A Multi-agent Cellular Automaton for Grapevine Growth and Crop Simulation

Subana Shanmuganathan *Senior Member, IACSIT*,
Ajit Narayanan, *Member, IEEE,* and Nicholas J L Robinson

***Abstract*—A multi-agent (MA) cellular automaton (CA) model framework for simulating grapevine growth and crop in *Chardonnay* cultivated in northern New Zealand is presented. Estimating or projecting grape crop (quantity of grapes in tons per hectare (ha) and berry quality in Brix (sugar content) is an extremely complex and challenging task as the crop depends on many factors that interact with each other at varying degrees and over different time intervals in a "chaotic" manner. These key factors and their influences are simulated using CA rules, MA behaviour and interactions. Two sets of CA lattices and rules are used to simulate individual grapevine growth and vineyard phonological dynamics. The results achieved show potential for simulating vine growth and yield in different grape varieties (*Pinot Noir, Pinot Gris, Merlot* and other wine styles) and scales, such as New Zealand's major wine regions and that of world's, in ways which that have not been explored previously.**

***Index Terms*—component; climate effects; yield; vineyard**

I. Introduction

Obtaining accurate estimations of grapes in quantity (tons/ha) and quality (sugar, aroma and other colour phenol contents) is an extremely complicated and challenging task that has operational and economic significance to viticulturists and vintners (1) (2). Traditionally, vineyard yield and must composition are measured in terms of tons per hectare (ha) and Brix (one degree Brix corresponds to 1 gram of sucrose in 100 grams of solution)/ p^H / acidity respectively. Over the years, there have been formulae developed to estimate the crop with vines/ha, clusters/vine, buds/vine and cluster/ berry weight values (sampled averages) for different varieties and some basic formulae currently in use are provided below. In any of the approaches used, inconsistencies observed between the estimated and real crop figures of a vineyard are considered to be resulting from two factors;

(i) 70% of the variation from year-to-year variability in the number of clusters, and
(ii) 30 % of it from the variability in cluster weight.

Celluar automata (CA) are a relatively old computational modelling technique based on a regular grid of cells that together perform a global calculation through local interaction/s. From the initial introduction of Van Neumann neighbourhood rules in the 1950s (3) (4) to recent satellite imagery grid quantification research (5), it is clear that significant advances have been made in the development and application of CA and other related hybrid model approaches to simulating spatial and temporal changes in a wide spectrum of disciplines. The second section of the paper briefly outlines a few CA frameworks specially developed for vegetation dynamics simulation. Details of a multi-agent CA framework being developed for simulating grapevine growth and yield in *Chardonnay* cultivated in northern New Zealand are also presented.

Manuscript received March 21, 2011. The research presented in the paper is being conducted at the Geoinformatics Research Centre and supported by the School of Computing and Mathematical Sciences and Kumeu River Wines, New Zealand.

S Shanmuganathan is with Geoinformatics Research centre (www.geo-informatics.org) Auckland University of Technology, New Zealand 1142 (e-mail: subana.shanmuganathan@aut.ac.nz).

A Narayanan was with the University of Exeter (UK). He is now Head of School of Computing and Mathematical Sciences, Auckland University of Technology, New Zealand 1142, (e-mail: ajit.narayana@aut.ac.nz).

N Robinson is with School of Computing and Mathematical Sciences, Auckland University of Technology, New Zealand 1142 (e-mail: nrobinso@aut.ac.nz).

II. Crop Estimation Issues in Viticulture

Adverse consequences of inaccurate grape crop estimation and related issues are well-known among viticulturists and vintners around the world, and this has led to increased demand for improved techniques to better estimate the crop (6). Currently used methods are:

(i) destructively harvesting whole vines or segments of vines or
(ii) randomly sampling and weighing bunches and then combining these with bunch counts.

Both methods require adequate sampling and data interpretation for more accurate crop estimation. Vineyard management is understandably unwilling to commit extra resources during the busy harvest season. This unwillingness to allocate more resources during harvest for proper sampling is a major problem associated with these conventional methods. Hence, there is an urgent need for less demanding and more automated methods for crop forecasting.

A. Conventional methods of crop estimation

Two commonly used conventional crop estimation methods are referred to as the 'traditional' and 'lag phase method' (1).

1) *Traditional Method:*

An average cluster weight of a season is obtained for use in the consequent harvest and the formula used for this is as follows:

$$PY = (ANV \times NC \times CW) / 2000 \quad (1)$$

where,

PY = *predicted yield (tons per acre)*
ANV = *actual number of vines / acre*
NC = *number of clusters per vine*
CW = *cluster weight (in pounds)*

The problem here is that, of the three parameters, two namely, number of clusters per vine and cluster weight, require resource-intensive measurement for accurate predictions at a critically busy time.

2) Lag Phase Method:

This method is based on cluster weights collected during the "lag phase", which refers to a period when seeds begin to harden and this occurs about 55 days after first bloom or corresponds to the accumulation of 1000-1300 growing degree days (GDD) or heat units. During this period berry growth slows temporarily and it is considered that at this lag phase the berries have reached about 50% of their final weight. Based on this theory, the cluster weight at harvest could be predicted by multiplying the lag phase weight by an "increase factor" of 2. However, the multiplier varies among varieties and seasons. Growers are advised to determine their own multiplier for each variety/ vineyard. GDD, on the other hand, can be obtained from any nearby meteorological station. The formula used for this method is as follows:

$$PY = (ANV \times NC \times Lag\ CW \times 2) / 2000 \quad (2)$$

Where

PY = predicted yield (tons per acre)
ANV = actual number of vines / acre
NC = number of clusters per vine
Lag CW = cluster weight at lag phase (in pounds).

3) More elaborate methods

Other methods exist, such as includeing average values for all possible variations from vine/ ha down to berry weight, as described in (7):

Predicted yield = (vines/ha) x (buds/vine) x (shoots/bud) x (bunches/shoot) x (berries/bunch) x (berry weight)

III. CA Framework in Vegetation Dynamics Simulation

CA frameworks designed and developed for vegetation dynamics simulation over the last six decades continue to gain popularity due to their ability to provide new information on the likely patterns in the spatiotemporal changes of complex natural habitats. Increasingly, the new knowledge gained via CA models is described as detailed enough for management decision making in certain specific problem domains. Spatial patterns and trends over time in the dynamics of forest tree population (8), alpine tundra vegetation (9), rain forest species composition (10) and weed population (11) are among some useful simulations in this domain and the publications (8-11) described how CA rules relating to a micro scale e.g., individual plant to plant relationships, could be applied to simulate changes at meso/macro scales influenced by different factors and at varying degrees e.g., field of plants under current and future /potential scenarios, as elaborated in the following example. As far as we are aware, there has been no previous reported research on the use of CA for predicting grapevine growth.

Research that is closest to our objectives is in (12), where the effects of future climate change scenarios were simulated using traditional approaches under different greenhouse gas emissions and then used to estimate future irrigation requirements for vineyards in Spain by combining global circulation and crop models. The scenarios for different greenhouse gas emissions were produced by perturbing the water generator based on Canadian climate change model (CGCM2) results for the areas studied in the north east corner of the Iberian Peninsula. The "LARS-WG" weather generator was run with historical data covering a 42 year period to generate some 100 possible local weather scenarios corresponding to years 2010, 2015 and 2025 for the simulation. Meanwhile, CropSyst was used to simulate vineyard water balance. The crop simulation for 2005 reflected the FAO-56[1] crop co-efficiencies and even though the weather model suggested early spring and hastened harvest, interestingly this was concluded to be causing lesser burden on future irrigation requirements than earlier anticipated.

IV. Multi-agent CA Framework for Simulating Grapevine Growth and Crop

This section presents details of a multi-agent CA framework designed with two different sets of lattices and rules for simulating an individual vine growth and *Chardonnay* grape yield within a vineyard. Cellular automata can be broadly described as discrete dynamical systems in which the individual cells are homogeneous (all of the same type). Through local interaction (as specified by common rules that all cells share) and a specified neighbourhood (one cell can only communicate with other cells in that neighbourhood), complex behaviour can arise over a number of generations or time-steps. A CA is deterministic if its next state (on or off for a simply binary CA) is fully determined by its own current state and the states of neighbouring cells, and probabilistic otherwise (the next state is probabilistic). In a synchronous CA all the cells update in parallel, whereas in an asynchronic CA a cell immediately updates to the next state depending on the states of its neighbouring cells. Agents, on the other hand, are characterised by their relative autonomy (they can perform actions independently of other agents) and partial views of the global system depending on their function and decentralisation (there is no designated agent that controls all other agents). Also, and perhaps most importantly for this work, an agent can be complex (i.e. an agent can itself consist of parts specialised to perform different sub-functions).

Merging CA with agents results in an interesting hybrid architecture where: (a) cells, in addition to communicating with other cells in their neighbourhood, also perform calculations and can receive input and send output independently of other cells; (b) cells can be grouped to perform functions specific to them (agent architectures independent of the cellular automaton architecture); and (c) cells or groups of cells can share information with each other to ensure that what is happening in one part of the system is communicated to other parts of the system. In other words, implementing CA cells as agents adds a degree of modelling power to the CA, and implementing agents as CA cells allows agents to be located in the CA architecture in such a way that basic communication and state updating processes are provided. For modelling plant growth, a multi-agent CA framework has many advantages, including allowing a cell to represent an individual plant which in turn is complex (the plant consists of leaves, trunk, roots, etc), each of which can update its state depending on the sub-

[1] Even though 56 (FAO-56) co-efficient is expected to provide a universally consistent methodology for obtaining reliable estimates of crop evapotranspiration it has its own limitations (13)

parts of neighbouring cells/plants. That is, a multi-agent CA can support many levels of interactions in a way that a simple CA cannot. Also, agents can themselves be grouped so that a population of cells can be identified to be of a specific type (e.g. chardonnay, cabernet) that have their own specific rules for interaction. Also, since cells are now relatively autonomous, they can get information and resources (such as nutrients and energy) from sources independently of any other cell or agent. Some element of competition can be introduced through agents possibly learning from their environment how to modify their interactions with other cells and how to survive on a potentially hostile environment.

There are currently no standard definitions for how to combine CA with multi-agent systems. The research design and experiments described below represent one possible hybrid approach to modelling grapevine growth and crop. Any model will ultimately be subject to the same checks as any other modelling technique: the model's ability to fit the data and/or to make predictions (using cross-validation or other train-test methods). The hybrid method adopted here is to use the CA for modelling grapevine growth both at the micro (individual plant) and macro (field of plants) levels, where agents represent the grapevine and global vineyard operations that can be broken down into components that map onto the various CA functions for updating states through local interaction. One of the purposes of agents in this hybrid architecture is to allow some degree of "non-intervention". In other words, once the hybrid architecture is set up and started for a specific number of time steps to simulate hours, days, weeks or months, the user should not have to interact with the system. Instead, the agents are responsible for collecting information from other agents and from information sources, which will include actual temperature and other environmental data that the model is attempting to fit.

A. Designing a CA lattice for grapevine growth

An individual vine growth can be divided into 1) budburst, 2) leaf growth, 3) clusters of inflorescence initiation, 4) flowering, 5) berry formation, 6) development and 7) ripening stages, and is simulated using a *l x l* set of lattice and vine growth rules. For this vine simulation, major growth factors (soil quality, water stress and exposure to solar radiation) and triggers (daily maximum, minimum and soil minimum temperatures, and GDD) are used to calculate a variable called "available energy" (AE), the driving force for vine growth (Fig 1).

The AE calculated using formulae (3) and (4) is in turn utilised for growth in five vine organs, namely trunk, bud, shoot, leaf and cluster, depending on the vine growth stage (1-7), which is defined by a term called "priority" in the vine CA cycle. In modern day viticulture, annual grapevine growth cycle is divided into the seven growth stages based on temperature/day length/ growing degree days (GDD/ heat units) hence, temperature and GDD are used in the vine CA cycle to define the "priority". Any excess energy from the current cycle will be stored in trunk as SE.

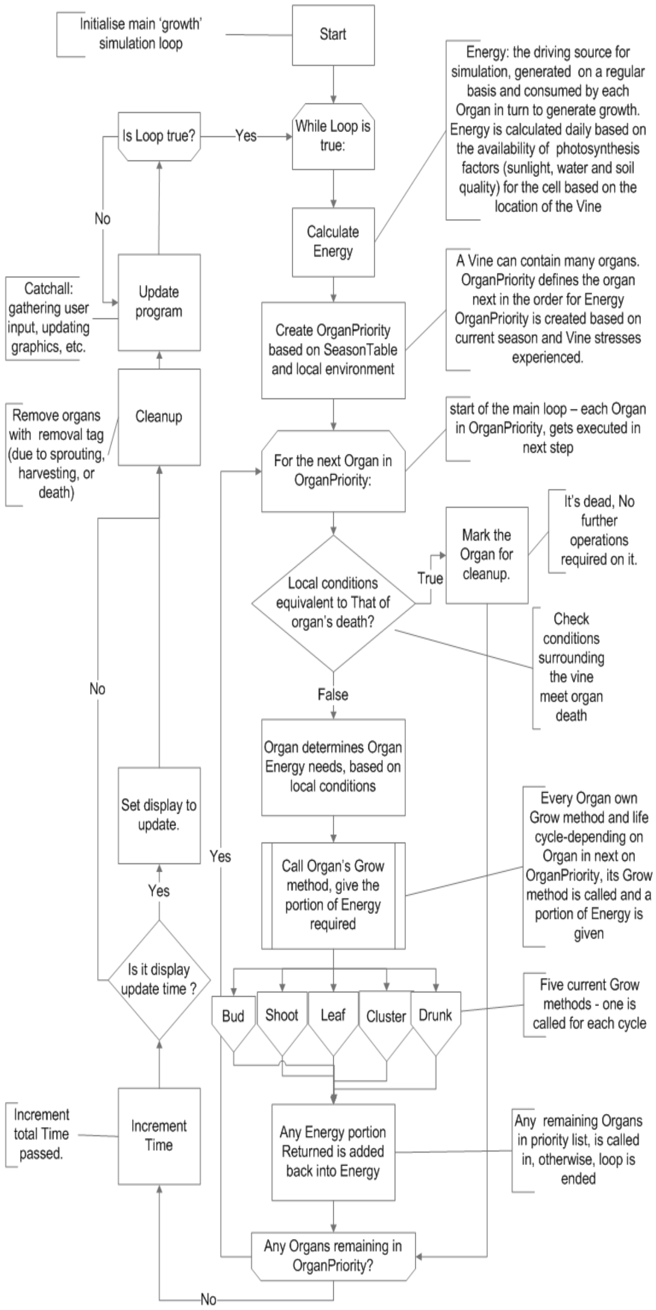

Figure 1. Schematic representation of the main processes relating to vine CA cycle. Each vine growth is displayed by a *l x l* lattice simulated by vine rules.

$$AE = ((GDD/DS) \ x \ AW \ x \ (TPV)) + SE \tag{3}$$

where,

AE	*=Available energy*
GDD	*= Growing Degree Days*
AW	*=Available water (1.0-0.0)*
DS	*=Day segments (morning, noon, twilight and night)*
TPV	*=Total photo synthesis value*
SE	*=Stored energy*

$$TPV \quad =(ALC \ x \ (A^2 \ x \ L) \ / \ LA^2 \tag{4}$$

where,

ALC	*=Active leaf cover in cell*
CA^2	*=Cell area ^{2} (in centimetre^2)*
L	*=Light (1.0-0.0)*
Cells	*=Cells in grid*

Vine organ initiation, growth, maturity and death vary based on the type of organ and are simulated using rules in the vine CA cycle. For example, organ “leaf” grows to become a full leaf after unfolding from a “shoot”. The leaf growth continues until it reaches maximum leaf blade length, stays alive for several days producing energy via photosynthesis and then eventually dies off; similarly, each organ has its own growth phases and rules in the vine CA cycle (see Fig 3 for bud growth rules).

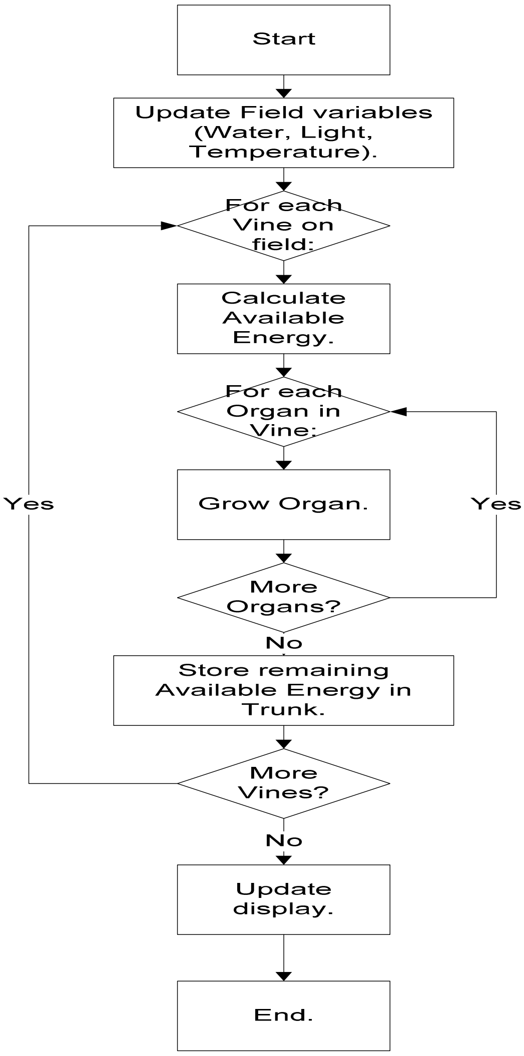

Figure 2. Schematic representation of the main processes on the crop simulation cycle displayed by *L x L* lattice based on vineyard rules

B. *Designing a CA lattice for grape crop (field) simulation*

The grape crop simulation displayed on a *L x L* set of lattice has its own set of rules. The field CA rules are applied to certain abstract vine and environmental parameters whereby yield at larger scales i.e., within a vineyard, is simulated. The key vine parameters, such as vine canopy structure (buds/ vine, clusters/ vine, cluster weight) and vital environmental factors, are used in the field CA cycle. Random values are used to simulate variability in environmental factors, such as soil quality, availability of water and light (solar radiation), temperature and humidity for each individual vine in the field CA cycle (Fig. 2).

C. *Designing agents for vineyard operations*

Global vineyard operations, such as pruning, fertiliser application, spraying (pesticide/ fungicide), irrigation and harvest are partly implemented through agents that will be responsible for maintaining and updating information on foliage and nutrient levels. Currently, rainwater and soil condition are globally available to all cells through agents that are provided with the information by environment models that simulate realistic growing conditions. These models will be replaced by in situ sensor data that will continuously monitor the environment.

V. The Results

The initial results of the multi-agent CA framework (Figs 3-5) show how grape crop simulation at micro and meso scales, such as a vineyard, are achieved using agent expertise, and how micro level issues (individual vine growth) are controlled by the cells and their interactions. The vine CA lattice simulates growth in vine organs (as explained in section IV) beginning with budburst, leaves, clusters (fluoresce and berry) to produce grapes for both in a vine (in berry weigh and berries/ cluster) and in a vineyard (in terms of grapes (tons/ha), Brix, p^H and acidity). The user interface has buttons, tabs and scroll bars to set/ change critical parameters relating to individual vine growth, such as buds/ shoot, shoots/ vine, clusters/ shoot, berries/ cluster and berry size. These parameters could be used to change values based on the grape variety being simulated.

All Organs
Variables:
Death Threshold = 0^oC

Standard Organ death rule:
IF Local Temperature < Death Threshold
    Organ is dead.
Bud rules
Variables:
Frost Threshold = 2 degrees centigrade.
Flower Daylength Threshold = 12 hours
Flower Temp Threshold = 10 degrees centigrade
Death:
IF Local Temperature < Frost Threshold
    Decrement Remaining Buds
IF Remaining Buds == 0
    Bud is dead
ELSE IF Local Temperature < Death Threshold
    Bud is dead.
Growth:
Add Growth to Total Growth
IF Total Growth >= Burst Threshold
IF Day Length > Flower Daylength Threshold AND Local Temperature > Flower Temp Threshold
Bud is dead
    Cause Vine to spawn new Shoot at Bud's location.
Remove any excess Total Growth beyond Burst Threshold, return to Growth
Return any remaining Growth

Figure 3. Field CA rules for budburst, death and growth

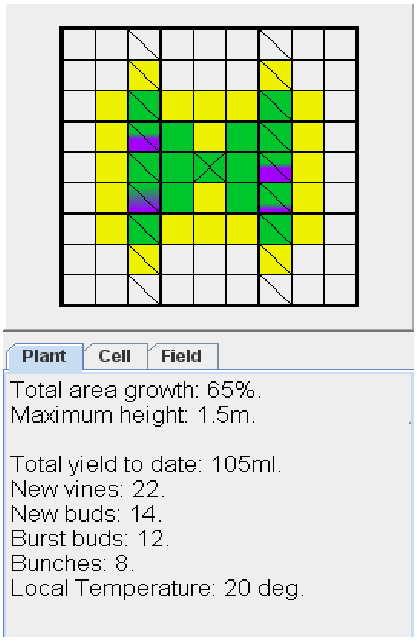

Figure 4. CA simulation showing vine growth with various grapevine organs that are incorporated in the vine CA cycle.

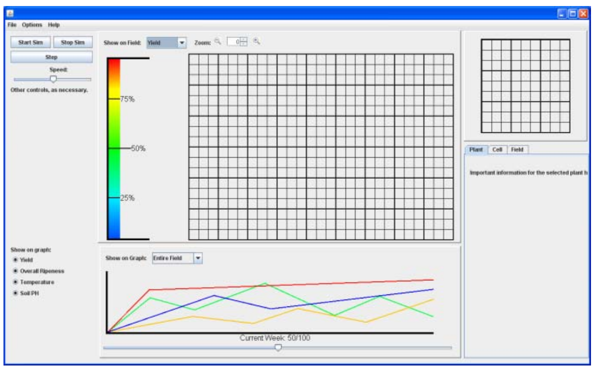

Figure 5. CA simulation of grape vine growth and yield at larger scales, such as vineyard, wine region. By changing the vine and field parameters it is possible to simulate growth and yield in diffident grapevine varieties, such as Chardonnay, Pinot Noir and Pinot Gris.

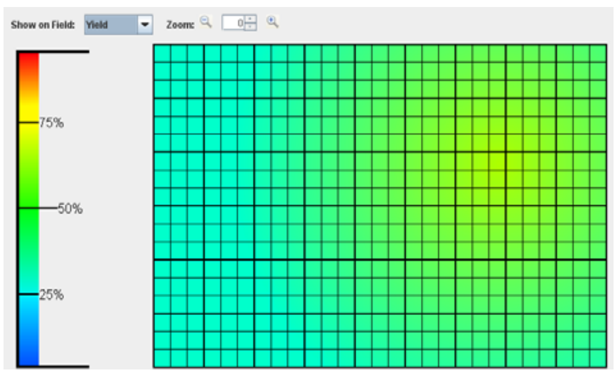

Figure 6. Screen display showing the CA simulation of grape crop. The variability in yield within a vineyard is simulated based on variations generated in soil, availability of water, nutrients, solar radiation, temperature and humidity created with random number generators.

The results thus far achieved with CA lattices (vine and field) and agents show potential for crop prediction in different grape varieties and at varying scales. By changing an individual vine growth parameters, such as (buds/vine) x (shoots/bud) x (bunches/shoot) x (berries/bunch) x (berry weight), users are able to predict the outcome from a vineyard, such as grapes in tons/ha under different scenarios. The ability to change vine parameters could be used for predicting crops in different grape varieties, such as *Pinot Noir* and *Pinot Gris*. By changing field variables soil quality, availability of water and light (solar radiation), temperature and humidity it is possible to create variability in field CA cycle and this is useful in creating within and among vineyard (different sites) variations.

VI. Conclusions

The paper described the initial investigation so far conducted on simulating vine growth and vineyard yield in Chardonnay cultivated in northern New Zealand. Even without a full multi-agent cellular automaton implementation of all vineyard operations, the preliminary results of CA simulations (vine and field) are promising. Future work will focus on a full implementation of the multi-agent based CA framework with an interface that will enhance viticulturists' ability to better predict their outcomes under different scenarios, such as pruning decisions; number of buds/ shoot to allow for full growth for that season, future climate change and at different scales.

The major benefit with the approach is that it provides an alternative method to estimating yield without incurring any additional cost as this approach can be simulated with historic and other model prediction data. As far as the authors are aware, this is the first attempt to contribute to 'precision viticulture' (14) through the use of a multi-agent cellular automaton that take into account detailed information concerning both resources (energy, water) as well as important botanical features (leaves, buds, etc). In the longer term, fitting the data and making predictions about growth will need to be related to quality of wine (15). With the inclusion of a wine quality module vintage ratings as well could be predicted under different possible weather and other atmospheric conditions.

Acknowledgments

The authors wish to thank Michel Brakavich, Kumeu River Wines. Prof Philip Sallis and Peter Sumich of AUT are acknowledged for their continuous support.

References

[1] Dami, Imed. Methods of Crop Estimation in Grapes. 2010. Cited in Wine-Grower News no 33 July 2, 2010, Midwest Grape & Wine Industry Institute
[2] Trought, Mike. Yield Management and Prediction. 2007, Science Reports 20-21 cited in p15 Vintage Research Marlborough Wine Research Centre annual report 2006/7, 1-78
[3] Bandini, S, Mauri, G and Serra, R. 2001 Cellular automata: From a theoretical parallel computational model to its application to complex systems. *Parallel Computing* 27 539-553 (2001).
[4] Geertman, S, Hagoort, M and Ottens, H. 2007 Spatial-temporal specific neighbourhood rules for cellular automata land-use modelling. International *Journal of Geographical Information Science*.21 (5) 547-568, (2007)
[5] Espínola, M., Ayala, R., and Leguizamón, Saturnino 2008 Classification of Satellite Images Using the Cellular Automata

Approach [Eds] M.D. Lytras et al. Springer-Verlag Berlin Heidelberg, 2008, The Open Knowledge Society. *A Computer Science and Information Systems Manifesto Communications in Computer and Information Science*, Vol. 19, 521-526, 2008

[6] Dunn, Gregory M and Martin, Stephen R. 2004 Yield prediction from digital image analysis: A technique with potential for vineyard assessments prior to harvest. *Australian Journal of Grape and Wine Research* 10, 96-198, 2004

[7] Keller, M, Tarara, J M and Mills, L J. 2010 Spring temperatures alter reproductive development in grapevines_105. Wiley Onlie Library, 2010, *Australian Journal of Grape andWine Research* Vol. 16 445–454, 2010

[8] Stamatellos, G and Panourgias, G. 2005 Simulating spatial distributions of forest trees by using data from fixed area plots. Institute of Chartered Foresters, 2005, F*orestry*, Vol. 78, No. 3,. 305-312. 2005

[9] Zhang, A Yanqing, et al. 2008 Cellular Automata: Simulating Alpine Tundra Vegetation Dynamics in Response to Global Warming. Arctic, *Antarctic, and Alpine Research* Volume 40, Number 1 / February. 56-263. 2008

[10] Alonso, David and Sole´, Ricard V. 2000 The DivGame Simulator: a stochastic cellular automata model of rainforest dynamics. *Ecological Modelling* 133 ,131–141. 2000

[11] Zhang, A. Y., Peterman, M. R., Aun, D. L., and Zhang, Y., 2003 Using CA model to obtain insight into mechanism of plant population spread in a controllable system: annual weeds as an example. *Ecological Modelling* 166, 277–286. 2003

[12] Marsal, Jordi and Utset, Angel. Vineyard Full Irrigation Requirements under Climate Change Scenarios for Ebro Valley, Spain. [book auth.] Valladolid, Spain) Angel Utset (ITACyL. Climate Variability, Modeling Tools and Agricultural Decision-Making 978-1-60692-703-8 261 pages. Nova Publishers, 2009, 119-126.

[13] Hunsaker, D J, Pinter, P J and Cai, H. Alfalfa basal crop coefficients for FAO-56 procedures in the desert regions of the southwestern U.S. [Online] 2002. ISSN 0001-2351.

[14] Best, S, León L and Claret, M. 2005. Use of precision viticulture tools to optimize the harvest of high quality grapes. Information and Technology for Sustainable Fruit and Vegetable Production FRUTIC 05, 249-258, 2005

[15] Shanmuganathan, S, Sallis, P and Narayanan, A. 2010. Modelling the effects of daily extreme weather on grapevine and wine quality. Proceedings of International Congress on Environmental Modelling and Software (iEMSs), David A. Swayne, Wanhong Yang, A. A. Voinov, A. Rizzoli, T. Filatova (Eds.), 5-8 July, Ottawa, Canada, 2010, 9pp

Dr Subana Shanmuganathan currently a Research Fellow at Geoinformatics Research Centre, School of Computing and Mathematical Sciences, Auckland University of Technology (AUT), returned to her fellowship at AUT (New Zealand) in 2007 after completing a two-year visiting postdoctoral fellowship at Ritsumeikan Asia Pacific University, Beppu, Japan awarded by the Japan Society for the Promotion of Science (JSPS). In 1997, Subana was appointed as Director of Studies, Institute of Printing, Sri Lanka. Prior to this, from 1986 to 1996, she held senior management positions in the Computer Branch of Department of Government Printing, Sri Lanka. Subana's research interests are in neural network computational modelling, Machine learning, ecological dynamics modelling, data/text mining, geo-spatial data analysis, design and development of integrated information systems in healthcare and office automation.

Prof Ajit Narayanan was appointed as Lecturer in the Department of Computer Science at the University of Exeter (UK) in 1980, became Head of Department during the late 1980s and was promoted to Senior Lecturer in 1990. He was elected Dean of Science in 1996 and Dean of the Undergraduate Faculty from 1998 to 2002. He was awarded a personal chair in 2001 for his research in artificial intelligence, quantum computing and bioinformatics. In July 2007 he joined AUT as Head of the School of Computing and Mathematical Sciences. His research interests are in intelligent bioinformatics, machine learning, nature-inspired computing and social impact of novel technologies.

Mr Nicholas (Nick) James Logan Robinson completed his Bachelor's Degree in Computing and Information Sciences at the School of Computing and Mathematical Sciences, AUT in 2010 and is now enrolled for a Master's degree in Forensic Information Technology at AUT. Nick research interests are in cellular automata, artificial intelligence, and computer forensics.