plant ecological models including IBM, is now refined by suggesting that plant individuals adapt to the local environment, e.g. by modifying the shoot-root ratio, so that plant growth is limited by several resources at the same time (Craine 2009). Ten years ago it was argued that the structure of simple mathematical fractals provided important insights into how real leaves were made, but the initial excitement has since then died out.

Grimm (2019) lists variation among individuals, local interactions, and adaptive behavior as the three main reasons for representing individuals in ecological models, of which the first two are relevant for plant populations. However, there exists plant population models where the variation among individuals and local interactions are modelled at the population level, and these models may be fitted to ecological data using standard statistical procedures. For example, the effect of the variation among plant individuals on demographic characteristics has been modelled using integral projection methods (e.g. Ellner and Rees 2006), Damgaard and Weiner (2008) have modelled the effect of size-asymmetric growth in a plant population using parameters at both the individual level and the population level, and Bolker and Pacala (1999) have outlined a plant competition model that take local interactions into account.

Ecological data play a dual role in IBM. First, ecological data are used as an integrate part of our collective ecological knowledge that guide the construction of the IBM from first principles, secondly, a selected subset of available and relevant data are used to parameterize the model, often by focusing on mean values of the parameters and ignoring covariation among the often many parameters that are needed in IBM. Moreover, even though ecological data often are collected by sampling individuals, hypothesis testing and statistical inferences of the underlying mechanisms that explain the observed patterns are made at the level of the population. It is a non-trivial and subjective task to formulate the underlying individual processes from inferences that are performed at the population level, where the sampled data is a complicated mixture of net interaction processes, e.g. facilitation and competition processes at different life stages, and emergent system properties (Lenton et al. 2021). Consequently, ecological data only play an *ad hoc* role in the construction of most IBM, i.e. the assumed mathematical descriptions are not formally linked to empirical ecological data, and there is no formal way to update the assumed mathematical descriptions when new data are collected.

Recently, there has been an important development in the statistical fitting of complicated models such as IBM. Hooten et al. (2020) has demonstrated how complicated IBM may be fitted to data using a Bayesian hierarchical model setup. Such a setup will effectively address many of the problematic issues discussed here (Hooten et al. 2020), but unfortunately, the robust statistical fitting of IBM has only rarely been done, and I know of no plant ecological study where such statistical methods have been applied. I highly recommend the IBM community to take a look of these important methodological developments.

It is sometimes argued that IBM may be tested by comparing the simulations to observed ecological patterns. However, different processes may lead to the same pattern, and reproducing patterns is no guarantee that underlying mechanisms are fully understood. Furthermore, the *ad ho*c role of ecological data in IBM, means that model uncertainties cannot be estimated by traditional statistical methods, which critically deter from their credibility and limits their practical use for making ecological prediction. Again, IBM are great heuristic tools for testing our general understanding of the mechanisms that leads to observed ecological patterns, but are less useful when making local ecological predictions to support management decisions.

It is recommended that IBM are constructed in an iterative way, starting from simple processes and then gradually increasing the ecological realism (e.g. Grimm 2019). This reasonable suggestion for the model building process, has the less-discussed consequence, that the more ecologically realistic IBM always tend to be work-in-progress and a scientifically unhealthy association develop between the modeler and the model. The development of a specific IBM may become the main activity of the modeler and thus tightly linked to the scientific carrier development of the modeler, and the tight association between modeler and the model may lead to a reduced ability to criticize the model. That is, it becomes more difficult to not let your prior belief influence the interpretation of the observed ecological patterns.

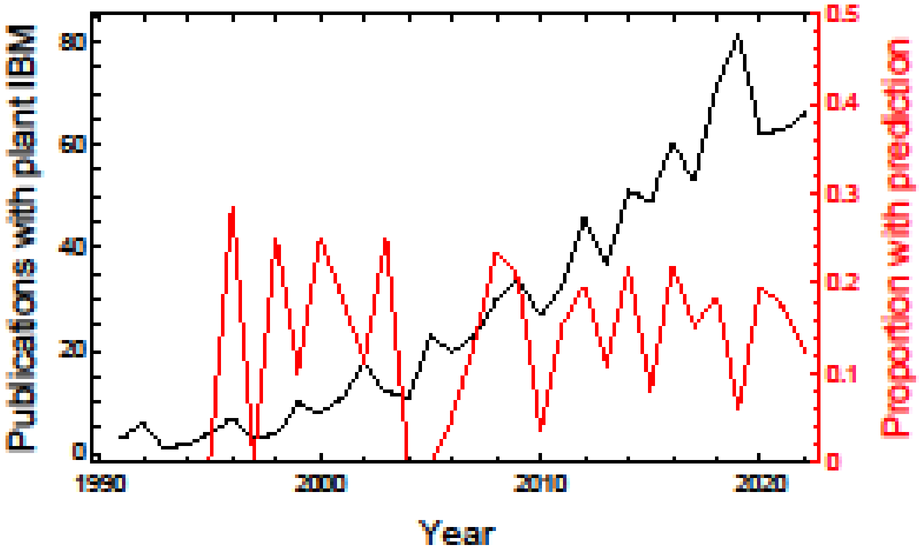
The use of IBM may be the only possible method for getting quantitative insights into some dynamical aspects of animal populations with complicated behavior, but this is surely not the case in plant ecology, where relevant plant population models have been formulated and used for decades (Harper 1977), and in my opinion, credible predictive plant ecological models is most effectively performed using population models where standard statistical methods can be used to estimate the different sources of uncertainty (Damgaard 2022).

Acknowledgements

Thanks to Jacob Weiner and Peter Borgen Sørensen for valuable discussions

Figures

Fig. 1. The number of publications with plant IBM in the title, abstract, or as a keyword since 1990, and in red the proportion of these that also mentioned the terms “prediction” or “forecast”. The number of publications with plant IBM increase on average with 2.4 each year.



Literature search in Web of Science:

1. ((ALL=((plant))) AND ALL=((("agent-based model*")) OR ((“individual-based model*”))))
2. ((ALL=((plant))) AND ALL=((("agent-based model*")) OR ((“individual-based model*”)))) AND ALL=(((prediction*) OR (forecast)))

References

Bolker BM, Pacala SW. 1999. Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal. American Naturalist. 153:575-602.

Craine JM. 2009. Resource strategies of wild plants. Princeton University Press.

Damgaard C. 2022. Processes and predictions in plant ecological models: Logic and causality. EcoEvoRxiv.

Damgaard C, Weiner J. 2008. Modeling the growth of individuals in crowded plant populations. Journal of Plant Ecology. 1:111–116.

Ellner SP, Rees M. 2006. Integral projection models for species with complex demography. Am Nat. 167(3):410-428.

Grimm V. 2019. Individual-based models. In: Fath BD, editor. Encyclopedia of ecology. 2 ed. Oxford: Elsevier. p. 65 - 73.

Harper JL. 1977. Population biology of plants. London: Academic Press.

Hooten M, Wikle C, Schwob M. 2020. Statistical implementations of agent-based demographic models. International Statistical Review. 88(2):441-461.

Kant I. 1781. Kritik der reinen vernunft.

Lenton TM, Kohler TA, Marquet PA, Boyle RA, Crucifix M, Wilkinson DM, Scheffer M. 2021. Survival of the systems. Trends Ecol Evol. 36(4):333-344.