Applying a deep learning-based approach for scaling vegetation dynamics to predict changing forest regimes under future climate and fire scenarios

Werner Rammer[1], Rupert Seidl

Abstract: The ability to anticipate future changes in terrestrial ecosystems is key for their management. New tools are required that bridge the gap between a high level of process understanding at fine spatial grain, and the increasing relevance for management at larger extents. Such a tool is SVD (Scaling Vegetation Dynamics), a scaling framework that specifically uses deep learning to learn the behavior of detailed vegetation models in response to different environmental factors. This trained deep neural network (DNN) is then applied within the framework on large spatial scales. In addition, SVD includes also explicitly modelled processes such as fire disturbances. Here we use the framework to simulate forest regime change in the 3 Mio. ha landscape of the Greater Yellowstone Ecosystem. We used four climate change scenarios and pre-defined fire events from statistical modelling, and analyzed whether prevailing forest types are able to regenerate after fire. Our results show that up to 60% of the area may undergo regime change until the end of the 21st century.

Keywords: SVD; vegetation dynamics; deep learning; Greater Yellowstone Ecosystem; fire; climate change

1 Introduction

Terrestrial vegetation is of crucial importance for human well-being and provide a wide variety of ecosystem services to society [As05]. However, vegetation is not static but changes dynamically, responding to drivers such as land-use change and climate change [Er18, Li10]. Thus, the ability to faithfully predict future trajectories of vegetation development is highly relevant for decision makers and society. Dynamic global vegetation models (DGVMs) are frequently used to simulate vegetation dynamics at large spatial scales. Such models increasingly include structural details (e.g., representing leaves or individual trees as entities of the simulation), but they typically assume that these structures represent the conditions of an entire grid cell (with a cell size usually between 10 and 250km). Therefore, biotic interactions such as seed distribution, mortality, or plant competition are neglected. Stand and landscape-level vegetation models – on the other hand - simulate vegetation demography on a detailed level and include biotic interactions (e.g., mortality and demography), as well as spatiotemporal controls (e.g., migration and legacies). However, they usually are limited to small spatial extents, which limits their application on larger scales. Since biotic interactions

[1] Ecosystem Dynamics and Forest Management Group, Technical University of Munich, 85354 Freising, Germany, werner.rammer@tum.de

doi:10.18420/inf2020_96

and the resulting demographic structures are crucial for many ecological questions such as carbon storage [Kö17], approaches that are able to capture the drivers of vegetation development at small scales and dynamically scale ecosystem dynamics across spatial domains are needed.

New methods in the field of artificial intelligence and machine learning excel at identifying structure in complex, nonlinear data, and generate accurate predictive models [GBC16]. Specifically, deep learning is an emerging machine learning technique at the core of recent breakthroughs in computer vision, speech synthesis, autonomous driving, and other fields [LBH15] and has been advocated for providing new opportunities in Earth Sciences [Re19] and Ecology [RS19a].

One approach for the scaling-up of detailed simulation models is the Scaling Vegetation Dynamics (SVD) framework [RS19b]. The modeling framework uses a deep neural network to learn the response of detailed vegetation simulation models to different environmental factors. This trained “meta model” can then be applied in a computationally efficient manner for projecting vegetation dynamics at regional to continental scale.

An important driver of future changes in forest ecosystems are disturbances, particularly as they are expected to increase in the future with a warming climate [Se14]. For example, the Greater Yellowstone Ecosystem (GYE) is an epitome of the complexities of climateand fire-driven vegetation changes. Situated in the Northern Rocky Mountains of the USA, it has received considerable attention due to 709,000 ha of wildfires affecting the system in 1988. Alarmingly, fire projections for the region suggest that the extreme event of 1988 could become the new normal at the end of the 21st century [We11]. It thus remains unclear whether changing climate and fire regimes will exceed the resilience of the GYE.

In this contribution we describe the deep learning-based modelling approach and present an example application. In this example, we assessed the probability of regime shifts in the forests of the Greater Yellowstone Ecosystem, given the expected future climate and fire regimes. Regime shifts were defined as changes in major forest types, with the inability of a current forest type to regenerate under future conditions used as an indicator for impending changes in forest type. Specifically, we asked how much of the currently prevailing vegetation experiences permanent regeneration failure.

2 Material and Methods

2.1 Scaling Vegetation Dynamics

Conceptually, the scaling vegetation dynamics (SVD) framework [RS19b] follows a state and transition approach, where vegetation is classified into discrete vegetation states and transitions between states are probabilistic. In SVD probabilities are conditional on environmental conditions and the local neighborhood of a cell and estimated by a deep neural network [LBH15]. In this example, the DNN at the core of SVD was trained on data generated by the process-based model iLand [Se12], which was used to simulate regeneration success after fire disturbance over a wide range of environmental conditions. Note that for other applications, the DNN can be trained to encapsulate the full dynamics of the simulated vegetation [RS19b]. The application of SVD consists of three distinct phases: the first phase comprises the generation of training data, and the second the training of the DNN, yielding a condensed meta-model of the post-fire response of the process-based model. The third phase applies the trained DNN for the dynamic simulation of vegetation transitions for the whole GYE within the SVD model.

The spatial scale of SVD is 1 ha, and the time step is annual. The SVD model is extended by modules to encompass other drivers of vegetation transitions such as natural disturbances or ecosystem management. These modules provide additional pathways of vegetation transitions and can interact with each other. Here we use the fire disturbance module that simulates fire spread dynamically on the landscape and uses data on fire size, ignition point and time as input.

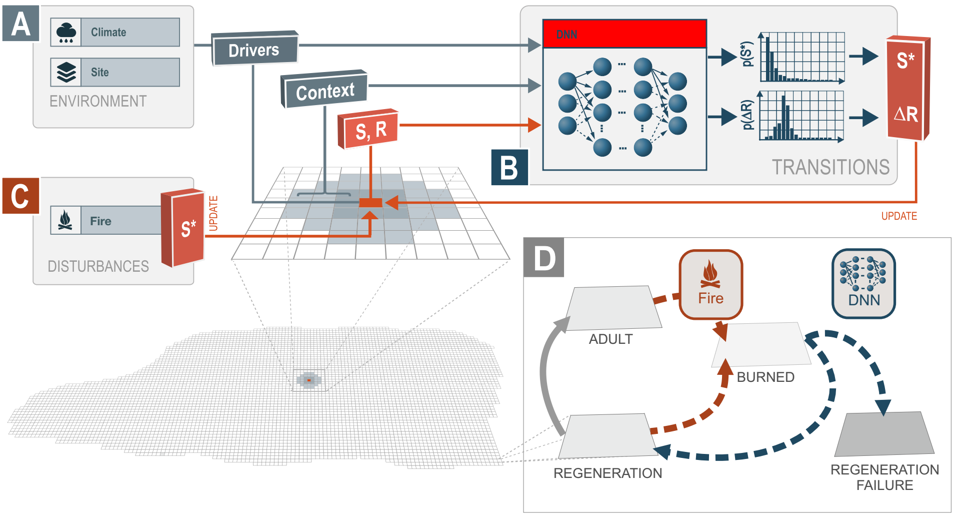

Fig. 1: Conceptual view of the scaling vegetation dynamics (SVD) framework (A-C), and the state and transition pathways (D) used in this study. In SVD, transitions on a single cell are predicted by a Deep Neural Network (B) and depend on environmental drivers (A), the current state (S) and residence time (R), as well as the spatial context (here: distance to seed source). Fire (C) as simulated by the SVD fire module adds an additional pathway of state change. States and transitions (D): Adult and regenerated cells transition to early seral states due to fire, and the regeneration success or failure is consequently determined by the DNN. The transition from regeneration to adult is deterministic.

Fig. 1 shows the conceptual pathways of vegetation transitions used in this study. Cells with vegetation of different forest type can be affected by fire. The regeneration success of recently burnt cells ("Early seral") is determined by the DNN, considering environmental conditions and distance to seed source within the dynamic simulation. Cells transition to a "Regeneration" state in case of success, or to the state "Regeneration failure" when regeneration fails to establish for more than 30 years. "Regeneration" cells become seed-producing "Adult" deterministically at a forest type specific maturation age.

SVD is a standalone software written in C++ that uses the C++ API of the deep learning library TensorFlow [Ab16] to perform DNN predictions (Fig. 2). By integrating the full TenorFlow framework into the model, the full functionality of TensorFlow can be used and the structure of the DNN can be tailored to the specific needs of each application. The DNN training itself facilitates the standard TensorFlow workflow and tools (e.g., Python, Keras, TensorBoard). When the training has finished, the full DNN (structure and weights) is saved and consequently used by the SVD model(Fig. 2) .

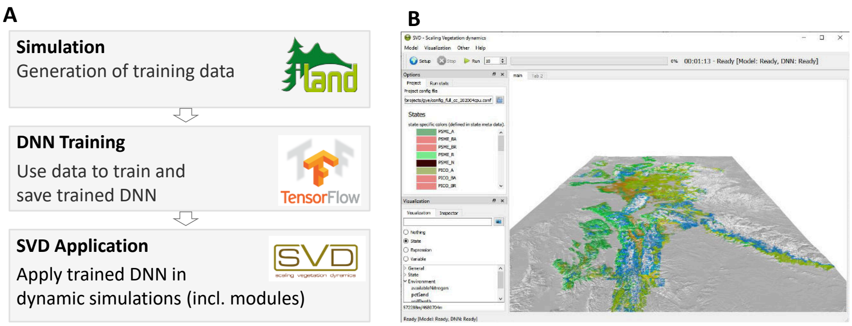

Fig. 2: Conceptual steps of a SVD application (A) and a screenshot from the SVD model (B). The training data for the DNN is generated with a detailed process -based model and then the DNN is trained with Te nsorFlow. The final model is saved and consequently used in the SVD model in the context of a dynamic simulation (including the fire module). DNN predictions are executed by TensorFlow which are triggered via the C++ API of the library.

2.2 Generation of training data

We followed the approach of [Ha18], who used the individual based forest landscape and disturbance model iLand [Se12] to analyze the conditions under which the most important forest types in Yellowstone National Park might fail to regenerate after stand replacing fire disturbance. iLand simulates individual trees within a stand and uses a hierarchical framework wherein broader-scale processes emerge dynamically from interactions among individual trees. The model represents tree growth, mortality, and competition in response to canopy light interception, radiation, thermal conditions, soil water, and nutrient limitation. While climate and soil conditions are considered at spatially homogeneous within a stand (1 ha), variation in light is simulated at 2 m horizontal resolution based on overstory structure and composition. Climate data (temperature, precipitation, radiation, vapor pressure deficit) is considered at a daily temporal grain. The model has been well tested and extensively used in the western United States [SRS14] and Europe [TRS17], and has recently been parameterized and evaluated for the Yellowstone area [Br18]. The model explicitly simulates tree regeneration based on seed production, seed dispersal, and effects of temperature, light, and soil-moisture conditions on seedling establishment and survival.

For the generation of the training data we set up a factorial simulation experiment focusing on factors that are likely to change under future climate and fire regimes [We11], and strongly affect regeneration success by altering either the availability of seeds or the establishment success of tree plants. We considered in the experiment fire return interval (FRI), distance to seed source, and variation in climate, and applied all combinations for four forest types on 1,296 representative sites across the region. We used four levels of FRI (11, 20, 50, 100 years) that cover the range from shortened future fire intervals [We11] to the upper end of historically observed FRI. The availability of seeds was represented by setting the distance to the nearest seed source between 50 m and 1250 m (50 m stepwidth). Finally, unfavorable climate conditions can hamper seedling establishment, predominantly due to drought in post-fire years [HT19]. We therefore repeated the simulations assuming either warmer, or warmer and drier conditions consistent with our climate change scenarios for the 21st century.

2.3 Training the deep neural network

The data derived from the iLand simulation experiment was then used to train the deep neural network that is used in dynamic SVD simulations. The network learned to predict the success or failure of regeneration contingent on FRI, distance to seed source, site and climate conditions (Tab. 1). We experimented with different DNN architectures and hyper-parameter settings and monitored the predictive performance of the network using a fixed subset of the training data set aside for evaluation (11% of the examples). The best-performing network was consequently used in the SVD model for dynamic simulations across the whole region.

2.4 Simulation setup

We simulated the vegetation development in GYE under four different climate change scenarios from 2005 to 2100. We selected scenarios from the RCP (Representative Concentration Pathways) 4.5 and RCP 8.5 emission scenario families; RCP 4.5 is considered as an intermediate scenario where CO2 emissions peak in the 2040s and decline afterwards, while

Tab. 1: Description of the data used for the training of the DNN.

Variable	**Description**
Predictors (input data)	
CLIM	Climatic conditions (mean monthly temperature (°C) and precipitation (mm) in the next ten years from the current year, i.e. 2 x 12 x 10 = 240 values)
CSS	Site conditions (Nutrient supply (with plant available nitrogen as proxy, in kg N ha^{-1} yr^{-1}), soil depth (m), % sand (proxy for soil texture))
DSS	Distance to seed source (m)
R	Residence time (yrs), the number of years the stand is already in state S
S	Current state (-)
Response variables (labels)	
S^*	Predicted next state (may be equal to S)
ΔR	Time until state change (up to 10 years)

RCP 8.5 represents rising emissions throughout the 21st century [Me11]. Global climate models still do not agree on future precipitation trends in the northern Rocky Mountains and we integrated this uncertainty by selecting two diverging scenarios for each RCP. Predictions of fire activity (fire locations and fire sizes) were derived from statistical fire modelling [We11] and were available as 20 replicated time series of fire for each scenario. Data on the current vegetation as well as site and climate conditions were compiled from available data sources for the entire area.

3 Results

In order to generate the training data for the machine learning algorithm in SVD, we ran the factorial simulation experiment in iLand, yielding a total of 2.59 Mio individual stand trajectories each including a stand replacing fire and post fire vegetation development from which a total number of 7.7 Mio training examples were extracted (see Tab. 1).

The final network architecture (i.e., the type, shape, and number of layers) used for this study was a feedforward network with 674,420 trainable parameters (Fig. 3). A series of fully connected layers compressed the information content of the climate input variables (CLIM, N=240) to a vector with ten elements. An embedding layer [5] was used to transfer the numerical state identifier to an embedding vector with six dimensions. Climate and state information were then combined with the remaining input data (site variables (CSS), distance to seed source (DSS), residence time (R)). From the concatenation layer two branches of fully connected and dropout layers led to the final Softmax layers with the classification result for the next state S^* (23 classes), and residence time ΔR (10 classes). See [An16, GBC16, Me11, RS19a] for more details on network architecture and the different types of network layers.

DNN performance metrics obtained for the evaluation data set. Response variable was the regeneration success or failure over 30 years.

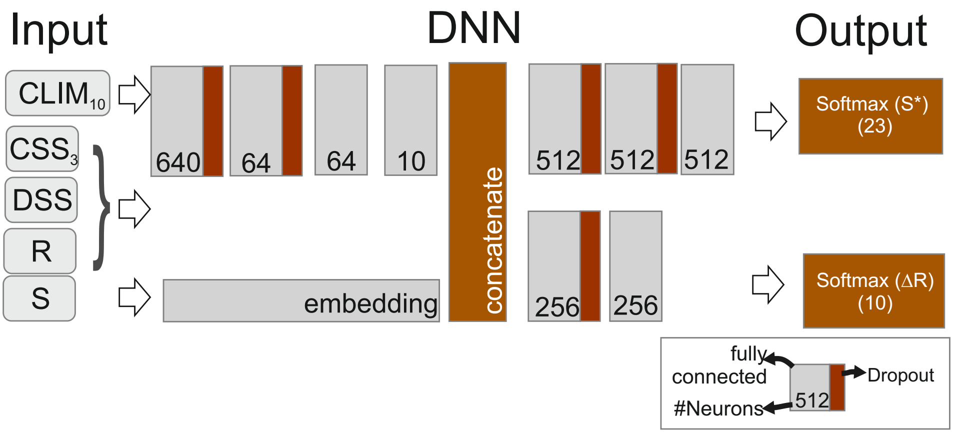

Fig. 3: Structure of the DNN.

Metric	Equation	Value
Accuracy	$\frac{tn+tp}{N}$	0.941
Precision	$\frac{tp}{tp+fp}$	0.948
Recall	$\frac{tp}{tp+fn}$	0.960
F1 Score	$\frac{2*precision*recall}{precision+recall}$	0.954
Conditional Kappa	$\frac{precision-\frac{tp+fn}{N}}{1-\frac{tp+fn}{N}}$	0.859
True skill statistic	$\frac{precision+tn}{tn+fn-1}$	0.878

Tab. 2: The DNN achieved an accuracy for predicting regeneration success or failure of 0.941 (F1 score 0.954) on an evaluation data set not used for network training (see Tab.2 for additional performance metrics).

The simulations with the SVD model for the GYE predicted a substantial proportion of area that failed to regenerate until the end of the century (Fig. 3). The proportion of the affected area was between 28 % under the moderate RCP 4.5 (wet) and 58 % under the RCP 8.5 scenario, indicating an increased probability of failure under hot and dry conditions. Forest types were also affected differently by fire. Historically well fireadapted forest types such as Lodgepole pine burned more frequently, but were also better able to regenerate compared to forest types less adapted to fire.

4 Discussion

Scaling has long been a central issue in ecology and remains a challenge for ecological modeling. The SVD framework combines the discretization of vegetation states with the predictive power of a deep neural network for estimating transition probabilities and

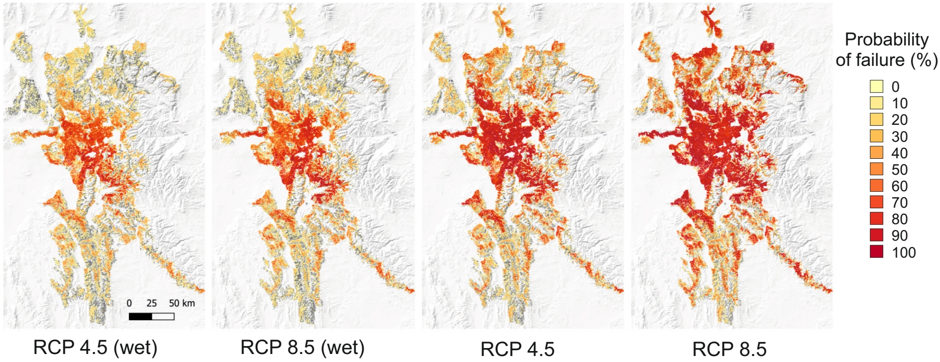

Fig. 4: Spatial distribution of the probability of regeneration failure under the four climate change scenarios for the year 2100. The probability is calculated as the average over the replicates per scenario.

pathways. We found DNNs to work well as the engine of such a meta-modeling approach. Specifically, the DNN accurately reproduced the complex responses of an underlying process-based model, yielding high prediction accuracies. The final DNN was well able to predict situations that were not included in the underlying training data, and thus showed high potential for generalization, which is an important ability in the context of upscaling [5, 9]. The application proved also the computational efficiency of SVD as simulation times were three to four orders of magnitude faster compared to the detailed process-based model. While simulations with the individual based model iLand are currently impractical for areas much larger than ~50,000 ha (with simulation times of hours), the whole GYE (2.9 Mio ha forested, 100 yrs simulation time) can be simulated with SVD in less than an hour on a standard PC (with a GPU to speed up DNN calculations).

The DNN that we used in our SVD simulations was the end point of many preparatory steps. Setting up a process-based model for simulations over a wide ecological gradient requires a considerable amount of model testing and evaluation. Here we were able to heavily build upon previous work [3, 6]. Furthermore, setting up and training a DNN require the modeler to make many design choices that can strongly affect the performance of the network. While the process of fine-tuning the structure of a network is potentially very time consuming, the availability of high-level abstractions with good default values (e.g., Keras) as well as powerful tools (e.g., TensorBoard) increasingly simplify this task.

In the current example we relied on synthetic data provided by the iLand model instead of empirical data as the “ground true” training data. A downside of this approach is that any biases present in the model are also transferred to the meta model. However, in many ecological domains the availability of empirical data is too sparse to allow an efficient training of notoriously data hungry DNNs . Moreover, process-based models are able to consistently consider also system responses under future “no analogous” conditions for which empirical data may not exist at all. A potential way to mitigate some of the problems of synthetic data would be to use an ensemble of data generating models, thus reducing the uncertainty related to the formulation of a single underlying PBM.

The prevalence of early-seral forests will increase drastically throughout the Greater Yellowstone Ecosystem. At the end of the 21st century the share of the current forest area stocked with less than 50 trees per hectare reached values of 60% for RCP 8.5. This suggests that the GYE will transition from its signature densely forested landscapes to predominately open conditions. The increasing share of early-seral systems will not only alter the visual impression of the landscape for visitors but will also change its habitat quality. An increasing share of sparsely stocked stands will also reduce the C storage potential of ecosystem, with negative consequences for the climate system.

We here focused on the regeneration phase after a disturbance and assumed static forest types. Since we neglected the adaptation capacity of forests, e.g. due to species migration, our results should not be interpreted as forest loss but rather as areas with forest loss or altered forest types.

Bibliography

[Ab16] Abadi, Martín; Barham, Paul; Chen, Jianmin; Chen, Zhifeng; Davis, Andy; Dean, Jeffrey; Devin, Matthieu; Ghemawat, Sanjay; Irving, Geoffrey; Isard, Michael et al.: Tensorflow: A system for large-scale machine learning. In: 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16). pp. 265–283, 2016.

[An16] Angermueller, Christof; Pärnamaa, Tanel; Parts, Leopold; Stegle, Oliver: Deep learning for computational biology. Molecular systems biology, 12(7):878, 2016.

[As05] Assessment, Millennium Ecosystem: Ecosystems and human well-being: wetlands and water. World Resources Institute, 2005.

[Br18] Braziunas, Kristin H; Hansen, Winslow D; Seidl, Rupert; Rammer, Werner; Turner, Monica G: Looking beyond the mean: Drivers of variability in postfire stand development of conifers in Greater Yellowstone. Forest Ecology and Management, 430:460–471, 2018.

[Er18] Erb, Karl-Heinz; Kastner, Thomas; Plutzar, Christoph; Bais, Anna Liza S; Carvalhais, Nuno; Fetzel, Tamara; Gingrich, Simone; Haberl, Helmut; Lauk, Christian; Niedertscheider, Maria et al.: Unexpectedly large impact of forest management and grazing on global vegetation biomass. Nature, 553(7686):73–76, 2018.

[GBC16] Goodfellow, Ian; Bengio, Yoshua; Courville, Aaron: Deep learning. MIT press, 2016.

[Ha18] Hansen, Winslow D; Braziunas, Kristin H; Rammer, Werner; Seidl, Rupert; Turner, Monica G: It takes a few to tango: changing climate and fire regimes can cause regeneration failure of two subalpine conifers. Ecology, 99(4):966–977, 2018.

[HT19] Hansen, Winslow D; Turner, Monica G: Origins of abrupt change? Postfire subalpine conifer regeneration declines nonlinearly with warming and drying. Ecological Monographs, 89(1), 2019.

[Kö17] Körner, Christian: A matter of tree longevity. Science, 355(6321):130–131, 2017.

[LBH15] LeCun, Yann; Bengio, Yoshua; Hinton, Geoffrey: Deep learning. Nature, 521(7553):436–444, May 2015.

[Li10] Lindner, Marcus; Maroschek, Michael; Netherer, Sigrid; Kremer, Antoine; Barbati, Anna; Garcia-Gonzalo, Jordi; Seidl, Rupert; Delzon, Sylvain; Corona, Piermaria; Kolström, Marja; et al.: Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. Forest Ecology and Management, 259(4):698–709, Feb 2010.

[Me11] Meinshausen, Malte; Smith, S. J.; Calvin, K.; Daniel, J. S.; Kainuma, M. L. T.; Lamarque, J-F.; Matsumoto, K.; Montzka, S. A.; Raper, S. C. B.; Riahi, K.; et al.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. Climatic Change, 109(1–2):213–241, Aug 2011.

[Re19] Reichstein, Markus; Camps-Valls, Gustau; Stevens, Bjorn; Jung, Martin; Denzler, Joachim; Carvalhais, Nuno et al.: Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743):195–204, 2019.

[RS19a] Rammer, Werner; Seidl, Rupert: Harnessing deep learning in ecology: An example predicting bark beetle outbreaks. Frontiers in plant science, 10:1327, 2019.

[RS19b] Rammer, Werner; Seidl, Rupert: A scalable model of vegetation transitions using deep neural networks. Methods in ecology and evolution, 10(6):879–890, 2019.

[Se12] Seidl, Rupert; Rammer, Werner; Scheller, Robert M.; Spies, Thomas A.: An individual-based process model to simulate landscape-scale forest ecosystem dynamics. Ecological Modelling, 231:87–100, Apr 2012.

[Se14] Seidl, Rupert; Schelhaas, Mart-Jan; Rammer, Werner; Verkerk, Pieter Johannes: Increasing forest disturbances in Europe and their impact on carbon storage. Nature Climate Change, 4(9):806–810, Aug 2014.

[SRS14] Seidl, Rupert; Rammer, Werner; Spies, Thomas A: Disturbance legacies increase the resilience of forest ecosystem structure, composition, and functioning. Ecological Applications, 24(8):2063–2077, 2014.

[TRS17] Thom, Dominik; Rammer, Werner; Seidl, Rupert: The impact of future forest dynamics on climate: interactive effects of changing vegetation and disturbance regimes. Ecological monographs, 87(4):665–684, 2017.

[We11] Westerling, Anthony L; Turner, Monica G; Smithwick, Erica AH; Romme, William H; Ryan, Michael G: Continued warming could transform Greater Yellowstone fire regimes by mid-21st century. Proceedings of the National Academy of Sciences, 108(32):13165–13170, 2011.