

# EXPLORING GEOGRAPHIC PARAMETER SPACE WITH A GIS IMPLEMENTATION OF CELLULAR AUTOMATA

*Vivienne Cole & Jochen Albrecht*

Department of Geography  
University of Auckland  
{vcol001, ja}@geog.auckland.ac.nz

Presented at SIRC 99 – The 11<sup>th</sup> Annual Colloquium of the Spatial Information Research Centre  
University of Otago, Dunedin, New Zealand  
December 13-15<sup>th</sup> 1999

## ABSTRACT

Many of the recent studies of vegetation spread and the invasion of introduced species incorporate GIS and some form of dynamic modelling in one- or two-dimensional space. Most work, however, concentrates on presenting a predetermined model in an abstract computer simulation environment. The focus has been on the system's ability to display the model (results), as opposed to its ability to develop or compose a model, acting as a creative tool in an applied situation. The research presented in this paper uses a GIS-based cellular automata (CA) framework to study and create an applied and ecologically significant model of spread for the invasive plant species *Rhamnus alaternus*. Instead of conventionally testing a previously created model, parameters such as generation time, fruiting phenology, fruigivore abundance, reproductive rate and vegetation cover are examined through simulation for their contribution to spread and the resulting spatial pattern. Initial simulation runs focus on the different effects of stochastic versus deterministic parameter values. Further studies examine the effect of different relationships, conditions and interactions between parameters. Upper and lower limits, the interactions between, and the relative importance of each parameter are examined in a number of combinations. Each parameter is then accepted or declined to create a conceptual model outlining the ecologically most significant features of the species. This method effectively examines and tests the ability for models of generic invasion and vegetation dynamics to be transposed into a contextually rich environment through a GIS-CA framework of model creation.

**Keywords and phrases:** biological invasion, dynamic modelling, hierarchical reasoning.

## 1.0 INTRODUCTION

Space and time have always been recognised as crucial components of ecological change (Colasanti & Grime 1993). Important spatial patterns such as aggregated distributions of plant community species (at one or more scales) and juxtaposed neighbourhood competition, has led to continued study of vegetation utilising spatial methodologies (Silvertown *et al.* 1992, Baltzer *et al.* 1998). Population processes such as dioecious reproduction and interspecies competition are affected by the initial spatial arrangement of a species. The history of a species in a landscape is influenced by the spatially and temporally autocorrelated nature of its agents.

Spatially explicit models are “expected to increase our ability to accurately model populations subject to complex processes” (Baltzer *et al.* 1998). The inclusion of space in vegetation models introduces more

complex rules into a simulation than those created from simple parameters. This can lead to unpredictable chaotic simulation and non-linear dynamics (Silvertown *et al.* 1992).

The spatial theory that many population models have been based on is the classic work of Skellam (1951). Skellam created a mathematical model of population dispersal with ideas from the random walk theory of probabilistic random displacement associated with several authors from the early 1900's. The model assumes random dispersal of a species over space, and a constant growth rate as represented by the muskrat invasion of Europe through the 20<sup>th</sup> century.

Many spatially explicit ecological models have focussed on the nature of dispersal based on internal, or population based conditions, such as density or extinction (Molofsky 1994, Tilman 1994). With increased computing power, researchers have continued to focus on technology, using abstract, generalised models of dispersion instead of applied modelling questions. These modelling questions continued to define the external environment as a homogeneous background to internal population dynamics. Variability over space however, points to variability throughout the population imposed by environmental conditions and resource division. Focussing on transient matrices that change over time will help to bridge the gap from equilibrium to temporal complexity (Baltzer *et al.* 1998).

Deterministic models generate a response to a given input by one fixed law or group of laws (Heylighen 1999). Ruxton & Saravia (1998) argue that biological realism and hence more deterministic ideas should be incorporated into modelling real world phenomenon. Itami (1994) argues for the use of deterministic based models in his simulation of spatial dynamics utilising Cellular Automata theory within a Geographic Information System.

The authors of this paper followed Itami's suggestion and implemented a Cellular Automaton (CA) within a Geographic Information System (GIS) to model and examine the spread of an invasive plant species in a heterogeneous environment. The result is a spatially and temporally explicit model that can be explored and 'played' with to examine a range of parameter values and potential dispersal and distribution relationships. Section two examines CA and their contribution to modelling in combination with GIS. From this introductory setting the creation of the spatio-temporal conceptual model, parameter formalisation, rule base and model assumptions are identified and explained. Finally, implementation assumptions and results are presented and discussed before concluding remarks.

## **2.0 DYNAMIC MODELLING IN GIS**

Geographic Information Systems (GIS) have been increasingly utilised within modelling projects this decade. A strength of incorporating GIS into ecological modelling has been integration of environmental and biological features across a range of spatial scales (Aspinall 1994) within a complex terrain and heterogeneous landscape (Steyaert 1993). Primarily this modelling adopted Tomlin's (1990) cartographic modelling approach, which involves analysis of spatial data with Boolean or mathematical operations to identify unique combinations of characteristics (Wade *et al.* 1998, Johnston 1998). Models in GIS can only operate in equal time steps (Aspinall 1994, Coleman *et al.* 1994, Wade *et al.* 1998). Because of this GIS can only produce simulation models when coupled with temporally explicit modelling software (Mann 1996). The best way to model in GIS is through integration of GIS with the strengths of another system. For the successful analysis of the parameters of complex systems, Cellular Automata (CA) is a useful and widely applicable tool.

### **2.1 Cellular Automata in Ecological Modelling**

Cellular automata theory is based on a set of cell states and rules to model a system over space and time. CAs are useful in ecological modelling because they very effectively deal with the issue of spatial autocorrelation and have - by GIS standards - superb temporal capabilities. Cellular automata are an ideal tool for the manipulation of ecological parameters in simulation models because of their "bottom-up" modelling approach, which mirrors the scaling from individual to community level in ecological research. Instead of complex, averaged algorithms or non-spatial statistical methods, CA utilise parameters, rules and cell states as model components.

Cellular automata have become important in ecological modelling because they are both spatially and temporally explicit. However, both time and space are measured in discrete blocks (Wolfram 1986, Itami 1994). Grid cells respond to the state of neighbouring cells through conditional constraints (Itami 1994). Sessile population growth, competition and interspecies interaction can all make use of this function.

However, because of this, classical cellular automata are limited to short range interactions, with little “mixing” (Ruxton & Saravia 1998). “Mixing” or long distance dispersal has been integrated into invasion models to represent spread and colonisation of new areas (Peng & Albrecht 1998).

In order to develop a cellular automata model that simulates the ‘real world’ an intermediary method of transformation needs to occur. Because a model is a representation of reality, all details can not be included within the final simulation. Instead parameters and relationships need to be defined and rationalised through assumptions before coding in GIS can occur. This step towards implementation is discussed below.

### 3.0 THE CONCEPTUAL MODEL

The conceptual model of *Rhamnus alaternus* is implemented by running a Cellular Automaton in a combination of ESRI’s ARC/INFO GRID and Arc Macro Language (AML). Definition and formalisation of parameter limits, a rule base and related assumptions act as a guide for implementation strategy.

#### 3.1 Conceptual Model of Spread

The conceptual model of *Rhamnus alaternus* spread is based on the plant life cycle, incorporating temporal phenological and spatial dispersal features at the level of the individual. This approach produces the highest population based resolution possible. Aggregation of the model cycle in space produces the behaviour of a population, the cyclical, feedback nature of the model indicating temporal dynamics. Local, focal and global spatial identity is inherent in bird and gravity based dispersal mechanisms, for example in pollination and dispersal relationships (local), vegetative growth and gravity dispersal (focal).

#### 3.2 Formalisation of Parameters

To further define the model components described within the conceptual model, a method of formalisation has been created. This method of formalising parameters aims to simplify and unify the modelling process by defining each parameter in the same terms. Explicit definition of each parameter acts as a transfer medium towards implementation of the model as a spatially dynamic simulation.

As shown in the template below (see *Figure 1*), each parameter is given a name and a brief qualitative definition. The measurements units and expressions of the defined parameter are also included where applicable. The assumed minimum and maximum values of this expression are noted where possible. These values are assigned to the parameter based on available literature, discussion with experts within ecology of invasive species and the author’s own field based investigation.

<b>NAME</b>
Definition
Measure(Units)
Minimum value
Maximum value
Spatial resolution
Spatial extent
Temporal resolution (time slice updates)

*Fig 1: Template for formalisation of conceptual model parameters.*

Spatial resolution and extent define the parameter in a spatial context. Variation of extent and resolution between parameters is valid as GIS can accommodate such differences. The temporal resolution of a parameter defines the time step within which changes to the parameter value can occur. This is predominantly a one-year interval equivalent to the phenology of a perennial plant species.

An example parameter definition is given below for gender ratios within a population. Relationships between parameters are not explored in this definition, hence no reference to determination of gender is given.

## **GENDER**

**Definition:** Dioecious species gender ratio assignment for a population.

**Measure:** Individual – male or female

**Min:** 35% female

**Max:** 50% female

**Resolution:** 2 x 2 m grid cell

**Extent:** Waiheke Island / Hauraki Gulf (at coarse grid resolution of 100 x 100 m grid)

**Temp. res.:** Yearly association at germination

Minimum and maximum values for GENDER come from work by Fromont (1996) and Herrera (1984). These values are intended as starting points for exploration in this area, however this ratio seems to be one congruent with dioecious reproductive patterns.

### **3.3 Rule Base**

The initial rule base follows the temporal cycle of the conceptual model, incorporating constraints such as substrate, parental and coastal proximity. It must be acknowledged that incorrect approximation of disperser behaviour parameters can be a major factor in simulation model error (Ruckelshaus *et al.* 1997). Because of this, the rule base incorporates several indicators of disperser movement, including favouring roosting perches such as the native coastal canopy species Pohutukawa (*Metrosideros excelsa*) and Karo (*Pittosporum crassifolium*) and temporal and spatial floral competition landscapes for pollinators and dispersers (Buchmann & Nabham 1996). The rule base variables are entered as a simple ASCII file that can be created by any domain scientist regardless of their GIS expertise.

Rulebase statements incorporate both spatial and temporal conditional statements to determine cell state change. Below is a pseudo code rulebase example describing the frequency of different dispersal events. As shown the greater the density of fruiting trees, the more bird dispersal, as mass fruiting attracts birds to an area.

```
Dispersal = if F > 5 in NEIGHBOURHOOD
then 2/3 * FRUITING = BIRD_DISPERSAL and 1/3 * FRUITING = GRAVITY_DISPERSAL
else 1/2 * FRUITING = BIRD_DISPERSAL and 1/2 * FRUITING = GRAVITY_DISPERSAL
```

## **4.0 IMPLEMENTATION**

### **4.1 Assumptions and Adaptations**

For the implementation of the model within a GIS, several additional assumptions have been made to allow for run times constraints and data model structure. Sub-cellular homogeneity is assumed for all constraints and events except seed dispersal. The spatial resolution is maintained at 2 m to simulate individual adult plant size (as per field research and Fromont 1996). Due to storage capacity and complexity, seed dispersal is set to one germinating seed per 2 x 2 m cell. This limit is imposed by GRID functionality, however it also fits into the rule base directive that assumes individual mortality within each cell over time due to resource pressures. Other assumptions for implementation include climatic homogeneity over the study period and no change in landcover within the simulation period.

Long distance dispersal, or the movement of bird dispersers to nearby islands or mainland, has been implemented as a hierarchical process. A coarse resolution grid of 100 x 100 m grid cells is created over the extent of the study area. At specified time intervals, a random neighbourhood coarse grid cell will be invaded by dispersal from the seeded coarse cell, which is divided into a 2 x 2 m lattice. Upon invasion of the neighbouring cell of the coarse grid a new grid entity will be formed from that grid cell, divided into 2500 two-by-two m cells.

## 4.2 GIS Implementation

Based on data gathered on Waiheke Island in Auckland's Hauraki Gulf (Cole 1999), a  $50 \times 50$  cells starting grid with  $2\text{ m}$  resolution was populated with a *Rhamnus alaternus* population. The gender ratio was set to 65:35 in accordance with GENDER parameter definitions, and members of the population randomly assigned a gender based on this ratio. Each cell was assigned a state as defined in the rule base. ARC/INFO Grid and AML were used to code the conceptual model described in section 3. One of the most important features of the algorithms is that they maintain biological realism. This is achieved by ordering of events within the one-year temporal resolution to mimic a phenological cycle. The one-year temporal resolution or time slice was chosen as ARC/INFO Grid can not work simultaneously on each member of the population within a grid to produce phenological plasticity over the course of a year.

A hierarchical approach strategy was developed to examine the effect of individuals spreading across a heterogeneous landscape. Because of the storage and runtimes requirements for a  $2\text{ m}$  resolution grid that extends across the whole of Hauraki Gulf (appr.  $10^4\text{ km}^2$ ), a quadtree-like algorithm has been implemented (Samet 1990). All operations occur at a local (fine grid) level. Only when the local population spills over its original extent, or if seeds are long-distance dispersed by some chance bird, does the coarse grid come into action and a new fine grid (as a sub set of a coarse grid cell) is created. Thus, while a coarse resolution grid ( $100 \times 100\text{ m}$  resolution) was created covering the entire study area, storage and runtime intensive fine grids are created only in areas of active spread and dispersal. By creating sub-cellular detail of coarse grid cells as individuals reach the cell, the model simulation can cover a large area, maintaining detailed behaviour of individual movement without wasting processing time. Each fine grid created is recognised within the coarse grid as 'invaded'. The initial fine grid models the spread of individuals through its extent until either a long distance dispersal event or a border of the fine grid is reached.

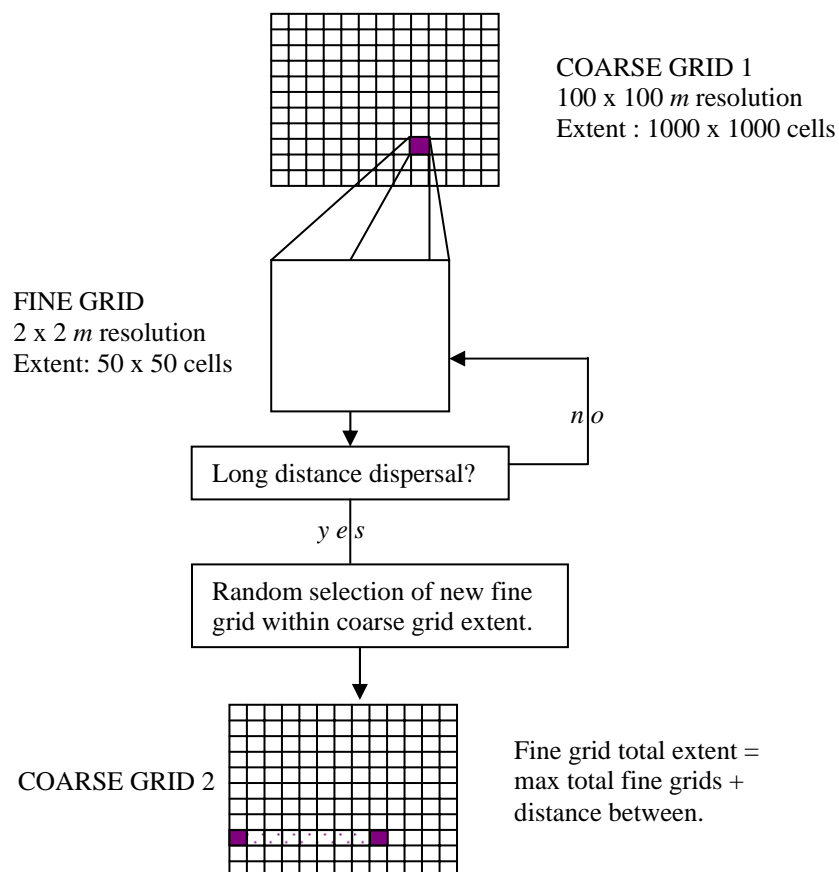
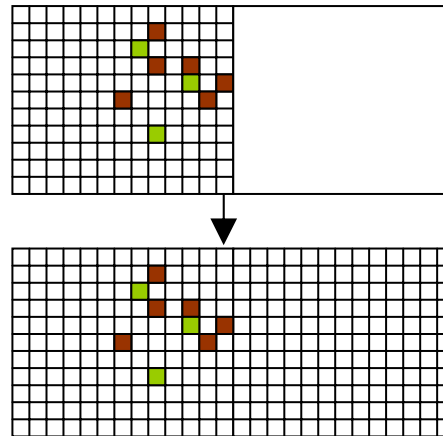


Figure 2: Process of long distance dispersal utilising hierarchical method of space subdivision and selection.

A long distance dispersal event forces the spread of the species outside the fine grid. This can be a random placement over the entire coarse grid, or random within a specified distance of the source fine grid. The newly selected cell within the coarse grid is subdivided into a fine resolution grid and a gender is assigned

randomly to a germinating seed within the new fine grid space. To maintain the distance and neighbourhood relationships inherent to the CA dispersal model, fine grids are appended together as shown by shading in COARSE GRID 2 in *Figure 1* above.

When the edge or border of a fine grid is reached by dispersal or vegetative spread, an adjacent coarse grid cell is 'opened', sub-divided into fifty by fifty cells and merged with the initial fine grid as shown in *Figure 2*. This process ensures no dispersal bias or increased densities occur at edges due to a model-induced, rather than environmental, lack of space.



*Figure 3: Adding to the extent of a fine grid due to an expanding population through subdivision of an adjacent coarse grid cell*

### 4.3 Effects of Parameter Error

Because models are abstractions of reality, error as deviance from reality is an inherent part of their functionality. At certain scales, however, parameter error may undermine a model's predictive ability.

Ruckelshaus et al. (1997) examine the consequences of parameter error for prediction of spatially explicit population models. Testing for error involved a generalised sensitivity analysis of patch size, shape and suitable habitat of an endangered species in a fragmented landscape. Parameter errors can be categorised as follows:

- misclassification of habitat suitability
- incorrect estimation of how far a disperser can travel
- incorrect estimation of the mortality rate of dispersers.

Spatially explicit population models are particularly error-prone, for example in substantiating input data and results from individual based models. Very few assessments of habitat availability have occurred at a landscape scale and information on dispersal parameters is scarce.

In an attempt to leave the choice of dispersal parameter values within the model up to the domain scientist, the authors developed an open, explorative environment. A 'scenario' ASCII file allows direct manipulation of parameter values by the domain scientist or end user. Outputs from each run of the model are recorded in a specified directory with a name coded to the combination of values used, allowing for easy comparison.

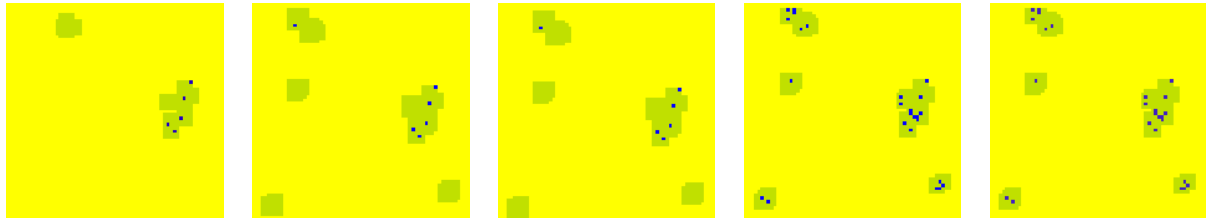
Parameters open for exploration include:

- input grid for initial population
- number of years in model run
- maximum distance between individuals for pollination
- disperser behaviour toward perches
- gender ratios
- frequency and nature of long distance dispersal
- effect of environmental constraints (distance to coast in this example)

## 4.4 Results

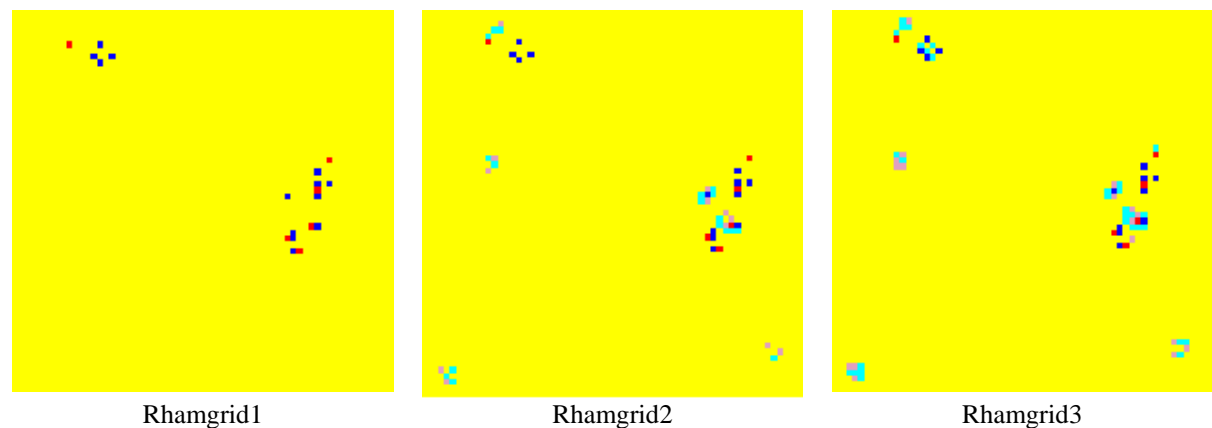
The ability to change parameter settings has obvious effect on the final distribution of a species. Two examples of spatial pattern as a direct result of parameter manipulation are shown below.

The spread of the species *Rhamnus alaternus* is defined by its neighbourhood or the maximum distance between male and female plants for successful pollination. *Figure 4* depicts a time-series of fruiting females, within a “pollination landscape”, where females are searched for by pollinators (shaded areas), and successfully pollinated adult females result in dark squares. Over time, the pollination landscape and population grows, allowing greater numbers of fruiting female trees.



*Figure 4: ‘Pollination landscapes’ and the corresponding pollinated females over a 5-year period.*

The attraction of bird dispersers to perches (large canopy trees used as roosts) is a defined disperser behaviour that affects the spatial distribution of the *Rhamnus alaternus* population over time, as shown in the time series in *Figure 5*.



*Figure 5: An example of the effect of perches as bird attractors. Clusters of plants occur quickly, here over a 3-year period.*

## 5.0 DISCUSSION

Species invasion is a prime example of a complex and not well understood problem in the realm of dynamic spatial modelling. It is also an example of complexity that can be difficult to recreate within the limits of current GIS technologies. Through the implementation of the *Rhamnus alaternus* invasion model, it became apparent that movement at different scales within a GIS was time consuming and difficult to implement. A series of imbedded loops and intermediary grids had to be created that allowed a small scale resolution grid to be moved to a randomly selected large scale cell site. In order to conserve run times and storage space, GIS models have typically been limited to small detailed grids, or large, aggregated data sets. Utilising a combination of both allows for unrestricted examination of pattern and distribution of the same model outputs at individual sites of interest or over large areas. In our specific example of species invasion modelling, unutilised space may become coveted or ‘invaded’ over time. Since the number and values of all parameters are kept variable, this approach can be applied to a wide range of diffusion processes in heterogeneous landscapes (Albrecht & Car 1999). Using a hierarchical approach allows for a detailed analysis of this space to occur only at the time of invasion, when the space becomes ‘important’. Hierarchical data structures such as quadtrees constitute a major improvement and allow for the easy integration of models at a range of scales.

## 6.0 REFERENCES

- Albrecht, J. and A. Car 1999. GIS Analysis for Scale-sensitive Environmental Modelling Based on Hierarchy Theory. In Dikau, R. and H. Saurer (eds.) *GIS for Earth Surface Systems*. Berlin, Gebrüder Bornträger, pp. 1-24.
- Aspinall 1994. GIS and spatial analysis for ecological modelling. In Mitchener, W., Brunt, J. and S. Stafford (eds) *Environmental Information Management and Analysis: Ecosystem to Global Scales*. Taylor & Francis. London.
- Baltzer, H., Braun, P. and W. Kohler 1998. Cellular automata models for vegetation dynamics. *Ecological Modelling*. **107**:113–125.
- Buchmann, S. and G. Nabham 1996. *The Forgotten Pollinators*. Washington, DC: Island Press/Shearwater Books.
- Colasanti, R. and J. Grime 1993. Resource dynamics and vegetation processes: a deterministic model using two-dimensional cellular automata. *Functional Ecology*. **7**:169–176.
- Cole, V. 1999. Modelling the invasion of *Rhamnus Alaternus* in GIS based Cellular Automata. *MSc Thesis (in preparation), Geography Department*. Auckland, University of Auckland.
- Coleman, M., Beraly, T. Burke, F. and W. Laurenroth 1994. Linking ecological simulation models to geographic information systems: an automated approach. In Mitchener, W., Brunt, J. and S. Stafford (eds) *Environmental Information Management and Analysis: Ecosystem to Global Scales*. Taylor & Francis. London.
- Fromont, M. 1996 *Ecological Research for Management of Rhamnus alaternus L.* Unpublished MSc Thesis. University of Auckland.
- Harré, M. 1998 – 1999 Biosecurity Officer, Auckland Regional Council. Personal communication.
- Herrera, C.M. 1984. A study of avian frugivores, bird-dispersed plants, and their interaction in Mediterranean scrublands. *Ecological Monographs*, **54** (1):1- 23.
- Heylighen, F. 1999. Web Dictionary of Cybernetics and Systems. In: Heylighen, F., Joslyn, C. and V. Turchin (eds): *Principia Cybernetica Web (Principia Cybernetica, Brussels)*, Accessed 25 September, 1999. <<http://pespmc1.vub.ac.be/ASC/IndexASC.html>>
- Itami, R. 1994. Simulation Spatial Dynamics: cellular automata theory. *Landscape & Urban Planning*. **30**:27–47.
- Johnston, C.A. 1998. *Geographic Information Systems in Ecology*. Blackwell Science. Oxford.
- Mann, S. (1996) Spatial Process Modelling for Regional Environmental Decision Making. *8th Annual Colloquium on Geographical Information Systems and Spatial Information Research*. University of Otago, Dunedin, NZ.
- Molofsky, J. 1994. Population Dynamics and Pattern formation in Theoretical Populations. *Ecology - Spatial Theory Special*. **75** (1):30–39.
- Peng, M. & J. Albrecht 1998. Integrating GIS with Dynamic Models of Biological Invasion. *Proceedings of the International Conference on Modeling Geographical and Environmental Systems with GIS*. Vol. **2**:659-64. Chinese University of Hong Kong. IGU/UGI.
- Ruckelshaus, M., Hartway, C. and P. Kareiva, 1997. Assessing the Data Requirements of Spatially Explicit Dispersal Models. *Conservation Biology*. **11**(6)1298–1306.
- Ruxton, G. and L. Saravia 1998. The need for biological realism in the updating of cellular automata models. *Ecological Modelling*. **107**:105–112.
- Samet, H. 1990. *Applications of Spatial data Structures: Computer Graphics, Image Processing and Geographical Information Systems*. Addison-Wesley, Reading, Mass.

Silvertown, J., Holtier, S., Johnson, J. and P. Dale 1992. Cellular automaton models of interspecific competition for space – the effect of pattern on process. *Journal of Ecology*. **80**:527–534.

Skellam J.G. 1951. Random Dispersal of Theoretical Populations. *Biometrika* **38**:196 – 218.

Steyaert, L.T. 1993 A Perspective on the State of Environmental Simulation Modeling. Pp 16 – 30. In: *Environmental Modelling with GIS*. (eds) M.F. Goodchild., B.O. Parks, L.T. Steyaert. 1993. Oxford University Press. Oxford.

Tilman, D. 1994. Competition and Biodiversity in Spatially Structured Habitats. *Ecology – Spatial Theory Special*. **75**(1):2–16.

Tomlin, D. C. (1990). *Geographic Information Systems and Cartographic Modeling*. Englewood Cliffs, N. J., Prentice Hall.

Wade, T., Schultz, B., Wickham, J. and D. Bradford 1998. Modeling the potential spatial distribution of beef cattle grazing using a Geographic Information System. *Journal of Arid Environments*. **38**:325 – 334.

Wolfram, S. 1986. *Theory and Applications of Cellular Automata*. Singapore: World Scientific Publishing Co Pte Ltd.