

UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI FISICA E ASTRONOMIA "GALILEO GALILEI"
CORSO DI LAUREA TRIENNALE IN FISICA

APPLICAZIONI DEL VOTER MODEL IN ECOLOGIA

Relatore interno: Prof. Samir Suweis

Tesi di Laurea di
Emanuele Pigani
Matricola n. 1096156

ANNO ACCADEMICO 2016 - 2017

Indice

1	Introduzione	3
2	Modelli deduttivi per la descrizione di ecosistemi	5
2.1	La visione niche-assembled e il modello di Lotka-Volterra	5
2.2	Stocasticità demografica e Teoria Neutrale	6
3	Voter Model	8
3.1	Mean field Voter Model	9
3.1.1	Voter Model con speciazione	10
3.1.2	Mean field senza speciazione	11
4	Relative species abundance per il Voter Model	12
4.1	Barro Colorado Island	14
5	Species-area relationship per il Voter Model	15
6	Species Persistence Time per il Voter Model	18
7	Conclusioni	22
	Bibliografia	23

Abstract

In questo lavoro presentiamo un'applicazione in ambito ecologico di un modello di particelle interagenti, noto come Voter Model, originalmente sviluppato per descrivere la dinamica della formazione del consenso in una popolazione di individui. Attraverso una approssimazione di "mean field", abbiamo ricavato una descrizione analitica di alcuni pattern emergenti macro-ecologici noti in letteratura e abbiamo verificato i risultati così ottenuti con simulazioni numeriche. Abbiamo infine anche confrontato le predizioni del modello con alcuni dati.

Il lavoro presentato è basata sull'articolo di Review of Modern Physics 88(2016) "Statistical mechanics of ecological systems: Neutral theory and beyond [1].

Capitolo 1

Introduzione

É interessante contemplare una rigogliosa ripa fluviale, coperta di molte piante appartenenti a molti tipi, con gli uccelli che cantano tra i cespugli, i diversi insetti che svolazzano intorno e con i vermi che strisciano nel terreno umido, e riflettere che queste forme dalla struttura così complessa, tanto differenti le une dalle altre e dipendenti le une dalle altre in modo talmente complicato, sono state tutte prodotte dalle leggi che operano attorno a noi.

Charles Darwin, L'Origine delle Specie

Nel celebre “L’origine delle Specie” Darwin esprime tutto il suo stupore per le leggi che sottostanno alla natura: nonostante la notevole diversità di specie e di forme, la natura mostra profonde analogie su scale temporali, spaziali e di complessità organizzativa diversissime tra loro. Da più di mezzo secolo, infatti, in tutto il mondo gli ecologi hanno censito popolazioni di moltissimi ecosistemi provenienti da comunità differenti, come ad esempio da barriere coralline o foreste tropicali, osservando qualcosa di inaspettato: per quanto le condizioni biologiche e ambientali possano essere profondamente diverse in queste comunità ecologiche, è possibile formulare alcuni modelli macroecologici che rivelano l’esistenza di caratteristiche sorprendentemente simili nelle varie comunità (vedi Tabella 1.1). Questo suggerisce che esistano meccanismi di auto-organizzazione e che questi siano insensibili ai dettagli del sistema ecologico in esame. Se, per molti aspetti, il funzionamento e le caratteristiche di un certo ecosistema dipendono dalle proprietà biologiche delle singole specie e dalle loro interazioni, è altrettanto vero che i dati mostrano come certe proprietà “emergenti”(si veda la Tabella 1.1) siano comuni a una varietà di ecosistemi e passano quindi essere considerati universali [1].

La questione è dunque quella di comprendere quali siano i pattern comuni a tutti gli ecosistemi e quali processi ecologici fondamentali e condivisi guidino tale auto-organizzazione nelle comunità ecologiche. Questa affascinante sfida intellettuale si sposa perfettamente con il modo e lo stile di indagine con cui i fisici si avvicinano ai problemi scientifici. Analogamente alla meccanica statistica, che fornisce una teoria per mettere in relazione le proprietà microscopiche di singoli atomi e molecole con le proprietà macroscopiche dei materiali, anche l’ecologia ha bisogno di una

Tabella 1.1: Descrizione di alcuni pattern macroecologici

Relative species abundance (RSA)	Probabilità che una specie abbia n individui in una determinata regione. Se moltiplicata per il numero totale di specie nella regione, dà il numero di specie con n individui in tale regione. Talvolta in letteratura viene chiamata anche Species-abundance distribution (SAD).
Species-area abundance (SAR)	Funzione che collega il numero medio di specie S alla zona A in cui vivono. In una gamma relativamente ampia di scale spaziali è ben approssimata da una legge di potenza. Si dimostra che la SAR nella scala log-log ha tre comportamenti qualitativamente diversi a seconda della scala spaziale: approssimativamente lineare per scale molto piccole o molto grandi, mentre per scale intermedie segue leggi di potenza.
Species persistence time (SPT)	Densità di probabilità dell'intervallo di tempo tra la formazione e l'estinzione locale di qualsiasi specie all'interno di una determinata area.

teoria per mettere in relazione le proprietà biologiche chiave a livello individuale con proprietà macroecologiche a livello comunitario. Questo passo, tuttavia, non si può ridurre a una mera generalizzazione dell'approccio meccanico-statistico standard: al contrario della materia inanimata, le comunità ecologiche sono composte da entità che si evolvono, mutano e possono cominciare o smettere di interagire tra loro nel corso del tempo.

Pertanto, il problema al centro della fisica statistica dei sistemi ecologici è quello di identificare quali siano i processi fondamentali che devono essere integrati nei modelli per riprodurre i pattern emergenti noti ed eventualmente, alla fine, scoprirne di nuovi.

Come vedremo nel prossimo capitolo, esistono diversi modelli per descrivere la dinamica di specie ecologiche interagenti: partendo dai classici modelli di Lotka-Volterra, negli ultimi anni l'approccio verso lo studio di tali sistemi è cambiato a poco a poco, fino ad arrivare alla Teoria Neutrale, introdotta nei primi anni 2000 da Hubbel [2] e tuttora oggetto di ricerca (si veda [3]). La Teoria Neutrale, in particolare, rappresenta una vera e propria teoria termodinamica degli ecosistemi: similmente alla teoria cinetica dei gas ideali in fisica, essa incarna appieno lo spirito dei fisici, ben sintetizzato dalla celebre affermazione di Einstein *"Tutto dovrebbe essere reso il più semplice possibile, ma non più semplice di così."*

Capitolo 2

Modelli deduttivi per la descrizione di ecosistemi

Ci apprestiamo ora a presentare i due principali tipi di modelli deduttivi per la descrizione di ecosistemi, mettendo in luce alcune lacune dei modelli di tipo Lotka-Volterra e introducendo i concetti di stocasticità demografica e neutralità, che sono alla base della dinamica stocastica di modelli per ecosistemi.

2.1 La visione niche-assembled e il modello di Lotka-Volterra

In ecologia, nicchia - niche in inglese - è il termine con cui si descrive la posizione spaziale relativa di una specie o una popolazione all'interno dell'ecosistema in cui vive: essa descrive cioè come un organismo risponde alla distribuzione di risorse e di competitori. L'insieme completo delle condizioni ambientali, sia biologiche che fisiche, nelle quali un organismo riesce a vivere costituisce la sua nicchia fondamentale. In accordo con il principio di esclusione competitiva, due specie non possono occupare la stessa nicchia nello stesso ambiente per lungo tempo. La visione che si è più diffusa a partire da Darwin considera le comunità ecologiche come un insieme di specie diverse che interagiscono e competono al fine di conquistare ciascuno la sua nicchia fondamentale per giungere ad una condizione di adattamento vicino all'equilibrio demografico: tale prospettiva prende il nome di *niche-assembly*. In accordo con questa visione, le specie coesistono con le altre specie nella comunità in un equilibrio interagente, la cui stabilità e resistenza alle perturbazioni esterne deriva dalla capacità di adattamento di ciascun individuo delle diverse specie.

Storicamente, i primi modelli che descrivono sistemi ecologici interagenti secondo questa prospettiva sono quelli preda-predatore di Lotka-Volterra, in cui la dinamica di due specie, di cui una è una preda e l'altra un predatore, è governata dalla seguente

$$\begin{cases} \dot{x} = \lambda x \left(1 - \frac{x}{K}\right) - p y x \\ \dot{y} = -\mu y + p y x \end{cases} \quad (2.1)$$

dove x è la popolazione della preda, y quella del predatore, λ , μ e K sono rispettivamente il tasso intrinseco di crescita, il tasso istantaneo di mortalità dei consumatori e la capacità portante della risorsa, mentre p è il coefficiente di predazione, che descrive il tasso di consumo della risorsa da parte di un singolo consumatore.

Il modello di Lotka-Volterra è stato la guida teorica per molti modelli successivi, che hanno fornito una base teorica importante per la comprensione degli ecosistemi.

Molti sono, però, i limiti di questi modelli, che condividono l'idea di fondo che le specie possano essere descritte da modelli deterministici derivanti dal concetto fondamentale di nicchia, che, tuttavia, può essere definito correttamente solo a posteriori:

- essi sono essenzialmente deterministici e non riescono in questo modo a tener conto nella dinamica delle specie degli effetti stocastici: la maggior debolezza di queste teorie consiste nella loro staticità e nel fatto che non emergono in modo chiaro come necessaria conseguenza dinamica dei processi naturali;
- al crescere del numero delle specie, tali sistemi diventano analiticamente intrattabili e computazionalmente espensivi: il numero di parametri necessari per descrivere il sistema è elevato e tuttavia generalmente è incapace di interpretare i risultati sperimentali di comunità reali;
- essi contengono molti parametri difficili da stimare a partire dai dati sperimentali;
- è difficile introdurre le dipendenze dai gradi di libertà spaziali.

2.2 Stocasticità demografica e Teoria Neutrale

La prima teoria deduttiva per sistemi ecologici che si basa sull'equilibrio dinamico e che introduce, quasi inconsapevolmente, anche il concetto di neutralità in ecologia è quella contenuta in una famosa monografia del 1967 di MacArthur e Wilson [4], "The theory of island biogeography".

In questo lavoro, essi proposero che il numero delle specie che costituiscono una comunità locale in un'isola dipendesse da due forze contrastanti: da una parte, specie presenti sulla terraferma ma non sull'isola possono raggiungere e colonizzare quest'ultima; dall'altra, specie esistenti sull'isola possono estinguersi.

Queste ipotesi (formulate in un'epoca in cui dominava la teoria statica delle isole, secondo cui le comunità insulari possono essere considerate strutture fisse su scale temporali ecologicamente rilevanti) hanno origine nel fatto che le specie insulari sono generalmente molto più ridotte di quelle nella terraferma e le popolazioni piccole sono soggette ad alcune complicazioni. In particolare:

- effetto Allee: a basse densità, il tasso di mortalità può essere più grande di quello di nascita, a causa di difficoltà di accoppiamento o della maggiore esposizione alla predazione;
- inbreeding, ovvero l'accoppiamento tra individui parenti tra loro, che può portare a un deterioramento del patrimonio genetico;
- forte influenza della stocasticità demografica: la probabilità che muoiano contemporaneamente, per effetti puramente casuali, tutti gli individui di una specie è maggiore se la specie è composta da pochi individui.

Secondo le supposizioni di MacArthur e Wilson, quindi, la dinamica delle specie è governata dalla seguente equazione

$$\dot{S} = I(S) - E(S) \quad (2.2)$$

dove $I(S)$ è il tasso di immigrazione di nuove specie ed $E(S)$ è il tasso di estinzione delle specie. Ipotizzando che $I(S)$ sia una funzione decrescente nel numero di specie, dal momento che tante meno specie sono presenti sull'isola tanto maggiore è la probabilità che una specie che arriva

dalla terraferma sia nuova, e che $E(S)$ sia crescente e si annulli per $S = 0$, esisterà un valore \bar{S} in corrispondenza del quale i due tassi si eguagliano: si raggiunge così una condizione di equilibrio stabile.

Il modello di MacArthur e Wilson implica un radicale allontanamento dalla visione niche-assembled predominante all'epoca per almeno tre ragioni:

1. La loro teoria mette in luce la possibilità che la stocasticità ambientale e demografica possano giocare di per sé un ruolo ugualmente se non più importante nel meccanismo di strutturazione della comunità ecologica di quanto lo siano le regole di assemblaggio dovute alle nicchie.
2. \bar{S} è un equilibrio dinamico, perchè quello che rimane costante è il numero di specie, non la loro identità: $I(\bar{S}) = E(\bar{S})$ non implica che i due tassi siano nulli e, dunque, c'è un ricambio delle specie presenti sull'isola.
3. In accordo con la teoria della biogeografia delle isole, non importa quali specie contribuiscono al bilancio tra immigrazione e estinzione in un'isola. Tutte le specie nella teoria, dunque, sono trattate come identiche.

Per quest'ultimo motivo, anche se questo termine è stato introdotto successivamente, si può parlare di *neutralità a livello delle specie*: tutte le specie, avendo la stessa probabilità di immigrazione, di estinzione e di speciazione, sono trattate come equivalenti.

Più in generale, riprendendo la definizione data dall'ecologo americano Hubbel [2], un modello sarà *neutrale a livello di individuo* se considera gli organismi nella comunità trofica come essenzialmente identici nella loro probabilità pro capite di nascita, morte, migrazione e speciazione, indipendentemente dalla specie cui appartengono. Ne consegue che, sebbene l'abbondanza di specie diverse possa essere differente, dal punto di vista competitivo esse sono equivalenti: tutti gli individui di tutte le specie obbediscono alla stessa dinamica ecologica. La specie a cui il singolo individuo appartiene è dunque irrilevante al suo destino o al destino della sua discendenza e all'impatto che questa ha sugli altri individui.

Notiamo infine che, contrariamente alla teoria di nicchia e all'approccio adottato da Lotka e Volterra, la Teoria Neutrale si occupa principalmente di modellizzare comunità monotropiche.

Capitolo 3

Voter Model

In questa sezione vogliamo concentrarci sullo studio dettagliato di un modello che vuole descrivere la dinamica della formazione del consenso in una popolazione di individui: il Voter Model [5].

Nella versione classica del modello, ogni individuo è situato in una cella di un grafico e si può trovare in uno dei q equivalenti stati d'opinione: ciascun individuo non ha un'opinione fissa, ma la può cambiare con una certa probabilità adottando quella di uno dei suoi vicini. Ad esempio, nel caso più semplice ($q = 2$), possiamo pensare che ogni individuo sia caratterizzato da uno spin *up* o *down* [6], che può cambiare con una probabilità pari alla frazione dei primi vicini di opinione opposta¹, come esemplificato in Figura 3.1: il consenso si raggiungerà quando tutta la popolazione avrà gli spin allineati.

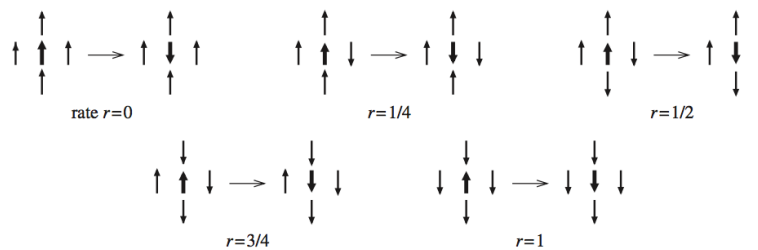


Figura 3.1: Probabilità r che l'opinione centrale cambi in funzione dei primi vicini su un reticolo quadrato.

Con questa dinamica, una popolazione finita di votanti arriverà ad uno stato di consenso in un tempo finito che dipende dalla grandezza del sistema, dalla dimensione spaziale e dal numero di opinioni q , come si può intuire dalla Figura 3.2. Sebbene a prima vista questo modello ricordi molto quello di Ising, tra i due emerge una profonda differenza: il Voter Model, al contrario di Ising, è un modello non all'equilibrio, perchè, raggiungendo l'assorbimento, non vale il bilancio dettagliato, ovvero non si ha reversibilità microscopica [6].

Il Voter Model è particolarmente interessante in ambito ecologico perchè può essere utilizzato per studiare la dinamica di specie che competono per le risorse: le celle, ognuna con la propria opinione, possono rappresentare le regioni spaziali di un ecosistema colonizzate ciascuna da un individuo appartenente a una delle q diverse specie (i.e. le q opinioni).

¹In letteratura i siti vicini con individui di opinione opposta sono definiti *active bonds*

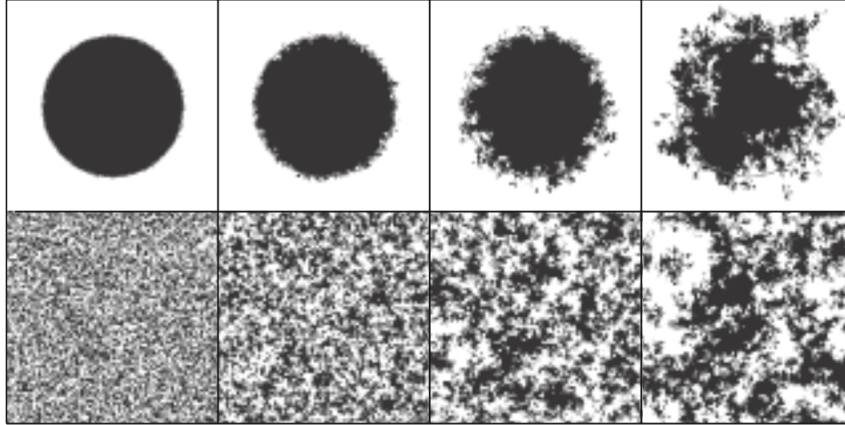


Figura 3.2: Evoluzione del Voter Model su un reticolo quadrato 256×256 [7]. La riga superiore mostra le istantanee a $t = 4, 16, 64$ e 256 , iniziando con una bolla circolare iniziale di raggio 180 avente lo stesso parere in un mare di opinioni opposte. La riga inferiore mostra la stessa evoluzione, ma con la condizione iniziale di avere i due stati di opinione distribuiti uniformemente.

3.1 Mean field Voter Model

Per semplicità, studieremo il Voter Model mean field, dove si perde la nozione di vicino: consideriamo le celle come se fossero tutto vicine l'una alle altre, di modo che ciascuna regione può essere colonizzata da un individuo di una qualsiasi altra regione, anche non fisicamente adiacente a quella colonizzata.

Consideriamo una griglia quadrata bidimensionale di lato L , con un numero J di siti (nodi). Ciascun nodo è un individuo rappresentato da un label che ne identifica la specie di appartenenza. Denotiamo con \mathbf{C} la collezione di tutti gli individui appartenenti alle S differenti specie e con $\vec{n} = \{n_1, n_2, \dots, n_S\}$ il vettore delle abbondanze per \mathbf{C} , dove n_j indica l'abbondanza di individui della j -esima specie.

Ci poniamo nell'ensemble microcanonico, imponendo che ogni individuo che muore venga sostituito immediatamente da un altro della stessa o di un'altra specie, ovvero

$$J = \sum_{i=1}^S n_i$$

La dinamica del mean field Voter Model a ogni istante di tempo è la seguente: un individuo selezionato casualmente muore e viene sostituito con un individuo di una nuova specie con una probabilità ν , mentre con probabilità $1 - \nu$ il sito viene colonizzato dalla prole di un individuo presente su un nodo selezionato casualmente nella griglia. A differenza della versione classica del Voter Model, dove l'interazione avveniva tra primi vicini, nel modello mean field essa avviene con qualsiasi nodo della griglia, senza che però ci sia dispersione sulla popolazione.

Il parametro ν , detto *diversification rate*, gioca un ruolo fondamentale nel Voter Model: a seconda che esso sia o meno nullo, la dinamica è completamente differente. Nel primo caso ($\nu = 0$) si parla di mean field Voter Model senza speciazione e in questo caso si avrà l'assorbimento delle specie: in un tempo finito tutti i nodi saranno occupati da individui della stessa specie (i.e. la popolazione dei votanti ha raggiunto un consenso). Se, invece, $\nu > 0$ si parla di mean field Voter Model con speciazione: in questo caso, in un tempo finito si raggiunge un equilibrio dinamico della distribuzione delle specie presenti.

3.1.1 Voter Model con speciazione

Notiamo innanzitutto che, per le ipotesi fatte, nel caso di Voter Model con speciazione la dinamica di ciascuna specie dell'ecosistema dipende esclusivamente dalla sua abbondanza e non da altre sue proprietà: le specie, dunque, vengono trattate tutte allo stesso modo e siamo quindi nell'ambito della Teoria Neutrale.

In accordo con la Teoria Neutrale, la dinamica dell'abbondanza delle specie è la stessa per ogni specie e viene descritta dalla *birth-death master equation* [8], che descrive la probabilità $P(n, t)$ per una data specie al tempo t di avere una popolazione di n individui

$$\frac{dP}{dt}(n, t) = b(n-1) \cdot P(n-1, t) - [b(n) + d(n)] \cdot P(n, t) + d(n+1) \cdot P(n+1, t) \quad (3.1)$$

dove $b(n)$ e $d(n)$ sono rispettivamente i parametri di nascita e morte, dati dalle seguenti

$$b(n) = (1 - \nu) \cdot \frac{n}{J} \frac{J-n}{J-1} \quad (3.2)$$

$$d(n) = (1 - \nu) \cdot \frac{n}{J} \frac{J-n}{J-1} + \nu \cdot \frac{n}{J} \quad (3.3)$$

Notiamo che queste definizioni sono consistenti con le due regole della dinamica del Voter Model:

- con probabilità $1 - \nu$ si ha la morte di un individuo della comunità non appartenente alla specie data (con probabilità $\frac{J-n}{J}$) accompagnata dalla nascita di un individuo della specie scelta ($\frac{n}{J-1}$);
- con probabilità $\frac{n}{J}$ si ha la morte di un individuo appartenente alla specie data accompagnata o dalla migrazione di un individuo proveniente da una delle altre specie presenti (con probabilità $(1 - \nu) \left(\frac{J-n}{J-1}\right)$) o dalla comparsa di una nuova specie (con probabilità ν).

La soluzione stazionaria o di equilibrio $P^*(n)$ per la (3.1) è la seguente [8]

$$P^*(n) = P(0) \prod_{z=1}^n \frac{b(z-1)}{d(z)} \quad (3.4)$$

dove $P(0)$ può essere dedotta dalla condizione di normalizzazione

$$\sum_n P^*(n) = 1$$

Per arrivare a questo risultato, imponiamo la condizione di equilibrio $\frac{dP}{dt}(n, t) = 0$ e notiamo che possiamo riscrivere l'equazione come

$$I(n+1) - I(n) = 0 \quad (3.5)$$

dove abbiamo definito

$$I(n) = d(n)P^*(n) - b(n-1)P^*(n-1)$$

Imponendo la condizione fisica che non si possa avere un numero di individui negativo, si ha che $b(-1) = d(0) = 0$ e, quindi, $I(0) = 0$. Inoltre, supporre che quando una specie si estingue contemporaneamente una nuova la rimpiazza è equivalente a imporre condizioni al contorno riflettenti alla (3.1): questa ipotesi è ragionevole sulle scale temporali di nostro interesse. Sommando su tutti gli n la (3.5) otteniamo

$$\sum_{z=0}^{n-1} [I(z+1) - I(z)] = I(n) - I(0) = 0 \quad \implies \quad I(n) = 0$$

da cui segue che $P^*(n)$ è ricorsiva, come si voleva: $d(n)P^*(n) = b(n-1)P^*(n-1)$.

3.1.2 Mean field senza speciazione

Nella trattazione del Voter Model mean field senza speciazione ($\nu = 0$), ci limitiamo al caso di spin *up* e *down*. In questa ipotesi, possiamo definire c_+ e c_- come le concentrazioni delle due opinioni che supponiamo siano distribuite uniformemente e in modo uguale ($c_+ = c_- = 1/2$). Queste quantità sono importanti perchè da esse dipende la dinamica (ad esempio, se $c_+ = 0$ abbiamo che $c_- = 1$ e dunque abbiamo raggiunto il consenso).

La frazione C_n di siti che hanno cambiato le loro opinioni n volte fino al tempo t si evolve secondo la legge [9]

$$\frac{dC_n}{dt} = C_{n-1} - C_n \quad (3.6)$$

con la condizione che $C_{-1} = 0$. Si può verificare che vale la condizione di normalizzazione $\sum_n C_n(t) = 1$. Risolvendo la (3.6) con la condizione iniziale $C_n(0) = \delta_{n0}$ (dove si è introdotta la δ di Kronecker, giustificata dal fatto che al tempo $t = 0$ la frazione di siti che non hanno mai cambiato opinione è evidentemente 1) si ricava che la funzione di distribuzione del cambio di opinione è poissoniana

$$C_n(t) = \frac{t^n}{n!} e^{-t} \quad (3.7)$$

In particolare, notiamo che la frazione di elettori persistenti, cioè elettori che non hanno cambiato spin fino al tempo t , diminuiscono in modo esponenziale: $C_0(t) = e^{-t}$. La probabilità che un elettore abbia il suo parere iniziale nel tempo t è invece $C_{pari} = \sum_n C_{2n} = (1 + e^{-2t})/2$: asintoticamente tale probabilità si avvicina esponenzialmente al valore $1/2$ e, dunque, gli elettori dimenticano la loro opinione iniziale.

Sfruttando la (3.7) si può dimostrare [9] che i valori asintotici delle concentrazioni dei due spin c_+ e c_- dipendono dal numero di dimensioni del sistema. In particolare

$$c_{\pm}(t) = \begin{cases} t^{-1/2}, & \text{se } d = 1, \\ 1/\ln t, & \text{se } d = 2, \\ \text{cost}, & \text{se } d > 2, \end{cases} \quad (3.8)$$

che mostra come fino a dimensione $d = 2$ la concentrazione di una delle due opinioni tende a 0 per tempi sufficientemente lunghi, raggiungendo così il consenso, mentre al contrario per dimensioni maggiori le concentrazioni assumono un valore costante e dunque non si raggiunge il consenso.

Capitolo 4

Relative species abundance per il Voter Model

Vogliamo ora addentrarci nello studio della *Relative species abundance (RSA)*, ossia della distribuzione di probabilità che una specie estratta a caso dall'ecosistema abbia n individui. Come già osservato in 3.1.1, le q specie (opinioni) nel Voter Model seguono tutte la stessa dinamica data dall'equazione (3.1). Essendo quindi nell'ambito della Teoria Neutrale, la $P(n)$ ricavata in (3.4) (cioè la probabilità che una data specie abbia una popolazione di n individui) corrisponde proprio alla probabilità che, scelta a caso un individuo qualsiasi della popolazione, questo appartenga a una specie con una popolazione di n individui, ovvero all'RSA.

Per studiare l'RSA del Voter Model possiamo dunque rifarci alla $P(n)$, di cui conosciamo la soluzione analitica, che è stata confrontata con una simulazione del Voter Model mean field con speciazione, in cui abbiamo scelto come parametri $J = 10^6$ e $v = 10^{-3}$.

Notiamo innanzitutto che, con questa scelta dei parametri, è conveniente, poichè $J \gg n$, approssimare il birth rate e il death rate, in modo da ottenere la soluzione all'ordine $1/J$. Infatti

$$b(n) = (1 - v) \cdot \frac{n}{J} \frac{J - n}{J - 1} \approx (1 - v) \cdot \frac{n}{J} \quad (4.1)$$

$$d(n) = (1 - v) \cdot \frac{n}{J} \frac{J - n}{J - 1} + v \cdot \frac{n}{J} \approx \frac{n}{J} \quad (4.2)$$

E dunque, nel limite di grandi ecosistemi, la soluzione stazionaria per la (3.1) è

$$P(n) = -\frac{(1 - v)^n}{n \log v} \quad (4.3)$$

Nei seguenti grafici è riportato il confronto tra la soluzione analitica della (4.3) e la RSA numerica del Voter Model simulato con i parametri sopraccitati.

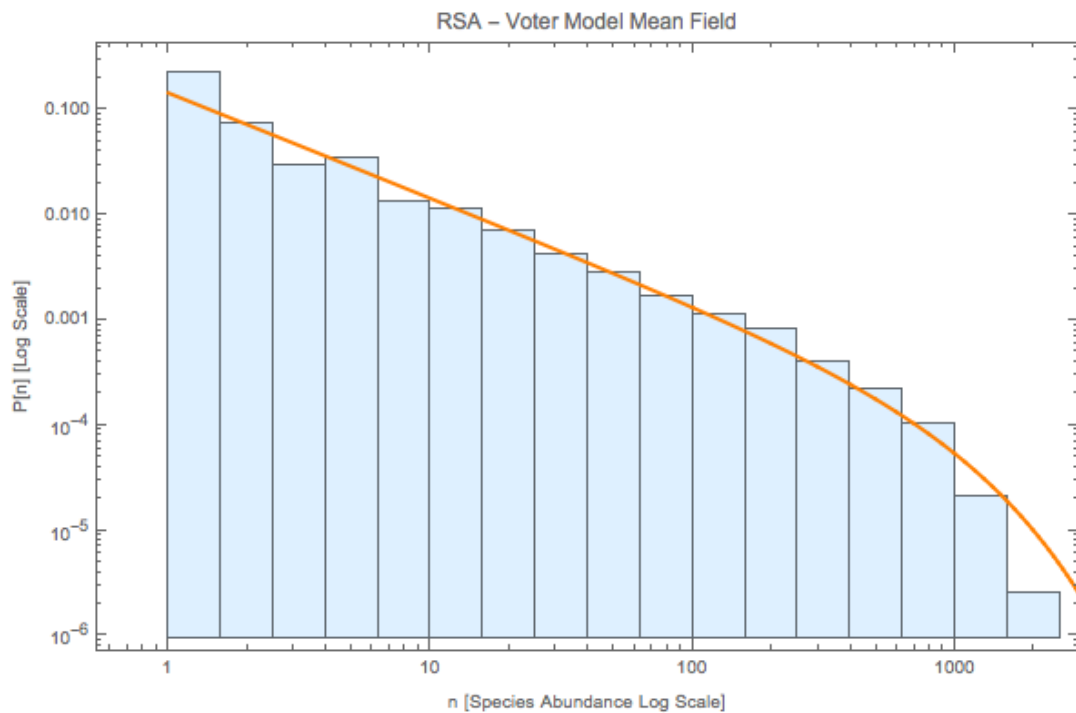


Figura 4.1: RSA per il mean field Voter Model.

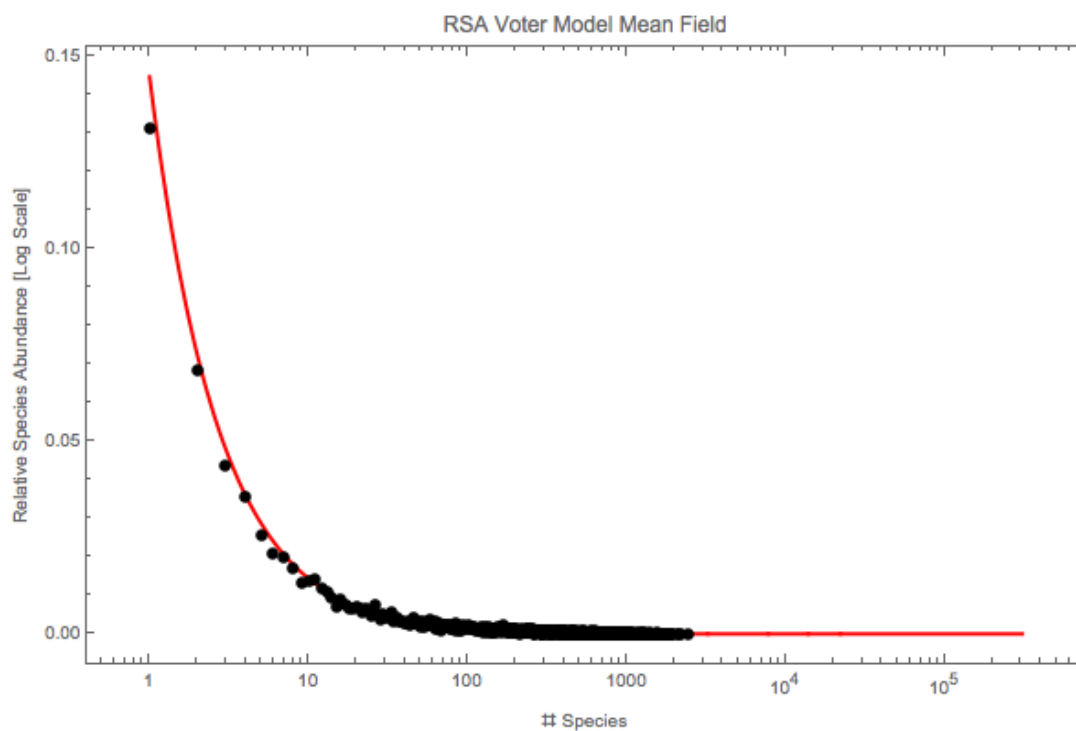


Figura 4.2: RSA per il mean field Voter Model.

4.1 Barro Colorado Island

La Barro Colorado Island è un'isola al largo di Panama da quasi un secolo oggetto di ricerche ecologiche [10]. Con i suoi 40 ettari di superficie, quest'isola è stata, a partire dal 1923, la base di una serie di ricerche sull'ecosistema di una foresta umida tropicale: la sua flora è conosciuta ormai meglio di qualsiasi altro sito di dimensioni comparabili nel mondo. Dal 1980 fino al 2000 ogni 5 anni sono stati eseguiti dei censimenti sulla fauna e sulla flora locale, che hanno mostrato un notevole dinamismo e un'instabilità delle popolazioni. Nel grafico seguente è riportata la RSA ottenuta per la fauna nel 2000, con il relativo fit, in cui si trova che $\nu = 10^{-3.1}$ e da cui emerge la bontà del nostro modello.

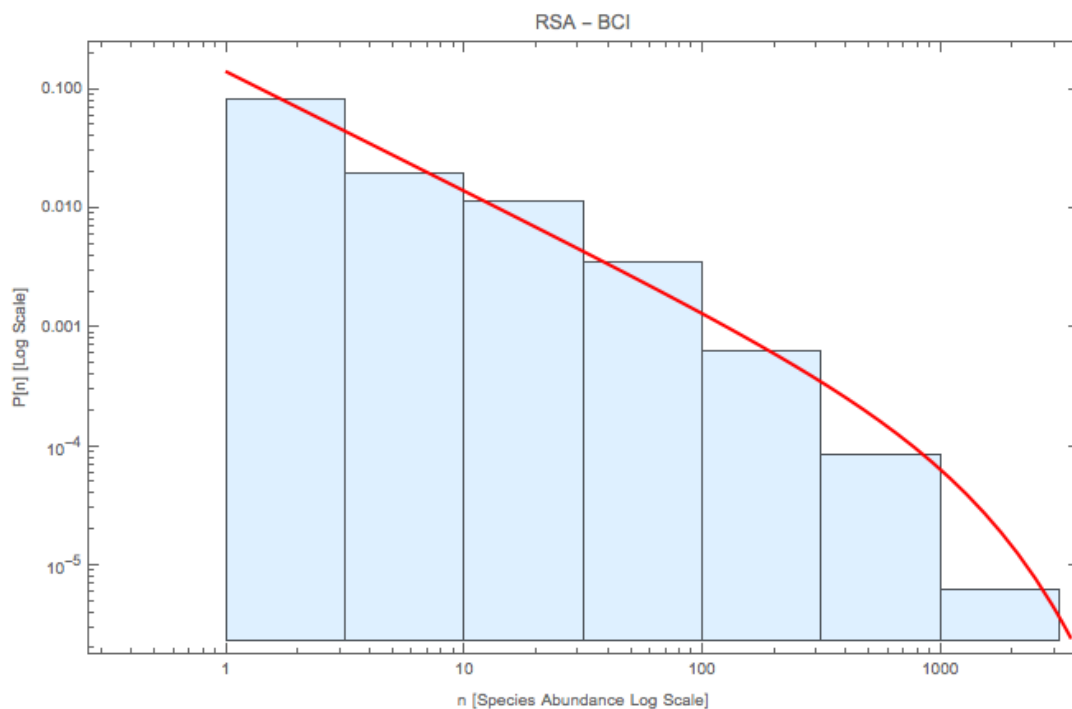


Figura 4.3: RSA relativa alla fauna della Barro Colorado Island.

Capitolo 5

Species-area relationship per il Voter Model

Vogliamo in questa sezione concentrarci sulla *Species-area relationship* (SAR) del Voter Model. Sotto alcune semplici ipotesi, verificate dal Voter Model, è possibile infatti esprimere analiticamente il numero medio di specie diverse al crescere dell'area considerata.

Seguendo il lavoro di Coleman [11], possiamo considerare il nostro modello come l'unione di K sottoregioni non sovrapposte¹ r_k ciascuna di area a_k , in modo che l'area totale sia $A = \sum_{k=1}^K a_k$. L'assunzione più semplice che si può fare sulla disposizione spaziale degli individui è quella della disposizione casuale e uniforme all'interno dell'area totale. Quindi, che sia o meno nota la collocazione degli altri individui di \mathbf{C} , la probabilità che un dato individuo della popolazione \mathbf{C} risieda in una data sottoregione r_k è uguale alla relativa area normalizzata $\alpha_k = a_k/A$.

Questa ipotesi, comunemente nota come *random placement* e valida per il mean field Voter Model con speciazione, può essere assunta anche per sistemi biologici, quando sono verificate le seguenti condizioni:

- l'interazione tra gli individui, sia della stessa che di specie diverse, ha un ruolo trascurabile sulla scelta della residenza;
- le diverse regioni r_k sono qualitativamente uguali: le chance che una di esse venga scelta da un individuo dipende solamente dalla sua estensione;
- parti della sottoregione r_k che hanno la stessa area contribuiscono ugualmente alla probabilità che un individuo scelga tale sottoregione.

Consideriamo ora una sottoregione r tra le K sottoregioni (r_1, r_2, \dots, r_K) . Dall'ipotesi di *random placement* segue che la probabilità che un particolare individuo di \mathbf{C} non sia risieda in r è $1 - \alpha$: per ogni specie i tra le S presenti la probabilità q_i che nessun membro della i -esima specie sia presente in r è

$$q_i = (1 - \alpha)^{n_i} \quad (5.1)$$

mentre la probabilità p_i che la specie i sia presente in r è

$$p_i = 1 - (1 - \alpha)^{n_i} \quad (5.2)$$

La probabilità $Q_s(\alpha)$ che s delle S diverse specie siano rappresentate in r corrisponde alla probabilità di ottenere s successi in S indipendenti e distinti prove di Bernoulli per i quali la probabilità

¹Queste regioni possono essere o meno topologicamente connesse, l'importante è che non siano sovrapposte.

di successo nell' i -esima prova è p_i . La funzione generatrice di $Q_s(\alpha)$ può essere espressa in modo analitico ed è

$$G(z) = \sum_{s=1}^S Q_s(\alpha) z^s = \prod_{i=1}^S (q_i + z p_i) \quad (5.3)$$

Per definizione, il numero atteso di specie presenti nella regione r sarà

$$\bar{s}(\vec{n}) = \sum_{s=0}^S s P_s = \left. \frac{d}{dx} \ln G(e^x) \right|_{x=0} = \sum_{i=1}^S p_i = S - \sum_{i=1}^S q_i = S - \sum_{i=1}^S (1 - \alpha)^{n_i} \quad (5.4)$$

mentre la varianza relativa a s sarà

$$\sigma^2(\vec{n}) = \sum_{s=0}^S (s - \bar{s})^2 P_s = \left. \frac{d^2}{dx^2} \ln G(e^x) \right|_{x=0} = \sum_{i=1}^S p_i q_i = \sum_{i=1}^S (1 - \alpha)^{n_i} - \sum_{i=1}^S (1 - \alpha)^{2n_i} \quad (5.5)$$

La media e la varianza del numero di specie presenti nella sottoregione r è dunque determinata una volta noti l'area relativa α di r e il vettore delle abbondanze \vec{n} di \mathbf{C} . Se \vec{n} fosse noto, ovvero se fossero note le abbondanze di ciascuna specie, la curva ottenuta graficando $\bar{s}(\alpha)$ in funzione di α prenderebbe il nome di *expected species-area curve*.

Nel caso del Voter Model, invece, è importante notare che, poichè le componenti del vettore delle abbondanze $\vec{n} = (n_1, n_2, \dots, n_S)$ sono variabili casuali, anche \bar{s} è una variabile aleatoria: in generale, sostituire a ciascun n_i la sua media \bar{n}_i in (5.4) e in (5.5) porterebbe a delle formule errate, in quanto le variabili casuali n_i non son distribuite uniformemente ma seguono la $P(n_i)$ della (4.3). Per ottenere la SAR per il mean field Voter Model, dobbiamo quindi fare la convoluzione tra $\bar{s}(\vec{n})$ e $P(\vec{n})$, che, per l'ipotesi di neutralità, può essere decomposta in

$$P(\vec{n}) = \prod_{i=1}^S P(n_i) = \prod_{i=1}^S \frac{(1 - \nu)^{n_i}}{n_i \log \nu} \quad (5.6)$$

da cui, ricordando che vale la condizione di normalizzazione $\sum P(n_i) = 1$

$$\begin{aligned} \langle S(\alpha) \rangle &= \sum_{n_1, \dots, n_S=1}^{J \rightarrow \infty} \bar{s}(\vec{n}) P(\vec{n}) \\ &= \sum_{n_1, \dots, n_S=1}^{J \rightarrow \infty} \prod_{j=1}^S P(n_j) \left(S - \sum_{i=1}^S (1 - \alpha)^{n_i} \right) \\ &= S \left(\sum_{n_1, \dots, n_S=1}^{J \rightarrow \infty} \prod_{j=1}^S P(n_j) \right) - \sum_{n_1, \dots, n_S=1}^{J \rightarrow \infty} \sum_{i=1}^S (1 - \alpha)^{n_i} \prod_{j=1}^S P(n_j) \\ &= S \left(1 - \sum_{n_i=1}^{\infty} \frac{(1 - \alpha)^{n_i} (1 - \nu)^{n_i}}{n_i \log \nu} \right) \\ &= S \left(1 - \frac{\log [1 - (1 - \alpha)(1 - \nu)]}{\log \nu} \right) \end{aligned} \quad (5.7)$$

Ricaviamo così la forma analitica per la SAR per il mean field Voter Model

$$\langle S(\alpha) \rangle = S \left(1 - \frac{\log [v(1 - \alpha) + \alpha]}{\log v} \right) \quad (5.8)$$

La simulazione numerica, riportata in 5.1, conferma questo risultato. Come si nota, l'accordo tra la curva teorica ottenuta dalla (5.8) con $\nu = 10^{-3}$, rappresentata in rosso, e i risultati della simulazione, in nero, è ottima.

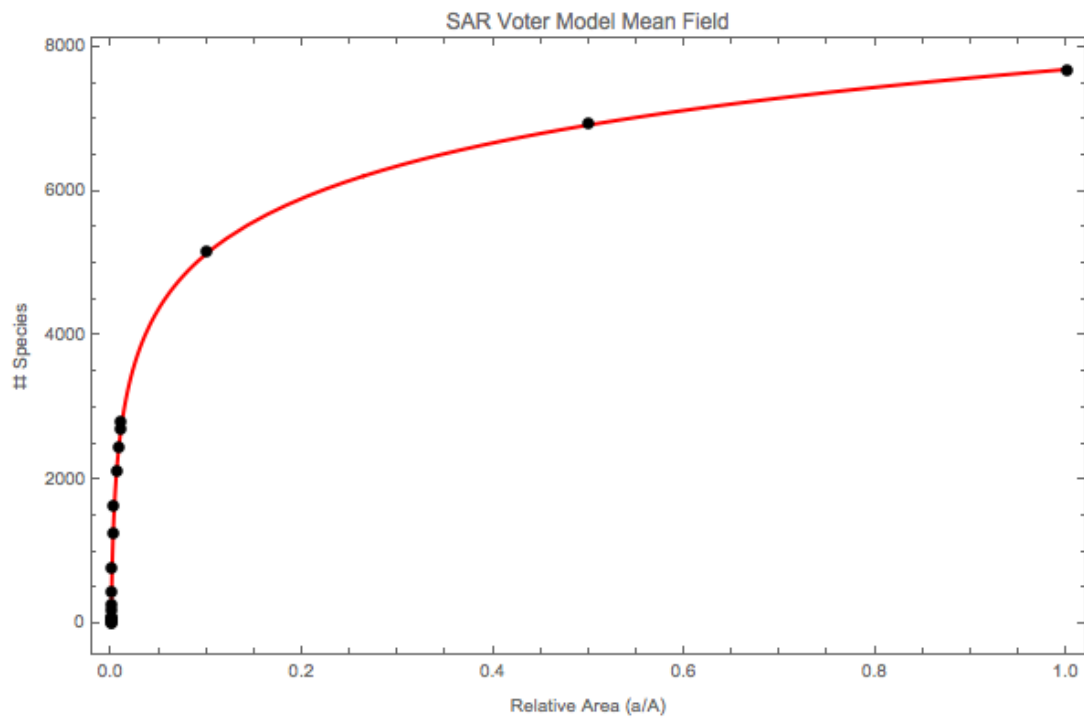


Figura 5.1: SAR per il mean field Voter Model

Species Persistence Time per il Voter Model

La terza distribuzione che vogliamo approfondire è la *Species Persistence Time (SPT)*, ovvero la distribuzione dei tempi che intercorrono tra la formazione e l'estinzione di una specie all'interno della comunità ecologica in una certa regione.

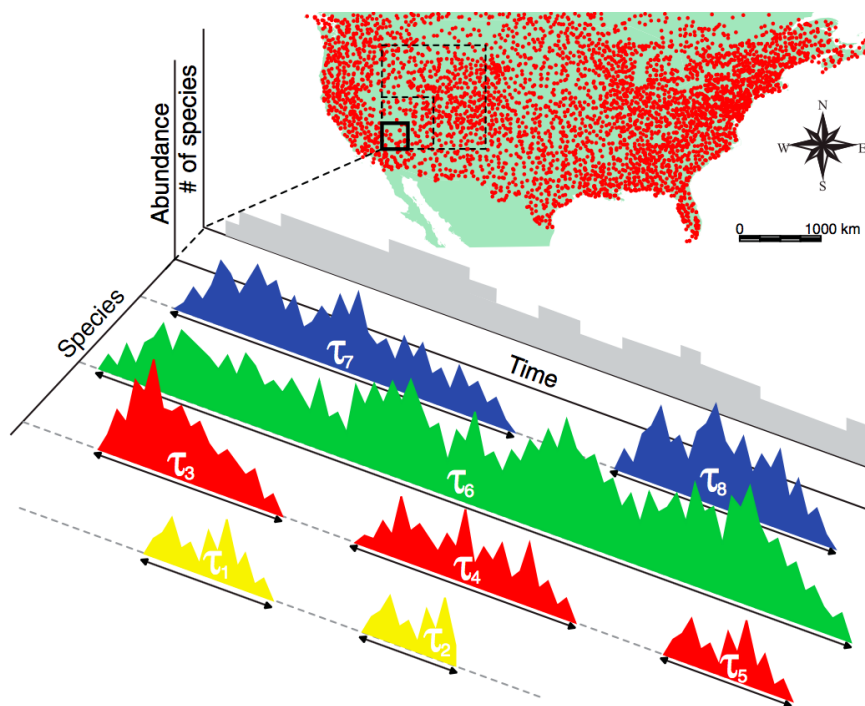


Figura 6.1: Species Persistence Time. Il tempo di persistenza all'interno di una regione geografica è definito come il tempo che si è verificato tra la nascita di una specie e la sua estinzione locale. Le colonizzazioni ricorrenti di una specie definiscono diversi periodi di pertinenza. Il numero di specie nell'ecosistema in funzione del tempo (area grigia ombreggiata) dipende essenzialmente dall'emergere di nuove specie e dai tempi di persistenza di ciascuna specie. In Figura sono rappresentati i tempi di persistenza di alcune specie di uccelli in America del Nord [12].

La *SPT* è, nello studio degli ecosistemi, una delle distribuzioni più utilizzate e allo stesso tempo dibattute: secondo alcuni studi [13] il tempo di sopravvivenza delle specie sembra essere

esponenzialmente distribuito, mentre secondo altri, invece, segue una legge di potenza con un esponente vicino a -2 o $-3/2$ (si veda [14] o [15]).

Sotto le ipotesi della Teoria Neutrale, tutti questi comportamenti possono essere riprodotti. Assumendo che gli individui in una comunità ecologica siano tutti equivalenti e che la popolazione di ciascuna specie sia regolata dalle stesse leggi, la dinamica è definita in modo univoco dai parametri di nascita e di morte definiti in (3.2), che dipendono esclusivamente dal numero n di individui della popolazione.

In generale, $b(n)$ e $d(n)$ racchiudono in sè, in modo addittivo, varie cause ecologiche. Pertanto, possiamo espandere in tutta generalità in serie di potenze di $1/n$ i parametri di nascita e di morte per individuo:

$$\frac{b(n)}{n} = b_1 + \frac{b_0}{n} + \frac{b_{-1}}{n^2} + \dots \quad (6.1)$$

$$\frac{d(n)}{n} = d_1 + \frac{d_0}{n} + \frac{d_{-1}}{n^2} + \dots \quad (6.2)$$

Nella Teoria Neutrale [2] i termini b_0 e d_0 , come abbiamo già visto, possono essere interpretati come il risultato del meccanismo di emigrazione e immigrazione, mentre i termini successivi all'ordine $1/n$ possono essere trascurati, in quanto rilevanti solo nei limiti di piccole popolazioni. In questo modo, il birth rate e il death rate che descrivono la $P(n, t)$ ricavata in (3.4) (che rappresenta la probabilità che una specie abbia una popolazione di n individui al tempo t) diventano

$$b(n) = b_0 + b_1 n \quad (6.3)$$

$$d(n) = d_0 + d_1 n \quad (6.4)$$

Con queste assunzioni, $P(0, t)$ rappresenta la probabilità per una specie di essere già istinta al tempo t .

La funzione di distribuzione della probabilità di estinzione al tempo t , $p(t)$, nota come *species persistence time equation*, sarà dunque la derivata nel tempo di $P(0, t)$ [16]

$$p(t) = \frac{d}{dt} P(0, t) \quad (6.5)$$

Esaminiamo cosa succede nei due casi limite in cui $b_1 = d_1 = 0$ e $b_0 = d_0 = 0$, limitandoci ad alcune osservazioni qualitative relative al caso generale.

Quando $b_1 = d_1 = 0$ perdiamo la dipendenza dalla popolazione e dunque il numero di individui appartenenti a una determinata specie compie un *random walk* nello spazio del numero n di individui, con b_0 e d_0 probabilità per unità di tempo di saltare "in avanti" o "indietro" sulla retta: la sopravvivenza di una specie corrisponderà dunque al tempo necessario al random walk per raggiungere $n = 0$, ovvero per uscire dall'asse positivo. Questo è un problema che è stato ampiamente affrontato e risolto analiticamente in letteratura (si veda, ad esempio, [17]). In particolare, nel caso critico $b_0 = d_0$ è facile verificare [16] che in questo caso asintoticamente vale $P(0, t) \simeq 1 - 1/\sqrt{t}$ e dunque il tempo di estinzione segue una distribuzione della forma $p(t) \sim t^{-3/2}$.

Analizziamo ora il caso $b_0 = d_0 = 0$, interessante per la trattazione degli ecosistemi, in quanto su scale molto grandi (ad esempio, continentali) questi termini sono effettivamente molto piccoli [18]. Notiamo che questo è proprio il caso del Voter Model: ricordiamo infatti che per questo modello possiamo approssimare i coefficienti di nascita e di morte come $b(n) = \frac{1-\nu}{J} n$ e $d(n) = \frac{1}{J} n$ e dunque abbiamo che

$$b_0 = \nu \approx 0 \quad b_1 = \frac{1-\nu}{J} \quad (6.6)$$

$$d_0 = 0 \quad d_1 = \frac{1}{J} \quad (6.7)$$

Analogamente al primo caso limite esaminato, la (6.5) può essere risolta analiticamente [16]: senza perdita di generalità, se imponiamo che $d_1 = 1$ e $b_1 = \alpha$, in modo che $d_1 - b_1 = 1 - \alpha$, avremo che

$$p(t) = \left(\frac{1 - \alpha}{e^{(1-\alpha)t} - \alpha} \right)^2 e^{(1-\alpha)t} \quad (6.8)$$

Questa distribuzione ha un andamento esponenziale per valori non troppo piccoli di $1 - \alpha$, ovvero quando $d_1 \gg b_1$; viceversa, quando $d_1 \sim b_1$ la distribuzione segue una power-law di esponente -2. Infatti, possiamo riscrivere la (6.8) nella forma

$$p(t) = \frac{1}{t^2} f\left(\frac{t}{t^*}\right) \quad (6.9)$$

dove $f(x) = [x / (1 - e^{-x})]^2 e^{-x}$ e $t^* = 1 / (1 - \alpha)$ è detto tempo caratteristico.

Per quanto concerne il caso generale, euristicamente possiamo affermare che le specie con un tempo di estinzione maggiore tipicamente hanno un grande numero di individui. Per queste specie, i coefficienti b_0 e d_0 possono essere ragionevolmente trascurati: ci aspettiamo dunque che a un certo tempo caratteristico ci sia un passaggio dall'andamento $t^{-3/2}$ a quello t^{-2} e infine, oltre un certo altro tempo caratteristico, un decadimento esponenziale.

Presentiamo di seguito i risultati della SPT ottenuta dalle simulazioni del Voter Model, con parametri $\nu = 10^{-3}$ e $J = 1000$: in Figura 6.2 è presentata la SPT cumulativa, mentre in 6.3 la SPT in scala log-log. Notiamo il buon accordo tra le simulazioni e i relativi fit. In particolare, in Figura 6.3 notiamo come l'andamento sia proporzionale a $t^{-3/2}$ per i tempi più bassi e a t^{-2} per quelli più alti, come ci aspettavamo.

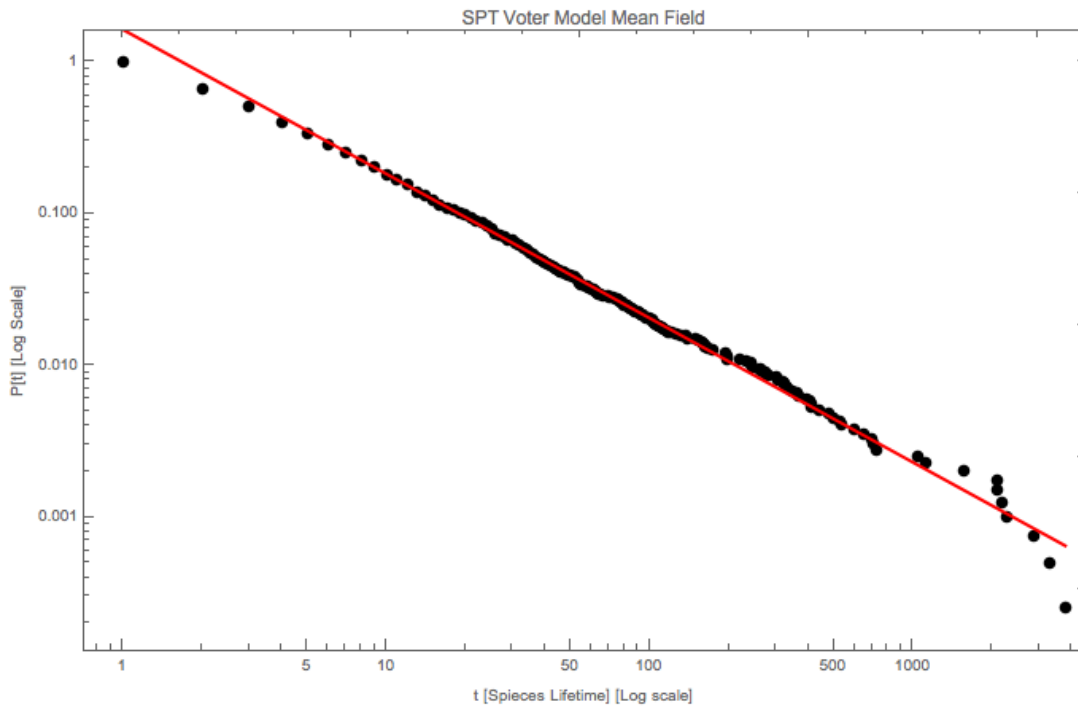


Figura 6.2: SPT cumulativa per il mean field Voter Model. In ascissa tempi di persistenza t , in ordinata probabilità che una specie sia sopravvissuta per almeno un tempo t prima di estinguersi. In nero sono rappresentati i risultati della simulazione, in rosso il relativo fit.

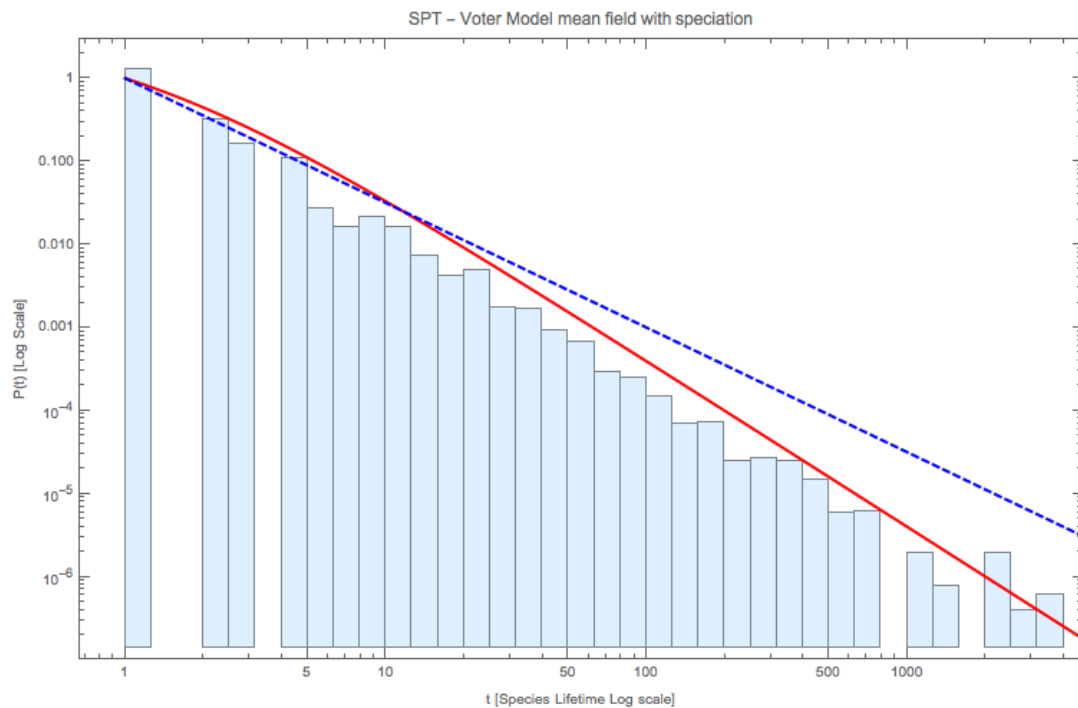


Figura 6.3: SPT per il mean field Voter Model. In ascissa i tempi di persistenza t , in ordinata probabilità che una specie sia sopravvissuta esattamente un tempo t prima di estinguersi. In rosso fit con la power-law t^{-2} , in blu quello con $t^{-3/2}$.

Conclusioni

Nello studio della dinamica della formazione del consenso in una popolazione descritta dal Voter Model mean field con speciazione abbiamo ricavato una descrizione analitica di alcuni pattern ecologici emergenti. In particolare, ci siamo occupati dei seguenti pattern:

- la *Relative species abundance*, che abbiamo mostrato essere descrivibile dall'equazione (4.3)

$$P(n) = -\frac{(1-\nu)^n}{n \log \nu} \quad (7.1)$$

Abbiamo inoltre confrontato i nostri modelli con i risultati sperimentali raccolti nella *Barro Colorado Island*.

- la *Species-area abundance*, che abbiamo mostrato essere descrivibile dall'equazione (5.8)

$$\langle S(\alpha) \rangle = S \left(1 - \frac{\log[\nu(1-\alpha) + \alpha]}{\log \nu} \right) \quad (7.2)$$

- la *Species persistence time*, che abbiamo mostrato essere descrivibile dall'equazione (6.8)

$$p(t) = \left(\frac{1-\alpha}{e^{(1-\alpha)t} - \alpha} \right)^2 e^{(1-\alpha)t} \quad (7.3)$$

Le simulazioni eseguite hanno consentito di verificare la consistenza dell'equazioni ricavate analiticamente nell'approssimazione mean field.

Alcune direzioni future di questo lavoro potrebbero essere:

1. aggiungere gradi di "realismo" ecologico nel modello, come per esempio introdurre la possibilità che specie nella comunità ecologica interagiscano tra loro;
2. andare oltre il mean field, studiando il ruolo dello spazio nei patterns macro-ecologici sopra descritti;
3. introdurre anche un noise ambientale che possa simulare effetti antropogenici o climatici nella dinamica.

Concludendo, con questa tesi abbiamo voluto mostrare gli interessanti e potenziali contributi che la meccanica statistica può dare all'ecologia.

Bibliografia

- [1] S. Azaele, S. Suweis, J. Grilli, I. Volkov, J.R. Banavar, and A. Maritan. Statistical mechanics of ecological systems: Neutral theory and beyond. *Rev. Mod. Phys.*, 88, 2016.
- [2] S. Hubbel. The unified theory of biodiversity and biogeography. *Princeton Univeristy Press*, 2001.
- [3] N. Goldenfeld and C. Woese. Biology's next revolution. *Nature*, 445, 2007.
- [4] R.H MacArthur and E.O. Wilson. The theory of island biogeography. *Princeton Univeristy Press*, 1967.
- [5] R.A. Holley and T.M. Liggett. Ergodic theorems for weakly interacting infinite systems and voter model. *Annals of Probability*, 3(4):643–663, 1975.
- [6] P.L. Krapivsky and S. Redner. *A kinetic view of statistical physics*. Cambridge University Press, 2010.
- [7] I. Dornic. Critical coarsening without surface tension: The universality class of the voter modell. *Phys. Rev. Lett.*, 2001.
- [8] C.W. Gardiner. *Handbook of Stochastic Methods*. Springer, 2003.
- [9] P.L. Krapivsky, E. Ben-Naim, and L. Frachebourg. Coarsening and persistence in the voter model. *Phys. Rev. Lett.*, 53(4), 1996.
- [10] D. Windsor. Climate and moisture variability in a tropical forest: Long-term records from barro colorado island. 1990.
- [11] B.D. Coleman. On random placement and species-area relations. 1979.
- [12] E. Bertuzzo, S. Suweis, L. Mari, I.R. Iturbe, and A. Rinaldo. Spatial effects on species persistence and implications for biodiversity. *PNAS*, 2011.
- [13] N.C Stenseth and J. Maynard Smith. Coevolution in ecosysyems. *J. Evol.*, 38:870–880, 1984.
- [14] T.H. Keitt and H.E. Stanley. *Nature*, 392:257–260, 1998.
- [15] P. Bak. *How Nature Works: The Science of Self-Organized Criticality*. Oxford University Press, 1997.
- [16] S. Pigolotti, A. Flammini, M Marsili, and A. Maritan. Species lifetime distribution for simple models of ecologies. *PNAS*, 2005.
- [17] S. Chandrasekhar. Stochastic problems in physics and astronomy. *Rev. Mod. Phys.*, 15, 1943.

- [18] T.E. Harris. *The Theory of Branching Processes*. 1989.
- [19] C. Darwin. *On the Origin of Species*. Signet Classic.