Received: 10 November 2022 | Accepted: 3 May 2023

DOI: 10.1111/gcb.16808

Global Change Biology WILEY

TECHNICAL ADVANCE

Climate–ecosystem modelling made easy: The Land Sites Platform

Lasse T. Keetz[1] | Eva Lieungh[2] | Kaveh Karimi-Asli[1] | Sonya R. Geange[3] | Emiliano Gelati[1] | Hui Tang[1,2,4] | Yeliz A. Yilmaz[1,5] | Kjetil S. Aas[1,6] | Inge H. J. Althuizen[7] | Anders Bryn[2,5] | Stefanie Falk[8] | Rosie Fisher[6,9] | Anne Fouilloux[10] | Peter Horvath[2] | Sunniva Indrehus[11] | Hanna Lee[7,12] | Danica Lombardozzi[9] | Frans-Jan W. Parmentier[1,5,13] | Norbert Pirk[1] | Vigdis Vandvik[3,14] | Ane V. Vollsnes[5,15] | Olav Skarpaas[2] | Frode Stordal[1,5] | Lena M. Tallaksen[1]

[1]Department of Geosciences, University of Oslo, Oslo, Norway

[2]Natural History Museum, University of Oslo, Oslo, Norway

[3]Department of Biological Sciences, University of Bergen, Bergen, Norway

[4]Finnish Meteorological Institute, Climate System Research, Helsinki, Finland

[5]Centre for Biogeochemistry in the Anthropocene, University of Oslo, Oslo, Norway

[6]CICERO Center for International Climate Research, Oslo, Norway

[7]Division of Climate and Environment, NORCE Norwegian Research Centre, Bergen, Norway

[8]Department of Geography, Ludwig Maximilian University of Munich, Munich, Germany

[9]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, Colorado, USA

[10]Simula Research Laboratory AS, Oslo, Norway

[11]Norwegian Geotechnical Institute (NGI), Oslo, Norway

[12]Department of Biology, Norwegian University of Science and Technology NTNU, Trondheim, Norway

[13]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden

Abstract

Dynamic Global Vegetation Models (DGVMs) provide a state-of-the-art process-based approach to study the complex interplay between vegetation and its physical environment. For example, they help to predict how terrestrial plants interact with climate, soils, disturbance and competition for resources. We argue that there is untapped potential for the use of DGVMs in ecological and ecophysiological research. One fundamental barrier to realize this potential is that many researchers with relevant expertize (ecology, plant physiology, soil science, etc.) lack access to the technical resources or awareness of the research potential of DGVMs. Here we present the Land Sites Platform (LSP): new software that facilitates single-site simulations with the Functionally Assembled Terrestrial Ecosystem Simulator, an advanced DGVM coupled with the Community Land Model. The LSP includes a Graphical User Interface and an Application Programming Interface, which improve the user experience and lower the technical thresholds for installing these model architectures and setting up model experiments. The software is distributed via version-controlled containers; researchers and students can run simulations directly on their personal computers or servers, with relatively low hardware requirements, and on different operating systems. Version 1.0 of the LSP supports site-level simulations. We provide input data for 20 established geo-ecological observation sites in Norway and workflows to add generic sites from public global datasets. The LSP makes standard model experiments with default data easily achievable (e.g., for educational or introductory purposes) while retaining flexibility for more advanced scientific uses. We further provide tools to

Lasse T. Keetz and Eva Lieungh should be considered joint first author.

[14]Bjerknes Centre for Climate Research, University of Bergen, Bergen, Norway

[15]Department of Biosciences, University of Oslo, Oslo, Norway

Correspondence
Eva Lieungh, Natural History Museum, University of Oslo, Oslo, Norway.
Email: evaleriksen@gmail.com

Funding information
Norges Forskningsråd, Grant/Award Number: 294948; Universitetet i Oslo, Grant/Award Number: UiO/GEO103920

visualize the model input and output, including simple examples to relate predictions to local observations. The LSP improves access to land surface and DGVM modelling as a building block of community cyberinfrastructure that may inspire new avenues for mechanistic ecosystem research across disciplines.

KEYWORDS
Application Programming Interface (API), Community Land Model (CLM), Docker container, Dynamic Global Vegetation Model (DGVM), ecological modelling, Functionally Assembled Terrestrial Ecosystem Simulator (FATES), Graphical User Interface (GUI), Land Surface Model (LSM)

1 | INTRODUCTION

Vegetation strongly influences energy, water and biogeochemical (e.g., carbon) exchanges between the land surface and the atmosphere (Bonan & Doney, 2018). The terrestrial biosphere and its interplay with the Earth system remain one of the largest sources of uncertainties for ecosystem and climate predictions (IPCC, 2022; Urban et al., 2016), and complex models are necessary to understand the numerous physical, biogeochemical and ecological processes that govern these interactions (Blyth et al., 2021).

Recent developments open up new avenues for integrating biological and Earth system sciences (Dietze et al., 2013; Kyker-Snowman et al., 2022): increasing computational capacity facilitates the use of more complex models on higher spatial resolutions, with more heterogeneity and more sophisticated ecological process representations (Fisher & Koven, 2020; Pettorelli et al., 2021; "Think Big and Model Small," 2022). This enables interdisciplinary research to address classical ecological and ecosystem-level research questions through a mechanistic modelling lens. For instance, what is the importance of functional diversity in ecosystem responses to climate? Which representations of physiological processes are robust, and how do different representations affect responses to environmental forcing? What biotic interactions influence the outcome of vegetation–climate feedbacks? Exploring broad and relevant questions through empirical, experimental, theoretical and modelling lenses opens up opportunities for cross-disciplinary learning, especially if the practitioners of these traditionally disparate approaches can communicate, share data and learn from each other.

Dynamic Global Vegetation Models (DGVMs) simulate emergent properties of ecosystems from the 'bottom-up' across space and time from physiological principles. DGVMs can simulate transient responses to climate and pressures on time scales at which ecological processes operate, and thereby circumvent some equilibrium assumptions of correlative approaches (Hartig et al., 2012). Advanced DGVMs, such as vegetation demographic models (see Fisher et al., 2018), can explore both the mechanistic underpinnings and the broader-scale consequences of community assembly, tradeoffs, life history strategies, disturbance and recovery, interactions and the role of functional biodiversity in the face of global change (Bugmann & Seidl, 2022; Chitra-Tarak et al., 2021; Rogers et al., 2022).

At the same time, increased complexity and realism in these models invariably present new uncertainties, fuelling a continuous need to strengthen the models' empirical foundations and expose process representations to rigorous testing (Collier et al., 2018; Fisher & Koven, 2020). A plethora of model parameters must be derived individually from empirical relationships or experiments. This poses limitations on the processes and locations we can currently model with high confidence (Hartig et al., 2012). Managing the comprehensive ambition and complexity of DGVMs thus requires the mobilization of knowledge from experts across diverse fields (Dietze et al., 2013). We argue that their use and development by a broader community will benefit both the science of DGVMs and enhance the role of trait-based ecology and process understanding in ecological science.

Widespread utilization and development of DGVMs by specialists in, for example, ecology, hydrology, snow and soil science faces numerous barriers. These models are often linked to and housed within the software architecture of Earth system models, which have some of the most complex scientific code in existence and significant technical requirements. Proprietary code or limited access to computational resources often impose additional hurdles. New modellers experience steep technical and theoretical learning curves to set up meaningful experiments. At the same time, some trained modellers lack experience with field- and laboratory-based approaches and whole-organism ecology. Arguably, the exchange of technical know-how, knowledge and empirical data could be accelerated if these barriers to communication and hands-on collaboration were alleviated. Enhancing integration between mechanistic ecosystem modelling and statistical-, laboratory- or field-oriented ecology requires community cyberinfrastructure tools to improve accessibility to complex model frameworks (Fer et al., 2020; Lombardozzi et al., n.d.).

To lower the technical barriers to modelling, we present the Land Sites Platform (LSP; Karimi-Asli, Keetz, Lieungh, Yilmaz, et al., 2022). It simplifies simulations with recent model versions of a demographic DGVM embedded in a land surface modelling framework (Section 2.1). The software is open-source and user-friendly: it includes a Graphical User Interface (GUI), an Application Programming Interface (API; Section 2.2) and self-guided analysis tools for input and output (Section 2.3). Owing to their complexity and driver requirements, many land model frameworks require users to choose and install the correct versions of source code, numerous external libraries and compilers (Fer et al., 2020). Setting this up on any computer is dependent on the operating system and requires substantial knowledge of software engineering. The LSP works on many operating systems and automates the installation of all software dependencies by virtualizing the software environment with Docker containers (Section 2.4; Table 1). A set of integrated sites (Section 2.5) provide a starting point with default atmospheric and land surface input data, and we provide instructions to add new sites. The user guide, technical documentation, site descriptions and analysis notebooks are written in a non-technical language (The NorESM-LSP Development Team, 2023). Educational workshops, interdisciplinary research, streamlined workflows for users without high-performance computing access and model development are possible with this open and collaborative tool (Section 3). Our goal is to make state-of-the-art process-based vegetation demographic modelling accessible to a broader community and to facilitate reproducible research on vegetation–climate processes.

2 | THE LAND SITES PLATFORM

The LSP software architecture relies on GitHub repositories for code management and on Docker containers to run the models and present easy user interfaces (Figure 1). Our software simplifies the interface between models and users, and between different software components, thereby reducing the number of technical challenges for running a simulation with the model framework.

2.1 | Model framework: NorESM, CLM and FATES models

The LSP's demographic DGVM relies on the software infrastructure of an Earth system model, such as a driver and case control system. It configures, compiles and executes the model, and orchestrates components providing the DGVM's boundary conditions. Earth system models numerically simulate the atmosphere, ocean, land, land ice and sea ice (Figure 2a). By tracing energy, water, carbon and other substances through these components, they represent the fundamental processes governing Earth's climate. The LSP's host model, the Norwegian Earth System Model (NorESM; Seland et al., 2020), contributed to the Coupled Model Intercomparison Project Phase 6 (CMIP6; Eyring et al., 2016), a core foundation for the Intergovernmental Panel on Climate Change's sixth assessment report (IPCC, 2022). The land component of the NorESM is the Community Land Model (CLM; Lawrence et al., 2019). The CLM solves process equations in discrete vertical layers for different land cover types, including natural vegetation (Figure 2b).

To model vegetation demographics within the natural vegetation land unit, we use the Functionally Assembled Terrestrial

TABLE 1 Glossary of technical terms.

Software engineering	
Container Container image Docker	*Containers* are isolated, virtualized computer environments based on a read-only *image* file with source code, libraries, dependencies and tools. *Docker* Inc. is a company that provides container solutions.
Application Programming Interface (API)	An interface for computer programs to communicate with each other efficiently.
Graphical User Interface (GUI)	A visual interface between humans and software or hardware, with clickable buttons.
Command line interface	Also called a terminal emulator; a computer program that receives commands from a user in the form of lines of text.
Jupyter notebook JupyterLab Jupyter Server	*Jupyter notebooks* combine computer code and rich text elements. They enable data analysis and viewing and plotting model output. *JupyterLab* allows users to develop and run notebooks interactively in a browser. *Jupyter Server* is responsible for storing and organizing data for Jupyter applications.
Git GitHub Repository	*Git* software enables version control by tracking changes in files. *GitHub* is an online host of *repositories*: data structures of files and their version history.
Earth system modelling	
Simulation Case Model experiment	A *simulation* is one realization of a model run, where the model advances forward in time from some initial state. A simulation is performed from a *case* folder containing one specific model configuration. Research questions may be answered by *model experiments* consisting of one or more simulations.
Model coupling Model stub Data model	A *coupled* Earth system model passes information dynamically between model components (atmosphere, land, ocean, etc.), allowing for direct interactions and feedbacks. Alternatively, model components can be deactivated *stubs* with placeholders that do not send or receive data, or they can be forced with *data* from observations or previous simulations.
Forcing data	*Forcing data* are not impacted by the model, but are provided as input to define the necessary boundary conditions to the model, for example, temperature and precipitation.

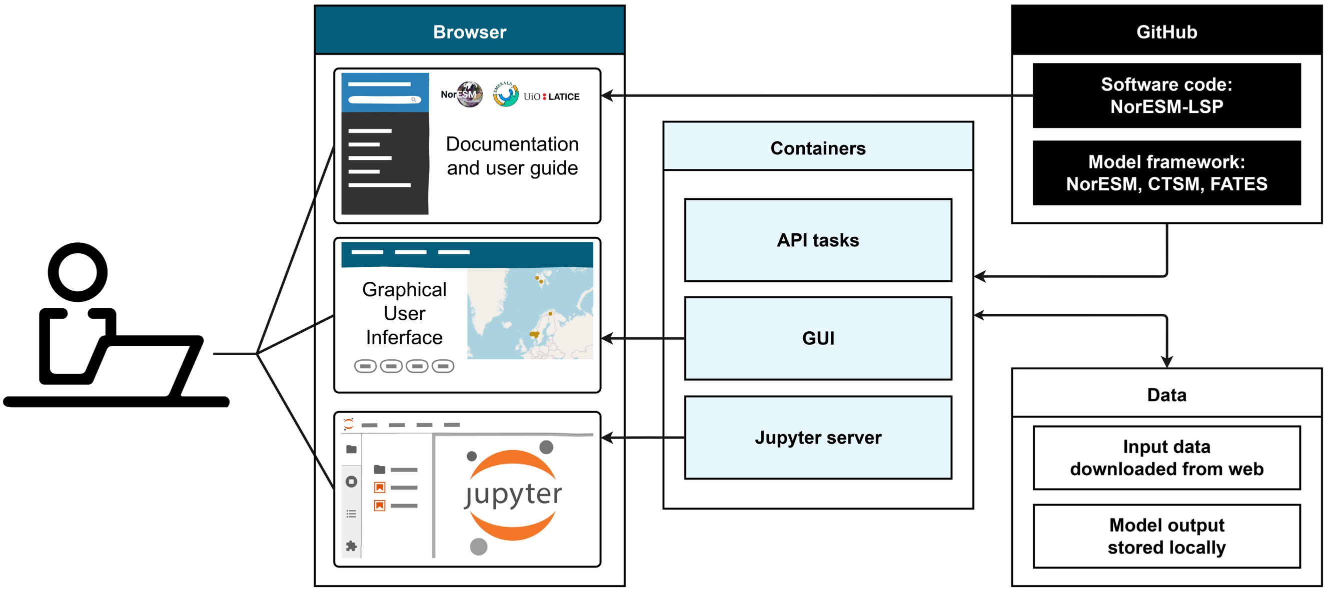

FIGURE 1 Simplified illustration of the main software structure of the Land Sites Platform (LSP) that serves the user interfaces. Users (left) access the user interfaces through their internet browser (Browser box): the documentation describes the workflow for customizing and running simulations with the Graphical User Interface (GUI), and for analysing the inputs and outputs with an integrated development environment (JupyterLab). Containers (centre) automate the installation of all software components, provide the interfaces (GUI/Jupyter server) and handle model configuration and simulations (API tasks). The code for the LSP, as well as the NorESM-CLM-FATES models, is managed in public GitHub repositories (top right). Model input data for the integrated field sites are automatically downloaded from external web storage, while the model output is stored on the user's computer (bottom right).

Ecosystem Simulator (FATES; Fisher et al., 2015; Koven et al., 2020; see Figure 2c). FATES is a vegetation demographic model (Fisher et al., 2018) that is an optional component of the CLM (hereafter CLM-FATES). In the LSP default setup, it models potential vegetation where distributions emerge via trait filtering. FATES condenses the tremendous functional diversity of the world's vegetation into cohorts of plant functional types (PFTs) and represents community dynamics by tracking cohort size and successional stage. Within each model grid cell's natural vegetation fraction, cohorts of different PFTs grow and compete for water and nutrient resources. Heterogeneity is created by differentiating patches of different ages since disturbance (Figure 2c illustrates four patches). Within a patch, cohorts of different PFTs compete with each other for light in the canopy and understory (Figure 2c illustrates cohorts as separate plant symbols). The number of PFTs and their functional traits can easily be modified, for example to better represent local ecological observations (Buotte et al., 2021; Koven et al., 2020). The 12 default PFTs represent trees, shrubs and grasses in broad non-phylogenetic categories based on, for example, leaf shape (needleleaf or broadleaf), phenology (cold deciduous, drought deciduous or evergreen) and photosynthetic pathway (C3 or C4). Grouping multiple individuals and species with similar functional properties into cohorts and PFTs endows FATES the computational efficiency for global simulations while also accounting for competition and recovery after disturbance. The default PFTs are experimental starting points for global simulations, and need regional or local tuning to represent specific ecosystems (e.g. Sulman et al., 2021).

2.2 | Accessible interfaces: API and GUI

We created an API as a user-friendly and scalable solution enabling communication between different components of the software (Karimi-Asli, Keetz, & Lieungh, 2022a). An API allows users to interact with a program by sending action commands with optional inputs (requests) to trigger processes. After performing a task, the API returns the results. The LSP's API handles creation of new cases, running cases and checking their status, deleting cases and downloading the outputs. Requests to create or run a case provide the desired model configuration. The API then generates and executes scripts that call the corresponding model functionalities via the model's original interfaces. By communicating with the model through an API, we hide the model's complexities and add validation of the user's input, making the process of setting up and running a case easier and less error-prone. The API allows us to automate parts of the workflow by bundling specific steps into fewer actions. For example, input data download and building a case are bundled into one action. Another advantage is that LSP developers can modify the code that processes the requests and responses for a specific version without changing the API's user-facing part.

Table 2 exemplifies an API request for creating a 1-year default simulation for site ‘BOR1’ with GSWP3 forcing starting from 1 January 1901. The API then retrieves other required information from a configuration file (e.g., the input data URL), executes the model tools to create and configure the case (e.g., ‘./create_newcase’ and ‘./xmlchange’), checks or downloads the input data and orchestrates the case settings and states via an internal database. Finally,

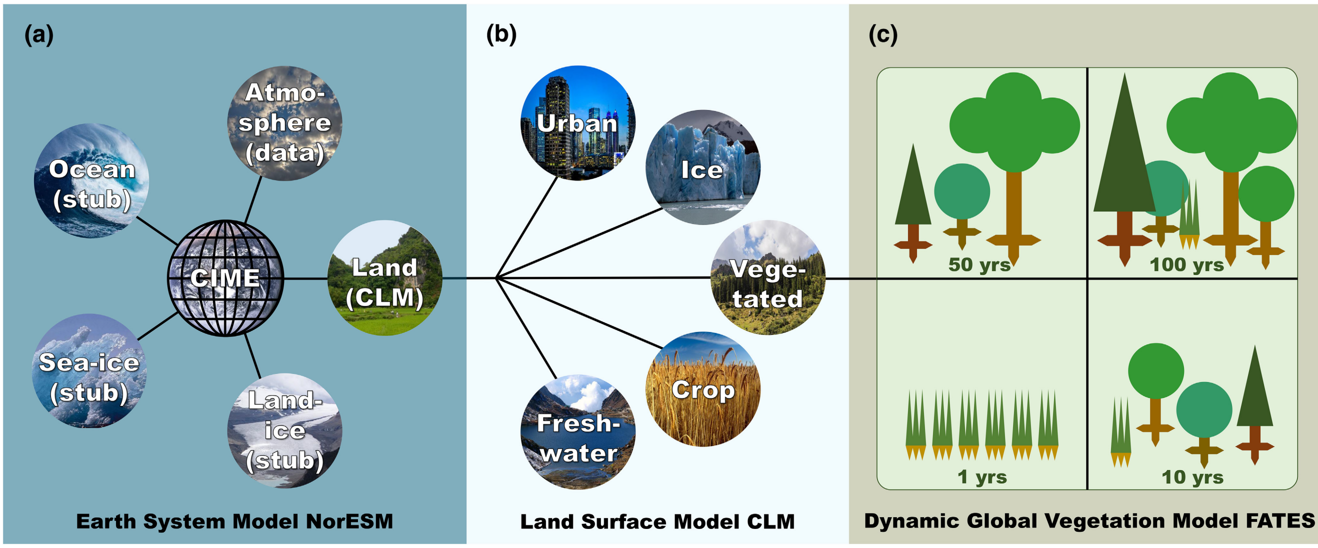

FIGURE 2 The Land Sites Platform model framework. a) the Earth system model (Norwegian Earth System Model; NorESM). The coupler—Common Infrastructure for Modeling the Earth (CIME)—connects components for the atmosphere, ocean, sea and land ice and land. The text in parentheses refers to the LSP model component setup: the Community Land Model (CLM) is the actively simulated land surface model, an observationally derived dataset replaces the atmosphere component (data mode), and the other components are deactivated ('stub' mode). b) the CLM divides the land surface into grid cells with urban, ice, vegetated, crop and freshwater fractions. c) the Dynamic Global Vegetation Model (Functionally Assembled Terrestrial Ecosystem Simulator; FATES) divides vegetated land fractions into age-since-disturbance patches, where cohorts of Plant Functional Types (PFTs) grow.

it returns a success or error response. Additional information about the API is in the LSP technical documentation (The NorESM-LSP Development Team, 2023), and all endpoints including usage examples are described under http://localhost:8000/api/v1/docs when the containers are running.

The LSP's GUI (Figure 3; Karimi-Asli, Lieungh, & Keetz, 2022) is an interactive tool to communicate with the API. Users can select among integrated sites, add new sites with their own data (Section 2.3), and see details of all existing model experiments (cases) they have created. After selecting an integrated site, the Create Case button triggers a pop-up with settings for a new simulation, such as the simulation period, CO_2 concentration, output types and a subset of FATES' parameters per PFT. Checkboxes allow disabling individual PFTs. An open text field adds flexibility for experienced users to change additional CLM parameters. Once a case is created, status information (pending, ready or failed) and new buttons appear to view the model settings and run, copy or delete the case. Note that the GUI presents a subset of the model framework's capabilities, which includes a large suite of optional process representations (e.g. plant hydraulics, nitrogen and phosphorus cycling, fire, logging and land use change). Advanced settings may require additional changes to the model, cases or input data, which experienced users can modify independent of the GUI or API.

2.3 | Data processing and analysis

The LSP by default performs simulations in 'land-only mode' where the land model is not dynamically coupled to the atmospheric model, but rather uses atmospheric datasets to force the land model (Figure 2). For version 1.0, we use the CLM default, observationally derived reanalysis dataset GSWP3 (Global Soil Wetness Project, Phase 3; Dirmeyer et al., 2006), which consists of gridded values with 0.5°×0.5° spatial resolution in three-hourly time steps from 1901 to 2014. It includes the incident (incoming) solar radiation ($W\,m^{-2}$), atmospheric temperature (K), precipitation ($mm\,s^{-1}$), atmospheric wind ($m\,s^{-1}$), atmospheric pressure (Pa) and atmospheric specific humidity ($kg\,kg^{-1}$). The CLM also requires background information describing the land surface—for example, the fraction of the grid cell occupied by different land cover types, soil properties or human population density over time. Finally, the model needs parameter files for CLM and FATES with values for parameters needed for the calculations.

To prevent users from having to download large global datasets, for example, climate forcing, we extracted grid-cell values from coordinates for the integrated sites. The data were compressed and uploaded to a web storage domain (Karimi-Asli, Keetz, & Lieungh, 2022b). When users set up a model experiment (case), the software downloads all required input data and configures the model accordingly. Workflows for adding new sites with custom data, based on model tutorials (https://ncar.github.io/CTSM-Tutorial-2022/, accessed 1 November 2022), are described in the LSP documentation (The NorESM-LSP Development Team, 2023).

The LSP also features Jupyter notebooks (Project Jupyter, 2014), interactive scripting environments with examples for visualizing the input data and analysing some of the many possible output variables (see Section 3). An additional container hosts Panoply (NASA

TABLE 2 An example API request and success response for creating a 1-year default case for an integrated site (BOR1). Long response strings were truncated and some entries removed for readability. See the NorESM-LSP Development Team (2023) for an overview of the API endpoints.

Request

```
curl -X 'POST' \
 'http://localhost:8000/api/v1/sites/' \
 -H 'accept: application/json' \
 -H 'Content-Type: application/json' \
 -d '{
 "site_name": "BOR1",
 "case_name": "MyCase",
 "variables": [
  {"name": "STOP_OPTION", "value": "nyears"},
  {"name": "STOP_N", "value": 1},
  {"name": "DATM_YR_START", "value": "2000"},
  {"name": "DATM_YR_END", "value": "2001"},
  {"name": "DATM_YR_ALIGN", "value": "2000"},
  {
   "name":  "RUN_STARTDATE",
   "value":  "2000-01-01"
  }
 ],
 "driver": "nuopc"
}'
```

Response body


```
{
  "id": "eda42eff3f4f29e3a37cb42e372e7fc4",
  "name": "MyCase",
  "model_version": "a5e48a...",
  "status": "INITIALISED",
  "date_created": "2022-10-31T17:59:23...",
  "compset": "2000_DATM%GSWP3v1_CLM...",
  "lat": 61.0355,
  "lon": 9.07876,
  "variables": [
    {"name": "STOP_N", "value": 1},
    {"name": "STOP_OPTION", "value":
"nyears"},
    ...
  ],
  "driver": "nuopc",
  "data_url": "https://.../BOR1.zip",
  "site": "BOR1",
  "create_task": {
    "task_id": "b30d000...", "status": "PENDING",
    "result": null, "error": null
  },
  "run_task": {
    "task_id": null, "status": null,
    "result": null, "error": null
  }
}
```


GISS, 2022), external software that facilitates model data exploration. Model outputs are stored in the local LSP repository.

2.4 | Containerization with Docker

Containerization with Docker (Docker, 2023) permits the LSP to work on current versions of macOS, Linux and Windows. Through this setup, we provide images and containers (Table 1) to configure a virtualized computing environment with source code, external libraries and compilers. Fixing the computing environment with the corresponding versions of the LSP, model and input data facilitates access to and reproducibility of model experiments. The LSP's containerization of CLM-FATES also allows easy integration into services such as cloud computing infrastructures, where users sign up or apply for computing resources without needing to install software locally. For example, a pilot version of a tool to run CLM-FATES is accessible on the open-source cloud infrastructure Galaxy (Fouilloux & Tang, 2021). See Melton et al. (2020) for another containerization example with the CLASSIC model, and Nüst et al. (2020) for best practice advice for writing Dockerfiles for reproducible data science.

2.5 | Integrated field sites

The LSP version 1.0 supports site-level simulations. As a starting point, it provides forcing data to run the model at 20 established geo-ecological observation sites in Norway (see the NorESM-LSP Development Team, 2023). Of these 20 sites, eight have flux tower (eddy covariance) measurements, and local climate and surface data (e.g., Pirk et al., 2017). The other 12 represent semi-natural grassland sites used in vegetation experiments examining climate change impacts on plants at the individual, population, community and ecosystem levels (Althuizen et al., 2018;

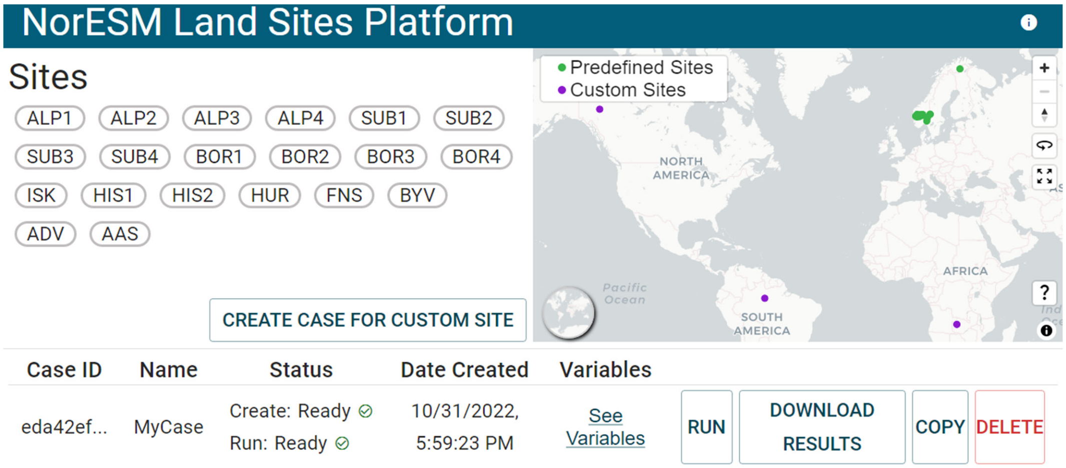

FIGURE 3 The LSP Graphical User Interface (edited screenshot). Users can choose among integrated sites (green markers), add custom sites (purple markers) and are guided through creating and running a case (i.e. simulation or model experiment) by customizing a selection of settings and model parameters. Existing cases are listed along with options to view model settings, run the case, copy and edit the case, or to delete it. The GUI sends requests to the API to operate the model. Background map by Carto. Map lines delineate study areas and do not necessarily depict accepted national boundaries.

Klanderud et al., 2015; Lynn et al., 2021; Töpper et al., 2018). Existing data from these sites include microclimate, weather, vegetation composition, plant traits and demography (Vandvik et al., 2022).

2.6 | Educational example application

We present a simplistic, educational application of the LSP to showcase its capabilities and the model outputs (Keetz et al., 2022). We also emphasize some possible pitfalls with land surface modelling at sites where the realized vegetation cover differs from predictions with a default model setup. Highlighting challenges serves to stimulate users to think carefully about how to set up and interpret model experiments, and to facilitate cross-disciplinary understanding along with some suggested ways forward.

We ran two simulations at the BOR1 site, a moderately grazed semi-natural grassland in Western Norway surrounded by boreal forest (approx. 61°N, 9°E; 589 m asl.; mean summer temperature 10.3°C; mean annual precipitation 600 mm; Klanderud et al., 2015). The first simulation included all 12 default PFTs to simulate potential natural vegetation, whereas the second only included the default 'cool C3 grass' and 'Arctic grass' PFTs. Each simulation ran for 1000 years while cycling the default GSWP3 climate forcing data from 2004 to 2014, with a constant atmospheric CO_2 concentration of 367 ppm. We kept the other CLM-FATES parameters at default values in both cases.

3 | PERSPECTIVES AND ADVANCES

Improvement of DGVMs and building process understanding in vegetation ecology requires a broad range of scientists to access, use, scrutinize and develop models like the CLM-FATES. Our educational example application (Figure 4) highlights several benefits and challenges.

The coarse-scale reanalysis climate data used to force the model, while capturing the seasonal variability, exhibit differences to the temperature measurements at the site (Figure 4a). This is not surprising, as the averaging of environmental variables across heterogeneous grid cells rarely represents the conditions at specific locations accurately. Considering input data uncertainties for model output interpretation is especially important in areas with high topographic variation, such as Norway. To adjust model experiments to local conditions, the forcing could be replaced with data products with higher spatial resolution or with local observations. While observations can also be valuable to adjust (e.g. bias-correct) coarse datasets (e.g. Yang et al., 2011), the default data provide continuous environmental information that can complement incomplete or short-term field observations.

Tracking the modelled growth of plant cohorts through time shows community assembly through trait filtering. The succession of simulated potential natural vegetation with all default PFTs (Figure 4b) predicts needleleaf evergreen trees as the only remaining PFT after reaching a steady state from initial bare-ground conditions. This is in line with the potential natural vegetation at this site (Bryn et al., 2013), and highlights how locally realized and modelled potential natural vegetation can diverge (Somodi et al., 2021). Previous DGVM studies over Scandinavia show limited ability to capture realized vegetation patterns (Hickler et al., 2012; Horvath et al., 2021). At the BOR1 site, an important explanation is likely the traditional farming practices, which are not currently included as explicit processes in the model.

To give a simplistic example of how changing the model configuration with the LSP allows emulating specific ecological conditions, we narrowed down the PFT list to two grass PFTs most closely

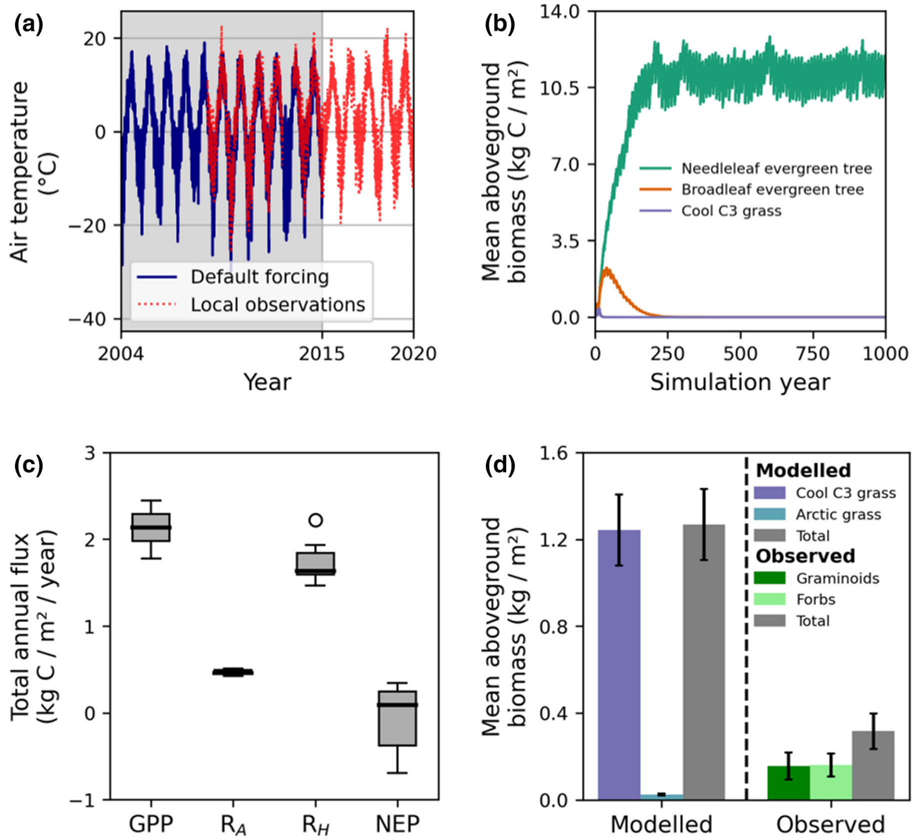

FIGURE 4 Plots from two example simulations illustrating the model climate forcing, model output and simple comparison to local observations at BOR1, Norway. (a) The daily mean temperature for default input data (GSWP3) plotted against 2 m logger data. Grey shade: forcing data period used for two example simulations, repeatedly cycled over 1000 simulation years. (b) Modelled aboveground biomass (AGB) per plant functional type (PFT) over the simulation period when all PFTs were allowed to grow (omitting PFTs with AGB < 0.3 kg C m^{-2}). (c) Modelled mean gross primary productivity (GPP), autotrophic (R_A) and heterotrophic (R_H) respiration and net ecosystem productivity (NEP) for the last 10 simulation years, only allowing 'Cool C3 grass' and 'Arctic grass' default grass PFTs to grow. (d) Mean AGB comparison of modelled grass PFTs from July of the last 10 years of the simulation only allowing 'Cool C3 grass' and 'Arctic grass' default grass PFTs to grow, versus observed for graminoids and forbs ($n = 16$; harvested in July 2015). Error bars represent one standard deviation.

resembling the realized vegetation (Figure 4c,d). CLM-FATES could also be configured to, for example, prescribe PFT distributions from satellite observations (Fisher & Koven, 2020). Once a suitable model experiment is set up, CLM-FATES yields, for example, productivity measures that are difficult to collect continuously in situ but are important for the ecosystem's impact on the global carbon budget under given climatic conditions (Figure 4c).

The biomass comparison to observed data in this simple example underlines that the global default PFTs are not expected to represent specific regional plant traits, nor to represent this semi-natural ecosystem (Figure 4d). The CLM-FATES currently models potential natural vegetation without domestic grazing, which would, for example, remove biomass from the vegetation. Adding such processes and calibrating model parameters with local observations could improve the model's fit to the site (Lambert et al., 2022). Both process understanding and data are required to disentangle biased predictions like overestimated biomass (Figure 4d). Biomass is an emergent property of productivity and turnover. Establishing which of these is biased high or low is the first task. For instance, new model sensitivity experiments where individual parameters are varied can be evaluated against data such as flux measurements of carbon assimilation and respiration.

While the basic capabilities of the LSP are targeted at new modellers, the software may serve as a shortcut to advanced model experiments, data integration and efficient cross-disciplinary communication that could inspire future model development (Figure 5). The advanced DGVM model framework (CLM-FATES) presents opportunities to test ecological hypotheses and scale up concepts through time and space in future research (Kyker-Snowman et al., 2022). We already know that ecological data and theory can improve DGVM parameters and structure (Nevalainen et al., 2022; Norby et al., 2016; Pastorello et al., 2020; Wullschleger et al., 2011, 2014). Observations, for example in plant trait databases, can inform the direct or inverse estimation of model parameters (Dietze et al., 2014; Hartig et al., 2012; LeBauer et al., 2013). Other modelling approaches using ecological assumptions can refine or validate DGVM parameterization, such as machine learning (Beigaitė et al., 2022) and distribution modelling (Horvath et al., 2021). Laboratory experiments in plant physiology can inform parameter values or build the foundations of process formulations (Ball

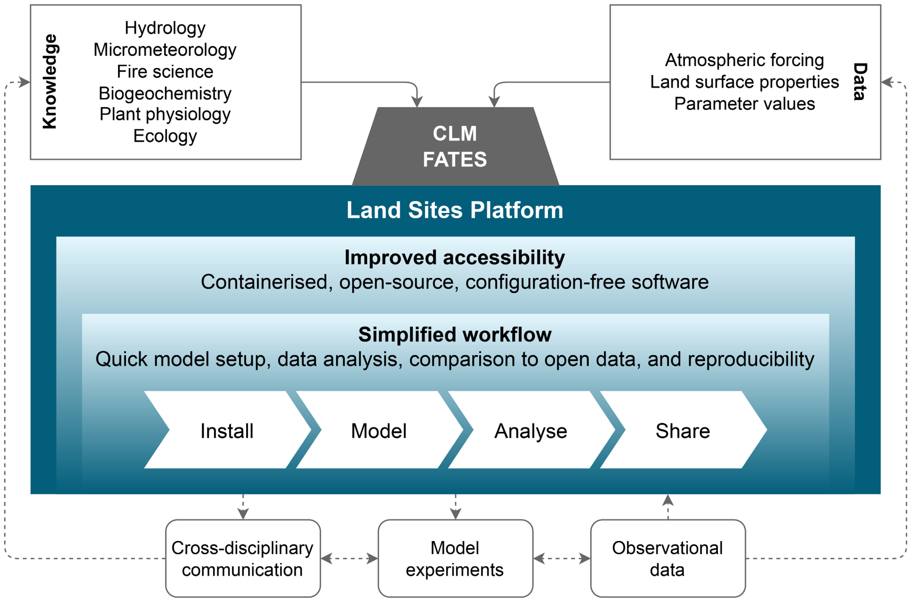

FIGURE 5 The LSP (centre blue box), adds layers of user-friendly software around the vegetation demographic model CLM-FATES (grey trapezoid). Technical facilitation reduces software and hardware barriers to modelling, while user-friendly interfaces simplify the workflow. Many disciplines contribute process understanding to the models (top left). Data (top right) can force the model and provide parameter estimates for model processes and plant functional types. Dotted arrows indicate possible outcomes of the LSP's facilitation: interconnections between cross-disciplinary communication, process understanding, new model experiments, improved understanding of model data needs, existing ecological data and how observations can be compared to model input and output.

et al., 1987). Concepts from community assembly theory may inform DGVMs' process representation (Scheiter et al., 2013). Vice versa, DGVMs may complement ecological hypothesis testing by tracing simulated plant growth from parameters through model processes to predicted states in a controlled environment. The LSP is well suited for interdisciplinary work on these issues because it lowers the technical thresholds and provides supporting materials for analysis and interpretation, thereby helping researchers focus on the scientific questions.

Modellers and field ecologists increasingly adopt open development and FAIR data sharing to facilitate community engagement, accessibility and data flow (e.g., Danabasoglu et al., 2020; Melton et al., 2020; Wilkinson et al., 2016). Tools and protocols conveniently relating observations to model outputs exist in different contexts and configurations (e.g. Fer et al., 2018; Hoffman et al., 2017). Several software-oriented efforts already aim to ease land surface model and DGVM usability and interpretation (Wang et al., 2014; Xu et al., 2017) and provide beginner-friendly instruments for education (e.g., LPJ-GUESS Education; Smith et al., 2001). Each solution comes with individual advantages but also limitations, such as missing cross-platform support, proprietary code or restricted possibilities to adapt the incorporated models and versions. The latter is crucial for collaborative efforts aiming at model advancement. The LSP is designed to be a flexible research tool, while simultaneously serving as a modern interface for education. We chose software solutions that are transferable to similar model frameworks, for example, with CESM instead of NorESM, or that could be merged with other initiatives to build on synergies. Further developments with the LSP may expand modelling capabilities with additional input datasets, climatic scenarios, model calibration workflows and alternative modes of the FATES model (see Fisher & Koven, 2020).

The LSP development was driven by needs to overcome technical challenges with land surface modelling, including knowledge transfer across institutional and geographical barriers, training resources for new modellers and setting up a shared software environment available to everyone. We combined ideas and experiences from biologists and geoscientists and involved research software engineers to pull the developments together and create the GUI, API and containers. Because the Earth system's complexity inspires increasingly complex scientific code structures, engaging software engineers is crucial to enhance the efficiency and scientific standards of model development (Purgar et al., 2022; Wieters & Fritzsch, 2018). The LSP was tested on many different computers by varied users, including students and professors at an interdisciplinary Master- and PhD-level course in Ecological Climatology (Bryn & Stordal, 2022; see also the NorESM-LSP Development Team, 2023, for performance and disc usage tests). Their feedback confirmed that they felt enabled to run the model and understand the results. They also stressed the need for training in Earth system science and DGVMs to set up robust simulations and interpret the outputs properly. This underlines the need for cross-disciplinary collaboration and training, and for creating community cyberinfrastructure beyond the mainly technical facilitation of the LSP.

4 | CONCLUSION

The scientific community needs interdisciplinary, open-access modelling tools and data, and stronger connections between geo- and biosciences to integrate global climate change and biodiversity science-policy agendas (Dietze et al., 2018; Pettorelli et al., 2021). Our experiences with building the LSP showed that this task requires lowering technical thresholds. Cross-disciplinary education with the LSP may stimulate collaboration and mutual understanding that lead to model development and improved process understanding. The full importance of the many aspects of natural diversity and ecological processes for global systems is still unknown (IPBES, 2019), and will remain unknown until collaborative research is strengthened. Dynamic, process-based models are increasingly important and sophisticated and can help understand the role of ecology within the Earth system, but require improved accessibility to be more widely adopted (Fer et al., 2020; Kyker-Snowman et al., 2022). While setting up and interpreting model experiments tailored to small-scale ecological research still requires additional model expertize, the LSP contributes to providing more accessible community cyberinfrastructure. It simplifies the modelling process for researchers and students without advanced programming skills while retaining the flexibility to set up custom and reproducible model experiments with a state-of-the-art model framework (NorESM-CLM-FATES). Engaging new users may ultimately improve our collective understanding of ecology within the Earth system, and of global environmental change as a result.

AUTHOR CONTRIBUTIONS

Hui Tang, Lasse Keetz, Eva Lieungh, Kaveh Karimi-Asli, Sonya Geange, Emiliano Gelati, Yeliz Yilmaz, Kjetil Aas, Inge Althuizen, Anders Bryn, Stefanie Falk, Rosie Fisher, Anne Fouilloux, Norbert Pirk, Vigdis Vandvik, Olav Skarpaas, Frode Stordal and Lena Tallaksen were involved in conceptualization. Kaveh Karimi-Asli, Lasse Keetz, Hui Tang, Emiliano Gelati, Eva Lieungh, Yeliz Yilmaz, Sunniva Indrehus, Danica Lombardozzi, Stefanie Falk, Rosie Fisher and Anne Fouilloux were involved in code development. Eva Lieungh, Lasse Keetz, Kaveh Karimi-Asli, Sonya Geange, Anne Fouilloux, Yeliz Yilmaz, Peter Horvath, Emiliano Gelati, Hui Tang, Anders Bryn, Olav Skarpaas, Frode Stordal and Sunniva Indrehus were involved in documentation and testing. Vigdis Vandvik, Sonya Geange and Inge Althuizen were involved in data contributions. Eva Lieungh, Lasse Keetz, Sonya Geange, Hui Tang, Kaveh Karimi-Asli, Yeliz Yilmaz, Frode Stordal, Kjetil Aas, Inge Althuizen, Anders Bryn, Rosie Fisher, Anne Fouilloux, Sunniva Indrehus, Hanna Lee, Danica Lombardozzi, Frans-Jan Parmentier, Norbert Pirk, Vigdis Vandvik, Ane Vollsnes, Olav Skarpaas and Lena Tallaksen were involved in writing. All authors critically revised the manuscript draft and approved the submitted version.

ACKNOWLEDGEMENTS

This work contributes to the EMERALD project (grant no. 294948) funded by the Research Council of Norway, and to LATICE funded by the Faculty of Mathematics and Natural Sciences at the University of Oslo (grant no. UiO/GEO103920). The Vestland Climate Grid sites and data collection were funded by the Research Council of Norway under the projects SEEDCLIM (NORKLIMA 184912) and FunCaB (KLIMAFORSK 244525). The CLM is primarily supported by the U.S. National Science Foundation, and FATES by the U.S. Dept. of Energy, Office of Science, and Office of Biological and Environmental Research (Next Generation Ecosystem Experiments-Tropics project). RF and KSA acknowledge funding by the EU H2020 research and innovation program, Grant Agreements #101003536 (ESM2025) and #821003 (4C). Thanks are due to EOSC-Nordic for support with Galaxy.eu and metadata creation; to the Vestland Climate Grid team for discussions and data insights; numerous students and collaborators for data collection; to the NCAR/NEON team, especially Will Wieder, for helpful discussions and code inspiration; to Rukaya Johaadien and Michal Torma for discussion and code contributions; and to three anonymous referees for valuable feedback.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The LSP code and data are available on GitHub and Zenodo, divided into repositories with the main code (Karimi-Asli, Keetz, Lieungh, Yilmaz, et al., 2022), Graphical User Interface (Karimi-Asli, Lieungh, & Keetz, 2022), the Application Programming Interface (Karimi-Asli, Keetz, & Lieungh, 2022a) and input data for the integrated sites (Karimi-Asli, Keetz, & Lieungh, 2022b). The example simulation data are openly available (Keetz et al., 2022). External resources: Code for the NorESM, CLM, and FATES models are openly available on GitHub in repositories https://github.com/NorESMhub/NorESM, https://github.com/ESCOMP/CTSM, and https://github.com/NGEET/fates respectively.

ORCID

Lasse T. Keetz https://orcid.org/0000-0002-8040-4122
Eva Lieungh https://orcid.org/0000-0003-4009-944X
Kaveh Karimi-Asli https://orcid.org/0000-0001-7782-6909
Sonya R. Geange https://orcid.org/0000-0001-5344-7234
Emiliano Gelati https://orcid.org/0000-0002-8972-8561
Hui Tang https://orcid.org/0000-0002-8745-3859
Yeliz A. Yilmaz https://orcid.org/0000-0001-7937-5056
Kjetil S. Aas https://orcid.org/0000-0003-4223-2267
Inge H. J. Althuizen https://orcid.org/0000-0003-3485-9609
Anders Bryn https://orcid.org/0000-0003-4712-8266
Stefanie Falk https://orcid.org/0000-0002-4552-5455
Rosie Fisher https://orcid.org/0000-0003-3260-9227
Anne Fouilloux https://orcid.org/0000-0002-1784-2920
Peter Horvath https://orcid.org/0000-0002-6017-5385
Sunniva Indrehus https://orcid.org/0000-0002-6647-4239
Hanna Lee https://orcid.org/0000-0002-2003-4377
Danica Lombardozzi https://orcid.org/0000-0003-3557-7929

Frans-Jan W. Parmentier https://orcid.org/0000-0003-2952-7706
Norbert Pirk https://orcid.org/0000-0002-8137-2329
Vigdis Vandvik https://orcid.org/0000-0003-4651-4798
Ane V. Vollsnes https://orcid.org/0000-0003-4903-4116
Olav Skarpaas https://orcid.org/0000-0001-9727-1672
Frode Stordal https://orcid.org/0000-0002-5190-6473
Lena M. Tallaksen https://orcid.org/0000-0002-8480-7842

REFERENCES

Althuizen, I. H. J., Lee, H., Sarneel, J. M., & Vandvik, V. (2018). Long-term climate regime modulates the impact of short-term climate variability on decomposition in alpine grassland soils. *Ecosystems*, *21*(8), 1580–1592. https://doi.org/10.1007/s10021-018-0241-5

Ball, J. T., Woodrow, I. E., & Berry, J. A. (1987). A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In J. Biggins (Ed.), *Progress in photosynthesis research: Volume 4 proceedings of the VIIth International Congress on Photosynthesis Providence, Rhode Island, USA, August 10–15, 1986* (pp. 221–224). Springer Netherlands. https://doi.org/10.1007/978-94-017-0519-6_48

Beigaitė, R., Tang, H., Bryn, A., Skarpaas, O., Stordal, F., Bjerke, J. W., & Žliobaitė, I. (2022). Identifying climate thresholds for dominant natural vegetation types at the global scale using machine learning: Average climate versus extremes. *Global Change Biology*, *28*(11), 3557–3579. https://doi.org/10.1111/gcb.16110

Blyth, E. M., Arora, V. K., Clark, D. B., Dadson, S. J., De Kauwe, M. G., Lawrence, D. M., Melton, J. R., Pongratz, J., Turton, R. H., Yoshimura, K., & Yuan, H. (2021). Advances in land surface modelling. *Current Climate Change Reports*, *7*(2), 45–71. https://doi.org/10.1007/s40641-021-00171-5

Bonan, G. B., & Doney, S. C. (2018). Climate, ecosystems, and planetary futures: The challenge to predict life in earth system models. *Science*, *359*(6375), eaam8328. https://doi.org/10.1126/science.aam8328

Bryn, A., Dourojeanni, P., Hemsing, L. Ø., & O'Donnell, S. (2013). A high-resolution GIS null model of potential forest expansion following land use changes in Norway. *Scandinavian Journal of Forest Research*, *28*(1), 81–98. https://doi.org/10.1080/02827581.2012.689005

Bryn, A., & Stordal, F. (2022). *GEO5915 – Ecological Climatology [university course]*. https://www.uio.no/studier/emner/matnat/geofag/GEO5915/v22/

Bugmann, H., & Seidl, R. (2022). The evolution, complexity and diversity of models of long-term forest dynamics. *Journal of Ecology*, *110*, 2288–2307. https://doi.org/10.1111/1365-2745.13989

Buotte, P. C., Koven, C. D., Xu, C., Shuman, J. K., Goulden, M. L., Levis, S., Katz, J., Ding, J., Ma, W., Robbins, Z., & Kueppers, L. M. (2021). Capturing functional strategies and compositional dynamics in vegetation demographic models. *Biogeosciences*, *18*(14), 4473–4490. https://doi.org/10.5194/bg-18-4473-2021

Chitra-Tarak, R., Xu, C., Aguilar, S., Anderson-Teixeira, K. J., Chambers, J., Detto, M., Faybishenko, B., Fisher, R. A., Knox, R. G., Koven, C. D., Kueppers, L. M., Kunert, N., Kupers, S. J., McDowell, N. G., Newman, B. D., Paton, S. R., Pérez, R., Ruiz, L., Sack, L., … McMahon, S. M. (2021). Hydraulically-vulnerable trees survive on deep-water access during droughts in a tropical forest. *New Phytologist*, *231*(5), 1798–1813. https://doi.org/10.1111/nph.17464

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., Mu, M., & Randerson, J. T. (2018). The International Land Model Benchmarking (ILAMB) system: Design, theory, and implementation. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2731–2754. https://doi.org/10.1029/2018MS001354

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., … Strand, W. G. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. https://doi.org/10.1029/2019MS001916

Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., Loescher, H. W., Lunch, C. K., Pijanowski, B. C., Randerson, J. T., Read, E. K., Tredennick, A. T., Vargas, R., Weathers, K. C., & White, E. P. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(7), 1424–1432. https://doi.org/10.1073/pnas.1710231115

Dietze, M. C., Lebauer, D. S., & Kooper, R. (2013). On improving the communication between models and data. *Plant, Cell & Environment*, *36*(9), 1575–1585. https://doi.org/10.1111/pce.12043

Dietze, M. C., Serbin, S. P., Davidson, C., Desai, A. R., Feng, X., Kelly, R., Kooper, R., LeBauer, D., Mantooth, J., McHenry, K., & Wang, D. (2014). A quantitative assessment of a terrestrial biosphere model's data needs across north American biomes: PEcAn/ED model-data uncertainty analysis. *Journal of Geophysical Research: Biogeosciences*, *119*(3), 286–300. https://doi.org/10.1002/2013JG002392

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., & Hanasaki, N. (2006). GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bulletin of the American Meteorological Society*, *87*(10), 1381–1398. https://doi.org/10.1175/BAMS-87-10-1381

Docker. (2023). *Docker [Computer software]*. Docker, Inc https://www.docker.com/

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fer, I., Gardella, A. K., Shiklomanov, A. N., Campbell, E. E., Cowdery, E. M., De Kauwe, M. G., Desai, A., Duveneck, M. J., Fisher, J. B., Haynes, K. D., Hoffman, F. M., Johnston, M. R., Kooper, R., LeBauer, D. S., Mantooth, J., Parton, W. J., Poulter, B., Quaife, T., Raiho, A., … Dietze, M. C. (2020). Beyond ecosystem modeling: A roadmap to community cyberinfrastructure for ecological data-model integration. *Global Change Biology*, *27*(1), 13–26. https://doi.org/10.1111/gcb.15409

Fer, I., Kelly, R., Moorcroft, P. R., Richardson, A. D., Cowdery, E. M., & Dietze, M. C. (2018). Linking big models to big data: Efficient ecosystem model calibration through Bayesian model emulation. *Biogeosciences*, *15*(19), 5801–5830. https://doi.org/10.5194/bg-15-5801-2018

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2018MS001453. https://doi.org/10.1029/2018MS001453

Fisher, R. A., Koven, C. D., Anderegg, W. R. L., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., Holm, J. A., Hurtt, G. C., Knox, R. G., Lawrence, P. J., Lichstein, J. W., Longo, M., Matheny, A. M., Medvigy, D., Muller-Landau, H. C., Powell, T. L., Serbin, S. P., Sato, H., Shuman, J. K., … Moorcroft, P. R. (2018). Vegetation demographics in earth system models: A review of progress and priorities. *Global Change Biology*, *24*(1), 35–54. https://doi.org/10.1111/gcb.13910

Fisher, R. A., Muszala, S., Verteinstein, M., Lawrence, P., Xu, C., McDowell, N. G., Knox, R. G., Koven, C., Holm, J., Rogers, B. M., Spessa, A., Lawrence, D., & Bonan, G. (2015). Taking off the training wheels: The properties of a dynamic vegetation model without climate envelopes, CLM4.5(ED). *Geoscientific Model Development*, *8*(11), 3593–3619. https://doi.org/10.5194/gmd-8-3593-2015

Fouilloux, A., & Tang, H. (2021). *Functionally assembled terrestrial ecosystem simulator (FATES) (galaxy training materials)*. https://training.galaxyproject.org/training-material/topics/climate/tutorials/fates/tutorial.html

Hartig, F., Dyke, J., Hickler, T., Higgins, S. I., O'Hara, R. B., Scheiter, S., & Huth, A. (2012). Connecting dynamic vegetation models to data - an inverse perspective. *Journal of Biogeography*, *39*(12), 2240–2252. https://doi.org/10.1111/j.1365-2699.2012.02745.x

Hickler, T., Vohland, K., Feehan, J., Miller, P. A., Smith, B., Costa, L., Giesecke, T., Fronzek, S., Carter, T. R., Cramer, W., Kühn, I., & Sykes, M. T. (2012). Projecting the future distribution of European potential natural vegetation zones with a generalized, tree species-based dynamic vegetation model. *Global Ecology and Biogeography*, *21*(1), 50–63. https://doi.org/10.1111/j.1466-8238.2010.00613.x

Hoffman, F. M., Koven, C. D., Keppel-Aleks, G., Lawrence, D. M., Riley, W. J., Randerson, J. T., Ahlström, A., Abramowitz, G., Baldocchi, D. D., Best, M. J., Bond-Lamberty, B., De Kauwe, M. G., Denning, A. S., Desai, A. R., Eyring, V., Fisher, J. B., Fisher, R. A., Gleckler, P. J., Huang, M., … Koch, D. (2017). *2016 International Land Model Benchmarking (ILAMB) Workshop Report* (DOE/SC-0186, 1330803). https://doi.org/10.2172/1330803

Horvath, P., Tang, H., Halvorsen, R., Stordal, F., Tallaksen, L. M., Berntsen, T. K., & Bryn, A. (2021). Improving the representation of high-latitude vegetation distribution in dynamic global vegetation models. *Biogeosciences*, *18*(1), 95–112. https://doi.org/10.5194/bg-18-95-2021

IPBES. (2019). *Global assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services* (p. 1144). In E. S. Brondízio, J. Settele, S. Díaz, & H. T. Ngo (Eds.), IPBES Secretariat. ISBN: 978-3-947851-20-1.

IPCC. (2022). *Climate change 2022: Mitigation of climate change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. https://doi.org/10.1017/9781009157926

Karimi-Asli, K., Keetz, L. T., & Lieungh, E. (2022a). *NorESMhub/ctsm-api: Release v1.0.0 (v1.0.0)*. Zenodo. https://doi.org/10.5281/zenodo.7310241

Karimi-Asli, K., Keetz, L. T., & Lieungh, E. (2022b). *NorESMhub/noresm-lsp-data: v1.0.0 (v1.0.0)*. Zenodo. https://doi.org/10.5281/zenodo.7308232

Karimi-Asli, K., Keetz, L. T., Lieungh, E., Yilmaz, Y., Tang, H., Torma, M., & Gelati, E. (2022). *NorESMhub/noresm-land-sites-platform: Release v1.0.0 (v1.0.0)*. Zenodo. https://doi.org/10.5281/zenodo.7310652

Karimi-Asli, K., Lieungh, E., & Keetz, L. T. (2022). *NorESMhub/noresm-lsp-ui: Release v1.0.0 (v1.0.0)*. Zenodo. https://doi.org/10.5281/zenodo.7310691

Keetz, L. T., Lieungh, E., & Karimi-Asli, K. (2022). *Climate-ecosystem modelling made easy: The land sites platform—simulation results. Version pre-submission*. Zenodo. https://doi.org/10.5281/zenodo.7305032 [dataset].

Klanderud, K., Vandvik, V., & Goldberg, D. (2015). The importance of biotic vs. abiotic drivers of local plant community composition along regional bioclimatic gradients. *PLoS One*, *10*(6), e0130205. https://doi.org/10.1371/journal.pone.0130205

Koven, C. D., Knox, R. G., Fisher, R. A., Chambers, J. Q., Christoffersen, B. O., Davies, S. J., Detto, M., Dietze, M. C., Faybishenko, B., Holm, J., Huang, M., Kovenock, M., Kueppers, L. M., Lemieux, G., Massoud, E., McDowell, N. G., Muller-Landau, H. C., Needham, J. F., Norby, R. J., … Xu, C. (2020). Benchmarking and parameter sensitivity of physiological and vegetation dynamics using the Functionally Assembled Terrestrial Ecosystem Simulator (FATES) at Barro Colorado Island, Panama. *Biogeosciences*, *17*(11), 3017–3044. https://doi.org/10.5194/bg-17-3017-2020

Kyker-Snowman, E., Lombardozzi, D. L., Bonan, G. B., Cheng, S. J., Dukes, J. S., Frey, S. D., Jacobs, E. M., McNellis, R., Rady, J. M., Smith, N. G., Thomas, R. Q., Wieder, W. R., & Grandy, A. S. (2022). Increasing the spatial and temporal impact of ecological research: A roadmap for integrating a novel terrestrial process into an earth system model. *Global Change Biology*, *28*(2), 665–684. https://doi.org/10.1111/gcb.15894

Lambert, M. S. A., Tang, H., Aas, K. S., Stordal, F., Fisher, R. A., Fang, Y., Ding, J., & Parmentier, F.-J. W. (2022). Inclusion of a cold hardening scheme to represent frost tolerance is essential to model realistic plant hydraulics in the Arctic-boreal zone in CLM5.0-FATES-Hydro. *Geoscientific Model Development Discussions*, *15*(23), 8809–8829. https://doi.org/10.5194/gmd-15-8809-2022

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., … Zeng, X. (2019). The Community Land Model Version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. https://doi.org/10.1029/2018MS001583

LeBauer, D. S., Wang, D., Richter, K. T., Davidson, C. C., & Dietze, M. C. (2013). Facilitating feedbacks between field measurements and ecosystem models. *Ecological Monographs*, *83*(2), 133–154. https://doi.org/10.1890/12-0137.1

Lombardozzi, D. L., Wieder, W. R., Sobhani, N., Bonan, G. B., Durden, D., Lenz, D., SanClements, M., Weintraub-Leff, S., Ayres, E., Florian, C. R., Dahlin, K., Kumar, S., Swann, A. L. S., Zarakas, C., Vardeman, C., & Pascucci, V. Overcoming barriers to enable convergence research by integrating ecological and climate sciences: The NCAR-NEON system version 1. *EGUsphere*. https://doi.org/10.5194/egusphere-2023-271 (preprint).

Lynn, J. S., Klanderud, K., Telford, R. J., Goldberg, D. E., & Vandvik, V. (2021). Macroecological context predicts species' responses to climate warming. *Global Change Biology*, *27*(10), 2088–2101. https://doi.org/10.1111/gcb.15532

Melton, J. R., Arora, V. K., Wisernig-Cojoc, E., Seiler, C., Fortier, M., Chan, E., & Teckentrup, L. (2020). CLASSIC v1.0: The open-source community successor to the Canadian Land Surface Scheme (CLASS) and the Canadian Terrestrial Ecosystem Model (CTEM)—Part 1: Model framework and site-level performance. *Geoscientific Model Development*, *13*(6), 2825–2850. https://doi.org/10.5194/gmd-13-2825-2020

NASA GISS. (2022). *Panoply* [computer software]. NASA Goddard Institute for Space Studies. https://www.giss.nasa.gov/tools/panoply/

Nevalainen, O., Niemitalo, O., Fer, I., Juntunen, A., Mattila, T., Koskela, O., Kukkamäki, J., Höckerstedt, L., Mäkelä, L., Jarva, P., Heimsch, L., Vekuri, H., Kulmala, L., Stam, Å., Kuusela, O., Gerin, S., Viskari, T., Vira, J., Hyväluoma, J., … Liski, J. (2022). Towards agricultural soil carbon monitoring, reporting, and verification through the Field Observatory Network (FiON). *Geoscientific Instrumentation, Methods and Data Systems*, *11*(1), 93–109. https://doi.org/10.5194/gi-11-93-2022

Norby, R. J., De Kauwe, M. G., Domingues, T. F., Duursma, R. A., Ellsworth, D. S., Goll, D. S., Lapola, D. M., Luus, K. A., MacKenzie, A. R., Medlyn, B. E., Pavlick, R., Rammig, A., Smith, B., Thomas, R., Thonicke, K., Walker, A. P., Yang, X., & Zaehle, S. (2016). Model–data synthesis for the next generation of forest free-air CO_2 enrichment (FACE) experiments. *New Phytologist*, *209*(1), 17–28. https://doi.org/10.1111/nph.13593

Nüst, D., Sochat, V., Marwick, B., Eglen, S. J., Head, T., Hirst, T., & Evans, B. D. (2020). Ten simple rules for writing Dockerfiles for reproducible data science. *PLoS Computational Biology*, *16*(11), e1008316. https://doi.org/10.1371/journal.pcbi.1008316

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., … Papale, D. (2020). The

FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, *7*(1), 225. https://doi.org/10.1038/s41597-020-0534-3

Pettorelli, N., Graham, N. A. J., Seddon, N., Bustamante, M. M. C., Lowton, M. J., Sutherland, W. J., Koldewey, H. J., Prentice, H. C., & Barlow, J. (2021). Time to integrate global climate change and biodiversity science-policy agendas. *Journal of Applied Ecology*, *58*(11), 2384–2393. https://doi.org/10.1111/1365-2664.13985

Pirk, N., Sievers, J., Mertes, J., Parmentier, F.-J. W., Mastepanov, M., & Christensen, T. R. (2017). Spatial variability of CO_2 uptake in polygonal tundra: Assessing low-frequency disturbances in eddy covariance flux estimates. *Biogeosciences*, *14*(12), 3157–3169. https://doi.org/10.5194/bg-14-3157-2017

Project Jupyter. (2014). *Jupyter* [computer software]. https://jupyter.org/

Purgar, M., Klanjscek, T., & Culina, A. (2022). Quantifying research waste in ecology. *Nature Ecology & Evolution.*, *6*, 1390–1397. https://doi.org/10.1038/s41559-022-01820-0

Rogers, A., Serbin, S. P., & Way, D. A. (2022). Reducing model uncertainty of climate change impacts on high latitude carbon assimilation. *Global Change Biology*, *28*(4), 1222–1247. https://doi.org/10.1111/gcb.15958

Scheiter, S., Langan, L., & Higgins, S. I. (2013). Next-generation dynamic global vegetation models: Learning from community ecology. *New Phytologist*, *198*(3), 957–969. https://doi.org/10.1111/nph.12210

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., … Schulz, M. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, *13*(12), 6165–6200. https://doi.org/10.5194/gmd-13-6165-2020

Smith, B., Prentice, I. C., & Sykes, M. T. (2001). Representation of vegetation dynamics in the modelling of terrestrial ecosystems: Comparing two contrasting approaches within European climate space. *Global Ecology and Biogeography*, *10*(6), 621–637. https://doi.org/10.1046/j.1466-822X.2001.t01-1-00256.x

Somodi, I., Ewald, J., Bede-Fazekas, Á., & Molnár, Z. (2021). The relevance of the concept of potential natural vegetation in the Anthropocene. *Plant Ecology & Diversity*, *14*(1–2), 13–22. https://doi.org/10.1080/17550874.2021.1984600

Sulman, B. N., Salmon, V. G., Iversen, C. M., Breen, A. L., Yuan, F., & Thornton, P. E. (2021). Integrating Arctic plant functional types in a land surface model using above- and belowground field observations. *Journal of Advances in Modeling Earth Systems*, *13*(4), e2020MS002396. https://doi.org/10.1029/2020MS002396

The NorESM-LSP Development Team. (2023). *NorESM Land Sites Platform technical documentation*. https://noresmhub.github.io/noresm-land-sites-platform/ https://doi.org/10.5281/zenodo.7310652

Think big and model small. (2022). Nature Climate Change, *12*(6), 493–493. https://doi.org/10.1038/s41558-022-01399-1

Töpper, J. P., Meineri, E., Olsen, S. L., Rydgren, K., Skarpaas, O., & Vandvik, V. (2018). The devil is in the detail: Nonadditive and context-dependent plant population responses to increasing temperature and precipitation. *Global Change Biology*, *24*(10), 4657–4666. https://doi.org/10.1111/gcb.14336

Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J.-B., Peer, G., Singer, A., Bridle, J. R., Crozier, L. G., De Meester, L., Godsoe, W., Gonzalez, A., Hellmann, J. J., Holt, R. D., Huth, A., Johst, K., Krug, C. B., Leadley, P. W., Palmer, S. C. F., Pantel, J. H., … Travis, J. M. J. (2016). Improving the forecast for biodiversity under climate change. *Science*, *353*(6304), aad8466. https://doi.org/10.1126/science.aad8466

Vandvik, V., Althuizen, I. H. J., Jaroszynska, F., Krüger, L. C., Lee, H., Goldberg, D. E., Klanderud, K., Olsen, S. L., Telford, R. J., Östman, S. A. H., Busca, S., Dahle, I. J., Egelkraut, D. D., Geange, S. R., Gya, R., Lynn, J. S., Meineri, E., Young, S., & Halbritter, A. H. (2022). The role of plant functional groups mediating climate impacts on carbon and biodiversity of alpine grasslands. *Scientific Data*, *9*(1), 451. https://doi.org/10.1038/s41597-022-01559-0

Wang, D., Xu, Y., Thornton, P., King, A., Steed, C., Gu, L., & Schuchart, J. (2014). A functional test platform for the community land model. *Environmental Modelling & Software*, *55*, 25–31. https://doi.org/10.1016/j.envsoft.2014.01.015

Wieters, N., & Fritzsch, B. (2018). Opportunities and limitations of software project management in geoscience and climate modelling. *Advances in Geosciences*, *45*, 383–387. https://doi.org/10.5194/adgeo-45-383-2018

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

Wullschleger, S. D., Epstein, H. E., Box, E. O., Euskirchen, E. S., Goswami, S., Iversen, C. M., Kattge, J., Norby, R. J., van Bodegom, P. M., & Xu, X. (2014). Plant functional types in earth system models: Past experiences and future directions for application of dynamic vegetation models in high-latitude ecosystems. *Annals of Botany*, *114*(1), 1–16. https://doi.org/10.1093/aob/mcu077

Wullschleger, S. D., Hinzman, L. D., & Wilson, C. J. (2011). Planning the next generation of Arctic ecosystem experiments. *Eos, Transactions American Geophysical Union*, *92*(17), 145–145. https://doi.org/10.1029/2011EO170006

Xu, Y., Wang, D., Janjusic, T., Wu, W., Pei, Y., & Yao, Z. (2017). A web-based visual analytic framework for understanding Large-scale environmental models: A use case for the community land model. *Procedia Computer Science*, *108*, 1731. https://doi.org/10.1016/j.procs.2017.05.181

Yang, Z., Hanna, E., & Callaghan, T. V. (2011). Modelling surface-air-temperature variation over complex terrain around Abisko, Swedish Lapland: Uncertainties of measurements and models at different scales. *Geografiska Annaler: Series A, Physical Geography*, *93*(2), 89–112. https://doi.org/10.1111/j.1468-0459.2011.00005.x

How to cite this article: Keetz, L. T., Lieungh, E., Karimi-Asli, K., Geange, S. R., Gelati, E., Tang, H., Yilmaz, Y. A., Aas, K. S., Althuizen, I. H. J., Bryn, A., Falk, S., Fisher, R., Fouilloux, A., Horvath, P., Indrehus, S., Lee, H., Lombardozzi, D., Parmentier, F.-J., Pirk, N. … Tallaksen, L. M. (2023). Climate–ecosystem modelling made easy: The Land Sites Platform. *Global Change Biology*, *29*, 4440–4452. https://doi.org/10.1111/gcb.16808