An improved cell based model to predict the spread and growth of *Caulerpa taxifolia* in southern Australian waters.

Dunn , J.M. [1] , L. McArthur[1], S. Schreider[1]

[1]Department of Mathematical and Geospatial Sciences, Royal Melbourne Institute of Technology University, Melbourne, Victoria
Email: jessdunn@bigpond.net.au

Keywords: *Grid cell based model, Caulerpa taxifolia, Lake Conjola, Environmental model, Invasive species*

EXTENDED ABSTRACT

This study explores increased complexity in cell models using a simplified, one dimensional interaction model as the foundation. Increased complexity leads to multiple dimensions and parameters. In turn this creates issues relating to integration. The questions *what to include?* and *how to include*? cell based interactions for an improved model become the focus of this study.

In exploration of this, we further develop a grid cell based model of the spread and growth of the green alga *Caulerpa taxifolia* (Vahl) C. Agardh for Australian waters by adding new levels of complexity relating to cell interaction. *C. taxifolia* is a marine, green, seaweed that is prevalent in tropical and sub-tropical waters throughout the world. However, invasive strains of the seaweed have been found in southern waters in Australia and they appear to be having some impacts on some native marine species. *C. taxifolia* is very difficult to contain and spreads at rapid rates. The main aim of the study is to model the spread of *C. taxifolia*, by taking into account biological and environmental factors of growth and in doing so predict the likely direction of spread in an estuary or lake.

Foundation models have been developed based upon the discrete form of the diffusion equation but which relied only on the current biomass. These models only considered average growth and spread and ignored environmental factors influencing the advancement of the seaweed in the waterway. This study implements ways to improve the predictive capabilities of models by including spatial and temporal aspects of growth and spread and by including *C. taxifolia* responses to seasonal weather patterns and temperature and nutrient changes. In addition, the spread of the plant due to the bathymetric and biological factors are considered.

The new dimensions of cell interaction, particularly the inclusion of the growth function, which varies seasonally, produced results that better reflected the observed data. Also, incorporation of vegetation distribution in the lake, used to indicate nutrient rich areas where *C. taxifolia* might thrive, wind direction and speed information and depth limitations have allowed for a significant improvement in the cell models predictive capabilities. In the case of seasonal fluctuations, variations outside the carrying capacity of cells created new problems with interpreting and comparing simulated results with the observed data.

The model was validated using the real growth data for several estuarine lagoons in New South Wales, Australia. We test the model on two waterways, Lake Conjola (including Berringer Lake) and Narrawallee Inlet, both of which are located in New South Wales, Australia.

Accurately representing the spread and growth using available information becomes particularly important in these waterways where locating outbreaks for eradication requires a large amount of resources. It is hoped that identifying the direction and distance of the spread will lower both the logistical and the financial aspects of eradication.

1. INTRODUCTION

1.1 Background

The previous cell-based model described by McArthur et al. (2006) was a simple model relying only on the current biomass. The principle aim of this work was to provide a simple predictive model which could also be used to identify the significant biological and environmental factors that needed further investigation in order to improve the predictive capabilities of the model.

The main outcome from this model was that the level of simplicity meant that the previous simulation could not be used for practical prediction of the *C. taxifolia* biomass. It was identified that the main problem with this model was that seasonal biomass fluctuations, in particular winter decay in biomass, and death rates were not included into the model. This previous research identified these two main areas as the future direction for simulation and the modelling of this problem.

Rather than reproducing sexually it has been observed that the invasive form of this species spreads by growth and fragmentation. Recent research by Wright et al. (2006) has shown that fragments of *C. taxifolia* have a higher chance of dispersal to new areas in summer, however in winter this is near zero. Hill et al. (1998) detailed the winter degeneration of *C. taxifolia*, due to the decline in water temperature, in the Mediterranean Sea. The observed data for Lake Conjola and the Narrawallee Inlet also confirmed that there was a significant decay in the total biomass and dispersion over the June to September period. This die back in winter can be attributed to the bleaching of the thalli, a part of the plant, observed during the colder months by Wright et al (2006) at Lake Conjola, where water temperatures in the NSW area can reach a low of 11°C. Growth of *C. taxifolia* is greatest in summer, Meinesz et al. (1995).

Meinesz et al. (1995) observed growth of *C. taxifolia* at depths of 100m in the Mediterranean Sea. However growth in these areas was slower due to low levels of light and colder water temperatures. In Lake Conjola it has been observed that no growth, over a two period year, in an established colony, is occurring at a depth below 10m. It is believed that Lake Conjola exceeds 10m in parts. Already established stochastic, discrete and deterministic models by Hill et al. (1998) and Ruesink et al. (2006) confirm these depth limitations and suggest a high risk of invasion for areas 5 to 10 m deep. To model the expected direction of spread wind directional data will be used to simulate water and fragmentation movement through the shallow areas of the waterway. Similarly, depth data will be used to indicate the areas where temperature and light changes are most influential. Spatial data is incorporated in such a way as to limit growth a certain distance from land and already established patches of *C. taxifolia*.

Interaction between the invasive *C. taxifolia* and other living organisms, particularly other seagrasses, result in *C. taxifolia* dominating the area. Incorporation of vegetation distribution in the lake will be used to indicate nutrient rich areas where *C. taxifolia* might thrive, although it has been recorded that this invasive species can grow in and on a variety of geological structures such as rock and sandy sea beds. In the Mediterranean Sea where *C. taxifolia* was first recognized by Meinesz et al. (1995) as a problem it has been recorded that *C. taxifolia* has eliminated entire areas of native seagrasses. Also, this species has been linked to a decline in total population numbers and varieties of fish species. This can be attributed to the loss of native seagrasses essential to the ecosystem. The same effect is a real possibility in Australia and the result could be the loss of marine biodiversity in the affected waterways.

The ideas below were found to be the main factors influencing *C. taxifolia* growth. We seek to integrate the following into the simple cell model:

- The seasonal fluctuations of weather including summer growth and winter decay.
- Bathymetric factors relating to depth.
- Growing conditions of other species in the waterway.
- Fragmentation spread direction and velocity.

Complexities neglected from the model include rainfall distribution, influencing flooding, and water nutrient loads. Certain dimensions, including the dimensions of rainfall and nutrient loading, are ignored for two reasons. The first is the unavailability of data where the four factors listed above can be integrated in waterways in NSW because of constant weather and seagrass distribution data. Secondly certain data varies randomly in the given waterway. For example nutrient loads in waterways depend on tidal waters, rainfall distribution and upstream behaviour. Because of the shear volume of dimensions, this single factor can not be included. Also this type of data is waterway specific and adaptation of the model to each waterway becomes unrealistic. Similarly certain dimensions, such as rainfall, cannot be broken down per cell. The computation time and exploration involved for dimensional interaction per cell is not possible and the effects on the overall modelling results are suspected to be negligible in comparison in this situation.

1.2 Objectives

Aside from addressing the environmental and biological concerns of the McArthur et al. (2006) model discussed above, this study will also address the fore mentioned studies prime objective – to model the spread of *C. taxifolia* to help focus search efforts by the New South Wales Department of Primary Industries. This is less important in Lake Conjola because the species has invaded a large area; however in other estuaries, particularly the Narrawallee Inlet, locating small outbreaks from fragmentation requires enormous resources. Specifically, this requires a model that can identify growth seasonally. It is hoped that an improved model will increase efficiency in locating outbreaks in infested waterways and hence lower the logistical and financial aspects of eradication.

We also seek to demonstrate that including some complexity can see marked improvements in modelling results.

1.3 Data

Two waterways, located in New South Wales, Australia, Figure 1, are considered for this model. Lake Conjola, Figure 2, is the primary source of data for model simulations. *C. taxifolia* in this waterway has essentially been allowed to grow naturally and is thus an ideal source of data to validate model simulation and output. The Narrawallee Inlet, Figure 3, will be used to test the models adaptability to different waterways.

Data collection was first undertaken in 2003 in areas of known infestations. Mapping of Lake Conjola was undertaken after the summer growth in March 2003, again after winter decay in August 2003 and finally in March 2004. Data from March 2003, August 2003, March 2004, September 2004 and April 2005 from the Narrawallee Inlet was available for initialisation of the algorithm and for comparisons. Distribution of other seagrasses in both waterways was also available. This data is used to indicate nutrient rich areas where *C. taxifolia* is likely to spread. For a more comprehensive explanation see McArthur et al. (2006).

The data is specified as latitude and longitude co-ordinates taken every 10m in north-south and east-west directions. For each grid cell coordinate in the estuary of study, data collectors have assigned categorical data: no biomass of *C. taxifolia* (B=0), sparse biomass (B=1) and dense biomass (B=2). Also assigned is land (B=-1). Throughout simulation the data becomes numerical and is then converted back to categorical for output. No depth data is available.

Wind data used in the algorithm is from the Ulladulla Climatological Station in NSW. Wind speed (km/hr) and direction (bearing) were measured four hourly, everyday, over a two year period from January 2003 until January 2005.

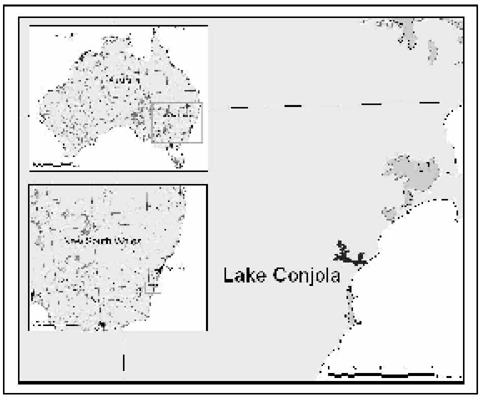

Figure 1. Location of waterways, NSW, Australia.

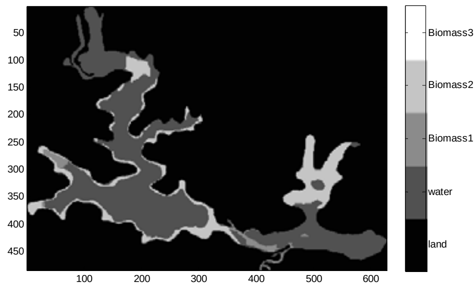

Figure 2. Lake Conjola, NSW, March 2004.

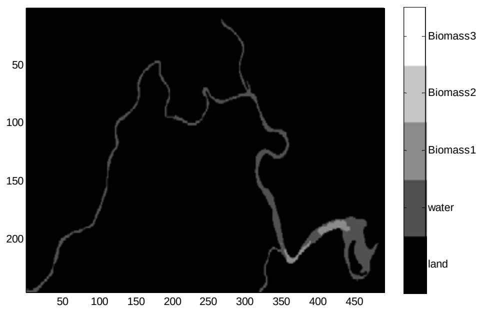

Figure 3. Narrawallee Inlet, NSW, April 2005.

2. DESCRIPTION OF MODEL

2.1 Methodology

The model structure was based on a previous cell based model by L. McArthur et al (2006). The model described a simple cellular model based on the diffusion equation that relied only on the current biomass. The approach that is used to model the spread and growth is related to the diffusion equation for biomass B in that model by:

$$\frac{\partial U}{\partial t} = a[\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2}] \qquad (1)$$

where $U = (t,x,y)$ denotes the biomass as time, t, in the cell with coordinates x, y. The model is in discrete time and therefore equation (1) at time, $t=n+1$, above can be represented as:

$$\frac{[B_{i,j}(n+1) - B_{i,j}(n)]}{\delta t} = \frac{a[B_{i+1,j}(n) - 2B_{i,j}(n) + B_{i-1,j}(n)]}{\delta l^2} + \frac{a[B_{i,j+1}(n) - 2B_{i,j}(n) + B_{i,j-1}(n)]}{\delta l^2} \qquad (2)$$

where the indices i and j represent the latitude and longitude co-ordinates, t is the discrete time step and l is the size of the grid cell.

Equation (2) can be modified to give the growth equation for a cell at time, $t = n + 1$ given by

$$B_{i,j}(n+1) = \alpha B_{i,j}(n) + \beta[B_{i+1,j}(n) + B_{i-1,j}(n) + B_{i,j+1}(n) + B_{i,j-1}(n)] \qquad (3)$$

Equation (3) demonstrates that α and β can be expressed as functions of a from equation (1), the time step δt, and grid size δl. In Section 2.2 the model will consider these coefficients as functions of the expected direction of the water current in shallow waters. In order to do this, biomass in diagonal cells will also be considered.

2.2 Description of Algorithm

In this section we look at ways to incorporate the four dimensions discussed in section 1.1. We also present ways to incorporate interactions between grid-cells arising from multi-dimensional integration challenges and the data available that can be translated across different waterways.

The data is stored in an $m \times n$ matrix of cells, A, where each cell $a_{i,j}$, $i=1, 2, 3,\ldots m$, $j=1, 2, 3,\ldots n$, contains the coordinates and the vector $B_{i,j}(t) = \{0,1,2\}$, contains the biomass matrix.

Functions within the algorithm govern the change in biomass. Biomass can either increase to a cell carrying capacity, K, during summer months as observed by Wright et al. (2006), or decay during winter months. K can take on numerical value between 0 and 3. The change from 2 to 3 in the maximum biomass is to allow greater freedom when modelling decay in the winter months. The time step, t, is determined by the seasonal fluctuations and hence is set in months (12 intervals per year). The growth, q, equation (4) and figure 4, is a function of the growth rate, g, and time, t.

$$q = g * sin((2 + \frac{t-1}{3})(\frac{\pi}{2}) + 0.4) \qquad (4)$$

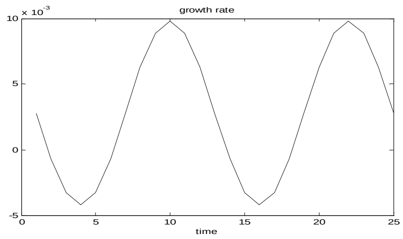

Figure 4. Growth Function, q, ($g = 0.007$).

The summer growth in each cell $a_{i,j}$ at time t, is given by, S = the sum of the surrounding cells:

$$S = B_{i,j}(n) + 1.04B_{i+1,j}(n) + 0.9B_{i-1,j}(n) + 2.28B_{i,j+1}(n) + 2.58B_{i,j-1}(n) + 1.84B_{i-1,j-1}(n) + 0.8B_{i+1,j+1}(n) + 0.25B_{i-1,j+1} + 0.31B_{i+1,j-1}(n) \qquad (5)$$

where the weights of the biomass of the surrounding cells are determined by the expected wind direction in summer. In Lake Conjola and the Narrawallee Inlet where *C. taxifolia* can be seen to spread only in shallow waters (Wright et al. 2006) wind directional data can be used to simulate fragmentation movement through these areas of the waterway. Wright et al. (2006) also observes that fragments of *C. taxifolia* have a higher chance of dispersal to new areas in summer; hence wind data is only tested for the summer months. In winter these wights are assigned a value of one with diagonal cells additionally weighted by $2^{-1/2}$. Where if the sum of the surrounding cells if is less than the threshold, d, and the biomass does not equal K, then the biomass at time $n + 1$ is given by,

$$B_{i,j}(n+1) = 0 \qquad (6)$$

This allows for cells near the edges of patches to decrease rather than densely populated cells located in the centre because water temperature is expected to affect sparsely populated cells with recent growth. Direction has been divided into the following categorical data w={1,2,3,4,5,6,7,8} according to each cell, where directions in towards the diagonal are weighted by half because of their distance and lack of facing edge.

300°< 5 ≤330°	330° < 1 ≤ 30°	30° < 6 ≤ 60°
240°< 3 ≤300°	$B_{i,j}(t)$	60°< 4 ≤120°
210°< 7 ≤240°	150°< 2 ≤210°	120°< 8 ≤150°

Figure 5. Categorical data for wind direction.

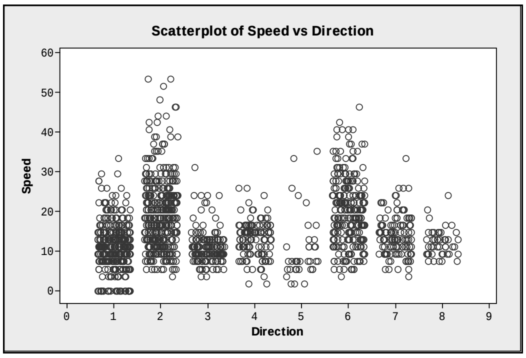

Figure 6. Scatterplot of wind direction and speed.

The expected frequency of each direction is then used to calculate the coefficients of the opposing cell. This then governs the amount of spread from fragmentation and growth from this cell onto the cell $B_{i,j}(t)$. Figure 6 shows the distribution of directions compared to speed.

If S is greater than the threshold of surrounding cells, p, and the biomass does not exceed the cell carrying capacity, K, then the biomass at time $t=n+1$ is given by,

$$B_{i,j}(t+1) = q * S + (q+1) * B_{i,j}(t) \tag{7}$$

Nutrient rich areas, indicated by an $m \times n$ matrix, S, of seagrass distribution for each waterway, allow for additional spread at each time step. This matrix incorporates a level of simplification. To avoid large scaled complexity, the interaction between *C. taxifolia* and differing species of seagrass, "seagrass" is though of as a collective, where if there is "seagrass" additional growth is considered. Coordinates, with surrounding cells of *C. taxifolia*, corresponding to areas of known seagrass presence are governed by an additional percentage increase in biomass per time step, r.

Additionally, growth and spread of *C. taxifolia* in the waterway is subject to rules defining depth. Hill et al. (1998) described growth in the Mediterranean Sea beyond 10m as slow if at all, and this principle is adopted in this model. This is possibly due to light and temperature limitations at these depths. Specific depth data for each coordinate is unavailable for these two waterways. For the Narrawallee Inlet this is not a problem since the waterway is shallow and it does not exceed more than 10m in depth. In Lake Conjola however, the waterway does exceed 10m in certain areas. New spread has been constrained to originate from edge cells (<10 cells). However growth occurs in all cells.

3. RESULTS

The objective is to optimise the parameters p, d, g and r, in a physically realistic sense, which will give the best results in comparison with both Lake Conjola and the Narrawallee Inlet. This was done by simulating the spread and comparing the data set with the observed results. The model is initialized with data from March 2003 for both waterways. Data was compared over the three time steps, March 2003 to August 2003 and August 2003 to March 2004. Lake Conjola data contains 62749 observations and the Narrawallee Inlet contains 4525 observations

Some simple metrics are introduced in the McArthur et al. (2006) paper. These metrics are used here to indicate growth and are defined by:

$$M_S = \frac{N_{SIM} - N_{OBS}}{N_W} \tag{8}$$

$$M_B = \frac{B_{SIM}}{B_{OBS}} \tag{9}$$

where N_{OBS} and N_{SIM} are the number of cells containing biomass for the observed data and the simulated data respectively and N_W is the number of cells contained in the waterway. B_{SIM} and B_{OBS} are the total biomass in the simulated data and the observed data respectively. Values of M_S close to zero and M_B close to 1 indicate the best fit. M_S compares the total number of cells containing biomass, the spread, and M_B compares the actual biomass. The term 'best fit' here only indicates the amount of cells containing biomass and does not pertain to the configuration of patches. Also, the carrying capacity, K, is taken into account. However, K is not uniform over the data sets. This variation in K from the initial data structure allows for additional modelling freedom but makes it difficult to obtain a suitable estimate value for M_B. Similarly the incorporation of additional growth arising from seagrass makes interpretation of simulated results difficult when compared to observational data. As the system increases in complexity the simple metric considered here becomes inconclusive.

Both visual and metric data will be used to gain the best fit of the simulated data. Table 1 gives the different values of the depth limitation, *d,* the threshold of surrounding densities, *p,* the percentage increase in growth due to seagrass, *r,* the cell carrying capacity, *K,* and time, *t,* which is set to 12 months and is compared to Lake Conjola observed data from March 2003 to March 2004.

Table 1. Values of metrics M_B and M_S for varying parameter values of *g, p, r* and *d.*

g	*p*	*r*	*d*	*K*	*t*	M_B	M_S
0.004	2	0.1	10	3	12	1.437	0.008
0.005	2	0.1	10	3	12	1.429	0.006
0.006	2	0.1	10	3	12	1.426	0.004
0.007	2	0.1	10	3	12	1.426	0.004
0.008	2	0.1	10	3	12	1.427	0.005
0.007	0	0.1	10	3	12	2.830	0.189
0.007	1	0.1	10	3	12	1.461	0.010
0.007	3	0.1	10	3	12	1.451	0.007
0.007	4	0.1	10	3	12	1.411	-0.003
0.007	4	0.1	9	3	12	1.593	0.014
0.007	4	0.1	11	3	12	1.375	-0.005
0.007	4	0.2	10	3	12	1.417	-0.004
0.007	4	0.3	10	3	12	1.427	-0.002
0.007	4	0.4	10	3	12	1.432	-0.001

Table 1 indicates that that as *g* is increased to 0.007 both M_B and M_s improve. Increases in *p* and *d* suggest that the *C. taxifolia* begins to spread, or decrease in the case of *d,* with increased awareness of surrounding cells. Variation in *r* increases the biomass of a cell by a proportion of its value. Based on these results the parameters chosen are listed in Table 2, with validation from Lake Conjola March 2003 and March 2004 data. Figure 7 gives a good indication of the winter decay and summer growth for a one year simulation at Lake Conjola.

Figures 8 and 9 map the simulated spread using the parameters found in Table 2 for Lake Conjola and Narrawallee Inlet respectively which can be compared with Figures 2 and 3.

Table 2. Parameter values suitable for simulation.

g	*p*	*r*	*d*	*k*	*t*
0.007	4	0.4	10	3	12

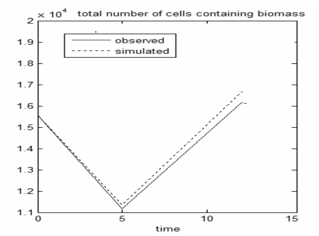

Figure 7. Decay in biomass during winter, Lake Conjola.

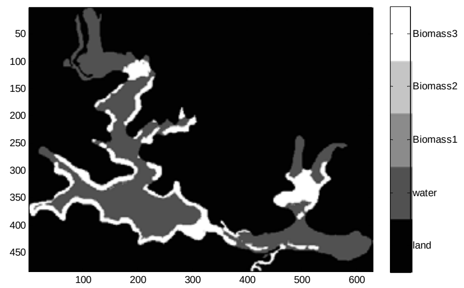

Figure 8. Lake Conjola, March 2004 Simulated Data.

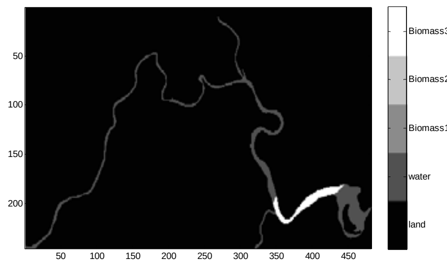

Figure 9. Narrawallee Inlet, April 2005 Simulated Data.

4. DISCUSSION AND CONCLUSION

We have presented the results of a grid cell model which describes the growth and spread of *C. taxifolia* in two waterways in coastal NSW, Australia. The model algorithm used information on prevailing weather conditions, seagrass distribution and some depth assumptions. The model was initialised using observed data from March 2003 and was validated by comparison between simulated and observed data from March 2004. The results indicate that the model describes the growth and spread of *C. taxifolia* in Lake Conjola very well according to simple metrics developed by McArthur et al. (2006). Further study will include adapting the model to different species of seagrass with similar properties of fragmentation and seasonal growth.

We have also presented increased complexity in cell models using a simplified, one dimensional interaction model as the foundation and discussion on appropriate growth and spread factors for these models. Increased complexity lead to multiple dimensions and parameters that in turn needed to be simplified to the available data and optimised using appropriate metrics. This however, created the levels of complexity, particularly in variational data sets between the observed and simulated data, which lead to inconclusive simulation results for particular metrics. In other words, it is very difficult to determine how to measure the simulated and observed data and to interpret the metrics. Further study will look into these issues.

Acknowledgements

Authors would like to thank the NSW Department of Primary Industries and the RMIT Emerging Researches Grant for funding. Data were supplied by NSW DPI.

5. REFERENCES

Colasanti, R.L., R. Hunt and L. Watrud (2007), A simple cellular automaton model for high-level vegetation dynamics, *Ecological Modelling,* 203, 363-374.

Glasby, T.M., R.G. Creese and P.T. Gibson (2004), Eradicating and preventing the spread of the invasive alga Caulerpa taxifolia in NSW, *NSW Fisheries,* 64.

Glasby, T.M., R.G. Creese and A.R. Davis (2005), Experimental use of salt to control the invasive marine alga Caulerpa taxifolia in New South Wales, Australia, *Biological Conservation,* 122, 573-580.

Glasby, T.M. and P.T Gibson and S. Kay (2005), Tolerance of the invasive alga Caulerpa taxifolia to burial by sediment, *Aquatic Botany,* 82, 71-81.

Hill, D., P.Coquillard, J. de Vaugelas and A. Meinesz (1998), An algorithmic model for invasive species: Application to Caulerpa taxifolia (Vahl) C. Agardh development in the North-West Mediterranean Sea, *Ecological Modelling* 109, 251-265.

McArthur, L., S. Schreider, A.A Shabani and T. Glasby (2006), Modelling the spread of *Caulerpa taxifolia* in Lake Conjola, Southern NSW. Proceedings of the 15th IASTED International Conference on Applied Simulation and Modelling, 2006. June 26-28, 2006, Rhodes, Greece.

Meinesz, A., L. Benichou, J. Blachier, T. Komatsu, R. Lemee, H. Molenaar and X. Mari (1995), Variations in the structure, morphology and biomass of Caulerpa taxifolia in the Mediterranean Sea. *Botanica Marina 38*, 499-508.

Meinesz, A. (1996, Killer Algae. *University of Chicago Press.*

Meinesz, A., T. Belsher, T. Thibaut, B. Antolic, K. Ben Mustapha, C.F. Bourdouresque, D. Chiaverini, F. Cinelli, J.M. Cottalorda, A. Djellouli, A. El Adeb, C. Orestano, A.M. Grau, L. Ivesa, A. Jaklin, H. Langar, E. Massuit-Pascual, A. Perino, L. Tunesi, J. de Vaugelas, N. Zavodnik and A. Zuljevic (2001), The introduced green algae Caulerpa taxifolia continues to spread in the Mediterranean, *Biological Invasions, 3,* 201-210.

Meinesz, A., and B. Hesse (2001), Introduction of the tropical alga Caulerpa taxifolia and its invasion of the northwestern Mediterranean, *Oceanologica Acta14*, 415-426.

NSW Department of Primary Industires, Caulerpa Distribution, Online: Acessed 1 July 2007, URL: http://www.fisheries.nsw.gov.au/threatened_species/general/species/caulerpa_distribution.

Ruesink, J.L. and L. Collado-Vides (2006), Modeling the increase and control of *Caulerpa taxifolia*, and invasive marine microalga, *Biological Invasions 8*: 309-325.

Wilson, J.R.U, D.M. Richardson, M. Rouget, S. Proches, M.A. Amis, L. Henderson and W. Thuiller (2007), Residence time and potential range: crucial considerations in modelling plant invasions, *Diversity and Distributions* 13, 11-22

Wright, J.T. and A.R. Davis (2006), Demographic feedback between clonal growth and fragmentation in an invasive seaweed, *Ecology* 87(7), 1744-1754.