Forest Ecology and Management 347 (2015) 107–115

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

An imputed forest composition map for New England screened by species range boundaries

Matthew J. Duveneck [a,*], Jonathan R. Thompson [a], B. Tyler Wilson [b]

[a] *Harvard Forest, Harvard University, Petersham, MA, USA*

[b] *Forest Inventory and Analysis, Northern Research Station, USDA Forest Service, Saint Paul, MN, USA*

ARTICLE INFO

Article history:
Received 19 December 2014
Received in revised form 3 March 2015
Accepted 7 March 2015
Available online 30 March 2015

Keywords:
Nearest-neighbor imputation
New England
Forest inventory
Tree species distribution

ABSTRACT

Initializing forest landscape models (FLMs) to simulate changes in tree species composition requires accurate fine-scale forest attribute information mapped continuously over large areas. Nearest-neighbor imputation maps, maps developed from multivariate imputation of field plots, have high potential for use as the initial condition within FLMs, but the tendency for field plots to be imputed over large geographical distances can result in species being mapped outside of their home ranges, which is problematic. We developed an approach for evaluating and imputing field plots based on their similarity across multiple spatial environmental variates, their species composition, and their geographical distance between source and imputation to produce a map that is appropriate for initializing an FLM. We used this approach to map 13 million ha of forest throughout the six New England states (Rhode Island, Connecticut, Massachusetts, New Hampshire, Vermont, and Maine). Using both independent state forest and, more extensive, ecoregion validation data sets, we compared the imputation map to field inventory data, based on the dissimilarity of tree community composition and the rank order correlation of tree species abundance. Average Bray–Curtis dissimilarity between the imputation map and field plots was 0.32 and 0.12, for the state forest and ecoregion validation data sets, respectfully. Average Spearman rank order correlation was 0.81 and 0.93 for the state forest and ecoregion validation data sets, respectfully. Our analyses suggest that this approach to imputation can realistically capture regional variation in forest composition. We expect the imputation map will be valuable for several regional forest studies and that the approach could be successfully used in other regions.

1. Introduction

Forest Landscape Models (FLMs) simulate succession and disturbance over meso-scales (generally 100–10,000 km^2) and incorporate spatially interactive processes represented using interacting raster map cells (pixels) (Scheller and Mladenoff, 2007). FLMs typically operate at 30–250 m cell resolution. Increasing use of FLMs to simulate forest change over large spatial and temporal scales is driving demand for fine grain forest attribute data that is mapped continuously over large areas for use as initial (or starting) conditions (He, 2008; Keane et al., 2004; Scheller and Mladenoff, 2007). Because these models simulate processes at the scale of individual trees or cohorts of trees, they require a level of detail that is typically only obtainable via field inventories. For example, the LANDIS-II FLM requires a spatial representation of tree species-age cohorts (Scheller et al., 2007).

Given a detailed representation of initial forest conditions, FLMs are frequently used to simulate the effects of natural and human processes on forests. They offer realistic spatial depictions of future forest conditions at a similarly high level of detail. The FLM approach is valuable for understanding how species distributions and ecosystem conditions may change over large areas and of long time frames in response to climate change, land-use and other environmental stressors (Bettinger et al., 2005; Duveneck et al., 2014a; Gustafson, 2013; Liang et al., 2014; Thompson et al., 2011).

The challenge of model initialization is pronounced for FLMs relative to other spatially explicit forest models. Unlike models used to examine aspects of ecological theory (e.g., gap models (Keane et al., 1996; Shugart et al., 2010)), FLMs are used to examine landscape changes that are specific to a certain place and time making a "spin-up to equilibrium" approach potentially unreliable. Also, because the focus of a FLM is typically on individual species—whether to infer differences in species distributions or to look at species-specific disturbances (e.g., pests, pathogens, timber harvest)—users cannot rely on simplified representations of plant

* Corresponding author. Tel.: +1 978 756 6167.
E-mail address: mduveneck@gmail.com (M.J. Duveneck).

http://dx.doi.org/10.1016/j.foreco.2015.03.016
0378-1127/

functional types, which is common practice for initializing dynamic vegetation models (Bonan et al., 2011; Haxeltine and Prentice, 1996; Moorcroft et al., 2001). Finally, due to the longevity of trees and the tendency for forests to have strong compositional inertia, the representation or choice of initial conditions can have significant consequences for FLM simulation outcomes.

While remote sensing platforms are able to survey large areas; they are generally unable to discern individual species or stand structure information (e.g., species age). Forest inventory data, such as data from the U.S. Department of Agriculture Forest Service Forest Inventory and Analysis (FIA) program (Bechtold and Patterson, 2005), are widely available to populate initial conditions within model simulations. However, inventory plots are sparsely distributed relative to the typical grain size used within a FLM. As a result, researchers often impute attributes from field plots (source plots) to each of the raster pixels within a landscape (Hudak et al., 2008; Ohmann and Gregory, 2002). Several methods for spatial imputation exist.

Imputation methods have been used to develop initial conditions for FLMs. "Landscape Builder" (Dijak, 2013) is software designed to stochastically impute source plots within landcover and landform spatial layers. This has the strength of using land

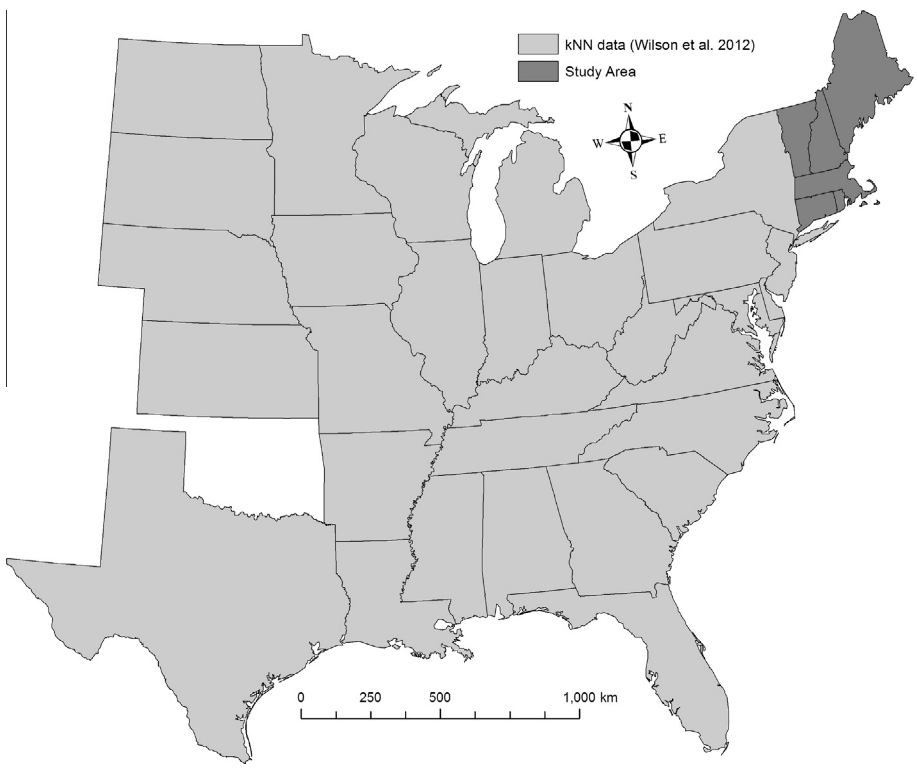

Fig. 1. Study landscape in New England (dark gray) within area where kNN data are available in the eastern United States (light gray).

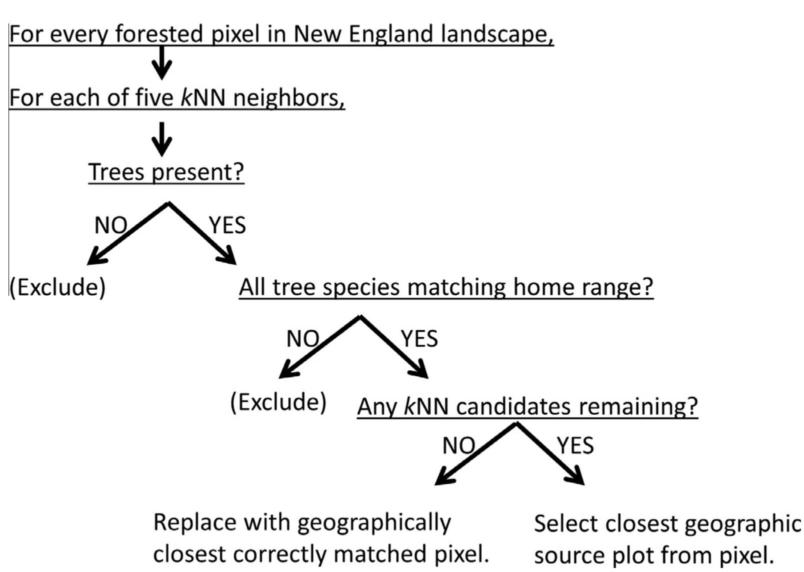

Fig. 2. Decision tree representing our source plot selection criteria for each pixel.