Modeling epidemiological disturbances in LANDIS-II

Francesco Tonini, Chris Jones, Brian R. Miranda, Richard C. Cobb, Brian R. Sturtevant and Ross K. Meentemeyer

Forest landscape simulation models (FLSMs) – often used to understand and project forest dynamics over space and time in response to environmental disturbance – have rarely included realistic epidemiological processes of plant disease transmission and impacts. Landscape epidemiological models, by contrast, frequently treat forest ecosystems as static or make simple assumptions regarding ecosystem change following disease. Here we present the Base Epidemiological Disturbance Agent (EDA) extension that allows users of the LANDIS-II FLSM to simulate forest pathogen spread and host mortality within a spatially explicit forest simulation. EDA enables users to investigate forest pathogen spread and impacts over large landscapes (> 10^5 ha) and long time periods. We evaluate the model extension using Phytophthora ramorum as a case study of an invasive plant pathogen causing emerging infectious disease and considerable tree mortality in California. EDA will advance the utility of LANDIS-II and forest disease modeling in general.

Keywords: LANDIS-II, forest landscape model, pathogen, Phytophthora ramorum, disturbance, epidemiological model

Introduction

Epidemiological disturbances, such as emerging pathogens and infectious disease outbreaks, are important agents of forest change around the world, causing tree mortality at scales ranging from individual trees of a single species to entire forest patches (Meentemeyer et al. 2008, Welsh et al. 2009). Beyond the complete loss of certain tree species, forest pathogens can significantly alter the functioning of forested ecosystems and the services they provide (Liebhold et al. 1995, Vitousek et al. 1997, Simberloff 2000). For example, pathogens can reduce the capacity of forests to sequester carbon, and can strongly interact with other types of disturbance such as fire, insects, and drought (Vitousek et al. 1997, Anderson et al. 2004, Dwyer et al. 2004, Dale et al.)
2009, Jactel et al. 2012). Developing a better understanding of how forest pathogens interact with other disturbances and changing environmental conditions to alter forest ecosystem dynamics is crucial for land managers, decision makers, and any stakeholder with multiple local interests involved (Cobb and Metz 2017).

Forest landscape simulation models (FLSMs) have been developed to specifically address management and research questions at landscape scales (> 10^3 ha) by projecting forest dynamics over space and time (Mladenoff 2004, Scheller and Mladenoff 2007). These models typically include details such as tree age, species and biomass, and are widely used to analyze the influence of disturbances over time as they affect large-scale forest ecosystem dynamics (Thompson et al. 2016). One of several FLSMs, LANDIS-II stands out as a process-based forest landscape model that can include variable time steps for different ecological processes (e.g. succession, disturbance, seed dispersal, forest management, carbon dynamics) and simulate their interactions as an emergent property of the independently simulated processed (Mladenoff 2004, 2005, Scheller et al. 2007). LANDIS-II continues to grow its user community and several extensions are available to simulate disturbances like wind, fire, insects, harvesting, or land-use change. To date, the representation of forest pathogen and disease spread in FLSMs including LANDIS-II has been lacking.

Landscape epidemiological models frequently treat forest composition and host density as static (Meentemeyer et al. 2011), meaning that the species do not age or experience effects of disturbance. This makes it difficult to understand how disease alters competitive interactions among species, a process known as apparent competition, which can alter species composition at a landscape level (Cobb et al. 2010). This lack of realistic changes in host community composition greatly impedes modeling the interactions of other landscape-level disturbances with disease spread (Cobb and Metz 2017).

In this paper, we fill this gap by introducing the Base Epidemiological Disturbance Agent (EDA) extension for LANDIS-II, which simulates forest pathogen spread and mortality in forested landscapes. The new extension is compatible with all LANDIS-II succession extensions and can be used in conjunction with other disturbance extensions (e.g. fire, insect, wind) to simulate their combined effects on forest landscape dynamics. In this paper, we provide an overview of the modeling framework behind Base EDA and an example application of the extension to simulate the expansion of the pathogen \( \text{Phytophthora ramorum} \) that causes 'sudden oak death' within the Big Sur area of California (USA).

**Model description**

LANDIS-II is a raster-based modeling framework consisting of a model core that links, parses, and validates data from multiple extensions (modules) and allows the user to 'plug in' a forest succession extension and any number of optional disturbance extensions (Scheller et al. 2007). EDA is a disturbance extension compatible with all LANDIS-II succession extensions. It is open source and freely available at the LANDIS-II website <www.landis-ii.org>. The download comes with an installer, user guide and sample data.

Base EDA requires the user to supply a raster map with location(s) of initial infection. The user must also supply agent-specific parameters such as host transmissivity, host susceptibility, climate tolerances and preferences, mean transmission rate, acquisition rate, maximum dispersal distance, and choose the appropriate dispersal kernel and exponent (see below). The user also provides parameters defining how other disturbances modify likelihood of infection. We demonstrate Base EDA with a case study of \( \text{Phytophthora ramorum} \), the pathogen which causes sudden oak death, a major forest disease in California (Meentemeyer et al. 2008, 2011, Metz et. al. 2017). For sudden oak death, fire kills the pathogen and slows reinfection for several years following fire (Beh et al. 2012).

Base EDA is specifically designed to simulate asymmetric weather-driven transmission of pathogen infection within a multi-host landscape. Transmission is modeled as a dynamic process affecting a meta-population comprised of \( N \) contiguous subpopulations represented by cells (sites) arranged on a grid. Cells contain forest tree species age cohorts, and (optionally) nonforest vegetation types. Tree mortality simulated by EDA is passed to the succession model that in turn handles vegetation response to that mortality (e.g. changes in light, water, and/or nutrients, depending on the succession extension used). Epidemiological disturbances within the EDA are probabilistic at the site level, where each site is assigned a probability of being in one of the following states: susceptible (S), infected (infectious non-symptomatic) (I), diseased (infectious and symptomatic) (D). Probabilities are compared with a uniform random number to determine whether the site becomes infected or, if already infected, to become diseased. Disease causes species- and cohort-specific mortality in the cell. The epidemiological model is similar to that in Meentemeyer et al. (2011) with adjustments made to fit the LANDIS-II framework and account for mortality. Additionally, the model can handle more than one EDA agent (pathogen), and is most compatible with aerial dispersal.

**Site host index**

Site host index (SHI) was adapted from the 'site resource dominance' concept in the LANDIS-II Biological Disturbance Agent Extension (Sturtevant et al. 2004). SHI accounts for the spatial distribution of known hosts of the EDA agent and is a combined function of tree species composition and the age cohorts present on that site. This approach allows the quantification of susceptibility for each non-infected cell to become infected, and the suitability of each infected cell to produce infectious spores. The relative host index value of a given species cohort is defined by its host competency class,
where low, medium, and high competency classes are user-defined using values ranging between 1 and 10, with non-hosts having a value of 0. The EDA extension compares a look-up table with the species cohort list at each cell generated by LANDIS-II to calculate SHI at time $t$ using one of two methods: 1) the host value from the maximum host competency class present, or 2) an average host value of all tree species present, where the host value of each species is represented by the one assigned to the oldest cohort. Species identified as ‘ignored’ do not contribute to the calculation of average resource value, while non-host species that are not ignored contribute a value of 0. Non-sporeulating hosts (i.e. hosts that do not contribute to pathogen or disease transmission) should not be included in the host index calculation.

**Site host index modifiers**

Site host index modifiers (SHIMs) are optional parameters used to adjust SHI to reflect variation introduced by both site environment (i.e. land type) and recent disturbances (Sturtevant et al. 2004). Land type modifiers (LTMs) and disturbance modifiers (DMs) range between −10 and +10, and are added to the SHI value of all affected sites where host species are present (SHI > 0). LTMs are assumed to be constant for the entire simulation, while DMs have a defined duration and decline linearly with the time since last disturbance ($t_{DST}$) as follows:

$$DM_{DST}(t) = DM_{max,DST} \star \frac{DM_{duration,DST}}{DM_{duration,DST}} - t_{DST}$$

Disturbances that may affect a given EDA agent include fire, wind, other EDA agents and insects, as well as timber harvest. SHI is then modified by LTM and the sum of all DMs:

$$SHIM(t) = SHI(t) + LTM + (DM_{wind}(t) + DM_{fire}(t) + \ldots)$$

The user should calibrate the two modifiers to reflect the relative influence of species composition/age structure, the abiotic environment, and recent disturbance on SHI. SHIM is normalized by its mean over the entire study area,

$$SHIM(t) = \frac{SHIM(t)}{SHIM_{mean}}$$

and modifies the disease transmission rate, $\beta$ (see section Weather). Normalization of SHI allows comparison of $\beta$ against homogeneous landscape conditions (where $SHIM=1$) and to interpret $\beta$ as the rate of secondary infection of cells by a single infected neighboring cell in an otherwise uninfected landscape.

**Weather**

An annual weather index, $w(t)$, is used to account for the effect of weather conditions on the probability of uninfected hosts becoming infected, and infected hosts spreading an individual EDA agent. Weather predictors (or transformations thereof) should be selected based on their relevance to the chosen EDA agent. The weather index is multiplied by a baseline transmission rate, $\beta_0$, to produce a time-dependent transmission rate, $\beta(t) = w(t)\beta_0$, where $\beta_0$ is defined by the user. The basic weather index for year $t$, $W(t)$, comprises the cumulative effect of $N$ weather predictors (e.g. rainfall alone, or rainfall and temperature) over a range of months, specified by the user (e.g. April to June), and is calculated as follows:

$$W(t) = \sum_{d[month_{1}(t)\ldots month_{2}(t)]} X_1 \times X_2 \times \ldots \times X_N$$

where $X_1 \times X_2 \times \ldots \times X_N$ represent the weather predictors and the cumulative sum runs over days included between two user-defined months ($month_1$ and $month_2$) for the current year $t$. If necessary, weather predictors in (Eq. 1) can be replaced by derived (e.g. aggregated, or transformed) versions. As an example, a predictor can be aggregated (summed or averaged) over $N$ consecutive days of a week or month (e.g. cumulative precipitation). Transformed predictors are expressed by a function, $\chi$. In the current version of the extension (version 1.0), only a polynomial transformation is available for the user, defined as:

$$f(X) = A + B + \exp\left(C \cdot \ln\left(\frac{X}{D}\right)\right)$$

where $A$, $B$, $C$, $D$, $E$, $F$ are constants specified by the user to adjust the shape of the polynomial (e.g. improving polynomial fit to empirical data on response of EDA agent to changes in temperature). As an example, such a transformation can reflect changes in rate of pathogen sporulation at increasing temperature values. The actual weather index, $w(t)$, is normalized by the mean $W_{mean}$ over the available time series of historical weather predictors:

$$w(t) = \frac{W(t)}{W_{mean}}$$

Normalization means that $\beta_0$ can be interpreted as the annual transmission rate under average (or under constant) weather conditions. The weather index built this way varies annually, but is spatially-uniform within each ecoregion.

**Epidemiological processes**

The epidemiological model shares features with spatially-structured metapopulation models and relies on a few important assumptions: first, only the presence/absence of infection in each cell is accounted for. This simplification ignores a transient effect (occurrence, spread and intensification) within the same cell, assuming that an effective level of inoculum is reached rapidly (but still below the maximum sporulating capacity of the cell). Improving this approximation would require a much larger computational effort in the parameter estimation procedure described in Filipe et al. (2012). Second, infected cells immediately become infectious, which is particularly true for an EDA with a small latent period across its host range. Third, infected sites remain infectious for an undetermined (i.e. long) period; in epidemiological terms the infectious period is considered indefinite.
and is left out of the model. The practical implication is that no cell can recover from infection throughout the simulation, for example by a within-host process such as a host defensive response. However, conversion from infected to uninfected status of a cell can occur due to 1) mortality of susceptible species by disease or other disturbances and/or 2) successional processes that result in a community with no hosts.

At every time step \( t \), a susceptible cell (site) \( i \) can become cryptically infected subject to a force of infection \( \Lambda_i(t) \) and, once infected, can become diseased at rate \( r_{i,D} \). Despite potentially containing dead hosts, symptomatically infected (diseased) cells have the same transmission rate, i.e. are as infectious as cryptically infected cells. The probabilities that cell \( i \) is in each of the possible states (susceptible, infected, diseased), \( P_{i,S} \), \( P_{i,I} \) and \( P_{i,D} \) respectively, are governed by a system of differential equations:

\[
\frac{\Delta P_{i,S}}{\Delta t} = -\Lambda_i(t) P_{i,S}
\]

\[
\frac{\Delta P_{i,I}}{\Delta t} = \Lambda_i(t) P_{i,S} - r_{i,D} P_{i,I}
\]

\[
\frac{\Delta P_{i,D}}{\Delta t} = r_{i,D} P_{i,I}
\]

The initial conditions for each cell, at the estimated time of onset of the outbreak, are \( P_{i,S} = 1, P_{i,I} = 0, P_{i,D} = 0 \), except at the cell estimated to be the location of the first infection, where \( P_{i,S} = 0, P_{i,I} = 1, P_{i,D} = 0 \). The force of infection, \( \Lambda_i(t) \), is given by:

\[
\Lambda_i(t) = \beta(t) \sum_{j \neq i} SHIM_j(t) \cdot SHIM_i(t) \cdot P_{j,i+D_{j,i}} \cdot K(d_{j,i}) \tag{2}
\]

where \( \beta(t) = w(t) \beta_0 \) is the transmission rate, with \( w(t) \) the annual index of weather fluctuation about a \( N \)-year average (see section Weather) and \( \beta_0 \) the baseline rate; \( K(d_{j,i}) \) is a dispersal kernel (see section Dispersal kernel) for a given distance \( d \) between target and source cells; \( P_{j,i+D_{j,i}} \) is the conditional probability that source cell \( j \) is infectious (either cryptic or symptomatic infection) given that target cell \( i \) is susceptible. To achieve a first order of approximation, we assume that \( P_{j,i+D_{j,i}} \approx P_{j,i} + P_{j,D} \), which we expect to be a reasonable approximation to the infection pattern, especially when dispersal is not too localized (e.g. within short distance from source of infection).

### Dispersal kernel

The dispersal kernel used in Base EDA is derived from, and shares code with, the seed dispersal kernel developed by Lichti and colleagues (N. Lichti, Purdue Univ., pers comm.). This dispersal function and associated distributions are especially suitable for aerially dispersed EDA agents that include a broad range of fungi and mistletoes. The probability that the agent disperses a distance \( d \) from the source was expressed by two main functional forms, often used in the literature: a power-law and a negative exponential. Their generic form can be defined as follows:

\[
K_{\text{PowerLaw}}(d) = d^{-a}
\]

\[
K_{\text{NegExp}}(d) = e^{-d/\alpha}
\]

An EDA agent produced in a source cell can only be deposited in a cell different from the source, i.e. transmission in force of infection (\( \Lambda \), see section Epidemiological processes above) is conditional on the agent being dispersed outside the source cell. The rationale for this choice is that infection processes within a cell are not tracked (no transient effect). In addition, the kernel must integrate to 1 within a chosen 2D spatial neighborhood window (excluding the source cell). The 2D window accounts for all possible pathways through which the target cell can become infected by a given source cell. A user-defined maximum radial distance is used to limit EDA agent dispersal within a chosen neighborhood size. For cases where only local, short-distance dispersal events are considered, this parameter becomes essential to reduce computational burden. Only isotropic dispersal (no wind-assisted directional spread) was considered for version 1.0 of the Base EDA extension.

### Tree species cohort mortality

Within each diseased cell, the mortality of individual tree species age cohorts is a probabilistic function of the mortality probability of the cohort’s vulnerability class. The user defines which species and ages fall into each vulnerability class (low, high), and the probability of cohort mortality for each class. Probabilities are compared with a uniform random number to determine whether an entire age-cohort dies (i.e. is removed) or not, where tree species cohort mortality is then passed to the succession extension which handles the removal of the cohort(s) and updates the cohort list. We acknowledge that complete cohort removal rather than a partial one may be a simplistic assumption in the current version of the model, but for many landscape-level processes or dynamics it should not cause significant changes in outcome. The Base EDA time step concludes updating the time since last disturbance, updating the time since last disturbance, outputting maps of cell states (1 = susceptible, 2 = infected, 3 = diseased) and cohort mortality, and by updating the Base EDA log file (Fig. 1).

### Case study

To demonstrate the capabilities of the Base EDA extension, we modeled 23 yr of *Phytophthora ramorum* spread within an 8017 km² area of central California, USA (Fig. 2). *Phytophthora ramorum* infects multiple hosts with some tree and shrub species experiencing non-lethal foliar symptoms known as ramorum blight, and oaks and tanoaks experiencing lethal stem cankers that lead to the disease sudden oak death.
The simulations were initiated with the best-known locations of initial infection in the study area in 1990 and simulated through 2013 (the last year for which plot level infection data are available) (Gaydos et al. 2017, Metz et al. 2017). We used LANDIS-II NECN Succession 1.0 (Scheller et al. 2011) to simulate forest growth and succession and the LANDIS-II Base EDA 1.0 extension to simulate spread of *P. ramorum* and mortality caused by SOD. Parameter values chosen for the EDA agent in this simulation are reported in Supplementary material Appendix 1 Table A1–A2. The simulations used a 30-m cell size. Base EDA used 1-yr time steps and NECN used 10-yr time steps. We compared the simulated disease spread in 2006, 2007, 2009, 2010, 2011, and 2013 with the subset of plots that were sampled in that year (i.e. plots sampled in 2006 were compared to model results in 2006 etc.) (Fig. 2) (Meentemeyer et al. 2008, Metz et al. 2017). We achieved a simulation accuracy of approximately 73.05 and 58.33% for infected and uninfected plots, respectively, for an odds ratio of 3.79 (Table 1). Calibration would allow for this to be further improved. Currently, the model is not predicting negative values as well as it does for positive values. Further calibration should improve this behavior. Moreover, it is partially due to the fact that the host data being used for the model are only 80% accurate at the landscape level.

The results also replicate the patchy nature of *P. ramorum* infection observed in the field (Meentemeyer et. al. 2008, Metz et al. 2017). This example illustrates the utility of being able to simulate disease spread and mortality with an existing FLSM to understand not only the spread of the disease, but also its potential impacts to the ecosystem through mortality of host trees.

We performed a sensitivity analysis of the model’s transmission rate (*β₀*) and α coefficient in the dispersal kernel. We choose to focus on both *β₀* and α coefficient as they are the parameters that will allow the user most flexibility when calibrating the model and they will have substantial impact on spread. For this analysis we focused on model accuracy as measured by the odds ratio. We ran 3 simulations of each model with a different random seed in order to account for stochasticity between model runs. *β₀* varied from 4.00 to 5.00 in 0.25 increments and α varied from 2.4 to 2.6 in 0.1 increments for a total for 15 different combinations of *β₀* and α and total number of model simulations of 45. On average decreasing *β₀* by 0.25 resulted in a 7.01% decrease in the odds ratio (a measure of accuracy) while holding α constant. On average a 0.1 decrease in α resulted in a 15.2% increase in the odds ratio while holding *β₀* constant.

More broadly, the Base EDA extension could be a suitable landscape modeling tool for a range of EDA agents. Across the globe, an increasing number of destructive pathogens have emerged as disturbance agents shaping forest structure and function at landscape scales. These events have substantial ecological and economic impacts, the understanding of which are important to designing management responses (Liebhold et al. 1995, Vitousek et al. 1997, Simberloff 2000). The default Base EDA data and parameterization is most suitable for aerially dispersed pathogens and those where a biologically-driven infectious period is not a significant factor. These conditions are met for the most destructive forest diseases in North America including chestnut blight, sudden oak death, and possibly Beech Bark Disease although the latter system involves an insect that may complicate the process of infection and spread. In practice, we emphasize the importance of parameterizing the dispersal kernel for application to a new system. Proper understanding of dispersal dynamics is critical to accurate forecasting of spread and disease dynamics (Meentemeyer et al. 2011, Filipe et al. 2012, Metz et al. 2017). Acquiring empirical measurements of dispersal at scales more than a few meters is challenging but we emphasize it is incumbent on users to overcome this difficulty in order to properly apply the model. Examples of confronting this problem for *P. ramorum* can be found in Meentemeyer et al. 2011 and Filipe et al. 2012. These examples integrated several datasets to estimate and validate dispersal parameters including spore trapping, molecular data, landscape-extent monitoring plot networks, and aerial tree mortality mapping from fixed-wing aircraft. We encourage further experimentation with alternative formulations of dispersal kernels and environmental (weather) dependencies as these could render the extension suitable for a greater range.
of epidemiological disturbance agents such as pathogens spread via insect vectors, movement of contaminated soil or plant material, and spread in waterways.

Table 1. Accuracy assessment of the model results at a landscape level comparing plot observations to model observations for the year the observations occurred (e.g., plots sampled in 2007 were compared to model results in 2007). The true positive rate is 73.1% and the true negative rate is 58.3% and total accuracy is 68.9%. Values are aggregates of all years considered in the model.

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>Modeled</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>225</td>
<td>50</td>
</tr>
<tr>
<td>Negative</td>
<td>83</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>73.1%</td>
<td>58.3%</td>
</tr>
</tbody>
</table>

Figure 2. 2013 most recent plot disease status compared to 2013 model results. For comparison we used modeled diseased status and most recent plot diseased status (not all plots are sampled every year so this comparison will tend to underestimate plot disease status). For simplicity and realistic comparisons, we treated both infected and uninfected model results as uninfected since infected non-symptomatic areas would be recorded as uninfected in the field due to no visible symptoms.

To cite the Base EDA LANDIS-II extension or acknowledge its use, cite this Software note as follows, substituting the version of the application that you used for ‘version 0’:


Software name: Base EDA for LANDIS-II.
Programming language: C#.
Available at: <www.landis-ii.org/extensions>.
Reproducible analysis repository: <https://github.com/f-tonini/LANDIS-II_EDA_CaseStudy>.
Acknowledgements – The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper. The authors thank all members of the Meentemeyer Landscape Dynamics Lab at the Center for Geospatial Analytics, Rob Scheller and Eric Gustafson for their feedback and valuable suggestions on the present work. We would like to thank N. Lichti for the use of his code for the dispersal component of the epidemiological process.

Funding – This research was supported by the National Science Foundation [grant numbers DEB-EF-0622677 and EF-0622770] as part of the joint NSF-NIH Ecology of Infectious Disease program. The authors also gratefully acknowledge financial support from the USDA Forest Service Pacific Southwest Research Station, the USDA Forest Service Forest Health Protection, State and Private Forestry, and the Gordon & Betty Moore Foundation. BS and BM are supported by the National Fire Plan of the US Forest Service.

Author contributions – All authors made substantial contributions to this work in the following areas: as first author, FT led and coordinated the study presented herein, alongside with code development for the epidemiological model. CJ and BM equally contributed to project development and methods. CJ helped with extensive editing, structuring, writing, and code development needed to parallelize the EDA extension; BM contributed code development with extensive efforts on the weather component of the EDA extension, alongside editing the manuscript. RC provided expertise on epidemiological processes and contributed to editing the manuscript; BS edited the manuscript and provided expertise on LANDIS-II and other ecological disturbances; and RM conceived the project and edited the manuscript.

References


