Prototype Biodiversity Digital Twin: Forest Biodiversity Dynamics

Bekir Afsar[‡,§], Kyle Eyvindson[|], Tuomas Rossi[¶], Martijn Versluijs[#], Otso Ovaskainen[‡]

‡ Department of Biological and Environmental Science, Faculty of Mathematics and Science, University of Jyväskylä, P.O. Box 35 (Survontie 9C), FI-40014, Jyväskylä, Finland
§ University of Jyvaskyla, Faculty of Information Technology, P.O. Box 35 (Agora), FI-40014 University of Jyvaskyla, Finland
| Norwegian University of Life Sciences, Ås, Norway
¶ CSC – IT Center for Science Ltd., P.O. Box 405, FI-02101 Espoo, Finland
# Altenburg & Wymenga Ecological Consultants and Research, Suderwei 2, 9269 TZ Feanwâlden, Netherlands

Corresponding author: Bekir Afsar (bekir.b.afsar@jyu.fi)

Academic editor: Marina Golivets

Abstract

Forests are crucial in supporting biodiversity and providing ecosystem services. Understanding forest biodiversity dynamics under different management strategies and climate change scenarios is essential for effective conservation and management. This paper introduces the Forest Biodiversity Dynamics Prototype Digital Twin (pDT), integrating forest and biodiversity models to predict the effects of management options on forest ecosystems. The primary objective is to identify optimal management strategies that promote biodiversity, focusing on conservation and adaptation to different climate conditions. We start with the case of Finnish forests and bird species and plan to expand to include more European countries and a variety of species as the pDT is further developed.

Keywords

forest management, biodiversity conservation, LANDIS-II, joint species distribution models, multiobjective optimisation

Introduction

Biodiversity conservation stands as a paramount objective in today's ecological discussions, supporting the preservation of ecosystem functions and services that are vital for human well-being (Cardinale et al. 2012). Forest ecosystems are important because they contain a wide variety of species and play a key role in climate regulation and adaptation (Oliver et al. 2015, Chazdon et al. 2016). However, the conservation of forest biodiversity is facing growing threats from human activities and environmental disturbances, highlighting the need for immediate intervention to minimise negative consequences (Brondizio et al. 2019, Sarre 2020).

Forest management significantly impacts biodiversity dynamics, as management strategies have direct and indirect effects on species composition, habitat quality and ecosystem resilience (Lindenmayer and Franklin 2002, Kroll et al. 2017, Jung et al. 2022, Konczal et al. 2023). In addition to management activities, climate change also affects biodiversity by altering temperature and rainfall patterns, which can change where species occur, their phenology or reproduction (Parmesan 2006). Hence, the integration of forest management strategies with climate change adaptation measures is critical for mitigating these dual threats and ensuring the long-term sustainability of forest ecosystems.

Modelling forest ecosystems and their biodiversity dynamics has historically been a difficult challenge due to the complexity of ecological interactions, the variability of environmental factors and the limitations of early computational tools. For example, early models struggled to accurately simulate the effects of climate change on species distribution and the impact of disturbances (Scheller and Mladenoff 2007). However, recent advancements in computational modelling and data analytics have facilitated novel approaches to comprehensively understand and predict the effects of diverse management strategies and environmental disturbances on forest ecosystems.

Introducing the Forest Biodiversity Dynamics Prototype Digital Twin (hereafter, Forest pDT), this paper presents an innovative initiative aimed at addressing the complexities of forest management and climate change impacts on biodiversity. Forest pDT provides a comprehensive approach to explore forest ecosystem dynamics by leveraging the capabilities of advanced modeling tools such as LANDIS-II (Scheller et al. 2007) for forest simulation and Hierarchical Modelling of Species Communities (HMSC) (Ovaskainen et al. 2017, Ovaskainen and Abrego 2020) for biodiversity modelling. By simulating forest succession, disturbances and species-environment relationships, the Forest pDT offers insights into how different management strategies and climate change scenarios could influence biodiversity dynamics over time.

In addition to exploring forest dynamics, we propose a digital twin application utilising interactive multiobjective optimisation methods (Miettinen et al. 2008) as a decision support system, which harnesses the predictive power of the Forest pDT. Interactive methods, allowing decision-makers to iteratively incorporate their expertise in decision-making, prove beneficial (Miettinen et al. 2008, Misitano et al. 2021). These methods suggest solutions reflecting decision-maker's preferences, finding balance among the conflicting objectives. Through the digital twin application, stakeholders (i.e. decision-makers, for example, government and ministry representatives and state-owned companies) can provide their preferences iteratively and can explore different forest management options under different climate scenarios and learn the feasibility of their preferences. This approach enhances stakeholder engagement and facilitates informed decision-making in forest management, ensuring that ecological, social and economic objectives are effectively addressed. Through the integration of advanced modelling tools and decision support systems, the Forest pDT represents a significant advancement in our ability to promote adaptive and sustainable forest management practices.

Objectives

The Forest pDT has two key objectives. The first is to provide a comprehensive platform for modelling and simulating forest ecosystems and their biodiversity dynamics under different management strategies and climate change scenarios. The second is to facilitate informed decision-making in forest management by integrating interactive multiobjective optimisation methods into the digital twin application. This decision support system enables stakeholders and decision-makers, including government and ministry representatives and state-owned companies, to identify optimal forest management strategies, considering ecological, social and economic objectives under different climate scenarios.

Ultimately, the Forest pDT and its application seek to promote adaptive and sustainable forest management practices by providing stakeholders with the necessary tools and information to make informed decisions that balance biodiversity conservation, ecosystem resilience and societal needs.

Workflow

The workflow of the Forest pDT application integrates LANDIS-II as the forest simulator, HMSC as the biodiversity model and an interactive multiobjective optimisation method. Fig. 1 illustrates the conceptual schema of the Forest pDT, showcasing the interconnected components and their interactions. Stakeholders who take decisions provide preference information (step 1) and the interactive method generates solutions (step 6) utilising the Forest pDT. To involve the Forest pDT in the optimisation process, we implement a configurator to set up the simulator, based on decision variable values (i.e. different management options) (step 2) and an interpreter to calculate objective and constraint functions (e.g. maximising timber production, carbon storage, deadwood and habitat suitability) (step 5). The interpreter combines various outputs from LANDIS-II, such as above-ground biomass, below-ground biomass and woody debris, to derive objectives like carbon storage. LANDIS-II predicts harvested biomass, deadwood, carbon storage and woody debris (step 3-b), while HMSC forecasts species distributions (step 4) using biomass volume and mean age predicted by LANDIS-II (step 3-a).

Data

The Forest pDT relies on a variety of crucial data streams to provide information for its simulations and predictions. These data streams encompass climate data, the Finnish Multi-source National Forest Inventory (MS-NFI) data, land-cover data, soil data, species occurrence data and species trait data. It is important to note that the data mentioned in this section pertain to the Finnish context, highlighting the necessity for its expansion to other countries.

Climate data obtained from the Earth System Grid Federation (ESGF) are essential for understanding the environmental conditions influencing forest dynamics and species (e.g. birds) occurrences, particularly under different Representative Concentration Pathways (RCPs), RCP 4.5 and RCP 8.5, as outlined by the Intergovernmental Panel on Climate Change (IPCC) (Van Vuuren et al. 2011). The MS-NFI data from the Natural Resources Institute of Finland (LUKE) provide wall-to-wall information regarding forest structure and composition, while land-cover data derived from CORINE Land Cover maps and soil data from the European Soil Data Centre (ESDAC) offer essential information on habitat characteristics and ecological regions, shaping the preparation of input files for LANDIS-II.

Species occurrence data, sourced from the Finnish Museum of Natural History (LUOMUS), record occurrences of 190 bird species from 2007 to 2019 across 2920 transects, while species trait data provide ecological traits and Red-List status for these bird species (Storchová and Hořák 2018) (the data is available on DRYAD). These data streams are integral to parameterising the HMSC model.

Table 1 outlines these critical datasets, their sources and reference links for further information. The data used in the Forest pDT follow community standards and are provided in non-proprietary formats, such as CSV for species occurrence and trait data, NetCDF for climate data and GeoTIFF for MS-NFI, land-cover and soil data.

Model

The Forest pDT seamlessly integrates two primary models: LANDIS-II for forest simulation and HMSC for biodiversity modelling. LANDIS-II is a spatially explicit, dynamic forest landscape model capable of simulating forest growth, disturbance and succession processes over large spatial and temporal scales (Scheller et al. 2007). It is openly available with extensive documentation and highly customisable, with dozens of extensions to choose from. It utilises input data such as climate, MS-NFI, soil and land cover to predict changes in forest composition and structure under different management scenarios and climate conditions. Outputs from LANDIS-II include maps representing forest attributes such as tree species distribution, age structure, above-ground biomass, below-ground biomass, deadwood and woody debris, generated over a simulation period of 100 years.

Forest simulations have been conducted with LANDIS-II PnET-Succession extension V4.0.1 (De Bruijn et al. 2014) for a combination of climatic scenarios (e.g. RCP 4.5 and RCP 8.5). The study area covers the entire forested area of Finland, totalling 33.8 million hectares, at a resolution of 100 m by 100 m. Currently, we consider seven management regimes applied on the forest stands in the study area that are either being implemented or considered for application in Finland by government agencies: Business as usual

(BAU), set aside (SA – no management), EXT10 (BAU with postponed final harvesting by 10 years), EXT30 (BAU with postponed final harvesting by >= 30 years), GTR30 (BAU with 30 green trees retained/ha at final harvest), NTLR (No thinning – final harvest threshold values as in BAU), NTSR (No thinning – minimum final harvest threshold values) (Triviño et al. 2017).

Complementing LANDIS-II, we use HMSC, which belongs to the class of joint species distribution models (Warton et al. 2015). HMSC models species-environmental relationships using a generalised linear model and is able to borrow information among species, based on their trait similarity and phylogenetic similarity. By integrating species trait data and occurrence records, HMSC provides insights into the potential impacts of forest management and climate change on biodiversity. The outputs from LANDIS-II, such as predicted tree biomass and mean age, serve as inputs to the HMSC model, which generates species distribution maps. In this study, our focus lies on boreal forest birds in Finland.

FAIRness

The Forest pDT maintains standards to ensure the FAIRness of its data sources, processing procedures and model output data. Input data from various repositories undergo thorough documentation, with careful attention to detail. Each input data point is referenced using persistent identifiers if previously published, enhancing its traceability and accessibility. Moreover, the pDT employs standard terms, controlled vocabularies and ontologies to ensure interoperability across datasets and models.

Scripts and tools developed for the pDT will be documented and made publicly available. The parameterisation process for the LANDIS-II and HMSC models in the pDT is ongoing and, once completed, the source codes for both scripts and models will be made publicly accessible on the BioDT GitHub organisation.

Furthermore, the data used by the models and the resulting output data will be openly available, pending compliance with licensing and data-sharing agreements, to ensure transparency and facilitating the reproducibility of results. These measures collectively contribute to the FAIRness of the pDT, promoting its accessibility, interoperability and reusability within the scientific community.

Performance

The primary challenge for Forest pDT is the computational demands of its underlying models. Specifically, the time required to run simulations using LANDIS-II and HMSC across the entire study area during each optimisation cycle poses a significant bottleneck due to its intensive computational nature. As a result, conducting these simulations in real-time alongside the interactive decision-making process with decision-makers becomes impractical.

Transitioning the pDT to an high-performance computing (HPC) setup is a pivotal step to overcoming computational bottlenecks and involves optimising the performance of the system. To address this, we plan to implement LANDIS-II simulations in parallel for Finland by subdividing input maps into smaller areas, processed concurrently and then integrating outputs for smaller areas to generate comprehensive outputs. Applying this strategy enhances efficiency by distributing computational workload across multiple cores, thereby accelerating model run-time. In preparation for HPC execution, the LANDIS-II model and required libraries have been containerised and test executions have been performed on the CPU partition of the LUMI supercomputer by using Singularity/Apptainer. Furthermore, the translation of HMSC into TensorFlow/Python is underway to leverage GPU acceleration, enabling enhanced performance on HPC systems (Rahman et al. 2024). This transition promises significant reductions in overall model run-time, potentially by up to three orders of magnitude compared to CPU-only implementations.

Moreover, to save decision-makers' time, we plan to simulate all necessary scenarios (combinations of considered management regimes and climate scenarios) and generate predictions before the actual decision-making, allowing for a more efficient process, based on pre-computed optimal solutions. This approach streamlines the workflow, ensuring timely and effective execution of simulations within the pDT framework, ultimately facilitating informed decision-making in forest management.

Interface and outputs

The Forest pDT application facilitates user interaction through the DESDEO framework, an open-source Python-based platform for interactive multiobjective optimisation methods (Misitano et al. 2021). End users (i.e. decision-makers) engage in iterative decision-making processes facilitated by various interactive optimisation methods available within DESDEO. Unlike other pDTs within the BioDT project, the Forest pDT has a specific user interface where the decision-makers can explore different management options under diverse scenarios, evaluate trade-offs among the objectives, assess the feasibility of their preferences and find the most appropriate forest management strategies.

The outputs of the pDT include optimised management strategies for individual forest stands, based on identified objectives and various climate scenarios. Visualisations generated by the pDT will enable stakeholders to comprehend the implications of different management options, assess their impacts on forest ecosystems and make informed decisions aligned with their preferences.

Integration and sustainability

In the context of integration and sustainability, we plan to pilot the incorporation of Destination Earth (DestinE) data into the Forest pDT. DestinE is an EU initiative to build a highly accurate digital twin of the Earth for the analysis of natural events and human activities. Currently, the Forest pDT relies on climate data sourced from the ESGF, which we aim to complement or replace with ClimateDT data from DestinE. This initiative serves as an initial step towards enhancing the pDT's capabilities and resilience. By leveraging ClimateDT's comprehensive climate information, we seek to improve the accuracy and robustness of our simulations and predictions. This piloting lays the groundwork for future developments, including the exploration of additional integration opportunities via DestinE’s data lake.

Application and impact

The Forest pDT can serve as a decision support tool for government and ministry representatives, as well as state-owned companies involved in forest management. From a scientific perspective, research and academic institutions can engage with us on projects related to biodiversity conservation and sustainable forest management, leveraging the pDT's capabilities. Additionally, they can reproduce and extend our work, given that the code and data will be made openly available. Looking ahead, there is potential to extend the pDT application's beyond academia through a business model approach. This could involve customising the pDT to address specific regional and individual client needs, thereby offering valuable support to forest companies in the private sector. Such expansion initiatives could facilitate access to advanced decision-making tools, encouraging collaboration among various stakeholders and promoting sustainable land management practices in forest ecosystems.

As the Forest pDT transitions towards broader applicability, its scalability across taxonomic, temporal and spatial scales presents both opportunities and challenges. Scaling up entails extending the pDT's capabilities to encompass a wider range of species, timeframes and geographical regions beyond its initial prototype scope. This expansion necessitates model enhancements to accommodate the increased complexity and diversity of ecological systems. For instance, within the LANDIS-II model, adaptations may be required to incorporate additional species interactions and ecological processes relevant to diverse taxonomic groups. Similarly, the HMSC model may need refinements to better capture the dynamics of species distributions across larger spatial extents and longer temporal scales.

Moreover, scaling up the pDT demands careful consideration of data storage, processing and flow. As the volume and variety of input data increase, efficient data management becomes important to ensure timely and accurate simulations. Adopting FAIR principles on data management partially addresses these challenges. Additionally, leveraging computational resources, such as the LUMI supercomputer within the BioDT project framework and other private or public cloud computing options, help mitigate some of the computational limitations. However, looking beyond the current funding period, long-term sustainability remains a concern. Additionally, advanced data flow mechanisms, such as parallel processing and distributed computing, may need to be implemented to facilitate seamless integration of diverse datasets and models across different spatial and temporal scales.

In conclusion, the Forest pDT represents a significant advancement in forest management decision support systems. By integrating models such as LANDIS-II and HMSC, coupled with interactive methods, the pDT offers a comprehensive platform for stakeholders to assess and navigate complex ecological dynamics. Its potential applications span various sectors, from government agencies and forest companies to research institutions, facilitating collaboration and informed decision-making. Ultimately, the pDT stands poised not only to address current challenges in forest management, but also to provide informtion for future policy decisions and promote sustainable practices for preserving global forest ecosystems.

Acknowledgements

This study has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101057437 (BioDT project, https://doi.org/10.3030/101057437). Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. We acknowledge the EuroHPC Joint Undertaking for awarding this project access to the EuroHPC supercomputer LUMI, hosted by CSC (Finland) and the LUMI consortium through a EuroHPC Development Access call.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Brondizio ES, Settele J, Díaz S, Ngo HT (Eds) (2019) Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. URL: https://www.ipbes.net/global-assessment
- Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, Naeem S, et al. (2012) Biodiversity loss and its impact on humanity. Nature 486 (7401): 59-67. https://doi.org/10.1038/nature11148
- Chazdon RL, Brancalion PH, Laestadius L, Bennett-Curry A, Buckingham K, Kumar C, Wilson SJ, et al. (2016) When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration. Ambio 45 (5): 538-550. https://doi.org/10.1007/s13280-016-0772-y
- De Bruijn A, Gustafson EJ, Sturtevant BR, Foster JR, Miranda BR, Lichti NI, Jacobs DF (2014) Toward more robust projections of forest landscape dynamics under novel environmental conditions: Embedding PnET within LANDIS-II. Ecological Modelling 287: 44-57. https://doi.org/10.1016/j.ecolmodel.2014.05.004

- Jung M, Lewis M, Lesiv M, Arnell A, Fritz S, Visconti P (2022) The global exposure of species ranges and protected areas to forest management. Diversity and Distributions 28 (7): 1487-1496. https://doi.org/10.1111/ddi.13582
- Konczal AA, Derks J, Koning JH, Winkel G (2023) Integrating nature conservation measures in european forest management - An exploratory study of barriers and drivers in 9 european countries. Journal of Environmental Management 325: 116619. https://doi.org/10.1016/j.jenvman.2022.116619
- Kroll AJ, Verschuyl J, Giovanini J, Betts MG (2017) Assembly dynamics of a forest bird community depend on disturbance intensity and foraging guild. Journal of Applied Ecology 54 (3): 784-793. https://doi.org/10.1111/1365-2664.12773
- Lindenmayer DB, Franklin JF (2002) Conserving forest biodiversity: A comprehensive multiscaled approach. Island Press, Washington, DC, USA.
- Miettinen K, Ruiz F, Wierzbicki AP (2008) Introduction to multiobjective optimization: Interactive approaches. In: Branke J, Deb K, Miettinen K, Slowiński R (Eds) Multiobjective Optimization: Interactive and Evolutionary Approaches. Springer https://doi.org/10.1007/978-3-540-88908-3_2
- Misitano G, Saini BS, Afsar B, Shavazipour B, Miettinen K (2021) DESDEO: The modular and open source framework for interactive multiobjective optimization. IEEE Access 9: 148277-148295. https://doi.org/10.1109/ACCESS.2021.3123825
- Oliver TH, Heard MS, Isaac NJ, Roy DB, Procter D, Eigenbrod F, Bullock JM, et al. (2015) Biodiversity and resilience of ecosystem functions. Trends in ecology & evolution 30 (11): 673-684. https://doi.org/10.1016/j.tree.2015.08.009
- Ovaskainen O, Tikhonov G, Norberg A, Guillaume Blanchet F, Duan L, Dunson D, Abrego N, et al. (2017) How to make more out of community data? A conceptual framework and its implementation as models and software. Ecology Letters 20 (5): 561-576. https://doi.org/10.1111/ele.12757
- Ovaskainen O, Abrego N (2020) Joint species distribution modelling: With applications in R. Cambridge University Press https://doi.org/10.1017/9781108591720
- Parmesan C (2006) Ecological and evolutionary responses to recent climate change. Annual Review of Ecology, Evolution and Systematics 37: 637-669. https://doi.org/10.1146/annurev.ecolsys.37.091305.110100
- Rahman AU, Tikhonov G, Oksanen J, Rossi T, Ovaskainen O (2024) Accelerating joint species distribution modeling with Hmsc-HPC: A 1000x faster GPU deployment. bioRxiv https://doi.org/10.1101/2024.02.13.580046
- Sarre A (Ed.) (2020) Global forest resources assessment, 2020: Main report. URL: https://www.fao.org/forest-resources-assessment/2020/en/
- Scheller RM, Mladenoff DJ (2007) An ecological classification of forest landscape simulation models: tools and strategies for understanding broad-scale forested ecosystems. Landscape Ecology 22: 491-505. https://doi.org/10.1007/s10980-006-9048-4
- Scheller RM, Domingo JB, Sturtevant BR, Williams JS, Rudy A, Gustafson EJ, Mladenoff DJ (2007) Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution. Ecological Modelling 201 (3-4): 409-419. https://doi.org/10.1016/j.ecolmodel.2006.10.009
- Storchová L, Hořák D (2018) Life-history characteristics of European birds. Global ecology and biogeography 27 (4): 400-406. https://doi.org/10.1111/geb.12709
- Triviño M, Pohjanmies T, Mazziotta A, Juutinen A, Podkopaev D, Le Tortorec E, Mönkkönen M (2017) Optimizing management to enhance multifunctionality in a boreal

forest landscape. Journal of Applied Ecology 54 (1): 61-70. https://doi.org/10.1111/1365-2664.12790

- Van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Rose SK, et al. (2011) The representative concentration pathways: an overview. Climatic Change 109: 5-31. https://doi.org/10.1007/s10584-011-0148-z
- Warton DI, Blanchet FG, O'Hara RB, Ovaskainen O, Taskinen S, Walker SC, Hui FK (2015) So many variables: Joint modeling in community ecology. Trends in Ecology & Evolution 30 (12): 766-779. https://doi.org/10.1016/j.tree.2015.09.007

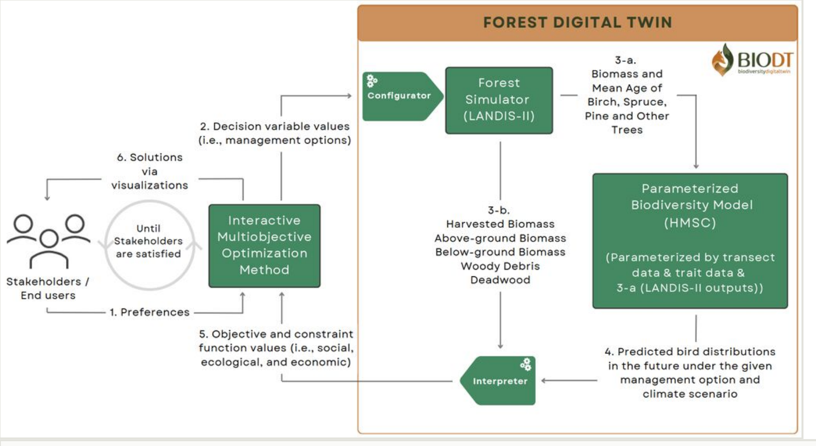

Figure 1.

The Forest pDT workflow diagram, linking forest simulation and biodiversity modelling with interactive multiobjective optimisation.

Table 1.

Data resources for the Forest pDT.

Data	Source	Link
Climate data	ESGF	https://aims2.llnl.gov/
MS-NFI	LUKE	http://kartta.luke.fi/opendata/valinta-en.html
Land-cover data	CORINE	https://land.copernicus.eu/en/products/corine-land-cover
Soil data	ESDAC	https://esdac.jrc.ec.europa.eu/
Species occurrence data	LUOMUS	https://www.luomus.fi/en
Species trait data	DRYAD	https://datadryad.org/stash/dataset/doi:10.5061/dryad.n6k3n