A Cellular Automata Model of Pine Wilt Disease Spread and Its Simulation Tool for Application

Mingxiang HUANG

Information Center of Ministry of Environmental Protection
of China
Beijing 10029China, 86-1086445168
Email:huang.mingxinag@mep.gov.cn

Chaohui YU

Information Center of Ministry of Environmental Protection
of China
Beijing 10029China

Dali ZHOU

Information Center of Ministry of Environmental Protection
of China
Beijing 10029China

Fang BEI

College of Surveying and Geoinformatics, Tongji University
Shanghai 200092 China

Abstract—In the study a cellular automata (CA) was applied to model the spread process of pine wilt disease (PWD), and a PWD-CASim tool was developed to simulate PWD spread based on the CA model using VC++6.0 platform. For the further understanding of PWD spread with the CA model, a case study was implemented to simulate the impacts of neighborhood size and deforestation rate on the disease spread. The simulation results of the case study show that PWD spread is very strong when the range of neighborhood is up to 90 m, and deforestation rate of 60% maybe is the better choice to balance the management cost and disease diffusion prevention. The study results also present that the CA is another effective way to model PWD spread.

Keywords- cellular automata; pine wilt disease; spread; model; Simulation

I. INTRODUCTION

An epidemic of pine wilt disease (PWD) is caused by the pinewood nematode *Bursaphelenchus xylophilus* which is vectored by with the pine sawyer *Monochamus alternatus*. Due to their mutualistic relationship, according to which the nematode weakens and makes trees available for beetle reproduction and the beetle in turn carries and transmits the nematode to healthy pine trees [1]. PWD has killed huge numbers of pine trees every year in the East Asian counties.

In the modeling of the spread of PWD, the previous studies mainly covered two types, one focusing on the modeling of population dynamics of the vector beetle and spatial expansion of the disease using an integro-difference equation with a dispersal kernel [2,3], the other focusing on the modeling of mechanistic interactions at the individual level [1,4]. Yoshimura present a mathematical model to describe the host-vector interaction between pines and pine sawyers carrying nematodes, on the basis of detailed data on the population dynamics of pine sawyers and the incidence of pine wilt disease at a study site located on the northwest coast of Japan [5]. Evans used current understanding of plant host interactions as the basis for assumptions incorporated into a mechanistic

model describing relevant process dynamics [6]. Takasu constructed an individual-based model based on mechanistic interactions at the individual level, and discussed the analysis of a realistic individual-based model for understanding better the mechanistic background of the Allee effect in pine wilt disease [1].

Nowadays, the mathematical models based on discrete systems, such as stochastic interacting particle models, cellular automata models have been widely used [7-10]. These simple models are especially suitable for computer simulations. Especially the cellular automata (CA) are simple models of computation capable to simulate complex physical, biological or environmental phenomena. A two-dimensional CA is formed by a two-dimensional array of identical objects called cells, which are endowed with a state that change in discrete steps of time according to a specific rule. As the CA evolves, the updated function (whose variables are the states of the neighbors) determines how local interactions can influence the global behavior of the system [11].

The objectives of this study are: (i) to construct a CA model of PWD according to the CA components; (ii) to develop a PWD-CASim tool for visible simulation of PWD spread based on the constructed CA model; (iii) to simulate the spread of PWD in different neighborhood sizes and deforestation rates and reveal the their impacts on PWD spread.

II. AN OVERVIEW ON CA

Cellular automata are dynamical systems which are discrete in space and time. They operate on a uniform, regular lattice and are characterized by local interaction. They are generally attributed to Ulam [12] and von Neumann[13], who introduced the concept in the late forties to provide a realistic model for the behavior of complex systems. They used discrete models to study computability theory, mathematics, and biological systems [14]. Formally, a CA can be defined as a 4-uplet. $A = (Z, S, N, f)$, where:

Z is the finite or infinite lattice, S is a finite set of cell states or values, N is the finite neighborhood, f is the local transition function defined by the transition table or the rule. The *lattice* is a finite or infinite discrete regular grid of cells on a finite number of dimensions. Each *cell* is defined by its discrete *position* (an integer number for each dimension) and by its discrete *value* (one of a finite set of integers). Time is also discrete. The future state of a cell (time T_1) is a function of the present state (time T_0) of a finite number of cells surrounding the observed cell called the neighborhood.

The state of each cell is an element of a finite state set, S, in such a way that the state of the cell (a, b) at time t is denoted by $s_{ab}^{(t)} \in S$. The matrix $Z^t = s_{ij}^{(t)}$ is called configuration of the CA at time t. Moreover, $Z^{(0)}$ is the initial configuration of CA. The neighborhood of a cell (a, b) is the set of cells whose states at time t determine the state of the cell (a, b) at time t + 1, by means of the local transition function. Depending on the process to be modeled, one can choose an appropriate neighborhood (von Neumann neighborhood, Moore neighborhood, Margolus neighborhood). As is mentioned above, the CA evolves deterministically in discrete time steps, changing the states of the cells by means of the local transition function f. As the cellular space is considered to be finite, boundary conditions must be considered in order to assure a well-defined dynamics of the CA. These boundary conditions depend on the process to be simulated; in this work, we will use reflective boundary.

III. A CA MODEL OF PWD SPREAD

A. Lattice(Z)

Two-dimensional raster type is chosen for the CA lattice for spatially simulating the PWD spread. The raster resolution is 10m × 10m grid to define a cell of the lattice according to the PWD spread distance.

B. Cell state(S)

Based on thematic forestry survey, five vegetation types denoted as $\{S_0, S_1, S_2, S_3, S_4\}$ are define as cell states: the non-forest land (settlements, rivers, and arable land), non-pine type (broad-leaved forest and shrubbery), healthy pine, pine infected by PWD under the incubation period, pine infected by PWD. In addition, the auxiliary variable (infected time: *PDTime*) is applied for recording the time of a cell infected by PWD.

C. Neighborhood(N)

The neighborhood of CA model for simulating PWD spread is Moore and extended Moore types. According to the intermediate hosts' (*Monochamus alternatus*) activity characteristic of PWD spread, the maximum for extended Moore neighborhood defines 11 × 11 cells, means the maximum distance of PWD spread is 110 m.

D. Transition function(f)

[1]. Set yearly time-step for the model, set α between [0~1] for the deforestation rate.

[2]. $S_{i,i}^t$ represents the cellular state at time t in the (i, j) cell, S^{t_0} represents the initial state of cellular space.

[3]. $T_{i,i}^t$ represents infected-time the at time t in the (i, j) cell, the maximum of $T_{i,i}^t$ is 4, that is natural death of the infected pine needs four years. $InT_{i,i}^t$ is incubation time, setting $InT_{i,i}^t \leq 2$. The adjacency matrix M in the (i, j) cell denotes in Eq.1, n is the tier number of neighborhood.

$$M = \begin{cases} S_{i-n,j-n}^t & \cdots & \cdots & \cdots & \cdots & \cdots & S_{i+n,j-n}^t \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & S_{i-1,j-1}^t & S_{i,j-1}^t & S_{i+1,j-1}^t & \cdots & \cdots \\ \cdots & \cdots & S_{i-1,j}^t & S_{i,j}^t & S_{i+1,j}^t & \cdots & \cdots \\ \cdots & \cdots & S_{i-1,j+1}^t & S_{i,j+1}^t & S_{i+1,j+1}^t & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ S_{i-n,j+n}^t & \cdots & \cdots & \cdots & \cdots & \cdots & S_{i+n,j+n}^t \end{cases} \tag{1}$$

[4]. If $S_{i,i}^t = S_0$ or $S_{i,i}^t = S_1$ then $S_{i,i}^{t+t}$ remains unchanged;

[5]. If $S_{i,i}^t = S_2$ then

♦ calculating the conversion probability ($p_{i,i}^t$) for healthy (i, j) cell according to Eq.2.

$$p_{i,j}^t = w_1 \times \frac{C_1}{8 \times 1} + w_2 \times \frac{C_2}{8 \times 2} + \cdots + w_n \times \frac{C_n}{8 \times n} \tag{2}$$

C_n denotes infected cell ($S_{i,j}^t$=3 or 4) number, w_n denotes the weight of the n-tier neighborhood, $8 \times n$ denotes cell number of n-tier neighborhood.

♦ Calculating the healthy cell number (TN_2), and generating TN_2 random number R_2^t between [0 1].

♦ Comparing $p_{i,i}^t$ and R_2^t, if $R_{2(i,i)}^t \geq p_{i,i}^t$ then $S_{i,i}^{t+1} = S_2$ else $S_{i,i}^{t+1} = S_3$ or S_4.

♦ Calculating the infected cell number (TN_{3or4}), and generating TN_{3or4} random number R_{3or4}^t b between [0 1]. If $R_{3or4(i,i)}^t \geq 0.2$ then $S_{i,i}^{t+1} = S_3$ else $S_{i,i}^{t+1} = S_4$.

[6]. If $S_{i,i}^t = S_3$ then

- If $InT_{i,i}^t > 2$ then $S_{i,i}^{t+1} = S_4$ .else calculating the latent infected cell number (TN_3), and generating TN_3 random number R_3^t between [0 1]. If $R_{3(i,i)}^t \geq 0.9$ then $S_{i,i}^{t+1} = S_3$ and $InT_{i,i}^{t+1} = InT_{i,i}^t + 1$ else $S_{i,i}^{t+1} = S_4$.

[7]. If $S_{i,i}^t = S_4$ then

- If the infected pine is cut down randomly then: calculating the infected cell number (TN_4), and generating TN_4 random number R_4^t between [0 1]. If $R_{4(i,i)}^t \leq \alpha$ then $S_{i,i}^{t+1} = S_1$ else $S_{i,i}^{t+1} = S_4$ and $T_{i,i}^{t+1} = T_{i,i}^t + 1$

- If the infected pines with higher density are given more priority to be cut down then: calculating the infected cell number (TN_4), calculating the infected cell density ($\rho(i,j)$) depending on the size of the statistics window, calculating the new deforestation threshold (R_a) based on $\alpha \times TN_4$, if $\rho(i,j) \geq R_a$ then $S_{i,i}^{t+1} = S_1$ else $S_{i,i}^{t+1} = S_4$ and $T_{i,i}^{t+1} = T_{i,i}^t + 1$.

IV. PWD-CASim TOOL DEVELOPMENT AND APPLICATION

A PWD-CASim tool development

Based on the CA model for simulating PWD spread in the previous section, the PWD-CASim tool was developed using Microsoft Visual C + + 6.0 platform. The PWD-CASim tool structure was showed in Figure 1. The PWD-CASim tool includes three parts, CA-Model module, graph module and statistic module. The tool interface was shown in Figure 2. The PWD-CASim tool determines each cell status in each step through CA model real-time calculation, presents real-time visualization of all cells in each step, counts the infected pine number and shows the 2D curve in the statistic window. The specific functions of PWD-CASim tool include:

(1) Cellular status visualization, and dynamic visualization of cellular evolution processes.

(2) Dynamic 2D curve display of PWD-CASim simulation results.

(3) Neighborhood type selection and custom neighborhood weight.

(4) Incubation period selection and its custom parameter.

(5) Custom rate of deforestation.

(6) Random or focused deforestation options, real-time calculation of PWD density according to custom statistics window size.
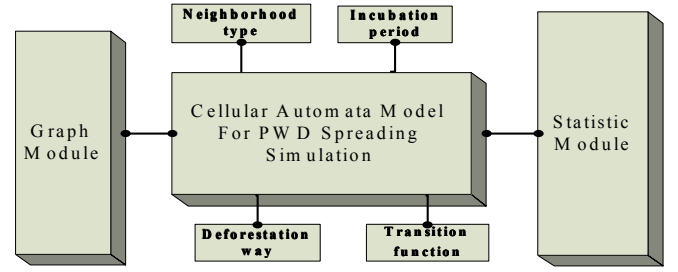
(7) Data output of PWD-CASim simulation results.
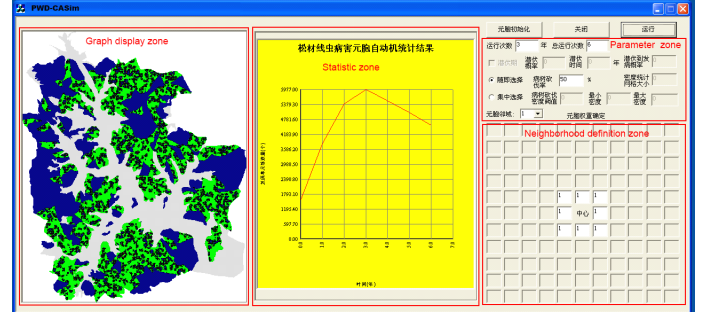

Figure 1.   PWD-CASim tool framework

Figure 2.   PWD-CASim tool interface

B PWD-CASim tool application for PWD spreading simulation

1) Case study area

The case study area is situated at southwest of Sizhoutou town, Xiangshan County, Zhejiang province of China, in which PWD was firstly found in the 1991. In the paper, the forest resource thematic map in Sizhoutou town was applied to define the land use type. According the cellular automata status, the forest resource thematic map was reclassified to non-forest, pine and other forest. Based on the forest pest survey data of the study area in 2000, 1600 pines as the un-health pine were randomly distributed in the pine type, of which 1400 pines for the infected pines, and 200 pines for latent infected pines. Thus the case study area has five types, non-forest, health pine, infected pine (1400 cells), latent infected pines(200 cells), other forest. The study area was re-sampled into 10m × 10m grid for easy to use in the PWD-CASim tool.

2) Simulation of the impact of neighborhood size on PWD spread

The strength of the PWD spread can be evaluated by the impact distance of the infected pine on the neighborhood. In the study, we can find the strength of the PWD spread by comparing the simulation results with different neighborhood sizes adopted by PWD-CASim tool. Five neighborhoods, from Moore 1 (10m for neighborhood) to Moore 5 (50m for neighborhood) neighborhoods, were chosen for neighborhood parameter in PWD-CASim tool, and other parameters remain unchanged. The assumptions of the unchanged parameters include 60% of deforestation rate, random deforestation way, no incubation period, 10 steps. Through separately running the PWD-CASim tool with 5 neighborhoods, the 5 simulation results of PWD spread were showed in Figure. 3. As can be seen from Figure 3, with increasing of the neighborhood size,

the strength of PWD spread is stronger and the infected cells are increasing in general. In particular, the infected cells slowly increased with from Moore 1(10 m) to Moore 3 (30 m). When the neighborhood size was Moore 4(40m), the infected cells significantly increased as compared with Moore 3. This phenomenon means that when the range of the PWD spread reaches 90 meters (Moore 4), the PWD spreads extremely rapid. According to the simulation results, if we can limit the scope of activities of the insect vector, we can greatly control the speed and extent of the PWD spread. From the simulation results, for reducing the PWD spread speed and infected pines, on the one hand, we can make the range decrease of the vector (*Monochamus alternatus*) activity that means the decrease of neighborhood by sprinkle Fenitrothion into the pinewood. On the other hand, The intermediate cutting way for the pinewood, which can increase the gap between pines and lower the forest density, can reduce the spread speed of PWD and make more pines uninfected.
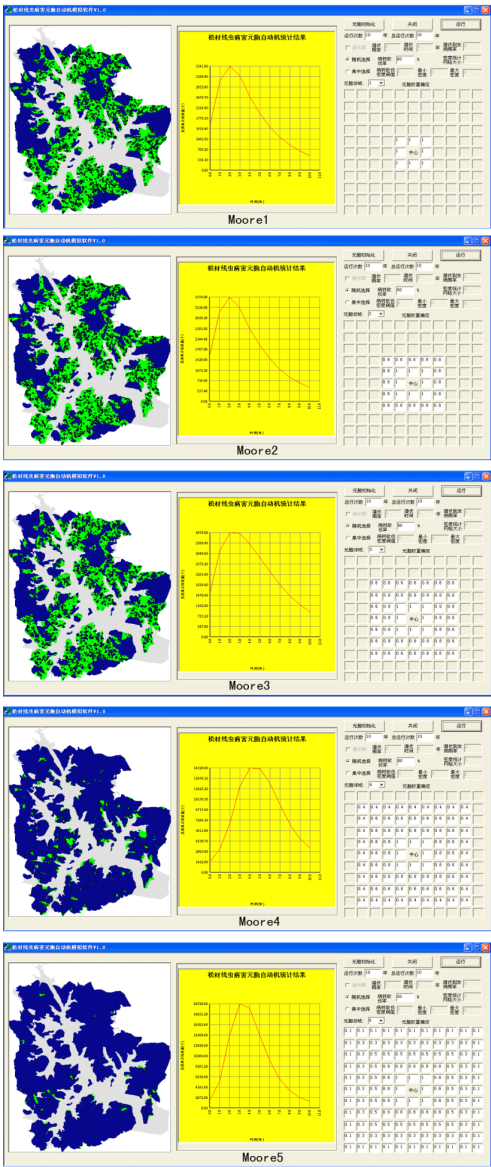

Figure 3.   Comparsion of different neighborhood size for PWD spreading

3) Simulation of the impact of deforestation rate on PWD spreading

The felling of infected pines is an effective way to prevent the diffusion of PWD. How to determine the deforestation rate under which the diffusion of PWD can be controlled and the felling time and cost are lowest. For the determination of the best deforestation rate, in the study the PWD spread processes were separately simulated in the PWD-CASim tool under 10 deforestation rates which are 0%,10%, 20%, 30%,40%,50%, 60%, 70%, 80%, 90%,while other parameters for the simulation remain unchanged. Those unchanged parameters are Moore type for neighborhood, ramdon deforestation way, no incubation period, 10 steps. The simulation results in different deforestation rates were showed in Figure 4. From the Figure 4, if the deforestation rate is less than 40%, the infected cells increase quickly with the increasing of time steps. The simulation results showed that if no considering other prevention and control measurements, the deforestation measurement which fell less than 40% infected pines cannot control the PWD diffusion. If the deforestation rate is up to between 50% and 70%, the infected cells increase in the first few years and decrease in the later years. However the deforestation rate reaches 80% or 90%, the infected pines decrease notably. From the simulation results, the deforestation rate to well balance the felling cost and PWD control should be at 60%. If no caring about the costs of infects pine felling, the deforestation rate of 80% maybe is a good choice.
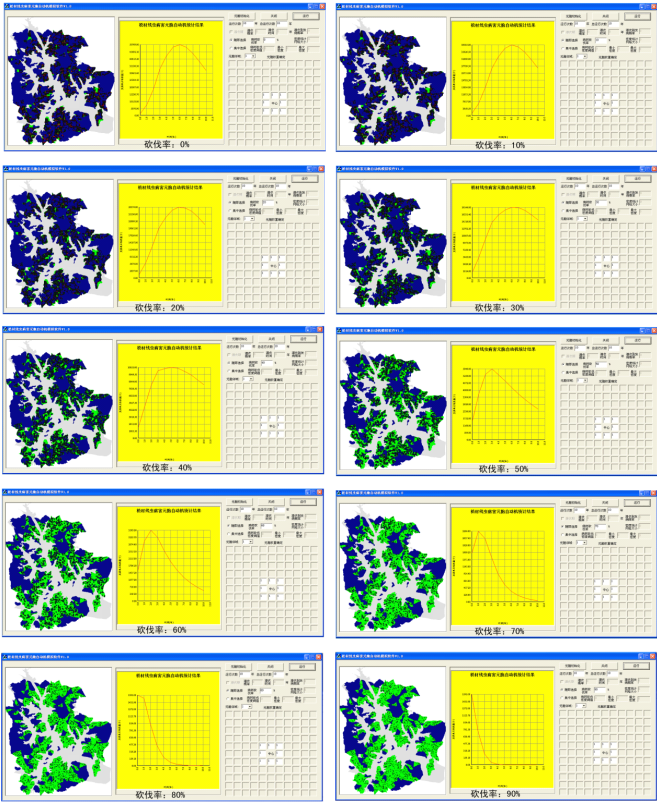

Figure 4.   Comparsion of different deforestation rates for PWD spreading

V. Conclusions

In the study, we constructed a CA model of PWD spread based on the overview of CA model, developed the PWD-CASim tool for simulating disease spread, and implemented the simulation experiment with different neighborhood sizes and deforestation rates to find out their impacts on PWD spread in a case study. From the simulation results of the impact of neighborhood size on PWD spread, the strength of PWD spread can be reflected by the range of neighborhood, and the PWD is a highly infectious if the range reaches 90m. The facts showed that the intermediate cutting way for the pinewood could an effective measurement to prevent the wide diffusion of PWD. The comparison among simulations with the different deforestation rates presented that how to choose the deforestation rate is very important for most effective control of PWD spread. Although the experiment results showed deforestation rate of 60% maybe is the better choice, the best deforestation rate should together consider other facts, such as the cost of PWD control, the special terrain, other measurements for PWD control. In the future work, the CA model of PWD spread should be improved in the transition function, and an agent-based model of *Monochamus alternatus* individual behavior can be developed to integrate with the CA model for more clearly understood on PWD spread process.

References

[1] F. Takasu. Individual-based modeling of the spread of pine wilt disease: vector beetle dispersal and the Allee effect. Population Ecology. 2009, 51(3):399-409.

[2] F. Takasu, N. Yamamoto, K. Kawasaki, K. Togashi, N. Shigesada. Modeling the expansion of an introduced tree disease. Biol Invasions. 2000, 2:141–150.

[3] N. Shigesada, K. Kawasaki. Invasion and the range expansion of species: effects of long-distance dispersal. In: Bullock JM, Kenward RE, Hails RS (eds) Dispersal ecology: the 42nd symposium of the British ecological society, Blackwell, 2002, pp 350–373.

[4] R. Durrett, S. Levin. The importance of being discrete (and spatial). Theor Popul Biol. 1994, 46:363–394.

[5] A. Yoshimura, K. Kawasaki, F. Takasu, K. Togashi, K. Futai, N. Shigesada. Modeling the spread of pine wilt disease caused by nematodes with pine sawyers as vector. Ecology. 1999, 80:1691–1702

[6] S. Evans, H. Evans and M. Ikegami. Modeling PWN-Induced wilt expression: a mechanistic approach. pine wilt disease: a worldwide threat to forest ecosystems. 2008, Part V, pp.259-278.

[7] C. Beauchemin, J. Samuel and J. Tuszynski, A simple cellular automaton model for influenza A viral infections, J. Theor. Biol., 2005, 232: 223 -234.

[8] M. L. Martins, G. Ceotto, S. G. Alves, C. C. B. Bufon, J. M. Silva and F.,F. Laranjeira, A cellular automata model for citrus variegated chlorosis, Physica A, 2001, 295: 42 - 48.

[9] J. Satsuma, R. Willox, A. Ramani, B. Grammaticos and A. S. Carstea, Extending the SIR epidemic model, Physica A, 2004, 336:369–375.

[10] G. Ch. Sirakoulis, I. Karafyllidis and A. Thanailakis. A cellular automaton model for the effects of population movement and vaccination on epidemic propagation, Ecol. Model., 2000, 133:209-223.

[11] S. Hoya White, A. Martín del Rey, G. Rodríguez Sánchez. Modeling epidemics using cellular automata. Applied Mathematics and Computation. 2007. 186(1):193-202.

[12] S. Ulam. Random Processes and Transformations. Proceedings of the International Congress of Mathematicians, 2, 1952, pp 85-87.

[13] J. Von Neumann, Theory of Self-Reproducing Automata, Univ. of Illinois Press. 1966.

[14] S. Wolfram. A new kind of science, Wolfram Media, Champaign, IL, , 2002, pp 48-57.