

¹The design of a spatially explicit stochastic model for the simulation of oceanic seed dispersal

Julio Martín-Herrero and María Calviño-Cancela

¹Applied Physics Dept., ²Ecology and Animal Biology Dept.

Faculty of Sciences, University of Vigo, SPAIN

julio@uvigo.es

Abstract

We propose a stochastic model for analyzing the influence of the physical environment on oceanic seed dispersal. The model requires a description of the spatial configuration of the region and of the dynamics of the physical processes involved (ocean and atmosphere dynamics). The degree of accuracy and resolution of these descriptions will determine the accuracy of the simulations. Parameters to characterize the species of interest will come from empirical knowledge about the species, via experiments and measurements in the field. Two modeling approaches could be used, a particle model or a density model. We discuss both options and provide some recommendations. The described model could be used to answer questions about the origin of the current spatial distributions of species dispersed by the ocean and to study metapopulation dynamics.

1 Introduction

Dispersal is a very important process in the life of most organisms, playing a major role in their geographical distribution (Brown, Gibson, 1983; Thorne, 1972) and having a direct effect on the dynamics of metapopulations (Hanski, Gilpin, 1997; Ims, 1995; Saunders *et al.*, 1991). In the South Pacific region, oceanic dispersal plays an important role in the distribution and dynamics of plants. Many species, especially those dwelling on or near the shoreline, have fruits, seeds, seedlings, or vegetative parts that float on salt water. These propagules are passively dispersed by the sea.

Although dispersal is continually occurring, not all dispersal events play a role in metapopulation dynamics or in the geographical distribution of a taxon. Besides being able to travel to a given location, a propagule has to be able to survive, grow and reproduce upon arrival. Moreover, in the case of colonization, it has to establish a viable population. Therefore, together with the capability of each species to establish and maintain populations (Ebenhard, 1991), a series of factors will determine the success in the spread of a species or the exchange processes involved in metapopulation dynamics. These are spatially explicit factors (Wiens, 1997), such as the geographical distribution of suitable habitats and ocean dynamics. As sessile organisms, plants cannot actively select their habitat or their breeding partners, whereas mobile organisms integrate heterogeneity over relatively broad scales, and therefore use a coarse scale in their interaction with habitat (Kotliar and Wiens, 1990; With, 1994). The importance of spatial specificity is higher in plants due to this sessile character. For this same reason, stochasticity is greater for plant populations than for actively mobile organisms. Furthermore, the number of pioneer propagules reaching a new location is often very small, increasing the risk of stochastic extinction (Given, 1994; Lande, 1993).

Phylogenetic molecular analysis can help identify the origin of different populations and the relations between them (Avice, 2000). However, genetic and geographical distances are not necessarily correlated. When this happens the question of why a species reached some places but not others arises. A spatially explicit stochastic model can show which are the more probable dispersal routes for a particular species. Here we propose a model that can be equally valid at different spatial scales, and for different species. The scale of the model is given by the input data sets describing the spatial characteristics of the region of interest. The species is modeled using a set of specific parameters. The spatial data sets required for such a model are as follows:

- (i) *The spatial distribution of land and ocean, as a digitized binary cell matrix in which each cell is of one of two types: ocean or land.*
- (ii) *A digitized vector field measuring drag in the ocean cells. This represents the displacement (distance and direction) per time unit that a seed will experience if adrift in the cell. Depending on the complexity and accuracy of the simulation, displacement can be a constant, a random variable characterized by mean and variance, or can follow a known variation with time. This will depend on available knowledge of oceanic currents and tides, and, depending on the species, on geostrophic and local wind patterns.*
- (iii) *A digitized scalar field measuring habitat suitability in the land cells, representing the probability of successful settlement of a healthy seed that arrives at that land cell. This primarily depends on substrate, but it could also be influenced by factors such as elevation, inclination, climate, predators or presence of competing species. This categorization of habitat suitability of the land cells depends on the species of interest.*

For the purpose of this model, plant individuals can be in one of three stages. For simplicity in the notation, we will call them seed (propagule or mobile form of the plant, usually seeds or fruits, but sometimes seedlings, as in *Rhizophora* spp.), immature (rooted seedling and juvenile; i.e. sessile plants that are not able to produce seeds) and mature (fertile). The parameters characterizing the species in the model are described later.

The working of the proposed model is straightforward. The objective is to simulate the process of oceanic seed dispersal in 'quick motion'. Using fast desktop computers we can simulate the dispersal of one species over decades or centuries in just a few hours or days, depending on the machine and the complexity of the model. We select a number of land cells as populated with given densities at the beginning of the simulation and, when the desired period of model time has elapsed, we obtain a final population distribution. We can watch a 'movie' of the evolution of the populations along time, where the model shows the most probable interaction routes between locations.

Oceanic seed dispersal is a chaotic process (another reason to use a stochastic model), and thus a slight variation in the initial conditions can induce a great variation in the final results. Therefore, several simulations would have to be run, and averaged, in order to obtain a probability measure of the expected final distribution and interaction routes. These can then be compared with the actual distribution of the species, or with available knowledge of the species' metapopulation dynamics.

Here we describe the input data sets and parameters that the proposed model will require, and the way that it will work. We also provide some remarks on the model and its applicability. The reader should bear in mind that this paper just describes the model, and no specific implementation or results of simulations are given here.

2 Input data sets

The model requires three main input data sets. The first of them we call the *Map* and is just a matrix of binary cells. Each cell represents a plot of surface in the region of interest. The size of the plot is determined by the number of cells we are going to define in the region (see the section below for the model parameters). Cells with a *Map* value of 0 are cells where the seeds can be moved by ocean currents, and those with a *Map* value of 1 are those where the seeds can settle and plants grow if the habitat is suitable. The easiest way to obtain the *Map* matrix is from a raster geographical map. We can scan a suitable map of the region, adjust the scale and assign the appropriate value to each pixel.

The second input data set is the *Drag*, which is a two-dimensional vector matrix whose element $Drag_{xy}$ is defined only if Map_{xy} is 0. The two-dimensional vector $Drag_{xy}$ is added to the position vectors of the seeds that are adrift in the cell x,y each time unit. Every seed has a position vector that determines its position. The $Drag_{xy}$ elements can be constant vectors, or a set of means and variances. If such, they are handled as four-dimensional random variables with known distribution (we use new values each time unit for the variables, according to their probability distributions). In the most complex form, these four-dimensional vectors (mean and variance of drag strength in x and y directions) will change with time, according to available knowledge of atmosphere and ocean

dynamics. These values can be obtained dynamically at run-time from a known set of equations describing the physical model (if any) or can be read from a previously calculated look-up table.

The surface current patterns can be obtained from sensors on oceanographic satellites. NOAA's AVHRR (Advanced Very High Resolution Radiometer) has a 1.1 km resolution covering any expanse of ocean between 81°N and 81°S. It is possible to obtain a Sea Surface Temperature (SST) map from infrared images of the region and compute a displacement vector field using sequential images. This method, of maximum cross-correlation (MCC), was devised by Ninnis *et al.* (1986) to track sea ice displacements, and Emery *et al.* (1986) demonstrated that it can be used to obtain surface displacement patterns, using shallow drifters and a CTD (Conductivity and Temperature Data-Logger) survey to confirm the patterns and velocities of the SST-infrared currents (Emery *et al.*, 1992; Kamankian *et al.*, 1990; Kelly and Strub, 1992).

Additionally, if the shape and buoyancy of seeds are such that wind has a significant drag effect on them, surface wind vectors can be computed, using geostrophic winds computed from atmospheric pressure fields of 500 or 800 mb surfaces. The magnitudes of the geostrophic winds are reduced by a factor of 0.7 and their direction rotated clockwise in the Southern Hemisphere by 15°. This method (Fofonoff, 1960) is still used by weather forecast systems in the Pacific, although the dynamics of the ocean-atmosphere boundary layer dynamics is not well understood, and there is no agreement about the appropriate value of the drag coefficient for wind speeds greater than 5 ms⁻¹. An alternative is the use of satellite microwave scatterometer data to reliably measure wind speed and direction. This has been demonstrated using SEASAT data and ship and buoy observations (Brown, 1983). Today the SEASAT mission is finished, but all data are archived; alternatively there are new and more accurate scatterometers on board satellites such as the ERS series.

The third input data set is the *Habitat*, describing the habitat suitability of the land cells. This is an integer valued matrix defined for cells in which Map_{xy} is 1. The values, from 0 to 100, are used to compute the spatial specific probabilities of a fresh seed becoming a seedling, to modulate the average number of seeds a mature individual can produce, and, together with the life span, the probability of death of each plant. The values are derived from type of substrate, elevation, inclination, climatic constraints, presence of competing species, and predators. *Habitat* will depend on the species of plant under study and on the data available for the region (such as geological charts or vegetation land cover maps). One simple way to construct the matrix is to divide the different major types of habitats found in the region into several categories, ranging from "very good" to "inadequate", assigning a corresponding probability of success to each category. A map can be drawn with the different types of habitat, using a raster image of *Map*, in which a land cell is printed on the screen using a black pixel and an ocean cell with a white pixel. The color of each black cell can then be changed to one of a set of previously selected colors to represent the different types of habitat. Converting the raster image to a matrix of integers (each color is assigned a different value) we obtain *Habitat*.

It is possible to merge *Map* and *Habitat*, in order to optimize the algorithm at the core of the model and

economize memory storage. Only one integer-valued matrix, *Map*, would be necessary if a negative value is used for the ocean cells (such as -1) and a value from 0 to 100 is used in the land cells to represent habitat suitability.

3 Parameters

The *spatial scale* of the model is defined by the cell size and the size of the study region. These affect computation time, the accuracy of the results, and the amount of data needed to run the simulation. For a given cell size, the bigger the region the greater the number of cells, and the more the computations to be done per time unit. For a given region, the smaller the cells, the greater the number of cells and the greater the accuracy of the results, but only while cell size is coherent with the resolution of the available spatial data. If we use a finer cell resolution than that of the spatial data, we may think that we have an improved accuracy in the results, but this will not be real, and we will only be increasing the computational overhead. It is similar to what happens when we apply a digital zoom to a raster image before applying a filter: we have more pixels to deal with, but not more information.

To define the *temporal scale* we have to decide how much time we want the simulation to cover and how much real time a model time unit represents. If the latter is one day, for example, then to simulate one year we need to run the simulation 365 model time units. The magnitudes of the values of *Drag* have to be adjusted to the magnitude of the time units. Notice that each time unit is an entire run throughout the main algorithm of the model. Thus, if we are going to make a simulation of 1,000 years with a time unit of one day, the computer will have to run the main algorithm 365,000 times. If we use a time unit of one week, the computational overhead will be reduced but we will lose temporal resolution, as reflected by the rescaling of the values in *Drag*. For example, when using a dynamic *Drag* we cannot take into account tidal flows if we use time units greater than 6 hours. Also seeds will travel clumped in "time packets" if we use long time units, instead of being set ashore with different time delays.

An extensive set of parameters characterizes a species, some of which can be hard to measure. It is up to the modeler to decide the degree of required accuracy. Although generic values could be used for some parameters, successful modeling of one species will be difficult without a good knowledge of its biology. The required parameters are as follows:

Seed dormancy: In some species, seeds have a period of inactivity, failing to germinate even under favorable conditions (Harper, 1977). This parameter is the average time that must elapse before the seed is able to become a seedling. If the seed's dormancy period has not finished when arrives at land, it will have to wait until the end of the dormancy period.

Immature mortality: Proportion of immature individuals that do not become adult plants (mean, μ , and variance, σ). A probability of death for each immature individual is computed. It determines the proportion of successfully dispersed seeds that become mature plants.

Maturity time: Average time for an immature to become a fertile individual, capable of producing an average number of viable seeds. Each surviving 'seed' begins to produce seeds when its life clock is greater than the sum of dormancy time (if any) plus maturity time (in dioecious species only if it is female).

Sex ratio (if dioecious): Determines the proportion of each sex in the initial populations. Only females produce seeds, and only if a male is available within the *mating range* (see below).

Mating range: An estimate of the maximum distance that pollen can travel. It depends on the pollinating strategy of the species (e.g. foraging distance of pollinator insects for entomophilous plants). If no mature plant is available within the mating range of a given plant, it cannot produce seeds. If the species is dioecious, the individual within the mating range has to be male. In monoecious self-compatible species a mating range of 0 is used, meaning that no other plant in the neighborhood is necessary for successful production of seeds.

Number of viable seeds per individual: Number of viable seeds a mature plant can produce per season (μ and σ). With this and *fruiting season* we obtain the probability of a plant producing a given number of viable seeds each time unit.

Fruiting season: Period of the year in which mature plants produce seeds. It is expressed as time from the beginning of the year (time 0) to that of the season, and duration of the season. If this is one year, seeds are produced all year round.

Number of seeds per dispersal unit: Some species will favor the simultaneous arrival of several individuals in order to increase the probability of successful colonization (especially in dioecious species, which need individuals of both sexes to found new populations). This can be achieved by clumping together several seeds in one dispersal unit (usually fruits).

Proportion of seeds from shoreline that reach the ocean: Proportion of seeds produced by an individual situated right on the shoreline that will reach the ocean (μ and σ). This depends on how far shoreline individuals can be from the water, and the characteristics of the terrain. In mangroves, for example, this parameter will be high, because almost all of the seeds fall directly to the water. This parameter is used together with the next.

Maximum distance from the shoreline for oceanic dispersal: Estimate of the average distance from the shoreline to the first inland individual whose seeds cannot reach the ocean (μ and σ). It can be a fraction of the cell size or it can be several cells long. This and the prior parameter are used to compute the distribution probability of a seed reaching the ocean relative to its distance from the shoreline.

Mature individual canopy radius: Average radius of the canopy of a mature plant (μ and σ), used to compute the probabilities of a seed falling at a given distance from the stem. This is used to simulate regeneration within the same spot of land with suitable habitat, i.e. viable seeds that do not reach the ocean. Asexual reproduction is not specifically included, but can be simulated using the probabilities of seeds falling to the ground that will not reach the ocean.

Viability time of floating seeds in seawater: Average time a seed remains viable while afloat in seawater (μ and σ) from which we compute the probability that a floating seed loses its viability. When it happens, the seed disappears (unsuccessful dispersion). If this parameter is unknown, we can use the next one as an approximation (we assume that as long as the seed remains buoyant it is viable).

Floating time of seeds in seawater: Average time a seed remains positively buoyant in seawater (μ and σ).

With this value we can predict the probability that a seed will sink. When this happens, the seed disappears (unsuccessful dispersion).

Maximum density of mature individuals: Maximum number of mature individuals per unit of surface (cell). Land cells with such a number of mature individuals are not suitable for seedling growth, and therefore this is the maximum number of individuals that one land cell can contain.

Life span: Average number of years (μ and σ) that an individual lives, from which we compute a probability of death for each plant. The probabilities are conditioned on suitability of habitat in each specific location. Once a plant is dead, it disappears, and is not considered further in the computation of density in the cell.

We can simulate the occurrence of local catastrophes, as many possible types as desired: e.g. forest fires, floods, draughts, earthquakes and temporal climatic alterations. Each is characterized by the mean and variance of three parameters: the *frequency of occurrence* is the average number of occurrences per 100 years of model time, the *spatial range* is the average affected area, and the *survival proportion* is the average proportion of individuals that will survive the catastrophe inside the affected area. Whenever a catastrophe of a specific type happens, the model assigns values to two random variables (affected area and proportion of survivors) according to their probability distributions. Then it randomly selects a location as the center of the affected area and kills the required number of plants using random selection.

4 The model

The proposed model is based on a subset of stochastic processes called branching processes (Asmussen, Hering, 1983). A branching stochastic process is the mathematical model of a population of individuals that live, give birth to a finite number of individuals and die. Each individual lives a random time L and generates a random number of individuals, ν , that behave in the same way. Models of branching processes differ depending on the process of reproduction and on the probability distribution of the vector (L, ν) . The simplest form is the Bienayme-Galton-Watson process (Bienaymé, 1845; Keyfitz, 1985), where each individual has a life length of one time unit ($L=1$) and produces the new individuals at the moment of its death. If L is independent of ν and exponentially distributed the process is a Markov branching process (Ethier, Kurtz, 1985). If L and ν are considered independent and identically distributed for different individuals, but with arbitrary distributions, the process is not Markovian, and is called a Bellman-Harris process (Chauvin, 1986; Harris, 1948). If in a Bellman-Harris process we allow dependence of ν on L , we have a Sevastyanov model, which is an age-dependent process (Kaj, Sagitov, 1998). If we assume that individuals produce children at randomly chosen instants during their live, then it is a Kramp-Mode-Jagers model (Jagers, 1975). The model has immigration if we allow new individuals, not born of current individuals, to appear in the population at different random times, and we have a branching process with disasters if at random times we allow the occurrence of disasters in which each individual has a given probability of surviving (Sankaranarayanan *et al.*, 1978).

Within this frame, our model is a set of concurrent Kramp-Mode-Jagers branching processes (i.e. populations)

with immigration, subject to disasters, evolving on a spatially explicit environment that conditions the probability distributions. The core of the model is the logic that is performed every time unit of the simulation. Here we describe the basics of the cycle. The flowchart in Figure 1 can be used as a roadmap to the explanation that follows.

Besides the input data sets and parameters described above, the model uses another data set that up to now has been implicit in the text: the list of individuals (seeds and plants), ideally implemented as a dynamic list. The simulation starts generating a list from the initial conditions of the simulation: populated cells and average density. With these data the model generates the required number of initial individuals, assigning them to specific random locations within the selected cells. Each individual is assigned a random age within the species' life span, and, if dioecious, a random sex according to the sex ratio. We also need a flag indicating if an individual is a seed (i.e. if it still can be 'moved') or has begun its sessile life (i.e. immature or mature plant). There are different methods to generate random values for the variables according to their probability distributions: inverse transformation, rejection, composition, table look-ups and specific algorithms. See, for example, Schmeiser (1980).

Once the clock begins to run, each mature female with a mature male within mating range will produce a random number (with known distribution, conditioned on the suitability of habitat in the specific location) of new individuals (i.e. seeds). If the species is not dioecious, each mature individual will produce that same random number of seeds if some other mature individual is available within the mating range (unless the species is self-compatible). Depending on the distance to the shoreline, each new seed will have a probability of reaching the ocean. The ones that fall on land will have to wait for the dormancy period (if any) to finish and then, depending on habitat suitability, will have a probability of becoming immature individuals. Therefore, some will live and some will die (as a seed, seedling or juvenile).

Individuals that reach the ocean will change their location each time unit (if they are in ocean cells with non-zero *Drag*). Each seed will have a probability of remaining afloat and a probability of remaining viable. If any of these fail, the individual disappears. If a seed reaches land, the habitat suitability of the land cell is checked and the seed will have a specific probability of rooting (becoming an immature individual), after the intrinsic dormancy period (if any). The seed may become a mature individual after the necessary time units (depending on the proportion of immature plants that do not reach maturity and on habitat suitability). When this happens, if the necessary conditions are met, it will begin to produce seeds, some of which will be dispersed through the ocean. The seeds that fall to the ground next to the mother plant can eventually become a new mature plant. Thus, with time, some plants will grow on new locations, produce seeds and thus start a new population, while the initial population keeps growing. A probability of death is computed for each mature individual based on its age, the life span of the species, and the habitat suitability.

All of these steps are accomplished each time unit through the entire list, checking and altering the status of each individual. The actions to be taken are functions of the parameters, the current status of the individual (i.e. age, sex, location), location (i.e. density of individuals and

habitat suitability in land cells, current drag in ocean cells), and model (i.e. fruiting season, catastrophes).

This model is a “particle” model; i.e. it models each individual as a separate unit, instead of a “density” model in which only densities of seeds and plants in each cell are considered. The particle model is more accurate: we follow the seed through the ocean waiting for it to lose its buoyancy or reach land, and see each plant grow and die on specific locations. However we need a given amount of memory for each individual in the model (about 50 bytes, depending on the implementation), and therefore several megabytes are needed to simulate a reasonably sized region. If the computer is not powerful enough, we can scale down the number of plants and seed production. The main (most probable) dispersal routes and final emplacements of the populations still should be shown in the final results. The simulation can be run with different number of seeds and plants to investigate the effect of this rescaling on the results.

A density model does not need a list of individuals. It simply has two additional fixed size matrixes, one of plant densities, *Plant*, and one of seed densities, *Seeds*. *Plant_{xy}* elements are two-dimensional vectors representing the number of plants in the land cell *x,y* in the different stages (number of immature and mature individuals in the cell). In the case of dioecious species, we need four values per element (male and female immatures, and male and female mature plants). *Seeds_{xy}* elements have as many dimensions as stages are distinguished in the seeds at cell *x,y* (land or ocean). As a minimum: number of buoyant seeds, and number of dormant (if applicable) seeds. Probability distributions are then used to transfer seeds and plants from one stage to another inside the same cell. Seeds are also transferred from one *Seeds* element to another according to *Drag*, and from *Seeds* to *Plants*. Seeds originate in land cells *x,y* in which *Plants_{xy}* has any number of mature plants.

A combination of both particle and density models can give a good degree of accuracy whilst keeping the computer requirements within manageable bounds. It uses a list of individuals only for seeds, and a density matrix for plants. We recommend this approach, because it preserves accuracy in the most delicate part of the process (the ocean travel and seed settlement), specifically regarding buoyancy times, whilst permitting realistic simulations (regarding population densities and seed production) with affordable hardware requisites. Simple discrete-time population models can then be used to simulate the population dynamics, and therefore some of the species’ parameters would not be necessary. For example just a growth rate, *r*, and an equilibrium density, *K*, are needed if the Ricker model is used (Hanski, 1999).

The simulation ends when the desired time units have elapsed, the population has reached a stable distribution, or the major routes of interaction have been established, depending on the objective. We can use the list of individuals or the *Plants* density matrix to plot the final distribution on the raster image of *Map*. The algorithm can compute run-time statistics such as successfully dispersed seeds versus seeds produced per population, or percentages of seed inflow and outflow between populations, and it is possible to draw flow lines on the raster map to show the main dispersal and interaction routes. This can help explain why some locations are populated and some are not, or why some populations interact and some do not.

Due to the stochastic nature of the model, the simulation should be run several times with the same parameters (the exact number depending on the variance of the successive results) in order to compute a final distribution for a given set of parameters, by averaging the results of the different simulations.

If different varieties of plants are labeled and used simultaneously, it is possible to allow individuals to cross and thus have in the final distribution “hybrid” individuals. This requires another field in the records of the list of individuals, a numeric label identifying the variety. At the end of the simulation, different varieties and their hybrids can be plotted in different colors. The final results will reflect the distribution of the different varieties conditioned on the presence and dispersal of the rest.

The model can also be used to study the effect of specific parameters (e.g. seed buoyancy, life span) in the dispersal capabilities of species. We just have to run several simulations with different values for some of the characteristic parameters of a species (even a theoretical one) and compare the results.

It is possible to perform spatial analysis on the results, using an algorithm to label the different patches, to be able to compute statistics regarding nearest neighbors, average distances and sizes, or connectivity. One such algorithm is described in Martín-Herrero and Peón-Fernández (2000).

5 Some remarks on the model

Whenever dealing with a model we should keep in mind that the model is just an approximation to reality. When we choose the factors we are going to use to act on the variables, we can only choose the main ones from an infinite set, i.e., those who will influence the final result above the noise threshold we are ready to allow. Sometimes we cannot include some factors in the model, because we do not know their relation with the output, or because we do not know how to include them, or just because their inclusion would mean taking the model to an undesirable degree of complexity. We should keep in mind that even when we are not including these factors in the model they still are acting on the real output, and thus we should interpret the results of a simulation using this knowledge. One such a factor in the model we have just described is the human influence on seed dispersal. Another is the tidal flows and seasonal variations in ocean dynamics if we use a constant (annual averaged) *Drag* matrix. Although a model is only an approximation to reality, we are choosing how good the approximation is.

Within these considerations it could seem that using probability distributions and random variables instead of deterministic parameters blurs the final picture. However, the stochasticity of the model plays a major role in focusing the results. When we use a random variable, with a given probability distribution, as a parameter to characterize some factor we want to include in the model, we are not only including the factor itself, but all the “fuzzy” set of causes that contribute to the randomness of the factor. Therefore, using stochastic modeling we are modeling not only the deterministic part of the result, but the noise too. Thus, we are implicitly modeling an infinite number of smaller factors together with the main factors we are explicitly including in the model.

Not every question concerning the actual distribution of a particular species can be answered with the proposed model. Apart from the non-deterministic nature of the

process, it is impossible to account for all the factors potentially involved. Many factors may change over time, and catastrophes happen. Although we can introduce in the model the occurrence of local catastrophes as a coarse approximation, we would probably have to use modern patterns of oceanic currents to simulate those of the non-recent past. Nevertheless, the proposed model can provide us with a better insight onto the problem of oceanic seed dispersal and, when used together with other information (genetic distance, sedimentology, historic records) it can complement the study of the present distribution of oceanic dispersed species in the Pacific. On the other hand, if the objective of the study is metapopulation dynamics, we only need the modern spatial patterns of ocean dynamics and habitat suitability, and therefore the actual interactions between populations can be obtained.

6 Acknowledgments

The visit of Maria Calviño-Cancela to the Department of Biology of the University of the South Pacific in Suva (Fiji) was partially funded by the *Secretaria Xeral de Investigación e Desenvolvemento da Conselleria de Presidencia da Xunta de Galicia* (Spain). We want to gratefully acknowledge the kindness and assistance of our host Dr. Shahina Ghazanfar and of the staff of the Department of Biology of the University of South Pacific, Suva, Fiji, and Dr. Martin Fisher for his patience and useful comments.

7 References

- Asmussen, S. & Hering, H. 1983. *Branching processes*. Birkhauser, Berlin.
- Avice, J.C. 2000. *Phylogeography: the History and Formation of Species*. Harvard University Press, Harvard.
- Bienaymé, I.J. 1845. *De la Loi de Multiplication et de la Durée des Familles*. Société Philomathématique. Paris.
- Brown, J.M. & Gibson, A.C. 1983. *Biogeography*. Mosby, St. Louis.
- Brown, R.A. 1983. On a satellite scatterometer as an anemometer. *Journal of Geophysical Research*, **88**, 1663-1673.
- Chauvin, B. 1986. Arbres et processus de Bellman-Harris. *Annales Institut Henri Poincaré, Probabilités et Statistiques*, **22**(2):209-232.
- Ebenhard, T. 1991. Colonization and metapopulations: A review of theory and observations. *Biological Journal of the Linnean Society*, **42**, 105-121.
- Emery, W.J., Fowler, C.W., Clayson, C.A. 1992. Satellite image derived Gulf Stream currents. *Journal of Oceanic and Atmospheric Science and Technology*, **9**, 285-304.
- Emery, W.J., Thomas, A.C., Collins, M.J., Crawford, W.R., and Mackas, D.L. 1986. An objective procedure to compute surface advective velocities from sequential infrared satellite images. *Journal of Geophysical Research*, **91**, 12865-12879.
- Ethier, S.N. & Kurtz, T. 1985. *Markov Processes: Characterization and Convergence*. Wiley, New York.
- Fofonoff, N.P. 1960. Transport computations for the North Pacific Ocean. FRB Canada. *Manuscript Report on Oceanography and Limnology*, 77-80.
- Given, D.R. 1994. *Principles and Practice of Plant Conservation*. Chapman and Hall, London. 13-35.
- Hanski, I.A. 1999. *Metapopulation Ecology*. Oxford University Press, Oxford.
- Hanski, I.A. & Gilpin, M.E., eds. 1997. *Metapopulation Biology: Ecology, Genetics and Evolution*. Academic Press, London.
- Harris, T.E. 1948. Branching processes. *Annales of Mathematical Statistics*, **19**, 474-494.
- Harper, J.L. 1977. *Population biology of plants*. Academic Press.
- Ims, R.A. 1995. Movement patterns related to spatial structures. In Hanson, L., Fahrig, L. & Merriam, G., eds. 1995. *Mosaic Landscapes and Ecological Processes*. Chapman and Hall, London.
- Kaj, I. & Sagitov, S. 1998. Limit processes for age-dependent branching particle systems. *Journal of Theoretical Probability*, **11**, 225-257.
- Keyfitz, N. 1985. *Applied Mathematical Demography* (2 ed.). Springer Verlag, New York.
- Jagers, P. 1975. *Branching Processes with Biological Applications*. J. Wiley Sons, New York.
- Kelly, K.A. & Strub, P.T. 1992. Comparison of velocity estimates from advanced very high resolution radiometer. *Journal of Geophysical Research*, **91**, 10725-10734.
- Kotliar, N.B. & Wiens, J.A. 1990. Multiple scales of patchiness and patch structure: A hierarchical framework for the study of heterogeneity. *Oikos*, **59**, 253-260.
- Lande, R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *American Naturalist*, **142**, 911-927.
- Martín-Herrero, J. & Peón-Fernández, J. 2000. Alternative techniques for cluster labelling on percolation theory. *Journal of Physics A: Mathematical and General*, **33**, 9, 1827-1840.
- Ninnis, R.N., Emery, W.J. & Collins, M.J. 1986. Automated extraction of sea-ice motion from AVHRR imagery. *Journal of Geophysical Research*, **91**, 10725-10734.
- Sankaranarayanan, G. & Krishnamoorthy, A. 1978. Branching processes with immigration subjected to disasters. *Journal of Mathematics and Physical Sciences*, **12**, 2, 165-176.
- Saunders, D.A., Hobbs, R.J. & Margules, C.R. 1991. Biological consequences of ecosystem fragmentation: A review. *Conservation Biology*, **5**, 18-32.
- Schmeiser, B. 1980. Random variate generation: A survey. In Oren, T.I., Shub, C.M. & Roth, P.F., eds. 1980. *Simulation with Discrete Models: A State of the Art Review* 79-104 pp. IEEE Press.
- Thorne, R.F. 1972. Major disjunctions in the geographic ranges of seed plants. *The Quarterly Review of Biology*, **47**, 365-411.
- Tokamamkian, R., Strub, P.T., McClean-Padman, J.M. 1990. Evaluation of the maximum cross-correlation method of estimating sea surface velocities from sequential satellite images. *Journal of Atmospheric and Oceanic Technology*, **7**, 852-865.
- Wiens, J.A. 1997. Metapopulation dynamics and landscape ecology. Chapter 3 of Hanski and Gilpin (1997).
- With, K.A. 1994. Using fractal analysis to assess how species perceive landscape structure. *Landscape Ecology*, **9**, 25-36.