Available online at www.sciencedirect.com
ScienceDirect
Ecological Modelling
Ecological Modelling 00 (2020) 1–52

Spread of Dutch Elm Disease in an urban forest

Nicolas Bajeux[a], Julien Arino[a,b], Stéphanie Portet[a,b], Richard Westwood[c]

[a] *Department of Mathematics, University of Manitoba, Winnipeg, MB, Canada*
[b] *Data Science NEXUS, University of Manitoba, Winnipeg, MB, Canada*
[c] *Center for Forest Interdisciplinary Research, University of Winnipeg, Winnipeg, MB, Canada*

Abstract

A complex network model for the spread of Dutch Elm Disease in an urban forest is formulated. American elms are the focus of the model. Each elm can be in one of five states, a combination of their life and epidemiological status. Each tree is also potentially a host to a population of elm bark beetles, the vectors of Dutch Elm Disease. The epidemiological dynamics of trees is governed by a stochastic process that takes into account the dispersal of spore-carrying beetles between trees and potential contacts between tree root systems. The model describes seasonal variations of beetle activity and population dynamics. Numerical simulations and sensitivity analyses of the model are carried out. In this introductory paper, we use data from the City of Winnipeg, where Dutch Elm Disease is prevalent, and focus on two neighbourhoods representative of a residential area and an area with urban parks.

Keywords: network epidemic model, plant pathogens, hybrid model, multi-scale model

1. Introduction

Urban forestry is a specialized branch of forestry that involves the cultivation and management of trees in urban, urban interface areas and greenspaces for the physiological, sociological and economic well-being of urban society (Jorgensen, 1974; Harris et al., 1999; Miller et al., 2015). The presence of trees in urban landscapes provides benefits by mitigating some of the negative impacts of urban development such as increased heating, energy and carbon dioxide levels, diminished air quality or loss of rain water (Dwyer et al., 1992; Grote et al., 2016; Livesley et al., 2016). Elm trees (*Ulmus spp.*) have traditionally been planted

widely in urban forests in North America due to increased tolerance to adverse environmental conditions, fast rates of growth and high survival as seedlings and juveniles (Rioux, 2003). American elm, (*Ulmus americana L.*), has been planted extensively in many North American cities as a park and boulevard tree due to its dense canopy and vase shaped form, which provides efficient shading in summer and its ability to tolerate cold winters and hot dry summers. These traits have made American elm the tree of choice for planting for many decades in urban environments.

However, urban trees are subject to many environmental stressors that includes attack by insects and diseases. North American elm species are highly susceptible to Dutch Elm Disease (DED), an Ascomycete fungal wilt disease caused by three species in the Genus *Ophiostoma*. In northern North America, the primary fungal strain is *Ophiostoma novo-ulmi Brasier* which is spread mostly by the Native elm bark beetle (NEBB) (*Hylurgopinus rufipes* Eichoff (Pines and Westwood, 2008)).

Once the fungus is introduced into the tree, it spreads through the cells of the phloem and xylem, which inhibits water and nutrient transport (Hiratsuka et al., 1987). American elm is the most susceptible North American elm species and infections are almost always fatal with death usually occurring over a period of weeks to months (Hubbes and Jeng, 1981; Hubbes, 1988; Hildahl and Jeffrey, 1980; Stipes and Campana, 1981). DED reached eastern North America in the late 1920's from Europe and the west coast of North America (Oregon) by 1973 and in the 50 years; after its introduction the disease is estimated to have killed 50 to 100 million elms (Soll, 2016). The disease eliminated American elm as a major component in the urban forests of many cities in eastern North America as well causing extensive mortality in natural stands of American elm (Sinclair and Campana, 1978).

World wide many bark beetle species are able to transmit *Ophiostoma* spores to elm with NEBB being the primary vector in the northern North America (Swedenborg et al., 1988). Bark beetles such as NEBB are an important symbiotic partner of these fungi and the primary vector of the disease among elms (Webber et al., 1984). Adult NEBB's mate and feed in twig crotches in the tree canopy during the spring. After mating, female beetles construct brood galleries within the cambium, laying eggs in the galleries. Brood galleries are formed in the larger diameter canopy branches and sometimes the upper trunk at the beginning of the infection (Kaston, 1939; Whitten, 1964; Thompson and Matthysse, 1972; Lanier et al., 1982; Swedenborg et al., 1988; Pines and Westwood, 1996). The larvae then feed and develop in the galleries, eventually pupating within individual chambers (Hiratsuka et al., 1987). Later in the summer and in fall, newly emerged adults usually move to new trees to feed. If the previously colonised tree was infected with *O. novo-ulmi*, NEBB will mechanically carry fungal spores upon emergence from brood galleries (Kondo et al., 1981). In the late fall the newly emerged adults will move to the base of a healthy tree or occasionally a still living infected tree to overwinter (Strobel and Lanier, 1981; Anderson and Holliday, 2003). Overwintering beetles burrow into the bark (at a height generally below the snow line) and overwinter below the phloem. After emerging from their overwintering diapause state in spring at the tree base adults carrying *Ophiostoma* spores will move into the canopy of healthy trees to feed and mate and begin forming brood galleries. It is during this period that inoculum can be introduced into the xylem of healthy elms (Gardiner, 1981).

A second infection route for DED may occur via interconnected root systems in areas of higher tree density (Brasier and Gibbs, 1978). In urban environments, where boulevard elms may be planted close together, root infections can quickly spread within streets causing

further mortality to elms within neighbourhoods.

In urban forests the disease may cost millions per year to keep the incidence economically manageable by spreading losses over a longer time period (Westwood, 1991). Dutch elm disease reached Manitoba in 1975 and in response a large scale integrated pest management program (IPM) was developed to manage the disease (Westwood, 1991). Control of Dutch elm disease is an expensive process due to the complex relationship between pathogen, host and vector, and its IPM program is based on reducing the probability of occurrence of new disease infections (Gibbs, 1978; Dreistadt et al., 1990; Westwood, 1991). Integrated control programs include sanitation, which involves the elimination of vector populations through insecticides and removal of elm wood that could provide potential habitat (Westwood, 1991). Other control methods include severing root grafts between adjacent trees to prevent fungal spread, chemical injections of fungicides to increase tree fitness, and increased propagation of resistant elm species. One missing aspect of DED control which requires more attention is improving disease management and reducing disease incidence through a better understanding of how DED spreads within urban forests. Studies that have focused on the spread of the disease are few, and mainly occurred in the United Kingdom (Sarre, 1978; Swinton and Gilligan, 1996, 1999, 2000; Harwood et al., 2011).

Winnipeg, Manitoba is located adjacent to the temperate shrublands and grasslands biome found over much of the southern portion of the prairie provinces in Canada. Winnipeg contains a large urban forest of over 8 million trees, including the largest remaining urban elm canopy in North America (Barwinsky, 2016; City of Winnipeg, 2018, 2019). The impact of DED in North America has been catastrophic in ecological and economical terms through the loss of tree diversity in urban forests and natural forest stands (Hubbes, 1999; Rioux,

2003; Pines and Westwood, 2008; Oghiakhe, 2014). Slowing the spread and incidence of the disease in urban forests is of major interest in order to save the remaining elm trees still present in cities of North America.

We propose a model to represent the mechanism of spread in a large urban forest by focusing on both infection modes. The pathogen and its dynamics are not explicitly modelled; only tree and vector dynamics are described. We consider a network model where vertices are the elms. Dispersal of beetles between elm trees defines the first network used in this study. The second network is defined by considering the connectivity between tree root systems. Beetle dynamics are represented by a stochastic matrix population model. A second stochastic process governs the life and infection dynamics of elm trees. The resulting model is complex and mathematically intractable. In order to study its behaviour, in this preliminary study, we perform computational and sensitivity analyses, focusing on the case of two urban forests in neighbourhoods in the City of Winnipeg.

2. Material and Methods

2.1. Modelling

In order to model the spread of DED in American elms, which dominate the urban forest in various Winnipeg neighbourhoods, we consider the spread via beetle vectors and the root systems of elm trees (Figure 1). Each tree is characterised by its health and epidemiological conditions. The population of beetles on each tree is also tracked. Dispersal of beetles between trees and interactions between root systems of different trees are modelled using two networks. The trees are the vertices of networks. The presence of an edge indicates that beetles could move between two trees or that the root systems are connected.

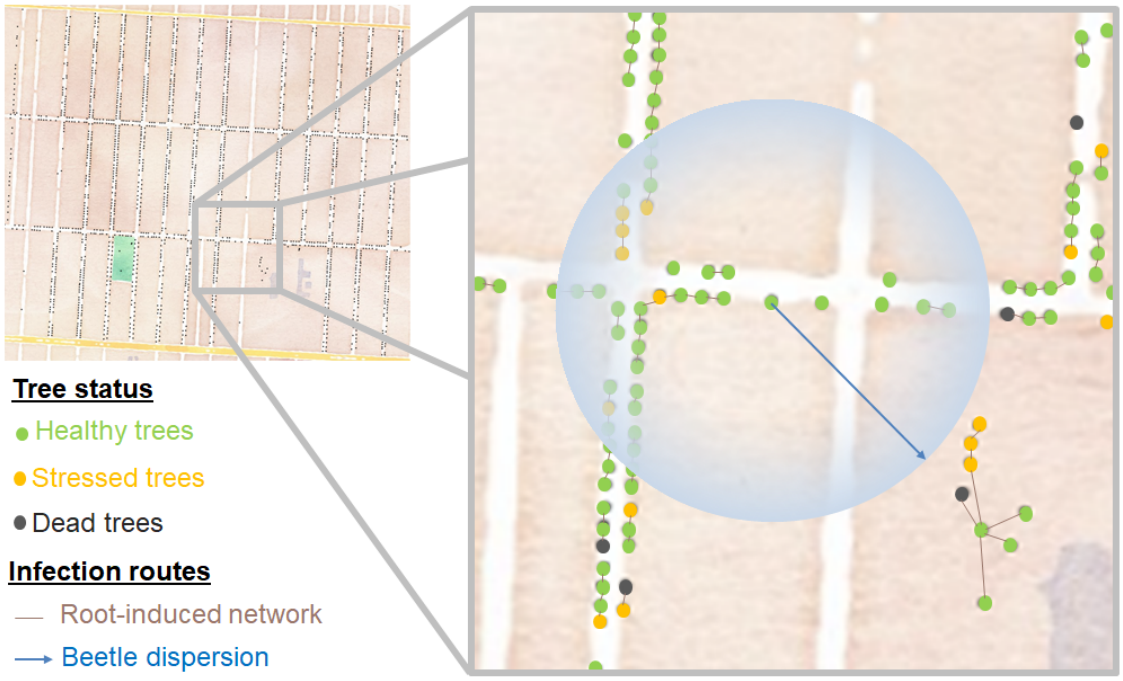


Figure 1: Conceptual model for the spread of DED in an urban forest following two routes of infection by vector dispersion or root systems interconnection. American elms (dots) are located along boulevards or in parks in an urban neighbourhood. Colour of dots codes for their life and epidemiological status. When trees are close enough, their root systems might be interconnected (brown network). The blue circle delimits the dispersal region for beetles present on the tree at the centre of the circle. All trees within the blue region are reachable by beetles when in their dispersal stages; all these trees are thus connected to the centre tree in the beetle dispersion network (not shown).

The model is formulated in discrete-time. Two time steps are used to depict the spread of the disease and the beetle dynamics.

- The time step required to update the life and epidemiological status of the trees is denoted Δt and represents one year in our simulations.
- The beetle lifespan in Manitoba is reported as one year in duration (Anderson, 1996); we denote δt the time step at which beetles grow and move (one week in the model).

2.1.1. Urban forest system

The spatial setting of the model is the population of American elm trees found in Winnipeg's urban forest. Elms are considered habitats for the beetles and can be represented as heterogeneous in age, size, environment, etc.

Inventory data for American elm trees on public owned land was obtained from the City of Winnipeg open data portal (`https://data.winnipeg.ca`). The dataset contains the locations (coordinates) of about 300,000 trees on public property in the city, including (at the time of writing) approximately 60,000 American elms. The location and diameter at breast height is available for each tree.

In this work, the set of elms is denoted $\mathcal{T}$. We refer to a given tree in the set as $T \in \mathcal{T}$ or using some index $k = 1, \ldots, |\mathcal{T}|$, where the cardinality $|\mathcal{T}|$ is the total number of American elms. For a given tree $T \in \mathcal{T}$, we denote $\varnothing_{bh}(T)$ its diameter at breast height.

2.1.2. Beetle dispersal network

Beetles disperse between trees when they seek food, a location to lay their eggs or an overwintering site. In order to keep the model computationally tractable, we create a network

$\mathcal{N}^B$ establishing, for each tree, what other trees are close enough that they can be targets of beetle movement. To do this, we consider the distances between each pair of trees in the database. If $T', T'' \in \mathcal{T}$ are two trees, then there is an edge between T' and T'' in $\mathcal{N}^B$ if

$$d(T', T'') \leq R_B,$$

where $d(T', T'')$ is the Euclidean distance between T' and T'' and the parameter R_B is the maximum beetle dispersal distance.

Weighting the beetle dispersal network. The network $\mathcal{N}^B$ lists the trees that a beetle can fly to if they are in a given tree $T \in \mathcal{T}$. Denote $\mathcal{D}^B(T)$ the set of trees that are directly connected to T (trees in the blue region in Figure 1), i.e., its neighbours within $\mathcal{N}^B$. When a beetle undertakes movement between trees, it is possible that it does not survive the trip; this is used later in the dynamical model. The probability of surviving dispersal is used as the weight of edges in $\mathcal{N}^B$ and is defined for $T \in \mathcal{T}$ as, for all $T' \in \mathcal{D}^B(T)$,

$$p_{\{d\}}(T, T') = \exp\left(-\frac{d(T, T')}{R_B}\right). \tag{1}$$

2.1.3. Root network

A different network, $\mathcal{N}^R$, is used to describe the potential connections between root systems of nearby trees. The edges of this network are weighted to indicate uncertainty in the extent of root systems.

The extent of the root system of a given tree is approximated from its diameter at breast height (DBH). If $T \in \mathcal{T}$ is a tree and $\varnothing_{bh}(T)$ is its DBH, we assume that the maximal extent of the root system of T is $3h(T)$, where $h(T)$ is the height of T obtained from $\varnothing_{bh}(T)$

by (A.1); see details in Appendix A. Let $T', T'' \in \mathcal{T}$ be two trees. Then there is an edge between T' and T'' in $\mathcal{N}^R$ if

$$d(T', T'') \leq 3\left(h(T') + h(T'')\right), \tag{2}$$

i.e., if the root systems of T' and T'' are in contact.

Pruning the root network – Removing roads and rivers. Root systems are greatly disrupted by the way roads are created, to the point that two trees separated by a cement road cannot have their root systems be in contact (personal communication of City of Winnipeg Urban Forestry Branch).

As a consequence, pruning of $\mathcal{N}^R$ is carried out as follows. For each edge $T_i \leftrightarrow T_j \in \mathcal{N}^R$, we instantiate the line segment T_iT_j as an `sf` object in `R`. Using the `R` library `openstreetmap`, we download the coordinates of all roads in Winnipeg and then remove from $\mathcal{N}^R$ all edges that are such that the line segment joining them intersects a road. The same procedure is used to remove edges between trees separated by one of the rivers that flow through the City of Winnipeg.

The pruned network, still denoted $\mathcal{N}^R$ and called root network, then has edges between any pair of trees whose root systems come into contact and that are not separated by a road or a river (brown network in Figure 1).

Weighting the root network. We assume that when the root networks of two trees are in contact and one of the trees is infected, the infection can be transmitted to the other tree. To model this root-driven infection, we add weights to the root network $\mathcal{N}^R$. These weights represent the probability of infection by roots. They take into account the distance between

trees and use a parameter $p_R \in [0,1]$ accounting for uncertainty in the extent of roots systems. Let $T', T'' \in \mathcal{T}$ be two trees. The weight of the edge between T' and T'' is given by

$$p_{\{r\}}(T', T'') = \begin{cases} p_R & \text{if } d(T', T'') \leq h(T') + h(T'') \\ \dfrac{3(h(T') + h(T'')) - d(T', T'')}{2(h(T') + h(T''))} p_R & \text{otherwise.} \end{cases} \quad (3)$$

Recall that there is no edge in $\mathcal{N}^R$ between T' and T'' if (2) is not satisfied. In the remainder of the paper, $\mathcal{N}^R$ refers to this weighted network. As for the beetle network $\mathcal{N}^B$, we define the root neighbourhood of tree $T \in \mathcal{T}$ as the set $\mathcal{D}^R(T)$ of trees that are directly connected to T in $\mathcal{N}^R$. The parameter p_R is the maximum probability of infection via roots.

Assessing topology of the root network. Some properties of networks are relevant for the propagation of the disease (Shirley and Rushton, 2005). To assess topological differences between different types of neighbourhoods, some properties of the root networks are investigated. We used the package `igraph` (Csardi and Nepusz, 2006) in the `R` environment. Here, we consider degree, strength, connected component, distance, eccentricity, betweenness and density of the root network; see details in Appendix B.

2.1.4. Elm dynamics

Hypotheses. Each tree $T \in \mathcal{T}$ can be in one of two epidemiological states and one of three life stages: susceptible (S) or infected (I) and healthy (H), stressed (W) or dead (D). The state W aggregates stressed trees and trees with excessive amounts of dead canopy wood. It is assumed that a tree cannot be infected and healthy, so elms can be in five states: healthy susceptible (S_H), stressed susceptible (S_W), dead susceptible (S_D), stressed infected (I_W) and dead infected (I_D).

The natural life cycle of trees in the absence of infection is to remain many years in S_H, then switch to S_W and finally die when they switch to S_D. Transitions between epidemiological states depend on the cumulative number of spore-carrying beetles that have travelled to a tree during a one year period (infection via beetles) and the distance that separates the tree from other infected and dead trees (infection via roots).

Mathematical representation. For each tree, transitions between states are governed by a discrete-time Markovian process on the state space $\mathcal{S} = \{S_H, S_W, I_W, S_D, I_D\}$ combining tree ageing and beetle- or root-driven infections. The flow diagram in Figure 2 gives transitions between states for a tree $T \in \mathcal{T}$. For beetle-driven infection, the cumulative number $B^T(t)$ of spore-carrying beetles that have travelled to tree T during $[t, t + \Delta t]$ is given by (5). For root-driven infection, the set $\mathcal{D}^R_{I_D}(T)$ represents dead infected trees (I_D) that are in $\mathcal{D}^R(T)$ at $t + \Delta t$.

Ageing and beetle-driven infection or root-driven infection are two independent processes; however, they impact the transition probabilities simultaneously. A selection process combining the outcomes of those independent stochastic processes is used to determine the future state of the tree; details are given in Appendix C. At time t, the outcome of stochastic process is a vector $Z_t = (Z_t^{T_1}, Z_t^{T_2}, ..., Z_t^{T_{|\mathcal{T}|}})$ representing states of all trees of $\mathcal{T}$ with $Z_t^{T_i} \in \mathcal{S}$. The update of tree states occurs during the week 21 of the year (middle of May), which is assumed to be the end of winter in Winnipeg. Parameters used in the tree dynamics are listed in Table C.3.

2.1.5. Beetle dynamics

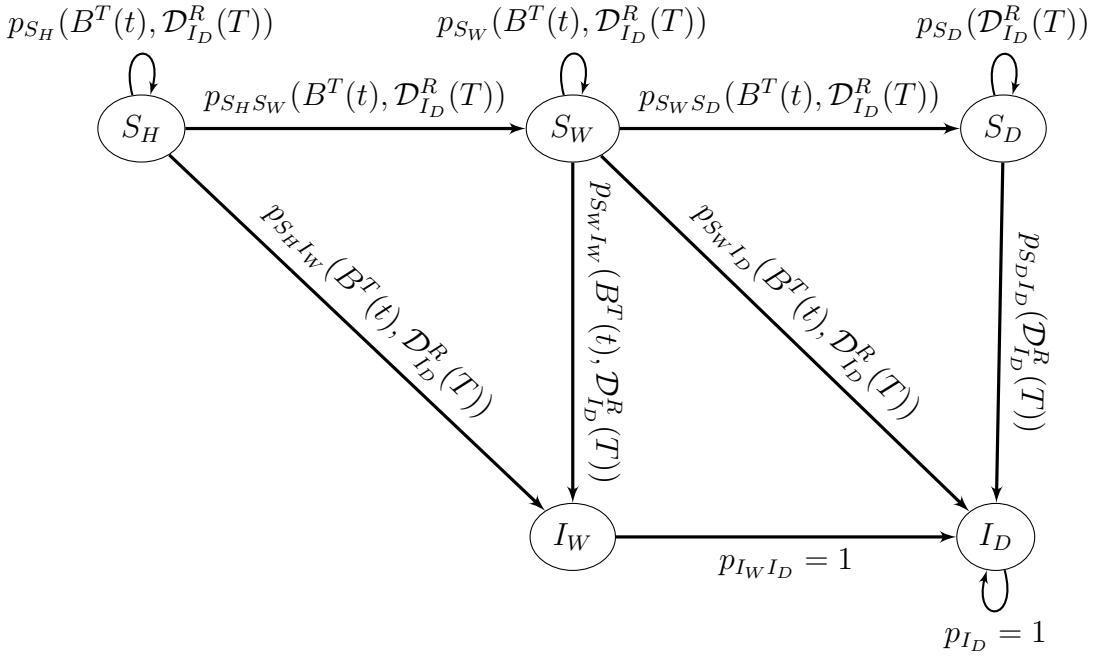


Figure 2: Flow diagram of states for a tree $T \in \mathcal{T}$. Nodes represent tree states and arrows are valued with transition probabilities indicating their dependence on beetle population or root system. The cumulative number of spore-carrying beetles that have travelled to tree T during $[t, t + \Delta t]$ is $B^T(t)$. The set $\mathcal{D}^R_{I_D}(T)$ is composed of dead infected trees (I_D) that are in $\mathcal{D}^R(T)$.

Hypotheses. In order to maintain computational tractability, our model for beetle dynamics focuses on the trees that insects are colonising and the season, rather than on the physiological states of beetles.

In the absence of DED, five stages for beetles are considered.

Juvenile J – the juvenile stage J is the aggregation of the biological stages eggs, larvae and pupae. Individuals J develop in live trees in summer and fall.

Dispersing F – newly emerged dispersing adults move in fall to the canopy of adjacent live trees to feed.

Callow adults C – callow adults C are adults feeding on canopy of live trees in fall, then entering winter diapause in the trunks of healthy trees and in spring leaving/emerging from tree trunks to their canopy.

Dispersing M – dispersing individuals that move in summer from overwintering sites (healthy or weak trees) to colonise new trees for laying eggs.

Adults A – adults A are mature adults that lay eggs during summer, resulting in the next generation of juveniles. After laying eggs, adults die.

The simplified life cycle of beetles considered in the present study is shown in Figure 3. Season lengths ought to vary as a function of the temperature; however, in this preliminary model we caricature a typical year in Winnipeg using data shown in Figure D.14. The seasons and corresponding beetle behaviour are as follows.

Winter (overwinter) – callow adults overwinter in healthy trees. If beetles do not reside on a healthy tree or if they are at any other stage of their life, they die when winter

starts. Winter lasts from week 45 one year to week 21 the next.

Spring (emergence) – callow adults go to feed at the canopy of trees in which they overwintered; they mature to dispersing individuals (M). Spring lasts one week (week 22).

Summer (breeding) – At the beginning of summer, callow adults that are still alive become dispersing individuals (M). Dispersing M fly to colonise new trees. Once they arrive, dispersing individuals become mature adults (A). Eggs are laid in branches or upper trunks of trees, resulting in the new generation of juveniles (J). It is assumed that beetles are univoltine; adult females lay on average 60 eggs (Kaston, 1939) and then die. New juveniles stay in this stage until the end of summer. Summer lasts from week 23 to week 38.

Fall (new generation of beetle) – the new generation of beetles (J) can now start emerging as dispersing individuals (F), which leave trees to seek food in healthy or weak trees. Once they have arrived in appropriate trees, dispersing individuals become callow adults (C) and begin feeding on the tree canopy. They stay there and prepare themselves for winter. This period lasts from week 39 to 44. After this period, the cycle starts again.

It is assumed that the *presence of disease* does not affect the beetle life cycle. However, in the presence of DED, individuals can become spore-carrier only by growing, as a juvenile, in an infected (stressed or dead) tree. If the tree is infected and dead, then the juvenile carries spores, whereas if the tree is infected and stressed, then the beetle has a chance to not become spore-carrier.

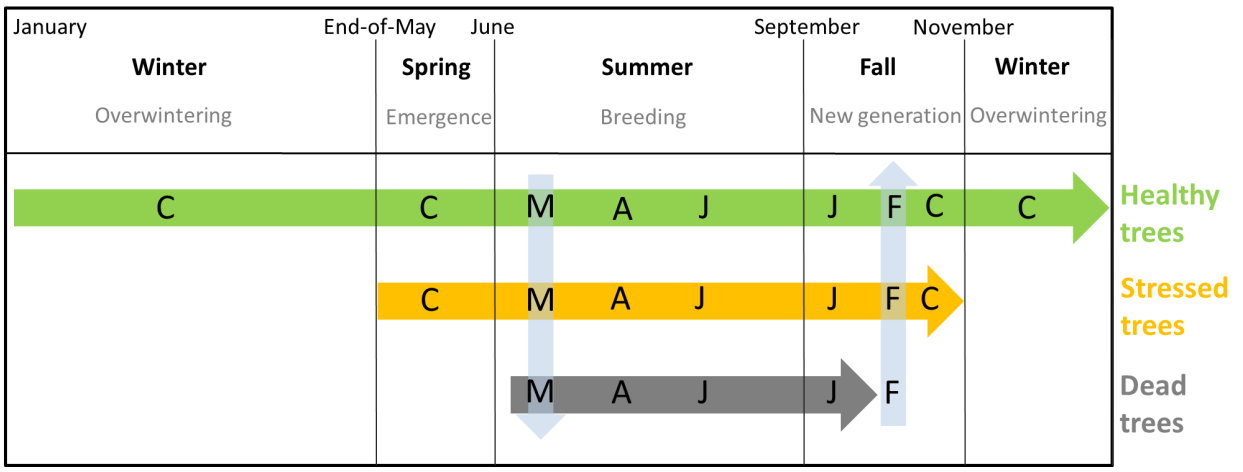


Figure 3: Simplified life cycle of beetles depending on tree of residence and period in a typical year in Winnipeg. C, A and J represent callow adults, adults and juveniles, respectively. M and F are the two dispersal stages as indicated by blue arrows. Overwinter, emergence, breeding and new generation refer to functional periods of beetles.

It is assumed that an individual becoming a spore-carrier as juvenile remains a spore-carrier all its life. We denote J_P, F_P, C_P, M_P and A_P the corresponding stages to indicate that individuals are spore-carrying. It is assumed that, of the spore-carrying states, only individuals in F_P can infect trees when they arrive to feed on the canopy.

Mathematical representation. The flow diagram for beetle dynamics is shown in Figure 4; its dependence on periods and tree states are detailed in Figure D.15. Beetle dispersal and growth are represented by a stage-structured matrix model (Caswell, 2001). The time step for the beetle dynamics is δt. In our simulations, δt is chosen as the minimum time of sojourn in stages, which is one week. As is customary in structured population models, only female beetles are considered in the model; a constant sex ratio is assumed (Caswell, 2001).

The evolution of the beetle population is governed by the following equation,

$$\mathbf{\Pi}(t+\delta t) = (s_{\delta t}\mathbf{S}(t, Z_t) + \beta(t, Z_t)\mathbf{P})\,\mathbf{D}(t, Z_t)\mathbf{\Pi}(t), \tag{4}$$

where $\mathbf{\Pi} = \left(\mathbf{\Pi}^{T_1}, \ldots, \mathbf{\Pi}^{T_{|\mathcal{T}|}}\right)$ with, for a given tree $T \in \mathcal{T}$,

$$\mathbf{\Pi}^T(t) = \left(J^T(t), J_P^T(t), F^T(t), F_P^T(t), C^T(t), C_P^T(t), M^T(t), M_P^T(t), A^T(t), A_P^T(t)\right),$$

the beetle population vector at time t. In (4) and wherever matrix products are involved, $\mathbf{\Pi}$ is assumed to be a column vector but we omit the transpose operator to simplify notation. The matrix $s_{\delta t}\mathbf{S}(t, Z_t)$ describes survival and state transitions, $\beta(t, Z_t)\mathbf{P}$ is the fertility matrix and $\mathbf{D}(t, Z_t)$ is the dispersal matrix. The mortality $1 - s_{\delta t} \in [0, 1]$ is a fixed proportion of beetles that die each week; it occurs irrespective of survival through other processes. The scalar-valued function $\beta(t, Z_t)$ indicates whether birth occurs depending on periods and tree states. The matrices $\mathbf{S}(t, Z_t)$ and $\mathbf{P}$ are detailed in Appendix D. Specific values for all parameters appearing in Figure 4 are given in Tables D.4 and D.5.

Beetle dispersal is encoded in the matrix $\mathbf{D}(t, Z_t)$. At each time step δt, this matrix, which acts before the processes of survival and fertility, potentially moves beetles between trees. Dispersal only occurs at the beginning of summer (breeding) and in fall (new generation) when beetles are in one of the four dispersal stages (M, M_P, F and F_P). In winter or spring, $\mathbf{D}(t, Z_t) = \mathbb{I}$, the identity matrix.

The matrix $\mathbf{D}(t, Z_t)$ is not shown explicitly for summer and fall. Its computational analogue is detailed instead here. All dispersal stages last one time step δt. When a beetle on tree $T \in \mathcal{T}$ enters one of the four dispersal stages, it looks for an appropriate tree T' in $\mathcal{D}^B(T)$:

- at the beginning of summer, during the breeding season, an M or M_P beetle seeks new trees (S_H, S_W, I_W, S_D and I_D) to colonise and to mate in;
- in fall, a new generation F or F_P beetle seeks healthy and stressed trees (S_H, S_W and

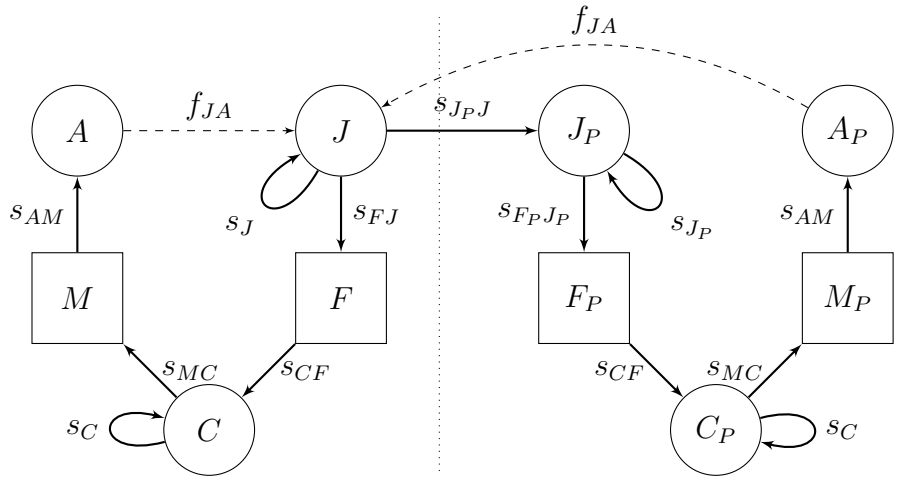


Figure 4: Flow diagram for the age- and epidemiological-structure of beetle population dynamics. Nodes are the life stages of the beetles, with those left of the thin dotted vertical bar being spores-free and those to the right spore-carrying. Solid arrows represent transitions between stages, with arrow labels indicating the probability rates at which these transitions occur. Dashed arrows represent fecundity. Squared nodes indicate where the stochastic dispersal process can take place.

I_W) to feed on the canopy.

If there is no appropriate tree in $\mathcal{D}^B(T)$, the beetle dies. Otherwise, the beetle chooses an appropriate tree T' at random (uniformly) within $\mathcal{D}^B(T)$. Once the destination is chosen, the probability for the beetle to survive dispersal is given by the weight $p_{\{d\}}(T, T')$ of the edge $T \leftrightarrow T'$ in $\mathcal{N}^B$, as given by (1).

Hence, the number $B^T(t)$ of spore-carrying beetles that have travelled to tree T during $[t, t + \Delta t]$ used for tree infection is computed as follows:

$$B^T(t) = \sum_{i=0}^{|\mathcal{W}|} F_P^T(t - i\delta t), \tag{5}$$

where $|\mathcal{W}|$ is the number of weeks in a given year.

2.2. Setting up simulations

As simulation of the model is computationally intensive, in this preliminary study, we focus on the two areas shown in Figure 5: a zone with two parks with high American elm densities (Figure 5a) and a residential block neighbourhood (Figure 5b). The park area (PCP) results from grouping two Winnipeg neighbourhoods: Pulberry and Crescent Park. The boulevard area, in which trees line the streets, is North River Heights (NRH).

(a) Park-type neighbourhood (PCP) (b) Boulevard-type neighbourhood (NRH)

Figure 5: The two Winnipeg neighbourhoods used in this study: Park-type (PCP) and Boulevard-type (NRH). Black dots are the American elms under consideration.

The initial date for simulations is July 31 of year 0, which falls within the breeding period. Furthermore, we assume that only spore-carrying juvenile beetles (J_P) are present at the beginning of simulations. The initial number of beetles in the juvenile stage per tree is set to 500 individuals per infected tree. Concerning initially infected trees, we use three approaches:

- *One cluster* has all trees in a zone delimited by a circle infected and dead (I_D); the

radii of the circles for the two neighbourhoods are given in Table 1. All other trees are healthy (S_H).

- *Two clusters* consists of two clusters as defined above, with the total number of infected and dead trees for the entire neighbourhood the same as in the one cluster case.

- *Randomised* has the same number of initially infected and dead trees, randomly located over the neighbourhood under consideration.

In order to have comparable initial conditions, the percentage of initially infected and dead trees is the same for the three approaches. This is done for clusters by adapting the radius. Thus, only the initial spatial layout of the infection changes.

Parameter values used for simulations are provided in Tables C.3, D.4 and D.5.

2.3. Main parameters – Sensitivity analysis

In addition to initial conditions, the study focuses on the effect of four parameters. (i) The maximal distance R_B that beetles can move during dispersal. Beetle dispersal is complex to measure and our present study explores a large range for beetle dispersal. Pines and Westwood (2008) found that beetles may travel up to 1000m from their original tree but most beetles move much shorter distances if sufficient healthy elms are available; based on this, we explore the beetle dispersal distance from 20 to 380m in the simulations. (ii) The probability p_i that a spore-carrying beetle successfully introduces the pathogen into a susceptible tree. This has been shown to be between 3 and 5 % in (Webber et al., 1984). (iii) The weekly survival rate $s_{\delta t}$ of beetles. (iv) The maximum probability p_R that infection is transmitted via the root system. Ranges of these parameters are given in Tables C.3 and D.5.

To assess the effect of these parameters on the prevalence of disease, i.e., the proportion of infected trees over time, a sensitivity analysis is carried out. The four main parameters above are the input factors for the sensitivity analysis. We consider ten different values for R_B (20 to 380m with 40m step), three for $s_{\delta t}$ (0.97, 0.98, 0.99), four for p_R (0,0.25,0.5,0.75) and three for p_i (0.01, 0.02, 0.03), leading to 360 different combinations of parameters. Each combination of parameter values is used for 100 simulations and the proportion of infected trees over time is the output. The sensitivity analysis is an analysis of variance (ANOVA) and is carried out using the `R` package `multisensi` (Bidot et al., 2018). Sensitivity indices (SI) of input factors are computed every year. Results represent the contributions of each input factor (main effect), of interactions between two factors (interactions) and of interactions of more than two factors (residual) to the total variability of the disease prevalence over time. Furthermore, global sensitivity indices (GSI) are calculated that represent the sum of the main effect and all the interactions for each input factor.

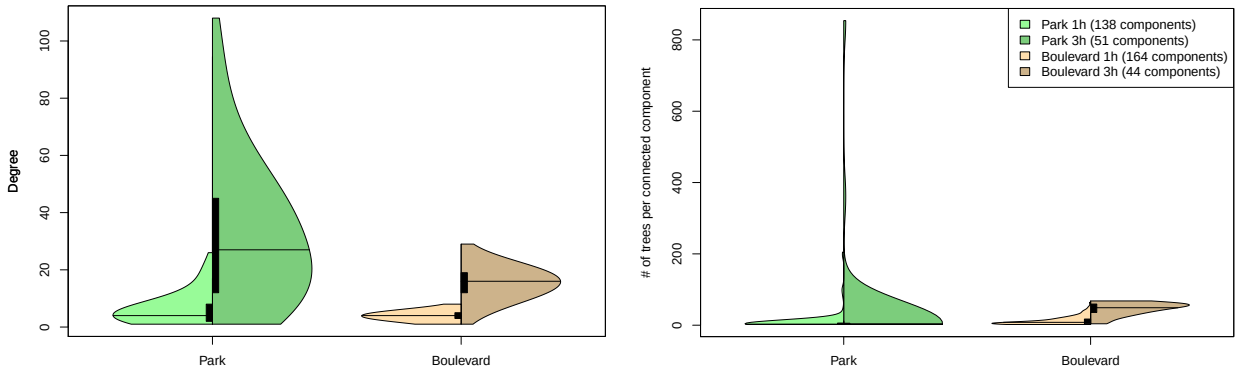
Type	Beetles in trees	Cluster radius
One cluster	500 (in I_D)	96m (PCP) and 100m (NRH)
	0 (otherwise)	
Two clusters*	500 (in I_D)	47.5m and 280m (PCP)
	0 (otherwise)	
Randomised	500 (in I_D)	–
	0 (otherwise)	

Table 1: Initial conditions used in simulations. *: Two cluster analysis only conducted in PCP.

3. Results and Discussion

3.1. Topology of the root network

The distribution of degrees with, for each neighbourhood, $1h$ (the network describing the extent of maximum transmission probabilities) on the left and $3h$ (the network describing the extent of positive infection probabilities) on the right are shown in Figure 6a. In a park neighbourhood, considering a larger extent for a tree's root system as is done in $3h$ has strong consequences, with some trees being in this case connected to more than 100 other trees.



(a) Degree (b) Trees per connected component

Figure 6: Violin plots of (left) the distribution of degrees in $\mathcal{N}^R$ for the two neighbourhoods used and (right) the distribution of trees per connected component in $\mathcal{N}^R$ for the two neighbourhoods used. See Appendix B for details.

Considering strengths, from Table 2, while the median is very similar for $1h$, it is almost twice as large for trees in parks for $3h$, showing a significant difference between the two topologies.

Another difference between parks and boulevards is observed when considering *connected components* in the graph. While the situation appears quite similar if referring to Table 2,

	Park 1h	Park 3h	Blvd 1h	Blvd 3h
Degree (M)	4	27	4	16
Strength (M)	8.00	14.50	8.00	8.37
Connected components (#)	138	51	164	44
Eccentricity (G,M)	10	14	7	4
Eccentricity (W,M)	–	1.52	–	0.45
Diameter (G)	30	21	32	8
Diameter (W)	–	3.42	–	2.04
Betweenness (G,M)	8	663	14	25
Betweenness (W,M)	7.81	201.00	14.00	40.00
Component density (%,M)	100	100	49.44	30.65
Component diameter (G,M)	1	1	3	5
Component diameter (W,M)	–	1	–	0.72

Table 2: Some network topology characteristics. Park and Boulevard refer to these two neighbourhoods, while 1h and 3h indicate whether the network under consideration is the one where root systems extend 1 or 3 times the height of the tree, respectively. "M" next to an indicator means the median value is given, while "G" and "W" refer, respectively, to indicators computed using geodesic (unweighted) and weighted graphs. "–" indicates that the value is automatically the same as the value in the line above. See Appendix B for details.

the situation is different when one considers $3h$ in Figure 6b (the right section of each violin plot). There are two main differences. In parks, the median number of trees per connected component is lower than in boulevards. However, park neighbourhoods also comprise some connected components that contain many trees (for instance, the two actual parks seen in Figure 5a). In boulevard neighbourhoods, connected components are more homogeneous with respect to the number of trees they contain.

The next three indicators in Table 2, eccentricity, diameter and betweenness, further highlight the organisation of the connected components. In terms of infection through root systems, eccentricity describes how far away in the graph the infection can travel in a connected component. Unsurprisingly, the median eccentricity is much higher for parks than it is for boulevards. Interestingly, while the mean eccentricity strongly differentiates both neighbourhood types, the diameter is quite similar for both. Finally, betweenness is lower in parks for $1h$ but much higher for $3h$.

The last three indicators listed in Table 2 are obtained by drilling down further into connected components. Clearly, density is higher in parks that in boulevards. Results are striking: the median density of components is 100% in parks, even for $1h$, while it is at most 50% in boulevards. Thus, half or more of the connected components in parks have all trees connected to each other through their root systems, which is translated in the diameter by a median value of 1. The diameter of the components is larger in boulevards.

3.2. Role of root-driven infection

The impact of the root network on the spread of the disease in both neighbourhoods (PCP and NRH) for different values of the beetle maximum dispersal distance R_B is investi-

gated in Figure 7. The presence of infection by root systems always favours the propagation of the disease. Increasing beetle dispersal R_B also augments the prevalence of the infection. Further, when both modes of infection are considered and beetle dispersal is low, the prevalence is higher in parks (PCP, dark gray). Then, as beetle dispersal increases past $R_B = 100$, the situation reverses and prevalence becomes higher in boulevards (NRH, light gray). Conversely, when there is no root-driven infection, the situation is different: there is no difference until $R_B = 220$, at which point the prevalence becomes higher in parks. Below $R_B = 260$, the topology of the neighbourhood does not seem to play a role in the spread of the disease.

This could be a consequence of the difference of topology of the root networks in both neighbourhoods (Section 3.1). There are two large connected components in the park neighbourhood and many small others, while the boulevard neighbourhood has connected components that are much more homogeneous in terms of the number of trees they comprise (Figure 6b). When beetles can disperse far enough, the infection is likely to spread faster in neighbourhoods with many medium-sized connected components (NRH) than in ones with few very large and many very small components (PCP). When beetles do not disperse very far, they can have trouble reaching and infecting the other connected components of the network; however, in the park-type neighbourhood, if they happen to reach one of the large connected components, then the root network takes over and drives the propagation.

From now on, all results consider both modes of transmission.

The maximum probability of infection through the root system is described by the parameter p_R; see Eq. (3). The influence of this parameter on the spread of the disease for one cluster and randomised initial conditions is shown in Figures 8 and 9, respectively. In

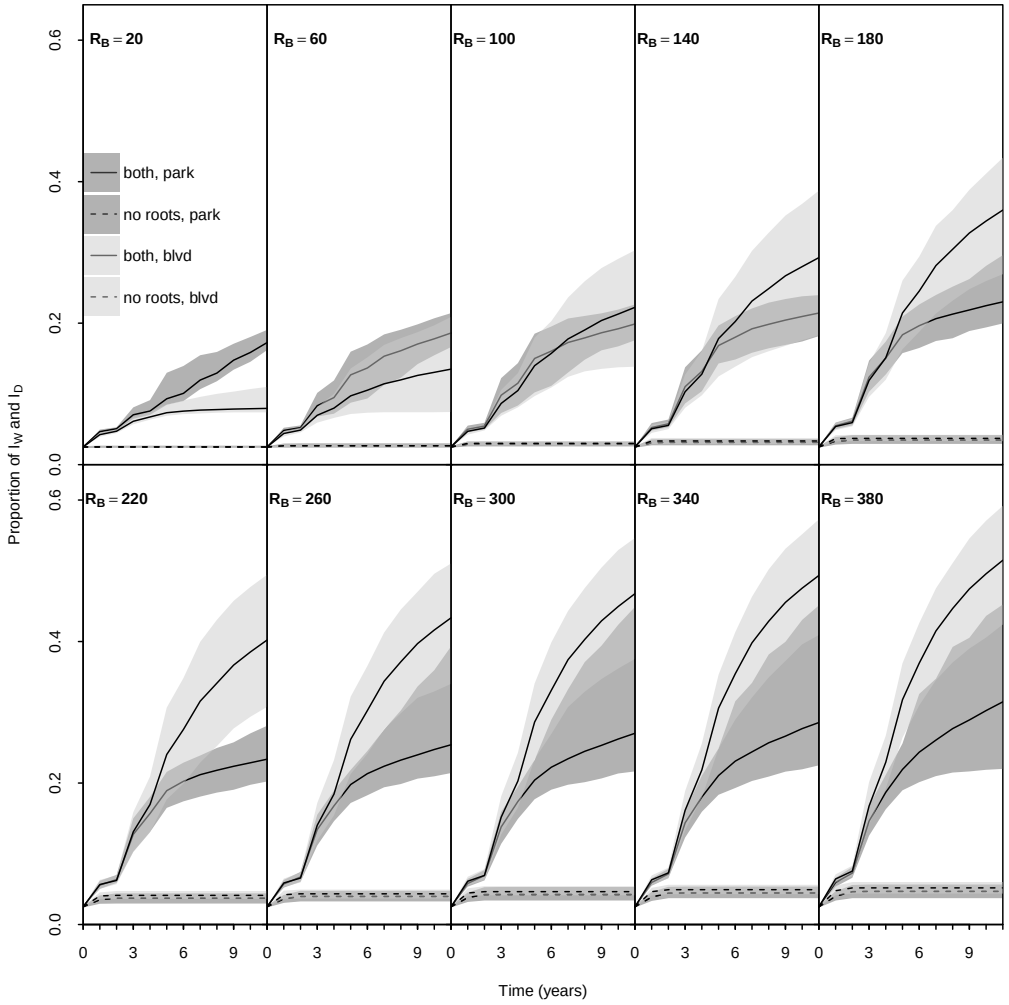


Figure 7: Influence of the root network $\mathcal{N}^R$ on the spread of DED infection: temporal evolution of the mean proportion of infected trees (I_W and I_D) over 100 simulations. Solid curves: both modes of infection are considered (beetle- and root-driven); dashed curves: no root-driven infection. Shaded zones represent 95% confidence bands. Dark gray is for PCP neighbourhood while light gray is for NRH. Beetle dispersal distance R_B varies from 20m to 380m by steps of 40m. Initial conditions are one cluster of infected trees with 500 spore-carrying juvenile beetles. Main parameter values are $p_R = 0.5$, $p_i = 0.02$ and $s_{\delta t} = 0.98$. Other parameters as in Tables C.3, D.4 and D.5.

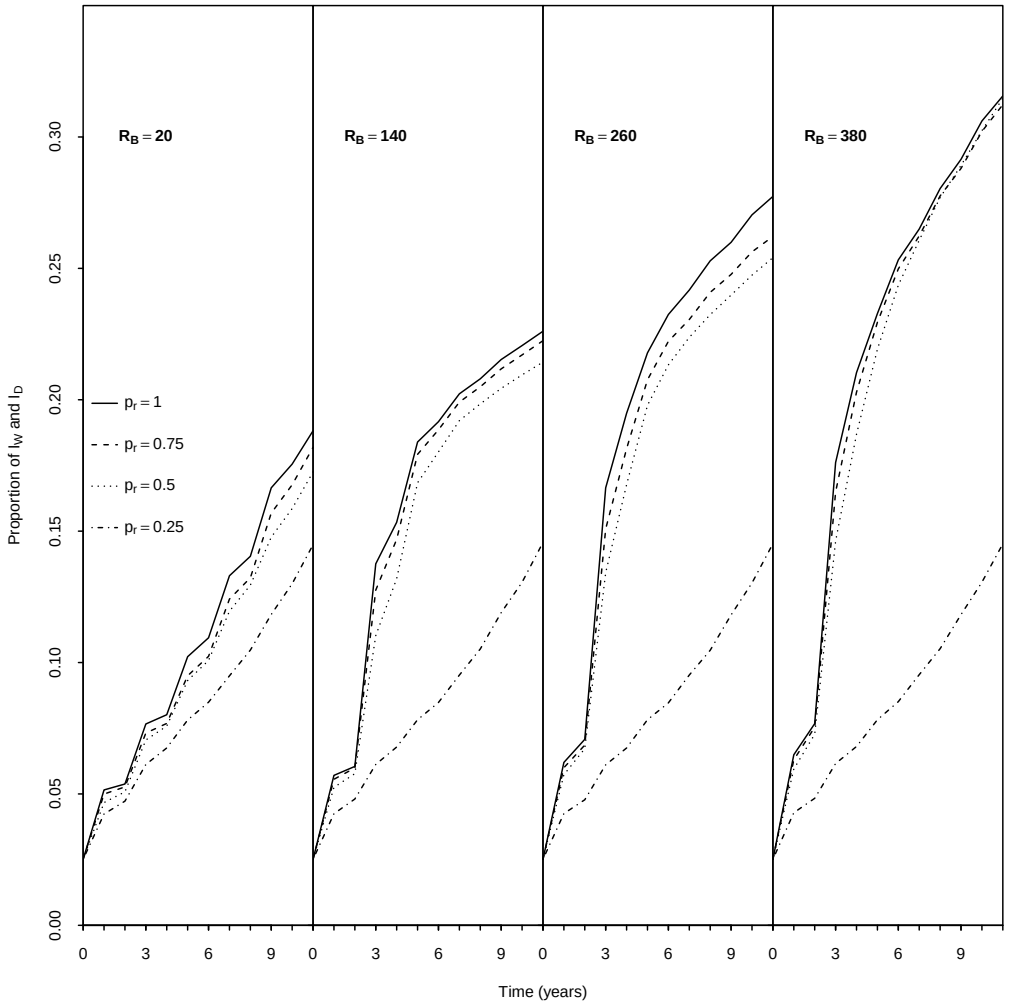


Figure 8: Influence of the maximum probability of infection via roots p_R on the proportion of infected trees in simulations for the park-type neighbourhood (PCP) with one cluster of initially infected trees, each with 500 spore-carrying juvenile beetles. R_B is increasing from 20m to 380m by steps of 120m. Curves are the mean proportions over 100 simulations. Main parameter values are $p_i = 0.02$ and $s_{\delta t} = 0.98$. Other parameters as in Tables C.3, D.4 and D.5.

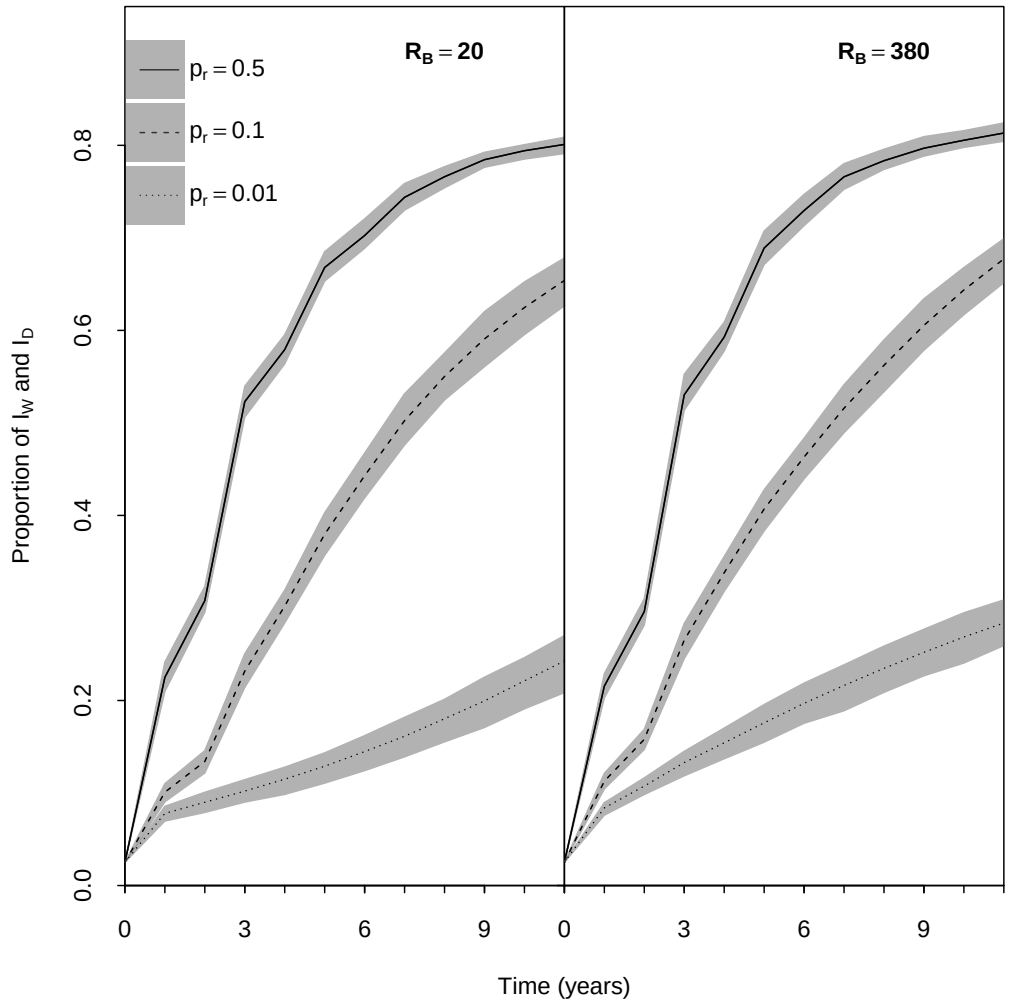


Figure 9: Influence of the maximum probability of infection via roots p_R on the proportion of infected trees in simulations for the park-type neighbourhood (PCP) with random initial conditions; each infected tree bears 500 spore-carrying juvenile beetles. (left) $R_B = 20$m. (right) $R_B = 380$m. Curves are the mean proportions over 100 simulations. Shaded zones represent the variability using the 95% confidence bands. Main parameter values are $p_i = 0.02$ and $s_{\delta t} = 0.98$. Other parameters as in Tables C.3, D.4 and D.5.

Figure 8, the prevalence of the infection is increasing as the maximum probability of infection via roots p_R becomes larger. As beetle maximum dispersal increases, so does the proportion of infected trees. Note that, once the probability of infection becomes larger than 0.5, the mean prevalence of the disease is qualitatively the same for all values of p_R. Note that there is a steep increase of the proportion of infected trees between years 2 and 3 (for $p_R > 0.25$); this occurs when the infection reaches one of the large-sized connected components of the root network. In Figure 9, in which initial conditions are random, the prevalence of infection is qualitatively the same for all values of R_B.

3.3. Effect of initial conditions

Initial conditions have an important effect on the prevalence of the disease. Based on Figure 8 (one cluster of initially infected and dead trees), we observe that increasing the maximum beetle dispersal distance R_B induces an increase of the number of infected trees. However, in Figure 9 (random initial conditions) and for a large enough value of p_R, the dispersal of the beetles has no influence on the propagation of the disease. This is also observed in Figure 10, in which the prevalence of the disease is compared for one cluster, two clusters and randomised initial conditions (see Section 2.2) in the park-type neighbourhood (PCP). The prevalence of the disease is higher when the initial distribution of infected and dead trees is randomised. In this case, the dispersal of beetles has little effect and we observe that the infection is severe from the first year of the epidemic. However, as R_B gets larger, the number of trees infected during the epidemic for the one- and two-clusters initial approaches increases.

The marked difference in Figure 10 between the behaviour of infection for randomised

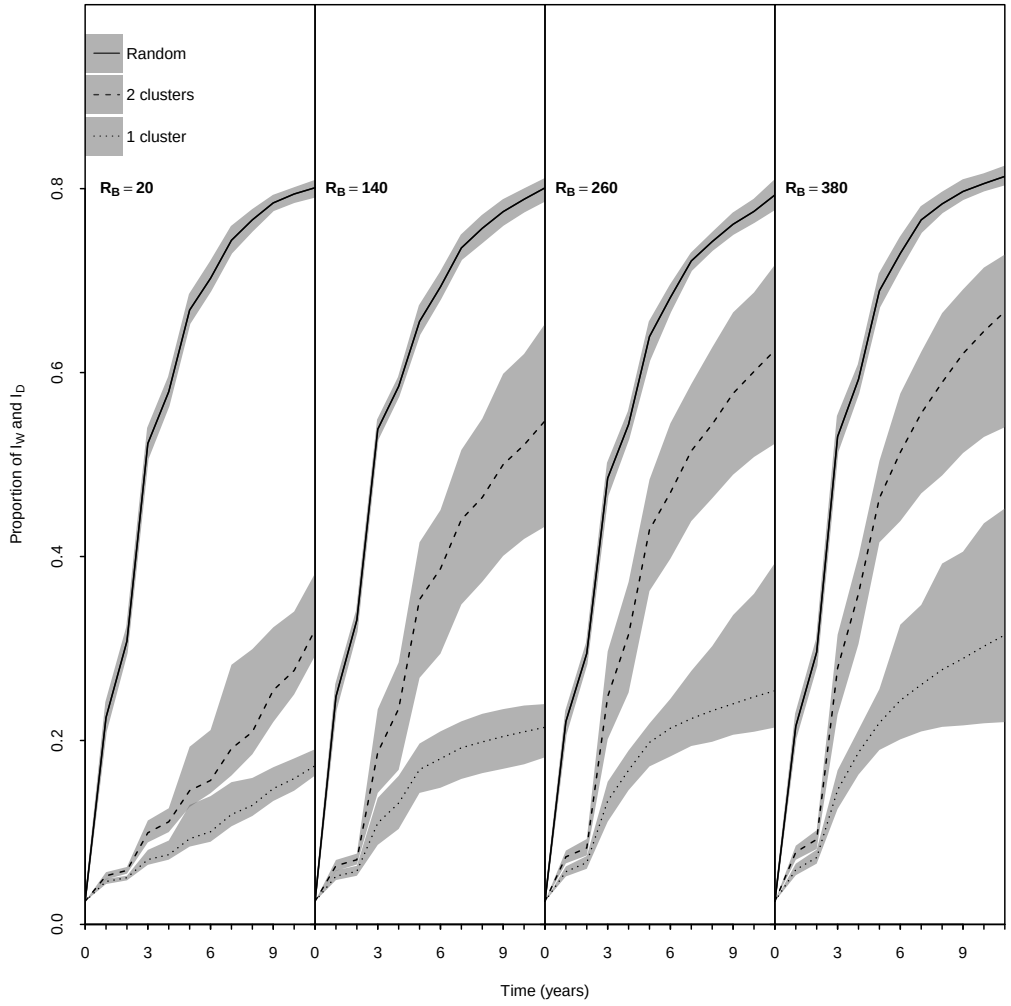


Figure 10: Influence of initial conditions to the proportion of infected trees in the park-type neighbourhood (PCP). Curves are the mean proportions over 100 simulations. Shaded zones are the 95% confidence bands. (Solid curves) *random* initial conditions, (dashed curves) *two clusters* and (dotted curves) *one cluster*. Each infected tree initially bears 500 spore-carrying juvenile beetles. R_B is increasing from 20m to 380m by steps of 120m. Main parameter values are $p_R = 0.5$, $p_i = 0.02$ and $s_{\delta t} = 0.98$. Other parameters as in Tables C.3, D.4 and D.5.

initial conditions and for clusters could be explained as follows. In clusters, only the root systems of trees at the periphery of the cluster can contribute to the root-driven infection process. By opposition, in randomised conditions, it is most likely that each initially infected and dead tree has only susceptible neighbours in $\mathcal{D}^R$ (see Section 2.1.3). Since the park-type neighbourhood has two large connected components in the root network $\mathcal{N}^R$, it is very likely that several of the initially infected and dead trees are within these components, implying that they have access to a very large number of susceptible trees. The same mechanism seems to operate when comparing one initial cluster and two initial clusters. We postulate that one initial cluster and randomised initial conditions represent the two extreme cases in terms of magnitude of infection and that increasing the number of clusters would increase the severity of the infection.

3.4. Impact of beetles dynamics

Using one cluster initial conditions for infected trees, the spatial and temporal spread of the disease is quantified using the wave front of infection. At the start of the simulation, all trees are within the cluster, so the maximum distance between any two trees is no larger than the radius of the cluster. As the infection starts to spread, we track the maximum distance separating all pairs of infected trees. Figure 11 shows the extent of disease propagation at each time step using the average over 100 simulations of this maximum distance. Parameters related to beetle dynamics ($s_{\delta t}$) and probability of beetle-driven infection (p_i) are varied for different values of R_B. For small values of R_B, varying the values of parameters p_i and $s_{\delta t}$ does not affect the wave front. When R_B gets larger, increases of both p_i and $s_{\delta t}$ have an effect on the wave front, substantially increasing the extent of the infection.

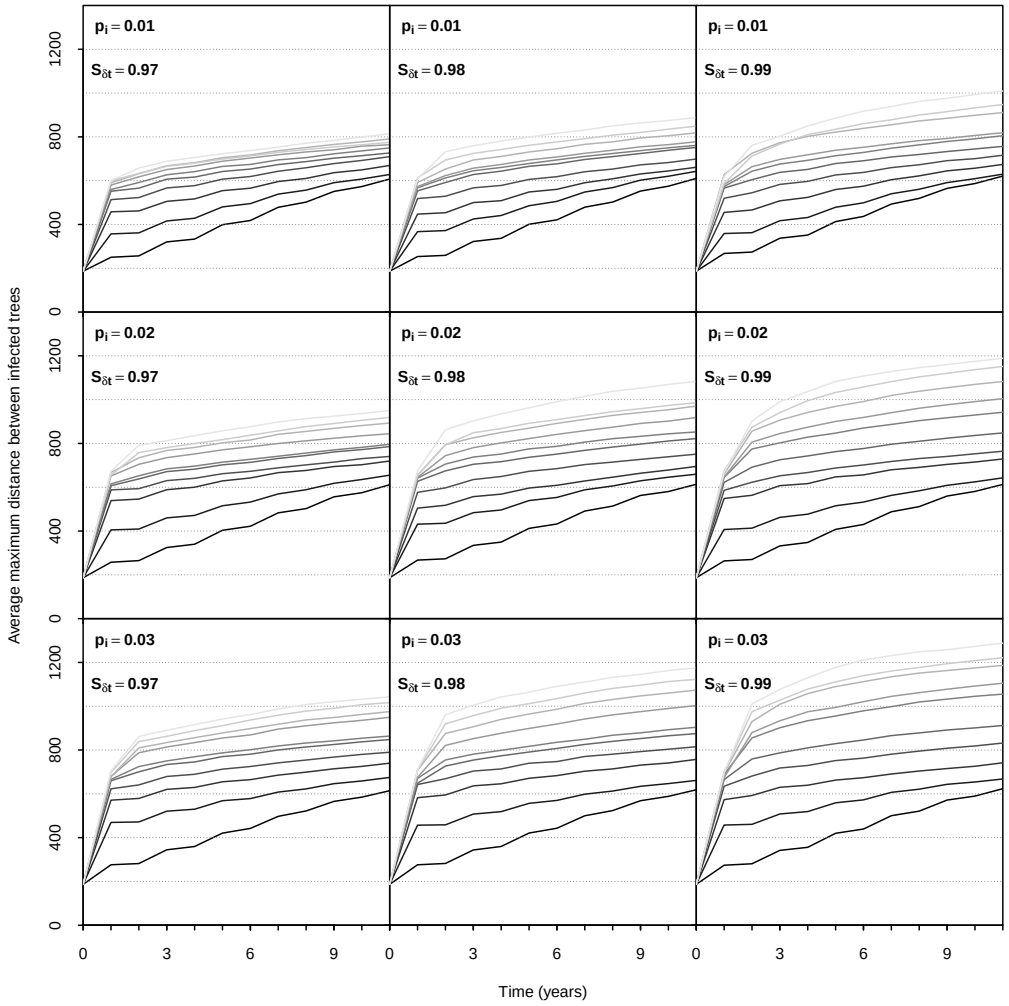
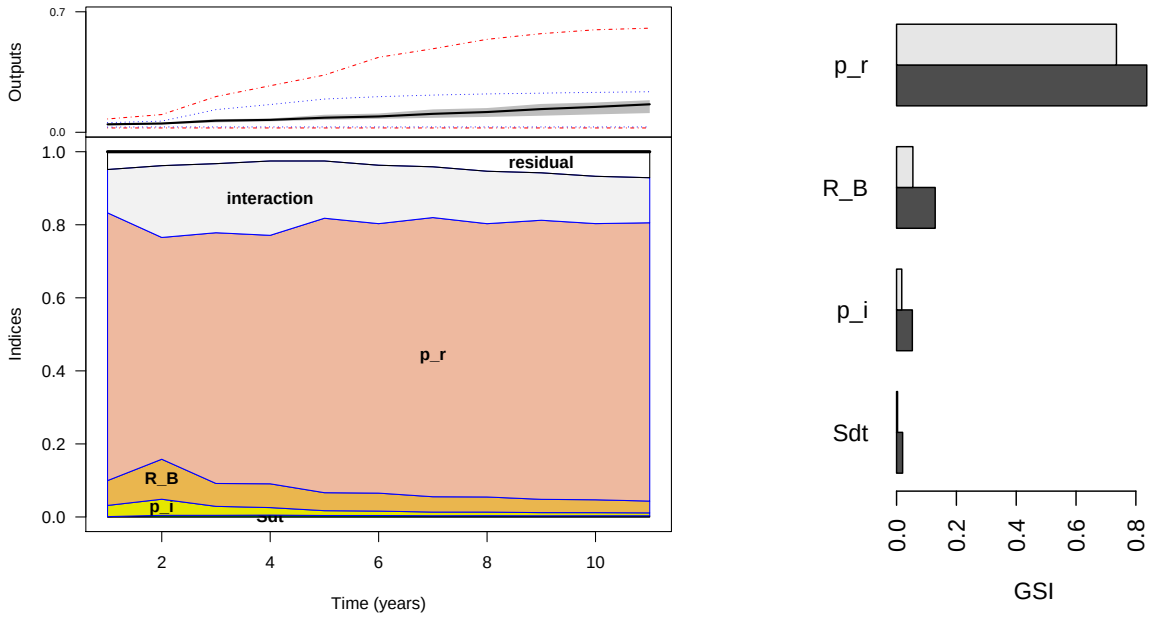


Figure 11: Disease propagation wave front. Average of the maximum distance between infected trees over time. In each panel, R_B varies from 20m (black) to 380m (light gray) by steps of 40m. Parameter $s_{\delta t}$ varies from 0.97 to 0.99, while p_i varies from 0.01 to 0.03. Other parameters as in Tables C.3, D.4 and D.5 with $p_R = 0.5$.

438 ### 3.5. Sensitivity analysis



(a) Sensitivities indices over time

(b) Global sensitivity indices

Figure 12: Sensitivity analysis on the proportion of infected trees over time over 360 scenarios. (Left) The upper part gives the overall trends of the outputs (proportions of infected trees). Extreme values (min and max) are in red, the grey zone is the interquartile and the median of all outputs is the solid black curve. The lower part gives sensitivity indices using ANOVA over time for input factors R_B, $s_{\delta t}$, p_R and p_i. (Right) Global sensitivity indices of the four parameters ranked in descending order of global effect. Main effect is in light gray while global effect is in black.

439 Results of the sensitivity analysis carried out on the four parameters, p_R, p_i, R_B and $s_{\delta t}$,
440 are summarised in Figure 12. The park-type neighbourhood (PCP) with one cluster initial
441 conditions is used for the analysis. Sensitivity indices are the contributions to the total
442 variability of each input factor or of interactions between two factors or more (residuals)
443 at time t. They are represented by the lengths of the different colour segments along a

vertical line at time t, in the lower part of Figure 12a. The prevalence of disease is mostly sensitive to p_r, i.e., the maximum probability that an infected tree infects adjacent ones. At the beginning of the infection, the maximum beetle dispersal distance R_B plays a role; it becomes less influential in the long-run. The dynamics of the proportion of infected trees is barely sensitive to the probability p_i that one beetle infects trees at the beginning of the propagation. The vector survival rate $s_{\delta t}$ does not influence prevalence. Global sensitivity indices show that p_r is the most influential input factor in disease propagation and acts mainly alone (Figure 12b).

4. Conclusion

We developed a spatially realistic model of Dutch Elm Disease spread in an urban context. The model is a complex network model that explicitly incorporates the actual locations of American elms in the City of Winnipeg. In a sense, it can be likened to a metapopulation model in which each tree would be a patch. The dynamics of the epidemiological state of trees is described using a stochastic process that takes into consideration both roots and beetles infection pathways. Beetle dynamics is modelled using a matrix population model that considers both beetle population growth (deterministic) and its dispersal (stochastic). The overall multi-scale model has three different time scales (week, period and year), making it extremely challenging to analyse theoretically.

Previous studies concerning the spread of DED did not *explicitly* take into account beetle dynamics and movement. In (Harwood et al., 2011), the beetle dispersal range is approximated to provide insights concerning the spread of the disease in the UK during an epidemic in the 1960s but explicit movement of the beetles is not represented. In the present study,

we attempt to provide a deeper understanding of how beetles spread the disease in an urban context by representing more precisely beetle development and movement through the trees of the city. Further, Harwood et al. (2011) omitted estimates of infection through the roots because they consider that infection via beetles was significantly greater than root infection. However, we show here that root infections should be considered to better understand disease spreads.

Based on the simulations and sensitivity analyses, it appears that at the early stage of infection of an entirely susceptible neighbourhood, preventing DED transmission through the roots would be the most efficient control measure in those situations where root grafting would be considered highly likely. In a neighbourhood where the disease is already established, beetle control would be more beneficial. However, as indicated above, our approach has some limitations. In order to address these limitations, it would be prudent to expand the present model to incorporate disease management practices. This would necessitate the investigation of the effects of sanitation and beetle control to more accurately predict spread rates. It would also be important to gather more information about two specific issues: the number and location of trees on private land and the number and location of removed DED infected trees, which would allow improved calibration of the model. The present study only takes into account the trees inventoried in the Winnipeg Open Data portal, which are the trees on public City of Winnipeg land. This represents roughly 56,000 American elms. However, American elms are also located on private property in the study neighbourhoods and their location and diameter at breast height cannot be accessed from the public database. The percentage of such trees was considered to be small, but nonetheless, their absence from the database means that this study does not consider interactions of private and public elms.

These interactions are likely to further densify the root network.

The present simulation framework, most of which we are making available at `https://github.com/julien-arino/DED-Ecological-Modelling-code`, could be adapted to other cities and tree diseases or pests, provided that an accurate and up to date tree inventory were available. For instance, the recent introduction of the emerald ash borer into urban forests across eastern North America could be studied using the same general framework, with adaptations of the model to account for differences between the life cycles of *Hylurgopinus rufipes* and *Agrilus planipennis*.

Acknowledgements. A preliminary version of this work was carried out with Matthew Murphy, now a student at the University of Toronto. Further work was carried out by Vladimir Nosov (currently at McMaster University). We acknowledge discussions with Martha Barwinski and Kerienne La France, City of Winnipeg Urban Forestry Branch. JA and SP are supported in part by NSERC.

Appendix A. Tree height as a function of DBH

We use data collected by one of the authors (RW) to obtain a relation between diameter at breast height and tree height. Data on 787 American elms in Winnipeg is shown in Figure A.13. Using nonlinear regression, we find that the relation can be approximated as

$$h(T) = 4.208 \log(\varnothing_{bh}(T)) - 1.707, \tag{A.1}$$

where $T \in \mathcal{T}$ is a tree, $\varnothing_{bh}(T)$ its diameter at breast height (in centimetres) and $h(T)$ its height (in metres). Tree height is then used to determine the extent of the root system.

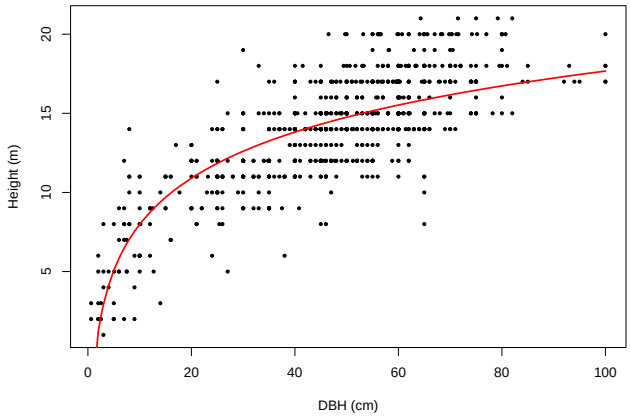


Figure A.13: Regression curve (red) for the relationship between diameter at breast height and tree height for 787 American elms in Winnipeg.

Appendix B. Description of graph measures

For the reader's convenience, we explain here some of the measures of graph topology used in this work. As much as possible, the vocabulary of graph/network theory is recast in the context of the root system networks considered in this paper.

We consider two root networks in each neighbourhood. The first one, which we call $1h$ for short, is the network that is generated if one considers the subgraph of $\mathcal{N}^R$ with edge weights equal to p_R, with p_R taken equal to 1. In other words, this is how trees are connected through their roots if the extent of a tree's root system is the same as its height. The other network, $3h$, is $\mathcal{N}^R$, the network obtained when using all trees satisfying (2), that is, the network describing the maximum extent of the "reach" of trees before the probabilities of infection through the root system vanishes.

The *degree* of a vertex in a graph is the number of vertices it is directly connected to. So, for trees, this is the number of other trees that a given tree is in contact with through roots and can therefore infect.

Strength is closely related to the degree and actually matches exactly in the case of an unweighted graph such as $1h$. Indeed, rather than consider the number of trees a given tree's root system is connected to, strength considers the weights of these connections (which are 1 in the case of $1h$). Thus, strength sums the probabilities that a given tree can infect its neighbours through their root systems. Note that the result is not a probability since probability of infection through roots is not normalised.

A *connected component* is a group of trees in which each tree has access to all other trees, directly or not. A connected component is thus a group of trees which, if left to itself in the presence of one pathogen-bearing tree in its midst, would end up completely contaminated.

Eccentricity, diameter and betweenness further help understand the organisation of the connected components.

Eccentricity is a tree-specific measure. Consider a tree $T \in \mathcal{T}$. The (geodesic) *shortest distance* to another tree $T' \in \mathcal{T}$ is the minimum number of root systems (i.e., edges) that must be traversed to get from T to T'. For instance, if T and T' are directly connected through their root systems, then the shortest distance between them is one, whereas if T and T' are not directly connected but are both connected to a third tree $T'' \in \mathcal{T}$, then the shortest distance between them is 2. Clearly, if two trees are in two different connected components, then there is no route from one to the other and we say that the shortest distance is ∞. Shortest distances can also be *weighted*; the mechanism is the same, but instead of returning the number of "jumps" needed to go from one tree to another it is connected to, one sums the weights of the edges involved. The eccentricity of a tree is then the largest finite shortest distance from that tree. In other words, in terms of infection through root systems, eccentricity describes how "far away" the infection can travel in a

connected component. The *diameter* of a graph is the maximum finite eccentricity for all trees in this graph.

Betweenness is a tree-centric measure that describes how "central" a tree is (it is sometimes also called *betweenness centrality*). The betweenness of a tree $T \in \mathcal{T}$ is the number of shortest paths going through T, excluding those shortest paths in which T is an endpoint. Thus a tree at the "end" of a line has a betweenness of 0, while one that is quite central has high betweenness.

The last three indicators are obtained by drilling down further into connected components. Indeed, some measures that apply at the network level fail to pick up on characteristics of connected components. We therefore compute these measures for each connected component. *Density* is a measure of how well connected trees are to one another. If all trees were connected to every other tree in the neighbourhood, we would say the graph is *complete*. Density is then the "percentage of completeness" of the graph, i.e., the percentage of the possible connections that is present.

Appendix C. Life and epidemiological dynamics of elms

Three mechanisms govern elm dynamics: ageing, infection via beetles and infection via roots. In order to describe these mechanisms, the submodels for ageing and beetle-driven infection on the one hand and for root-driven infection on the other, are discrete-time Markov chains in the absence of the other process. Denote these processes $\{X_t\}$ and $\{Y_t\}$, respectively. A selection process $\{Z_t\}$ is then used at each time step Δt to decide on the actual tree status, which then used as the initial state of both $\{X_t\}$ and $\{Y_t\}$. All the processes are defined on the same state space $\mathcal{S} = \{S_H, S_W, I_W, S_D, I_D\}$.

Appendix C.1. Markov chain for ageing and beetle-driven infection

In the absence of root-driven infection, the processes of ageing and beetle-driven infection are described, for a given tree $T \in \mathcal{T}$, by a single Markov chain $\{X_t^T\}$ defined by the transition matrix $\mathbf{T}^T_{\{a,b\}}(t)$ representing the probabilities to switch from one state to another given by

$$\mathbf{T}^T_{\{a,b\}}(t) = \begin{pmatrix} (1-p_a)\left(1-p_b(B^T(t))\right) & 0 & 0 & 0 & 0 \\ p_a\left(1-p_b(B^T(t))\right) & (1-p_a)\left(1-p_b(B^T(t))\right) & 0 & 0 & 0 \\ p_b\left(B^T(t)\right) & p_b\left(B^T(t)\right) & 0 & 0 & 0 \\ 0 & p_a\left(1-p_b(B^T(t))\right) & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}. \tag{C.1}$$

In the latter matrix, states are ordered as in the state space $\mathcal{S}$, with the (i,j) entry representing the probability of switching from state $\mathcal{S}_j$ to state $\mathcal{S}_i$.

The probability of ageing is a constant p_a and affects only states S_H and S_W. By our modelling assumptions, infection by beetles only occurs when a spore-carrying beetle feeds at the canopy of a susceptible healthy or stressed tree, i.e., trees in states S_H and S_W (Webber et al., 1984). Each spore-carrying beetle that arrives in a tree $T \in \mathcal{T}$ has a probability p_i to infect this tree; this probability is the same for all trees. The probability of becoming infected via beetles is $p_b(B^T(t))$ for a tree $T \in \mathcal{T}$ at state S_H or S_W, where $B^T(t)$ is the cumulative number of spore-carrying beetles having travelled to tree T between t and $t+\Delta t$. A binomial distribution $\mathcal{B}(B^T(t), p_i)$ is used to decide if tree T becomes infected at $t+\Delta t$. Then, the probability $p_b(B^T(t))$ of becoming infected at time $t+\Delta t$ is the probability that at least one spore-carrying beetle successfully infects the tree during $[t, t+\Delta t]$ so that

$$p_b(B^T(t)) = 1-(1-p_i)^{B^T(t)}. \tag{C.2}$$

Parameter values are given in Table C.3. Furthermore, it is observed that an infected tree dies within a year, giving that the probability to switch from I_W to I_D is equal to one. Finally, note that states S_D and I_D are absorbing.

Hence, for the urban forest as a whole in the absence of root-driven infection, the process of ageing and beetle-driven infection is described by a Markov chain $\{X_t\}$ with transition matrix taking the form

$$\mathbf{T}_{\{a,b\}}(t) = \bigoplus_{T \in \mathcal{T}} \mathbf{T}^T_{\{a,b\}}(t).$$

Appendix C.2. Markov chain for root-driven infection

In the absence of tree ageing and beetle-driven infection, the evolution of infection of a tree $T \in \mathcal{T}$ via the root system is driven by a Markov chain $\{Y^T_t\}$.

Let $T \in \mathcal{T}$ be a susceptible tree, i.e., a tree in state S_H, S_W or S_D. Let $\mathcal{D}^R_{I_D}(T) \subseteq \mathcal{D}^R(T)$ be the set of dead infected trees (I_D) in $\mathcal{D}^R(T)$. We assume that infection via the roots of trees in state I_W is not possible, since these trees will become I_D the next year but are not yet infectious. In order to determine if T becomes infected via the root network of one of its infected neighbours $T_1, \ldots, T_k \in \mathcal{D}^R_{I_D}(T)$, we use a Poisson binomial distribution (Wang, 1993) with parameters $p_{\{r\}}(T, T_1), \ldots, p_{\{r\}}(T, T_k)$, i.e., the weights given by (3) of the edges in $\mathcal{N}^R$ between T and all its infected neighbours.

We then have

$$\mathbb{P}(Y^T_{t+\Delta t} = \{I_W, I_D\} | Y^T_t = \{S_H, S_W, S_D\}) = \sum_{A \in \mathcal{F}_x} \prod_{T' \in A} p_{\{r\}}(T, T') \prod_{T'' \in A^c} \big(1 - p_{\{r\}}(T, T'')\big) \quad \text{(C.3)}$$

599 and 0 otherwise. Here, $\mathcal{F}_x$ is defined as

$$\mathcal{F}_x = \{A : A \subseteq \mathcal{D}^R_{I_D}(T), |A| = x\},$$

600 with $k = |\mathcal{D}^R_{I_D}(T)|$, $x \in \{0, \ldots, k\}$ and A^c the complement of A.

601 To illustrate the role of (C.3), suppose that a tree $T \in \mathcal{T}$ has three infected trees it is
602 connected to in the weighted pruned network $\mathcal{N}^R$. Then $\mathcal{D}^R_{I_D}(T) = \{T_1, T_2, T_3\}$ and

$$\mathcal{F}_0 = \emptyset, \quad \mathcal{F}_1 = \{\{T_1\}, \{T_2\}, \{T_3\}\},$$

603

$$\mathcal{F}_2 = \{\{T_1, T_2\}, \{T_1, T_3\}, \{T_2, T_3\}\} \quad \text{and} \quad \mathcal{F}_3 = \{T_1, T_2, T_3\}.$$

604 By considering each of these sets, the Poisson binomial (C.3) sums the probabilities that no
605 infection occurs ($\mathcal{F}_0$), or T is infected by exactly one of its neighbours ($\mathcal{F}_1$), or T is infected
606 by exactly two of its neighbours ($\mathcal{F}_2$) or T is infected by all its neighbours ($\mathcal{F}_3$).

607 Hence, the second Markov chain $\{Y_t\}$ representing the root-driven infection is defined by
608 the transition matrix $\mathbf{T}_{\{r\}}$ built using (C.3) for all trees in $\mathcal{T}$.

Parameter	Description (unit)	Range	Source
p_R	Max. prob. infection by roots	$[0, 1]$	
p_i	Prob. infection by beetle	$[0.01, 0.05]$	(Webber et al., 1984)
p_a	Prob. of ageing	0.01	
R_B	Maximum beetle dispersal distance (m)	$[20, 380]$	(Pines and Westwood, 2008)

Table C.3: Parameters used for tree dynamics.

609 *Appendix C.3. Selection process*

610 Define an order $\prec$ on the state space $\{S_H, S_W, I_W, S_D, I_D\}$ by

$$S_H \prec S_W \prec I_W \prec S_D \prec I_D. \tag{C.4}$$

611 The overall stochastic process $\{Z_t^T\}$ governing the dynamics of a tree $T \in \mathcal{T}$ is defined by
612 combining the processes for ageing and beetle-driven infection and root-driven infection as

$$Z_{t+\Delta t}^T = \begin{cases} \max\limits_{\prec}(X_{t+\Delta t}^T, Y_{t+\Delta t}^T) & \text{if } (X_{t+\Delta t}^T, Y_{t+\Delta t}^T) \neq (S_D, I_W) \\ I_D & \text{if } (X_{t+\Delta t}^T, Y_{t+\Delta t}^T) = (S_D, I_W). \end{cases} \tag{C.5}$$

613 This overall stochastic process thus selects, at each time step Δt, for each tree, the "worst"
614 possible outcome in terms of ageing and beetle-driven infection and root-driven infection.

To summarise, when both ageing and beetle- and root-driven processes are considered simultaneously, we can caricature the processes as operating as follows:

$$X_{t+\Delta t} = g_X(Z_t)$$

$$Y_{t+\Delta t} = g_Y(Z_t)$$

$$Z_{t+\Delta t} = g_Z(X_{t+\Delta t}, Y_{t+\Delta t}).$$

615 Thus, in the description of $\{X_t\}$ in Appendix C.1 and $\{Y_t\}$ in Appendix C.2, the states
616 $X_{t+\Delta t}$ and $Y_{t+\Delta t}$ depend on Z_t instead of X_t and Y_t, respectively. As all processes operate
617 on the same state space $\mathcal{S}$, $\{Z_t\}$ is also a Markov process.

Appendix D. Components of beetles dynamics

The transition matrix for beetles from tree $T \in \mathcal{T}$ is given by

$$\mathbf{S}^T(t, Z_t^T) = \begin{pmatrix} s_J & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ s_{J_P J} & s_{J_P} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ s_{FJ} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & s_{FJ_P} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & s_{CF} & 0 & s_C & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & s_{CF} & 0 & s_C & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & s_{MC} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & s_{MC} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & s_{AM} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & s_{AM} & 0 & 0 \end{pmatrix}, \tag{D.1}$$

where all entries depend on t and Z_t^T. Over a year, the four seasons depending on temperatures can be also defined in term of main events in the beetle dynamics : overwinter (winter), emergence (spring), breeding (summer) and new generation (fall) (Figure D.14). Combining these four periods and five tree states yield twenty different versions of the matrix $\mathbf{S}^T(t, Z_t^T)$ whose positive entries are given in Table D.4. Their twenty associated life cycles are given in Figure D.15.

The fertility matrix $\mathbf{P}^T$ consists mostly of zeros, with only the penultimate and last entries on row 1 being f_{JA} and corresponding to births into J from, respectively, beetles in states A and A_P. It is independent of t and Z_t^T, since this dependence is incorporated into

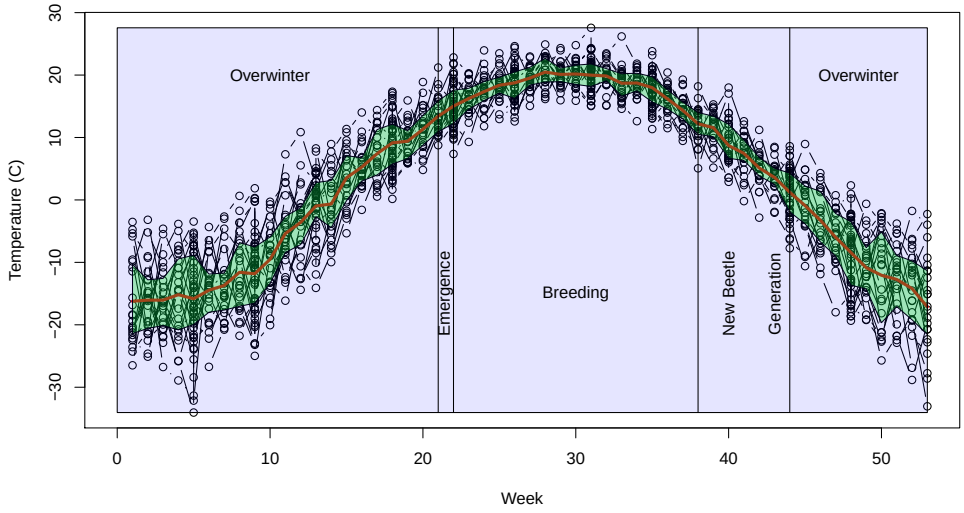


Figure D.14: Average weekly temperatures in Winnipeg since 1995. The red curve is the mean weekly temperature over all years and the green area shows the inter-quartile range. The boxes represent the succession of periods used in the model.

the scalar function $\beta(t, Z_t^T)$,

$$\beta(t, Z_t^T) = \begin{cases} 0, & t \notin \text{Breeding}, \\ 1, & t \in \text{Breeding and } Z_t^T \neq S_H. \end{cases} \tag{D.2}$$

More details on the beetle simplified life cycle are now provided.

Overwinter (30 weeks) During the winter period, only callow adult beetles can survive inside healthy trees. So, $s_C > 0$ for S_H trees. Otherwise, beetles die when overwintering.

Emergence (1 week) After the winter period, callow adults C or C_P emerge from the tree trunks. They might remain callow adults ($s_C > 0$) or mature to become M or M_P ($s_{MC} > 0$). However; as the emergence period only lasts one week; the newly M and M_P will only fly to new trees in the next period.

Period/Tree	s_J	s_{J_PJ}	s_{FJ}	s_{J_P}	$s_{F_PJ_P}$	s_{CF}	s_C	s_{MC}	s_{AM}	β
W/S_H	0	0	0	0	0	0	**+**	0	0	0
W/"All\S_H"	0	0	0	0	0	0	0	0	0	0
E/S_H	0	0	0	0	0	0	**+**	**+**	0	0
E/S_W	0	0	0	0	0	0	**+**	**+**	0	0
E/I_W	0	0	0	0	0	0	**+**	**+**	0	0
E/S_D	0	0	0	0	0	0	0	0	0	0
E/I_D	0	0	0	0	0	0	0	0	0	0
B/S_H	**1**	0	0	0	0	0	0	**1**	**1**	**1**
B/S_W	**1**	0	0	0	0	0	0	**1**	**1**	**1**
B/I_W	**+**	**+**	0	**1**	0	0	0	**1**	**1**	**1**
B/S_D	**1**	0	0	0	0	0	0	0	**1**	**1**
B/I_D	0	**1**	0	**1**	0	0	0	0	**1**	**1**
N/S_H	**+**	0	**+**	0	0	**1**	**+**	0	0	0
N/S_W	**+**	0	**+**	0	0	**1**	**+**	0	0	0
N/I_W	**+**	**+**	**+**	**+**	**+**	**1**	**+**	0	0	0
N/S_D	**+**	0	**+**	0	0	0	0	0	0	0
N/I_D	0	**1**	0	**+**	**+**	0	0	0	0	0

Table D.4: Type of values of the survivals in the matrix $\mathbf{S}^T(t, Z_t)$. **+** indicates the value is positive, while **1** indicates the value is 1. In the first column, Period/Tree type, the following abbreviations are used: W(inter), E(mergence), B(reeding) and N(ew generation). Column β indicates if the fecundity may occur or not.

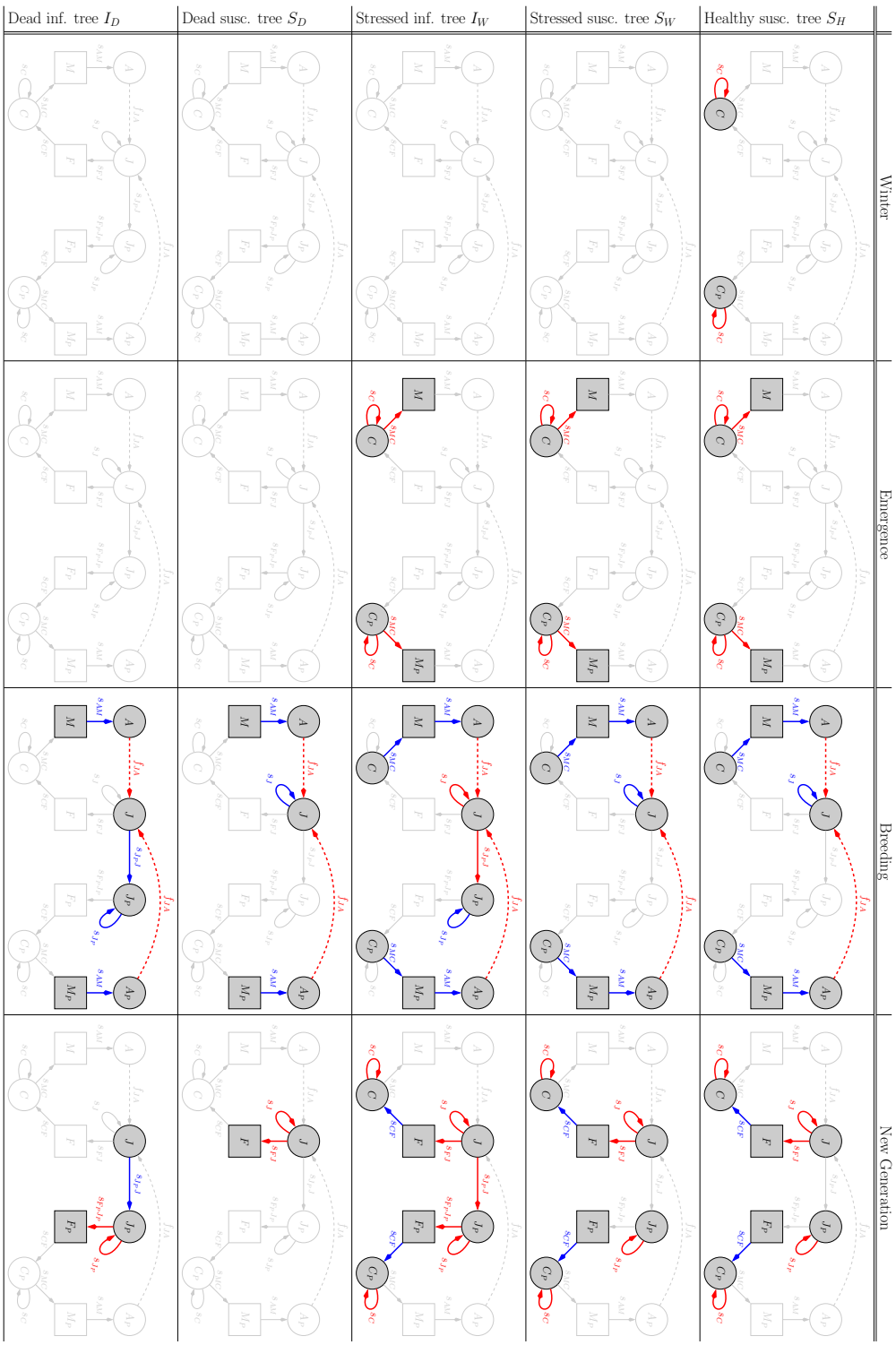


Figure D.15: Dependencies of beetle life cycle on tree states and periods.

Parameter	Description	Value
	Survival rates	
$s_{\delta t}$	Beetle survival rate	$[0.97, 0.99]$
s_J	Stays J	B:0.95/N(S):0.8/N(I_W):0.75
$s_{J_P J}$	Becomes spore-carrying J_P	N(I_D):0.8/O:0.05
s_{J_P}	Stays J_P	B:0.8/N:0.75
s_{FJ}	Becomes disperser F	0.2
$s_{F_P J_P}$	Becomes disperser F_P	0.25
s_C	Stays C	W:0.99/O:0.85
s_{MC}	Becomes disperser M	0.15
	Fertility	
f_{JA}	Female progeny of an adult female (number of eggs)	30

Table D.5: Beetle dynamics parameters. Values of the survival rates in the matrix $\mathbf{S}^T(t, Z_t)$ are those used when parameters are not equal to 0 or 1; the latter are listed in Table D.4. When parameters depend on the period or the type of trees, the period is indicated first then, where relevant, the type of tree is indicated between parentheses. The following abbreviations are used: W(inter), E(mergence), B(reeding), N(ew generation) and O(ther periods). The notation S means all susceptible trees (S_H, S_W and S_D). If "O" is specified, this means that the associated value is for all other periods and for all type of trees.

These processes only occur on healthy (S_H) and stressed (S_W, I_W) trees. Note that callow adults overwinter only in healthy trees; however, in the model, they might

emerge from stressed trees too. This is a feature of the model as the tree state update occurs during the last week of winter just before the emergence week.

Breeding (16 weeks) In the breeding period, beetles disperse and look for one tree to mate (any type of tree).

At the beginning of the breeding period, if beetles are still callow adults (in a healthy or stressed tree S_H, S_W, I_W) then they are forced to mature and become dispersal M or M_P ($s_{MC} = 1$). Then, dispersing individuals look for a tree (S_H, S_W, S_D, I_W or I_D trees). If they survive the travel (via (1)) and arrive at destination, then they become adults A or A_P. The following week, adults lay f_{JA} eggs (described as juveniles J in the model) and then die.

Juvenile beetles are not carrier of the pathogen so they are all born as J. Then, they may become spore-carrier J_P in function of the tree in which they grow. If the tree is I_D, then juveniles J become J_P with a rate $s_{J_PJ} = 1$. If the tree is I_W, then juveniles become J_P with a rate $s_{J_PJ} < 1$. Beetles stay spore-free if the tree is not infected (S_H, S_W or S_D). Juveniles J or J_P stay at this life stage until the end of the breeding season.

New generation (6 weeks) In the new generation period, juveniles J and J_P can emerge from trees (S_H, S_W, S_D, I_W or I_D) and fly to feed at the canopy of healthy or stressed trees (S_H, S_W or I_W). Thus, juveniles become F or F_P ($s_{FJ} > 0$) and then fly to a destination tree with a survival probability as defined in (1). Once arrived, beetles become callow adults C or C_P at rate $s_{CF} = 1$ and stay at this stage until the end of the period.

References

Anderson, P., 1996. Overwintering behavior of the native elm bark beetle, Hylurgopinus rufipes (Eichoff)(Coleoptera: Scolytidae). Master's thesis. University of Manitoba.

Anderson, P., Holliday, N., 2003. Distribution and survival of overwintering adults of the dutch elm disease vector, hylurgopinus rufipes (coleoptera: Scolytidae), in american elm trees in manitoba. Agricultural and Forest Entomology 5, 137–144.

Barwinsky, M., 2016. Canadian urban forest network: Prairies region update. URL: https://treecanada.ca/blog/canadian-urban-forest-network-prairies-region-update/.

Bidot, C., Lamboni, M., Monod, H., 2018. multisensi: Multivariate Sensitivity Analysis. URL: https://CRAN.R-project.org/package=multisensi. r package version 2.1-1.

Brasier, C., Gibbs, J., 1978. Origin and development of the current Dutch elm disease epidemic. Blackwell, Oxford, UK.

Caswell, H., 2001. Matrix Population Models. Second ed., Sinauer.

City of Winnipeg, 2018. URL: https://winnipeg.ca/sustainability/PublicEngagement/ClimateActionPlan/pdfs/WinnipegsClimateActionPlan.pdf.

City of Winnipeg, 2019. URL: https://data.winnipeg.ca/Parks/Tree-Inventory/hfwk-jp4h.

Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. InterJournal Complex Systems, 1695. URL: http://igraph.org.

Dreistadt, S.H., Dahlsten, D.L., Frankie, G.W., 1990. Urban forests and insect ecology. BioScience 40, 192–198.

Dwyer, J.F., McPherson, E.G., Schroeder, H.W., Rowntree, R.A., 1992. Assessing the benefits and costs of the urban forest. Journal of Arboriculture 18, 227–227.

Gardiner, L., 1981. Seasonal activity of the native elm bark beetle, hylurgopinus rufipes, in central ontario (coleoptera: Scolytidae). The Canadian Entomologist 113, 341–348.

Gibbs, J.N., 1978. Intercontinental epidemiology of dutch elm disease. Annual Review of Phytopathology 16, 287–307.

Grote, R., Samson, R., Alonso, R., Amorim, J.H., Cariñanos, P., Churkina, G., Fares, S., Thiec, D.L.,

689 Niinemets, Ü., Mikkelsen, T.N., et al., 2016. Functional traits of urban trees: air pollution mitigation
690 potential. Frontiers in Ecology and the Environment 14, 543–550.
691 Harris, R., Clark, J., Matheny, N., 1999. Integrated Management of Landscape Trees, Shrubs, and Vines.
692 Pearson College Div.
693 Harwood, T., Tomlinson, I., Potter, C., Knight, J., 2011. Dutch elm disease revisited: past, present and
694 future management in great britain. Plant Pathology 60, 545–555.
695 Hildahl, V., Jeffrey, C., 1980. The elm's enemy. winnipeg, manitoba: Department of natural resources.
696 Forestry Branch 5.
697 Hiratsuka, Y., et al., 1987. Forest tree diseases of the prairie provinces. Canadian Forest Service.
698 Hubbes, M., 1988. Pathogen virulence and host reaction in dutch elm disease. Nat Can 115, 157–162.
699 Hubbes, M., 1999. The american elm and dutch elm disease. The Forestry Chronicle 75, 265–273.
700 Hubbes, M., Jeng, R., 1981. Aggressiveness of ceratocystis ulmi strains and induction of resistance in ulmus
701 americana. European Journal of Forest Pathology 11, 257–264.
702 Jorgensen, E., 1974. Towards an urban forestry concept. Environment Canada.
703 Kaston, B., 1939. Native elm bark beetle, Hylurgopinus rufipes (Eichhoff), in Connecticut. Technical Report
704 420. Connecticut Agricultural Experiment Station.
705 Kondo, E.S., Hiratsuka, Y., Denyer, W., et al. (Eds.), 1981. Proceedings, Dutch Elm Disease Symposium
706 and Workshop, Manitoba Dept. of Natural Resources, Forest Protection and Dutch Elm Disease.
707 Lanier, G., Kondo, E., Hirastsuka, Y., Denyer, W., 1982. Behavior-modifying chemicals in dutch elm disease
708 vector control, in: Proceedings Dutch Elm Disease Symposium and Workshop, p. 371–394.
709 Livesley, S., McPherson, E., Calfapietra, C., 2016. The urban forest and ecosystem services: impacts on
710 urban water, heat, and pollution cycles at the tree, street, and city scale. Journal of environmental quality
711 45, 119–124.
712 Miller, R.W., Hauer, R.J., Werner, L.P., 2015. Urban forestry: planning and managing urban greenspaces.
713 Waveland press.
714 Oghiakhe, S., 2014. Biology and Management of the Dutch Elm Disease Vector, *Hylurgopinus rufipes* Eichhoff
715 (Coleoptera: Curculionidae) in Manitoba. Master's thesis. University of Manitoba.

716 Pines, I., Westwood, A., 1996. Evaluation of monosodium methane arsenate for the suppression of native
717 elm bark beetles, hylurgopinus rufipes (eichhoff)(coleoptera: Scolytidae). The Canadian Entomologist
718 128, 435–441.
719 Pines, I., Westwood, R., 2008. A mark-recapture technique for the dutch elm disease vector the native elm
720 bark beetle, hylurgopinus rufipes (coleoptera: Scolytidae). Arboriculture and Urban Forestry 34, 116.
721 Rioux, D., 2003. Dutch elm disease in Canada: Distribution, impact on urban areas and research. Technical
722 Report. Food and Agriculture Organization, World Forestry Congress XII.
723 Sarre, P., 1978. The diffusion of dutch elm disease. area , 81–85.
724 Shirley, M.D., Rushton, S.P., 2005. The impacts of network topology on disease spread. Ecological Com-
725 plexity 2, 287–299.
726 Sinclair, W.A., Campana, R., 1978. Dutch elm disease. perspectives after 60 years. Search Agric. 8, 1–52.
727 Soll, D., 2016. Dutch elm in st. paul and minneapolis: A tale of two cities. Minnesota History 65, 44–53.
728 Stipes, R., Campana, R.J., 1981. Compendium of elm diseases. Technical Report. APS Press,.
729 Strobel, G.A., Lanier, G.N., 1981. Dutch elm disease. Scientific American 245, 56–67.
730 Swedenborg, P.D., Jones, R.L., Ascerno, M.E., Landwehr, V.R., 1988. Hylurgopinus rufipes (eich-
731 hoff)(coleoptera: Scolytidae): attraction to broodwood, host colonization behavior, and seasonal activity
732 in central minnesota. The Canadian Entomologist 120, 1041–1050.
733 Swinton, J., Gilligan, C., 1999. Selecting hyperparasites for biocontrol of dutch elm disease. Proceedings of
734 the Royal Society of London. Series B: Biological Sciences 266, 437–445.
735 Swinton, J., Gilligan, C.A., 1996. Dutch elm disease and the future of the elm in the uk: a quantitative
736 analysis. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 351,
737 605–615.
738 Swinton, J., Gilligan, C.A., 2000. A modelling approach to the epidemiology of dutch elm disease, in: The
739 Elms. Springer, pp. 73–101.
740 Thompson, H., Matthysse, J., 1972. Role of the native elm bark beetle, hylurgopinus rufipes (eichh.), in
741 transmission of the dutch elm disease pathogen, ceratocystis lmi (buisman) c. moreau. Search .
742 Wang, Y.H., 1993. On the number of successes in independent trials. Statistica Sinica , 295–312.

Webber, J., Brasier, C., et al., 1984. The transmission of dutch elm disease: a study of the process involved. Invertebrate-microbial interactions , 271–306.

Westwood, A., 1991. A cost benefit analysis of manitoba's integrated dutch elm disease management program 1975-1990. Proceedings of the Entomological Society of Manitoba 47, 44–59.

Whitten, R., 1964. Elm bark beetles. Technical Report. U.S. Department of Agriculture, Forest Service, Division of Forest Insect Research Leaflet 185.