Ecohydrological Footprints: Quantitative Response of Ecosystems to Changes in their Hydrological Drivers

THÈSE N° 5239 (2011)

PRÉSENTÉE LE 6 DÉCEMBRE 2011
À LA FACULTÉ DE L'ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT
LABORATOIRE D'ÉCOHYDROLOGIE
PROGRAMME DOCTORAL EN ENVIRONNEMENT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Samir SUWEIS

acceptée sur proposition du jury:

Prof. M. Parlange, président du jury
Prof. A. Rinaldo, directeur de thèse
Prof. J. S. Arey, rapporteur
Prof. A. Maritan, rapporteur
Prof. I. Rodriguez Iturbe, rapporteur

Toutes choses méritent d'étre étudiées avec le souci le plus scrupuleux de l'exactitude, le souci de les voir telles qu'elles sont, de se mettre á l'écoute de leur vérité délicate, de toujours étre prét á remettre en cause les représentations que nous en avons pour déchiffrer avec encore plus de fidélité leur langage muet, parce que toutes, étant créées par Dieu, disent quelque chose de leur Créateur, qui est infiniment plus grand que nous et qui ne nous trompe pas.

Laurent Lafforgue

To my parents... and Agnese, Anna and Elisa...

Acknowledgements

This research project would not have been possible without the support of many people.

Firstly, I would like to thank my supervisor, Prof. Andrea Rinaldo, for his support and encouragement and his continual interest in my work. His advice, guidance and critical reviews have made the the work presented in this thesis possible.

For daily scientific discussions and even more for their non-scientific friendship, I would like to special thank: Lorenzo Righetto, Lorenzo Mari, Ludovico Nicotina, Serena Ceola, Francesco Carrara, Amilcare Porporato e Paolo D'Odorico. A big thanks also for Enrico Bertuzzo and Amos Maritan: the intensive scientific discussions with them and their friendship not only significantly enhanced my technical understanding but also built an attitude towards science that goes beyond the pure fulfilment of duties. I would like to thank Ignacio Rodriguez Iturbe from Princeton University for his engagement in the project, for his hospitality during my visits in Princeton, and for his friendly and always helpful advice. I would like to thank my dear colleagues Cara Tobin, Bettina Schaefli, Xavier Beuchat, Colin Schenk , Bastien Roquier, Anna Rothenbuhler and Bernard Sperandio.

These three years have not only been professionally intense, but also very rich for my human experience. I'll be always grateful for the gift of these years. For that I am deeply thankful also to my friends Jacky, Beppe, Boez, Anto, Gio, Nick, Coppa, Lemon, Marghe, Diana, Greg, Pascal, Benni, Fodda and Elo. Last, but certainly not least, I want to thanks Brisa, Coccio and Matte. Rewording Gaudí: "We need to experience a piece of New World, in order to keep up the struggle inside the world".

Lausanne, October 5, 2011 S. S.

Preface

This thesis is submitted in fulfilment of the requirements of the degree of Docteur És Sciences at the Ecole Doctorale in Civil and Environmental Engineering (EDCEE) of the École Polytechnique Fédérale de Lausanne (EPFL). It contains the candidate's scientific work at the Laboratory of Ecohydrology (ECHO) within the School of Architecture Civil and Environmental Engineering (ENAC) that was carried out under my supervision over a period of three years (October 2008 − October 2011).

The thesis blends and integrates the material published in a series of peer-reviewed publications, and is organized in seven Chapters. A general introduction outlines the conceptual thread joining the various issues studied. The Chapters are tailored from the published material as outlined in detail below. A set of conclusions, putting forth perspectives and further possible developments, and the Appendices are integral part of the thesis work. Although the contents of each Chapter rely on published material, they are revisited and edited for coherence and are displayed in revised form for consistency and continuity. The original references, almost entirely attributable to the leading role of the candidate and yet fully integrated into the scientific production of the ECHO group at EPFL, are:

Chapter II S. Suweis, E. Bertuzzo, G. Botter, A. Porporato, I. Rodriguez-Iturbe and A. Rinaldo, Impact of stochastic fluctuations in storage-discharge relations on streamflow distributions, *Water Resources Research*, 46 (3), 2010;

Chapter III S. Suweis, A. Rinaldo, S. Van der Zee, E. Daly, A. Maritan and A. Porporato, Stochastic modeling of soil salinity, *Geophysical Research Letters*, 37, L07404, 2010;

Chapter IV and V 1) E. Bertuzzo, S. Suweis, L. Mari, A. Maritan, I. Rodriguez-Iturbe and A. Rinaldo, Spatial effects on species persistence and implications for biodiversity, *Proceedings of the National Academy of Sciences*, 108 (11), 4346-4351, 2011;
2) S. Suweis, E. Bertuzzo, L. Mari, A. Maritan, I. Rodriguez-Iturbe and A. Rinaldo, On the universality of species persistence time distribution, in review;

Chapter VI and VII 1) S. Suweis, M. Konar, C. Dalin, N. Hanasaki, A. Rinaldo and I. Rodriguez-Iturbe, Structure and controls of the global virtual water trade network, *Geophysical Research Letters*, Vol. 38, pp. L08405, 2011;
2) M. Konar, C. Dalin, S. Suweis, N. Hanasaki, A. Rinaldo and I. Rodriguez-Iturbe, Water

for food: The global virtual water trade network, *Water Resources Research*, 47, W005520, 2011.

Appendix A S. Suweis, A.M. Porporato, A. Rinaldo and A. Maritan, Prescription-induced jump distributions in multiplicative Poisson processes, *Physical Review E*, 83 (6), 061119, 2011.

All issues addressed in the thesis are referred to the current frontiers of ecohydrological research. They cover a broad spectrum of open problems, using a probabilistic approach based on stochastic processes. The results are noteworthy, as witnessed by the publication output and by the breadth of methods and subjects involved.

Originality

The present thesis complies with the requirements of originality and relevance demanded by the stringent standards of EPFL and of the Doctoral School EDCEE. The linkage of the various issues, the search for ecohydrological footprints, serves as a manifesto for a new and emerging discipline, environmental statistical mechanics, seen from the hydrological perspective – and for this alont worth considering with special attention. Of particular importance in the thesis work submitted is the ubiquitous strive to achieve exact solutions to standing stochastic problems, like occurred here in: i) the case of the stationary solutions of soil salinity under stochastic hydrologic forcings; ii) the solution of a longstanding problem in multiplicative Poisson processes; and iii) the proper framework for species' persistence time distributions, as a function of topological constraints provided by hydrologic networks and as providers of important ecological spatial effects that allow a linkage with other macroecological laws. The original analysis of form and function of the global virtual trade network, and of its global ecohydrological footprints, literally opened a new field of great momentum.

In conclusion, the methodological basis for the treatment of the problems studied in the thesis is impeccable. The scientific interest for the body of problems of ecohydrological footprinting addressed here is remarkable. Overall, I deem outstanding the quality of the original work leading to the present thesis[1].

Lausanne, October 6, 2011

Prof. Andrea RINALDO
Thesis Director

Abstract

Ecohydrological footprints are defined as the response of ecosystem functions or services to changes in their hydrologic drivers. In this thesis, several diverse footprints are addressed: noise-driven effects on storage-discharge relations and catchment streamflow distributions, that are important drivers of biodiversity; soil salinization and its ecohydrological implications; topological effects of the ecological interaction networks on living communities (e.g. on their species persistence); and form and function of the global virtual water trade network. The coherence of the conceptual framework is provided by the study of drivers and controls of ecohydrological variability using methodological approaches based on statistical mechanics. In fact, this thesis work outlines a significant portion of environmental statistical mechanics, an overarching discipline that is emerging in recent years, which applies mathematical tools from statistical mechanics to model several ecohydrological processes.

The proposed relevance of this thesis lies in the major effects of hydrologic drivers on ecological process. The view that emerges from current research in ecohydrology, that this thesis supports, is that there exists a definite need for an integrated understanding of ecological and hydrological processes. Because stochasticity is intrinsic to environmental and ecohydrological variability, noise plays an important and constructive role in ecohydrological processes. In this thesis, a stochastic approach is applied to analyze different ecohydrological processes, ranging from green and blue water flows in river basins (part I), ecosystem dynamics affected by the directional dispersal provided by river networks (part II) to water footprints of human society (part III). Methods range from novel exact solutions to stochastic differential equations to random graph theory applications, and imply the analysis of suitable field data.

An analytical framework for quantitative analysis is laid out to tackle complex problems and to estimate the effects of environmental change on the interaction of the hydrologic processes with the biota. The main results of this thesis are: i) the achievement of exact solutions for the probability distribution of catchment streamflow, that takes in account stochastic fluctuations in the storage-discharge relation and for the condition of a noise induced phenomena to the streamflows regimes; ii) the stationary solutions of soil salinity under stochastic hydrologic forcing; iii) a novel solution of the Ito−Stratonovich problem in multiplicative Poisson processes; iv) the proper framework for species' persistence time distributions, as a function of topological constraints on the ecosystem, and its connection with other important macroeco-

Abstract

logical laws. A related length-bias sampling problem is also solved. v) A statistical analysis of the global virtual trade network and a semi-analytical model that is able to describe most of the observed properties.

x

Sommario

Le impronte ecoidrologiche sono definite come la risposta delle funzioni o dei servizi ecosistemici a cambiamenti nei loro driver idrologici. Questa tesi presenta lo studio di molteplici e diverse impronte ecoidrologiche: gli effetti del rumore stocastico sulla relazione tra volume invasato e portata del bacino e sulla distribuzione di deflusso, processi idrologici di vitale importanza per la biodiversitá all'interno del bacino; la salinizzazione dei suoli e le sue implicazioni ecoidrologiche; gli effetti della topologia delle interazioni ecologiche (per esempio della topologia della rete fluviale) sulle comunitá ecologiche, e in particolare sui tempi di persistenza delle specie all'interno del ecosistema; infine la forma e le funzioni della rete mondiale di commercio dell'acqua virtuale tra nazioni. La coerenza del quadro concettuale é fornita dallo studio dei drivers e dei controlli della variabilitá ecoidrologica utilizzando lo stesso approccio metodologico basato sulla meccanica statistica. Infatti questo lavoro di tesi delinea una parte significativa della meccanica statistica ambientale, una disciplina che sta emergendo negli ultimi anni, e che applica strumenti matematici tipici della meccanica statistica per modellare diversi processi ecoidrologici.

La rilevanza di questa tesi risiede nella grande importanza che ricoprono i processi idrologici sui sistemi ecologici. Il panorama che emerge dall'attuale ricerca in ecoidrologia, e che questa tesi sostiene, é che esiste una forte necessitá di una comprensione integrata dei fenomeni ecologici e idrologici. Poiché la stocasticitá é intrinseca alla variabilitá ambientale ed ecoidrologica, il rumore stocastico ha un ruolo importante e costruttivo nei processi ecoidrologici. In questa tesi, per analizzare i diversi processi ecoidrologici, che variano dai flussi d'acqua nei bacini fluviali (parte I), alla dinamica spaziale degli ecosistemi, dovuta alla diversa struttura spaziale dell'ambiente in cui l'ecosistema vive (parte II), agli effetti dello scambio di acqua virtuale sulla societá umana (parte III). I metodi comprenderanno soluzioni esatte di equazioni differenziali stocastiche, applicazioni della teoria dei grafi, e l'analisi di relativi dati empirici.

Un quadro analitico per l'analisi quantitativa di problemi complessi é quindi stato delieato, anche in modo da poter stimare gli effetti dei cambiamenti ambientali e climatici sull'interazione tra processi idrologici e biota. I principali risultati di questa tesi sono: i) il raggiungimento di soluzioni esatte per la distribuzione di probabilitá di deflusso in un bacino, che prende anche in considerazione le fluttuazioni stocastiche nella relazione tra volume invasato e portata, e per la condizione di attivazione di un fenomeno, indotto dal rumore stocastico,

Sommario

di cambiamento di regime del deflusso; ii) la soluzione stazionaria della concentrazione di sale nel terreno in presenza di eventi di pioggia casuali distribuiti Poissonianamente; iii) una nuova soluzione per il dilemma di Ito−Stratonovich nel caso di equazioni differenziali stocastiche con rumore Poissoniano e moltiplicativo; iv) la formulazione di un quadro teorico adeguato per la definizione della distribuzione dei tempi di persistenza delle specie, anche in funzione dei vincoli topologici presenti nell' ecosistema, e la sua connessione con altre pattern macro−ecologi. É stato risolto anche un connesso problema relativo al campionamento dei tempi di persistenza misurati dai dati empirici. v) Un'analisi statistica della rete del commercio di acqua virtuale tra nazioni e la presentazione di un modello semi−analitico che é in grado di descrivere la maggior parte delle proprietá osservate della suddetta rete.

Parole chiave: Equazioni differenziali stocastiche, fenomeni indotti dal rumore stocatico, relazione volume invasato−portata del bacino, salinizzazione del suolo, macro−ecologia, distribuzione dei tempi di persistenza delle specie, teoria neutrale, acqua virtuale, reti complesse, modello di fitness.

Contents

List of Figures

Chapter 1

Introduction

Ecohydrology is the field of scientific research that at its inception brought a first quantitative hydrologic perspective of climate-soil-vegetation dynamics [Rodriguez-Iturbe et al., 1999b; Rodriguez-Iturbe, 2000]. The stochastic treatment of soil moisture dynamics, the theme of the original work, set the tone and the domain of subsequent developments aimed at stochastically-driven mass balance of water resources, with a view to related controls on basic ecologic patterns and processes. On broader scales in space and time, the study of the water cycle has always been and is one of the main focus in water resources assessments [e.g. Manabe et al., 1991; Brutsaert and Parlange, 1998; Jackson et al., 2001; Oki and Kanae, 2006] largely because renewable freshwater, though comprising only a tiny fraction of the global water pool, is the foundation for life in terrestrial and freshwater ecosystems [Jackson et al., 2001]. If in the coming century climate change will produce a growing imbalance among freshwater supply, arguably consumption and population increase will alter the water cycle in a significant manner. This a problem hardly overestimated, especially in the regions of the world already limited by the amount and quality of available water. It has been suggested that in the next 30 years alone, accessible runoff is unlikely to increase more than 10% whereas the earth's population is projected to rise by approximately one-third. Unless the efficiency in water uses rises, therefore, such imbalance will impair freshwater ecosystem services, increase the number of aquatic species facing extinction, and further fragment wetlands, rivers, deltas, and estuaries [Jackson et al., 2001]. The growing demand on freshwater resources thus urgently calls for stronger links of ecohydrologic research with improved water management.

Evidence for intensification of the global water cycle is mounting [Huntington, 2006] and, with

that, the concern about ecohydrological footprints. Indeed, if the climate warms in the future, there will arguably be an intensification of the water cycle, with remarkable consequences. In fact, such intensification may lead to changes in local and global water resources availability, increases in the frequency and intensity of storms, floods, and droughts, and an amplification of warming trends through the water vapor feedback. The current state of science regarding historical trends in hydrologic variables, including precipitation, runoff, tropospheric water vapor, soil moisture, glacier mass balance, evaporation, evapotranspiration, and growing season length, provides evidence indicating an ongoing intensification of the water cycle [Huntington, 2006] . For instance, analysis of streamflow, snow mass temperature and precipitation in snowmelt-dominated river basins in the western United States indicate an advance in the timing of peak spring season flows over the past 50 years [Regonda et al., 2005]. The observed trends in hydroclimatology over the western United States can indeed have significant impacts on water resources planning and management therein.

The sensitivity of the global water cycle to the water-holding capacity of the plant-root zone of continental soils has been estimated [Milly and Dunne, 1994] by simulations using a mathematical model of the general circulation of the atmosphere, with prescribed ocean surface temperatures and prescribed cloud. With an increase of storage capacity, evaporation from the continents rises and runoff falls, because a high storage capacity enhances the ability of the soil to store water from periods of excess for later evaporation during periods of shortage. In addition to this direct effect, atmospheric feedbacks associated with the resulting higher precipitation and lower potential evaporation drive further changes in evaporation and runoff. Global evaporation from land was found to increase by about 7 cm for each doubling of storage capacity in the range from less than 1 cm to almost 60 cm [Milly and Dunne, 1994]. Such results may have implications for global hydrology and climate dynamics, including the effects of water resource development on global precipitation, climatic control of plant rooting characteristics, climatic effects of tropical deforestation, and climate-model errors induced by errors in land-surface hydrologic parameterizations. However, issues exist on global indicators of accelerations of the water cycle. For example, the evaporation of water, measured using evaporation pans, has been decreasing in the past few decades over large areas with different climates. The common interpretation was that the trend relates to increasing cloudiness, and that it provides an indication of decreasing potential evaporation and a decreasing terrestrial evaporation component in the hydrologic cycle, but it was shown [Brutsaert and Parlange, 1998] that pan evaporation has not been used correctly as an indicator of climate change.

Development of models that describes the circulation of water between atmosphere, soil, and plant life is important for a general understanding of the hydrologic cycle and the way ecosystems operate and survive [Chow, 1988; Falkenmark et al., 2004]. Terrestrial ecosystems play a crucial role in the water cycle. In particular, vegetation causes evaporative losses, modulating and increasing precipitation. Therefore, in absence of vegetation the water cycle would be much slower [Fraedrich et al., 1999]. The investigation of the spatiotemporal linkage between the hydrologic and ecologic dynamics, from both theoretical and empirical perspectives, are thus a crucial challenge for understanding the environment in which we

live [Rodriguez-Iturbe, 2000]. The study of the interaction of the water cycle with the biota has been termed ecohydrology [Zalewski, 2000; Rodriguez-Iturbe, 2000] and it is drawing the attention of on an increasingly broad segment of the science community [Baird A. J. and Wilby, 1999; Rodriguez-Iturbe et al., 1999a; Gordon and Folke, 2000; Porporato and Rodriguez-Iturbe, 2002; Settin et al., 2007; Botter et al., 2008a]. In fact, ecohydrology is contributing to the development of unifying concepts and new, testable theories for the understanding of the interplay between vegetation, climate and soil [D'Odorico et al., 2010a].

A first step in this direction is in the work by Rodriguez-Iturbe et al. [1999b] who provided an exact solution to a key probabilistic description of hydrologic controls on biota. In that work, in particular, the steady state probability function of soil moisture has been exactly obtained for rainfall of random and intermittent characteristics when both the infiltration as well as the losses depend on the state of the system in a nonlinear manner. The feedback of vegetation, in that context, proves crucial. The final product in Rodriguez-Iturbe et al. [1999b] is the probabilistic description of soil moisture at a point as a function of climate, soil and vegetation parameters. Shortly afterwards, a series of notable papers [Rodriguez-Iturbe et al., 2001; Porporato et al., 2001b,a; Laio et al., 2001] followed and extended the same probabilistic framework by analytically deriving the crossing properties of special relevance in the temporal evolution of soil moisture (i.e. the soil moisture threshold below which the plant is under water-related stress) and by applying the theoretical machinery to real-life applications. These works, expanded by new analysis, resulted in the book *Ecohydrology of Water-Controlled Ecosystems* [Rodriguez-Iturbe I. and Porporato, 2004] that is now a milestone in ecohydrology and from which part of the research of the present thesis has been inspired. Soil moisture affects the biota directly, by controlling the availability of resources for organisms, and indirectly, by modifying abiotic process that, in turn, affect ecosystem dynamics.

Owing to these and other studies, a quantitative analysis of the so called *green water flows*, (i.e. flows of water vapor or the rainwater lost from land in the form of transpiration, interception and evaporation from the soil and vegetation) has been laid out. Green water supports most terrestrial vegetation and food production because of its major impact on all physiological processes and because of the relatively large quantity of water required for the proper functioning of plants. Three main classes of stochastic models have been proposed in the last years. A first group of model describes the vertically-lumped water balance at the daily time scale within the root zone and it is suited for system of deep water tables. A second class of models includes vertically distributed models, which describes the vertical profile of soil moisture in the root zone [Laio et al., 2006]. Finally, a quantitative framework has been developed to investigate the influence of shallow aquifers on soil moisture dynamics [Rodriguez-Iturbe et al., 2007; Pan and Nakagoshi, 2008; Tamea et al., 2009]. The determination with these models of the analytical expression for the steady-state probability distribution of soil moisture in different hydrologic and climatic conditions have elucidated the complex and highly nonlinear interplay among climate, soil and vegetation, and their impact on vegetation water stress.

In the very recent last years, also a theoretical framework for the blue water flows dynamics,

based on the same stochastic analytical approach, has been developed Botter et al. [2007c,b,a]. *Blue water flows*, that refer to runoff and groundwater fluxes [Falkenmark et al., 2004], provide water and other resources to a number of organisms and societies. They also enhance landscape connectivity , which is fundamental to the transport of nutrients and sediments [Botter et al., 2010], the spread of species and the maintenance of biodiversity [Bertuzzo et al., 2009; D'Odorico et al., 2010a]. Botter et al. [2007c] presented a stochastic approach that allows to link the probabilistic structure of slow components of runoff with simple (pluviometric, soil, vegetation, and geomorphologic) macroscopic parameters, with implications for the ecohydrology of fluvial systems and for drought prediction in ungauged basins. The derived analytical solution for the seasonal probability distribution function of daily streamflows has allowed a physically based linkage between flow duration curves and the distribution of annual minima through geomorphic, climatic, and ecological features of the basin, with evident important ecological consequences [Botter et al., 2008a,b; Botter, 2010; Ceola et al., 2010].

Hydrologic controls on the ecological quantity is also on the connectivity structure of the space where the ecosystem lives. For instance river networks provide important resources and services that are crucial to the functioning of ecosystems and societies. Fluvial networks enhance the ecohydrological connectivity of the landscape. The commonalities existing among all types of river basins and their drainage networks suggest the possible existence of general rules commanding the role that their structure plays in ecological patterns. Moreover the ecological corridors provided by river networks induce anisotropy in the spreading of species, pathogenes or other agents along the waterways [Muneepeerakul et al., 2007; Bertuzzo et al., 2007]. In fact, the last years a general theory on the effects of dendritic geometries on ecohydrological processes operating at the scale of river basins has been developed (see Rodriguez-Iturbe et al. [2009] for a review). Among its notable applications, the theory was able to explain the spatial pattern of species dispersal in riparian corridors [Muneepeerakul et al., 2008a] or of cholera epidemics [Bertuzzo et al., 2008, 2011; Mari et al., 2011].

Ecohydrology also promises to help our understanding of the complex interactions of biosphere and hydrosphere with the "anthroposphere". The most critical interactions of ecohydrological processes with human societies are associated with food production. In fact food production is by far the most freshwater-consuming process (80% of the total world water resources [Rost et al., 2008]) and is largely performed to sustain the productivity of agroecosystems (croplands and rangelands) [Falkenmark et al., 2004]. While the societal demand is growing, water availability to ecosystems and societies is also changing as an effect of global and regional climate change and land use dynamics [Hoekstra and Chapagain, 2008]. To study how ecohydrological dynamics affect human societies, the concept of virtual water, defined as the amount of water used to produce a certain food commodity, was first introduced by Allan [1993], and then elaborated especially in the hydrologic milieu [Hoekstra, 2002; Chapagain et al., 2006; Hanasaki et al., 2010]. International trade links the fortunes and resources of countries, providing potentially important conduits for geographically limited water resources to be transferred to water-stressed regions. The virtual water trade between regions [Hoekstra

and Hung, 2005; Yang et al., 2006; Hanasaki et al., 2008b] and the gross virtual water flow of nations [Hoekstra and Chapagain, 2008] have been quantified. These studies have primarily focused on agricultural commodities [Liu et al., 2007; Rost et al., 2008; Hanasaki et al., 2008a], including those used for bio-fuel production [Gerbens-Leenes et al., 2009], but the concept has also been extended to include industrial products [Hoekstra and Chapagain, 2008].
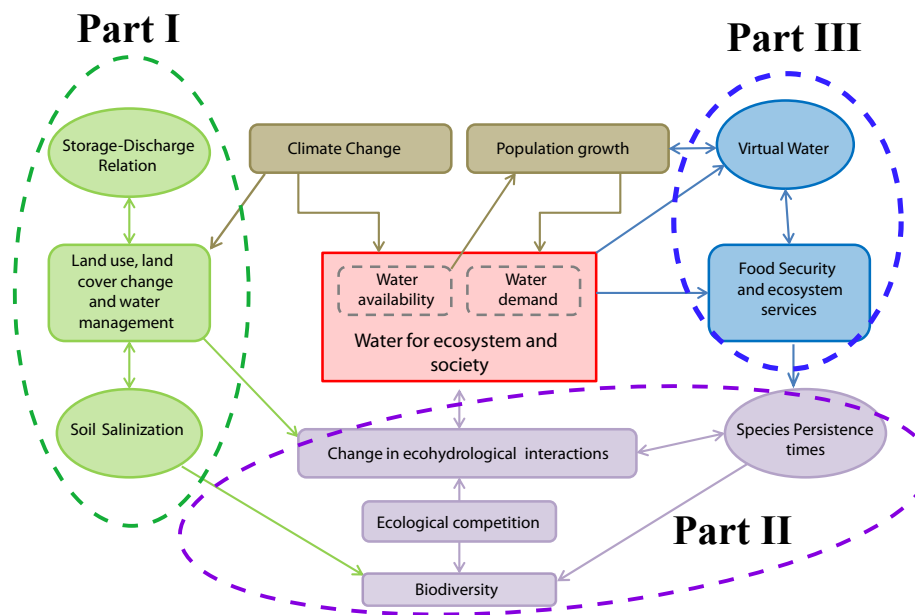
Figure 1.1: Correlation among major factors affecting water availability to ecosystems and societies considered in the present work. The complex relations among biosphere, hydrosphere and anthroposphere call for an integrated understanding of ecological and hydrological processes and their interaction with human society. Population growth contributes to the increase in water demand by humans. Climate variability and land-use change affect water availability either directly (e.g. by modifying the rainfall regime), or indirectly, through their impact on vegetation−water resource interactions. The competing needs for water by ecosystems and societies affect food security, human health, as well as the amount of water, land, and other resources available for natural ecosystems (freely redrawn from D'Odorico et al. [2010a]).

As we have just seen, ecohydrology studies the relation between hydrologic mechanisms and the underlining ecologic patterns and processes. This broad definition includes a wide variety of both ecological and hydrological processes, and therefore the signature of ecohydrological footprints can be found in several complex systems very different among each other. Ecohydrological footprints can be thus defined as the response of ecosystem functions or services, caused by the change in their hydrologic drivers. They are rooted in the analysis of the natural capital and environmental impacts [Wackernagel et al., 1999].

The view emerging from current research in ecohydrology is that there is a need for an integrated understanding of ecological and hydrological processes [Rodriguez-Iturbe et al., 2009;

D'Odorico et al., 2010a]. In spite of the large work produced in the last decade, there still exist open issues that need be addressed [D'Odorico et al., 2010a]. In particular, the present work is focused on the following major challenges:

- the effect of stochasticity in the analysis of the blue water flow. Existing works, in fact, have considered deterministic constitutive relations between catchment storages and streamflows, while theoretical and practical considerations suggest that noise and variability in this constitutive relation may be an important factor and the interplay among several nonlinear phenomena call for more complex assumptions than strict determinism between subsurface storage and discharge;

- Because most forms of terrestrial life depend directly or indirectly on soils, it is crucial to understand the long term and long range effects of soil loss. However, we have just begun to investigate how changes in soil moisture, fires, soil salinity or shifts in plant community composition might affect soil susceptibility to wind and water erosion;

- The interaction between biodiversity loss, ecohydrological process and spatial structure of the environmental matrix has only started to be appreciated and remains and important focus for future research;

- Although the stochastic frameworks developed in ecohydrology contribute to the understanding of the variability in natural resources, the understanding of the coupling with societal dynamics remains a major research task for the study of the ecohydrological interactions with human societies.

In the present thesis, steps in each of the above directions are pursued.

A major conceptual thread joining the various issues is the relation between ecohydrological footprints and the role of stochasticity. In general, the dynamics of complex ecohydrological system is affected by environmental variability. Stochasticity is inherent to weather patterns, climate fluctuations, evapotranspiration, to name a few relevant processes, and therefore a significant part of environmental variability is random [Lyon, 1973; Sole and Bascompte, 1996; D'Odorico et al., 2000; Rodriguez-Iturbe I. and Porporato, 2004; Raser and O'Shea, 2005]. Moreover, because stochasticity is intrinsic to environmental variability, noise can play an important and constructive role in ecohydrological processes as it may induce new states, enhance diversity, or lead to pattern formation [Ridolfi et al., 2011].

There exist three broad classes of approaches to handle process complexity: a) work at small scale and fully resolve the complex processes (this usually doesn't work because of the significant data requirements) b) upscale the mathematical relationships from the small scale to develop new deterministic relationships that work at a larger scale (macrodispersion is the prototypical example of this for groundwater) c) employ a stochastic approach in which the complexity is captured not explicitly but its effect is described probabilistically. In the
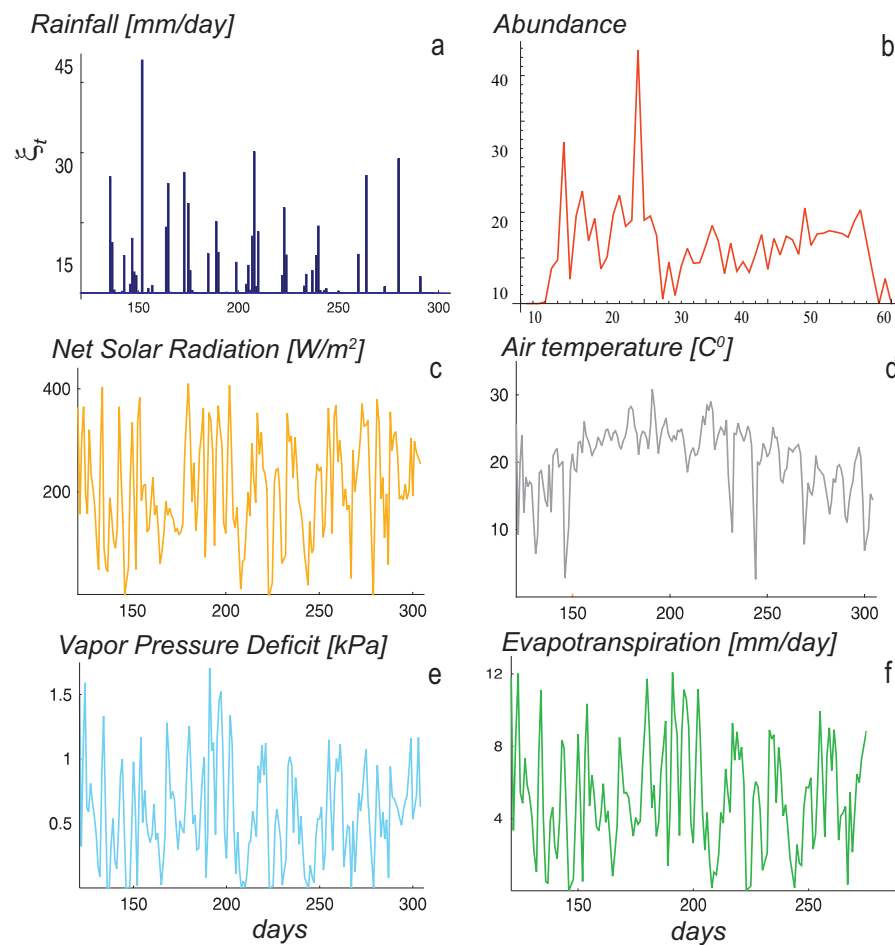
Figure 1.2: Examples of environmental fluctuations in ecohydrological systems dynamics: a)Precipitation recorded during the 2001 growing season at the Duke Forest site, North Carolina [Laio et al., 2001] b) Time series of abundance Rivul population. Data are from Balata et al. [2007]. d,e,f,g) Climatic time series from the same database of a).

thesis, the last approach has been confidently used to show how, in this way, analytical or semi-analytical models are built that may help to take further steps forward in wide range of different problems. Indeed, complex systems are characterized by a very high number of degrees of freedom (the system can be achieve many possible states), and yet their emergent behavior can be often largely independent of microscopic details of the system [Rodriguez-Iturbe and Rinaldo, 1997; Barabási and Albert, 1999; Stanley, 2000].

In this thesis, a variety of new models are proposed to study ecohydrological footprints through stochastic approaches where the intrinsic complexity is not captured explicitly but rather described probabilistically. In this way analytical quantitative framework analysis are laid out to facilitate progress in the estimation of the effects of environmental change on the interaction of the hydrologic process with the biota. Tools provided by statistical mechanics are necessary

in order to walk along this challenging path. In fact stochastic differential equations, master equations, statistical ensembles and network theory, which will be the building bricks of our models, are all mathematical tools familiar in this branch of theoretical physics. In particular stochastic differential equations (SDE) will be a recurring theme in our methodology and the first section in the Appendix is dedicated to give a short introduction on SDE to those that are not familiar with stochastic processes.
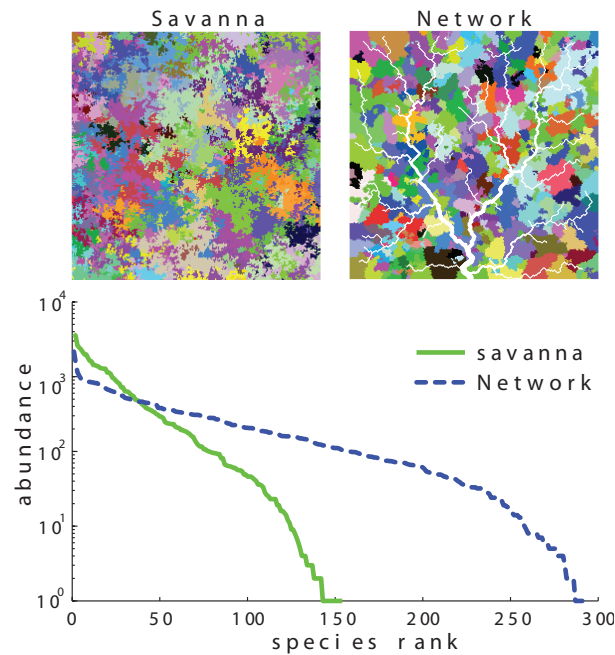
Figure 1.3: Comparison between biodiversity patterns in a savanna and in a network landscape represented by species spatial patterns (upper insets) and species rank-abundance curves (extracted from Rodriguez-Iturbe et al. [2009]).

In the present work, three different type of ecohydrological footprints will be studied. In part I of the thesis we will study the effect of blue water flows on water controlled ecosystem. In particular we will focus our attention on two topic: the relation between storage and discharge in river basin, and the effect of the salinization of the root zone on the basin vegetation. In part II we will focus on effects of the spatial structure of the environmental matrix on ecosystems dynamics. For instance, Figure 1.3 shows how hydrologic controls on the ecosystem dynamics is also on the connectivity structure of the space where the ecosystem lives. In particular we will investigate these effects on a particular ecological quantity, namely the species persistence times distributions. Finally in part III we will analyze the water footprints of human society. In fact, by trading food, nations continuously make exchange of water. These trades are well described by the global virtual water network, that takes in account all the fluxes of water associated with the trades of foods among nations. In this part we will build, analyze and model this network and also study its implication for human society. Each of these three parts is introduced by a short overview that summarizes and contextualizes its content.

Effect of green and blue water flows on water controlled ecosystem

Overview

Two problems are studied where hydrologic fluxes are crucial drivers of their ecosystem dynamics. In both cases, the problem is tackled using a probabilistic approach based on stochastic differential equations. In particular, it is shown from two different perspectives how green and blue water flows provide important ecohydrological footprints.

In the analysis of the blue water flow, previous works had considered deterministic constitutive relations between catchment storages and streamflows. Theoretical and practical considerations, suggest that noise and variability in this constitutive relation may be an important factor [e.g. McDonnell, 1990; Beven, 2001; Berne et al., 2005; Morbidelli et al., 2006]. If runoff-production schemes, for instance, account for spatial variability in the parameters characterizing the individual hydrologic processes, fluctuations inevitably emerge. For catchment scale predictions, point-scale representations of infiltration have been challenged by numerical approaches that have accurately explored infiltration processes including runoff-runon, rainfall variability, spatial correlation in infiltration and rainfall fields [e.g Beven, 2001; Berne et al., 2005; Morbidelli et al., 2006]. Moreover, connectivity patterns within a hillslope, or within any runoff-generating volume, change erratically as complex activation of preferential flowpaths may result in noise in the outflow discharge [e.g McDonnell, 1990; McGuire et al., 2005; Fiori and Russo, 2008]. For all these reasons the impact of stochastic fluctuations in the storage-discharge relation was deemed worth investigating. To this aim is devoted Chapter 2.

In Chapter 3, by means of similar mathematical treatments, another ecohydrological problem of great societal and environmental importance is studied: the problem of soil salinization. Salinity refers to excess concentrations of easily soluble salts present in soil water. It is one of the major problems in agriculture, especially in semi-arid regions. Salt accumulation in the root zone may be due to natural factors (primary salinization) or to irrigation (secondary salinization). Detailed numerical models have been developed to model soil salinization [e.g. Nour eldin et al., 1987; Schoups et al., 2006; Suarez and Simunek, 1991; Shah et al., 2011]. Generally, these models simulate unsaturated soil water flow via the Richards and solute transport equations. These models are suitable for local and short-term simulations, as they require precise soil characterization and are computationally demanding. Moreover, it is often difficult to identify clear cause-effect relationships or to synthetically compare the effects of different parameter scenarios from their numerical simulations. For this reason the proposed exact solution to a minimalist stochastic model of primary soil salinity, is deemed of interest. The long term probability distributions of salt mass and concentration are obtained by reducing the coupled soil moisture and salt mass balance equation to a single stochastic differential equation. The novel analytical solutions provide insight on the interplay of the main soil, plant and climate parameters responsible for long-term soil salinization. Mathematical spinoffs, published independently, are in Appendix A.

Part I

The Impact of Stochastic Fluctuations in Storage-Discharge Relations on Streamflow Distributions

He used statistics as a drunken man uses lampposts; for support rather than illumination.

A. Lang

2.1 Why a Stochastic Approach?

The complete probabilistic characterization of streamflows in river basins is a because of the noteworthy implications on water resources availability and management for human needs and ecological services related e.g. to riparian plant nutrition, preservation of fish habitat, irrigation, or storage management [e.g Brutsaert, 2005]. Streamflows at the closure of whole river basins are the byproduct of many intertwined eco-hydrological and climatic processes, such as infiltration from rainfall, evapotranspiration, runoff production and transport phenomena occurring in channeled and unchanneled regions of the basin. The intrinsic temporal variability embedded in the fluctuations of recorded runoff series thus reflects the vagaries of rainfall patterns in space and time and the random character of several related hydrologic processes and landscape morphologies [e.g Chow, 1988; Lamb R. and Beven, 1997; Brutsaert, 2005; Eng and Milly, 2007].

Stochastic fluctuations of streamflows have long been the subject of hydrologic and statistical investigations, in particular through the synthetic generation of time series obtained by

deterministic models of the hydrologic response driven by stochastic climate forcings [e.g Kottegoda N.T. and Horder, 1980]. Effective rainfall series, seen as trajectories of a stochastic process with prescribed statistical features [e.g Xu et al., 2002] or censored through soil-water balances applied to stochastic point rainfall processes [e.g Rodriguez-Iturbe I. and Porporato, 2004], have also been studied. Specific assumptions or observations are needed down to the scale of individual hydrologic processes, like recharge due to infiltration from rainfall, losses (evapotranspiration and discharge) and storage variations related to residence time distributions in channeled and unchanneled states [e.g Rodriguez-Iturbe I. and Porporato, 2004]. The temporal variability of streamflows thus reflects the stochastic nature of several underlying processes, which induces complex causal relations [e.g Lamb R. and Beven, 1997; McDonnell, 1990; McGuire et al., 2005]

In this general context, Botter et al. [2007c] have analyzed the linkage existing between streamflow distributions and the relevant soil moisture dynamics in catchments. A probabilistic model of streamflow dynamics was developed therein, where the steady-state probability distribution of the subsurface contribution to streamflows is analytically expressed in terms of a few macroscopic rainfall properties, soil-vegetation parameters, and key geomorphological features. The approach was initially structured in a spatially lumped framework by assuming average properties, as in the related literature on soil moisture dynamics [see e.g Rodriguez-Iturbe et al., 1999b; Rodriguez-Iturbe I. and Porporato, 2004], and it has been later extended to tackle spatially distributed soils, vegetation and morphological features and to derive flow-duration curves [Botter et al., 2007b, 2008a].

Key to the exact solutions described above was the assumption of a linear, deterministic relationship between the instantaneous outflow discharge and the water volume stored in the subsurface. Such an assumption is often assumed in practice coupled with nonlinear net precipitation schemes [e.g Chow, 1988; Beven, 2001], and is equivalent to assuming an invariant exponentially-distributed residence time.

Nonlinear storage-discharge relations have also long been considered in conceptual models of the hydrologic response [see, for a review Brutsaert, 2005]. Recent studies have revamped the interest on the subject by applying the method of Brutsaert and Nieber [1977] to characterize catchments as nonlinear dynamic systems, concluding that linear storage-discharge relations are not expected in general at the scales of interest [Kirchner, 2009]. Interestingly, this applies for the class of catchments in which discharge strictly depends on the volume of water stored in the subsurface [Ceola et al., 2010].

These previous works however have all considered deterministic constitutive relations between storages and streamflows, while stochasticity may play a crucial role in runoff formation.

Recently, Harman and Silvapalan [2009] examined how heterogeneity affects the flow response of a hillslope as a whole through numerical simulation. Overall, hydrologic storages are seen as random variables proportional to the probability of the travel times of rainfall particles to their exit boundaries [e.g. Fiori and Russo, 2008], thus sustaining the idea that random

fluctuations must affect the process of runoff formation.

The ratio of computed instantaneous discharge and stored volume for a heterogeneous system mimicking a hillslope transport volume clearly suggests that a noisy relation is indeed the norm rather than the exception, whether or not the mean behavior of the response of the system may be approximated by a linear storage-discharge relation. Note that more than to a whole watershed, the current control volume refers to a hillslope where a local storage-discharge relation indeed makes sense. The catchment-scale complexity arises from the combined effect of serial and parallel arrangement of runoff-generating control volumes like geomorphically linked sources areas. At the chosen scale it is believed that indeed stochastically affected hydrological processes result in greater variance of streamflows. Hydrologic connectivity [e.g. Gomi et al., 2008] also complicates matters, because only a fraction of the runoff generated in a catchment actually connects with the outlet during the rainfall event. As a result, hydrologic processes are highly variable in space and time, and dynamic changes in the spatial extent and the timing of runoff-runon phenomena call for more complex assumptions than strict determinism between subsurface storage and discharge. Moreover, it can be shown theoretically that the effects of spatial heterogeneity in material properties can be accounted for by adding a time-varying noise to the dynamic equations. Briefly, consider a system where particles are traveling along many pathways, each with a mean velocity $\bar{v}$, perturbed by a spatial random noise $\varepsilon(x)$, so that $v = \bar{v} + \varepsilon(x)$. In the limit of a large ergodic cloud of particles, the approximation $\varepsilon(x) = \varepsilon(vt) \approx \varepsilon(\bar{v}t) \propto \varepsilon(t)$ can be done, and thus incorporate the effect of the spatial variability using a time−dependent noise term.

2.2 Stochastic Model for Soil Moisture and StreamFlows Dynamics

In this section it is briefly reviewed the modeling scheme presented by Botter et al. [2007c,b] based on stochastic water balance equations, which provides a linkage between the probabilistic structure of streamflows and underlying eco-hydrological, climate and transport processes in relatively small vegetated catchments (see Figure 2.1).

Following Rodriguez-Iturbe et al. [1999b] and Rodriguez-Iturbe I. and Porporato [2004], rainfall is assumed as a marked point process with frequency λ_p $[T^{-1}]$. This assumption implicitly postulates catchment sizes (the spatial scale of the control volume *sensu* Kirchner [2009]), say A, smaller than the spatial correlation scale of rainfall events and timescales of the process of interest greater than or equal to daily (e.g., larger than the temporal characterization of rainfall events). Furthermore, daily rainfall depths are assumed to be exponentially distributed with mean $1/\gamma_p$ $[L]$. The catchment-averaged soil moisture dynamics is modeled assuming constant (spatially and temporally averaged) soil and eco-hydrological parameters: effective soil depth, Z_r, porosity, n, maximum evapotranspiration rate, ET_{max}, which are assumed as representative of a prescribed season [see e.g. Rodriguez-Iturbe et al., 1999b; Rodriguez-Iturbe I. and Porporato, 2004]. The temporal evolution of spatially-averaged relative soil moisture, $s(t)$, is given by the mass balance equation within the topsoil layer of the catchment [for a
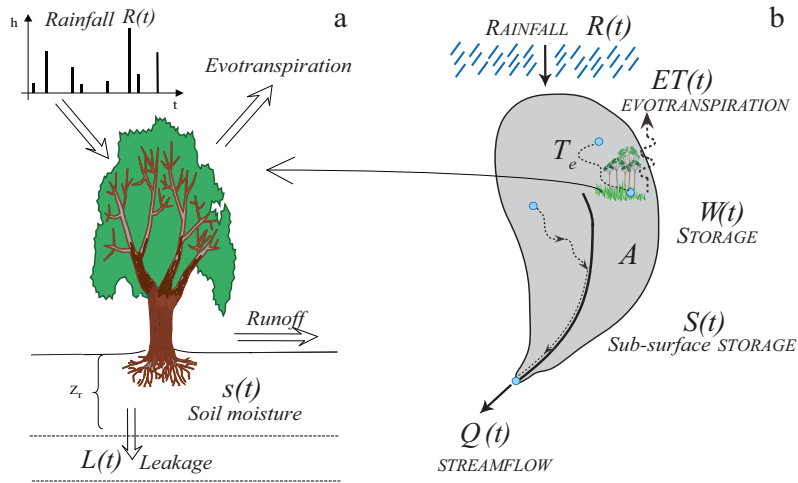
Figure 2.1: Schematic representation of a)The vertically-lumped water balance at the daily time scale within the root zone on which the stochastic model for soil is based b)The water mass balance equation over the whole catchment that allow to extend the stochastic approach also on the streamflow dynamics.

review see Rodriguez-Iturbe I. and Porporato, 2004]:

$$nZ_r \frac{ds}{dt} = -ET - L_t + R_t,$$ (2.1)

where R_t represents the inflow, i.e., the stochastic increments due to infiltration from rainfall, while the losses are the evapotranspiration ET (assumed to be linear in the range of soil moisture comprised between the wilting point, s_w, and a suitable soil moisture threshold s_1 for leakage to occur), and the leakage, L_t, due to deep percolation towards deeper layer and ultimately streamflow. This model follows the minimalist probabilistic soil moisture model described in detail by Rodriguez-Iturbe I. and Porporato [2004]. Figure 2.2a,b shows a typical realization of the stochastic rainfall model and the resulting temporal evolution of soil moisture $s(t)$.

Note that the timescales chosen require to focus only on subsurface contributions to stream-flow, thus neglecting fast surface runoff possibly triggered by intense storms. In the absence of pronounced topographic effects and of impervious areas, however, the surface contribution to runoff is usually not significant with respect to the corresponding subsurface contribution at large timescales [e.g. McDonnell, 1990; McGuire et al., 2005; Rinaldo et al., 2006]. The latter is linked to percolation from the top soil layer, which in turn is assumed to be triggered by the exceedance of the soil moisture threshold s_1 (Figure 2.2b), whose value lies typically between the field capacity and soil saturation. The temporal evolution of the water storage of the whole catchment can thus be schematized as sum of the variation of the top soil layer, described by

equation (3.1), and that of the catchment subsurface water storage, say $S(t)$, described by:

$$\frac{dS(t)}{dt} = -Q(t) + A \cdot L_t,$$

(2.2)

where Q is the streamflow subsurface discharge and A the area of the catchment.

Under the given assumptions, the spatially-averaged percolation process L_t can be well approximated by a marked Poisson processes $\xi_t(\lambda, \gamma_P)$ with frequency parameter λ $[T^{-1}]$ and percolation depths exponentially distributed with parameter γ_P. The percolation interarrival frequency can be expressed in terms of the underlying rainfall, soil and vegetation properties as follows [Botter et al., 2007a]:

$$\lambda = \eta \, \frac{\exp(-\gamma) \, \gamma^{\frac{\lambda_P}{\eta}}}{\Gamma(\lambda_P/\eta, \gamma)} \; ;$$

(2.3)

where $\Gamma(a, b)$ is the incomplete gamma function of parameters a, b; $\eta = ET_{max}/(nZ_r(s_1 - s_w))$ and $\gamma = \gamma_P nZ_r(s_1 - s_w)$. The mass balance equation for the catchment subsurface water storage (2.2) can thus be expressed as

$$\frac{dS(t)}{dt} = -Q(t) + A\xi_t(\lambda; \gamma_P)$$

(2.4)

where $\xi_t(\lambda; \gamma_P)$ represents the time series of percolation inputs.

Water pulses deeply infiltrating into soil are assumed to be released toward the stream network as subsurface or groundwater flow with a rate proportional to the instantaneous subsurface water storage and thus S is connected to the streamflow discharge by the linear relation

$$Q(t) = kS(t),$$

(2.5)

where k $[T^{-1}]$ is the inverse of the mean residence time in subsurface. Under the above assumptions the steady-state pdf $p(S)$ of the subsurface water storage is expressed by a Gamma distribution with shape and scale parameters λ/k and $1/\gamma_P$ respectively [Botter et al., 2007c]. By expanding the left hand side of equation (2.4) as $dS/dQ \cdot dQ/dt$, a stochastic differential equation is derived for the temporal evolution of the streamflow discharge $Q(t)$
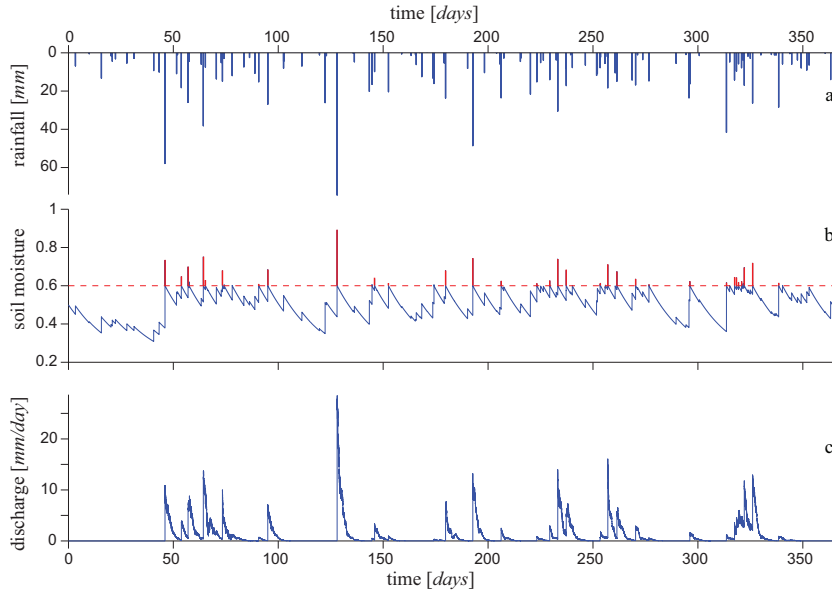
Figure 2.2: Schematic representation of the soil moisture and streamflows dynamics: (a) Temporal evolution of the overall rainfall depths (simulated data). The interarrivals and the rainfall depths are exponentially distributed with a frequency $\lambda_P = 0.3\,d^{-1}$ and with normalized mean intensity $1/\gamma_P = 16.7\,mm$, respectively. (b) Temporal evolution of the (catchment averaged) relative soil moisture, $s(t)$, which is commanded by the intermittent rainfall forcing shown in Figure 2.2a and by the deterministic decay due to evapotranspiration process, according to equation (3.1). The dash-dotted line represents the threshold s_1, whose up crossing determines the triggering of runoff events. Temporal sequence of rainfall excess, driven by the exceedence of the threshold s_1 for the soil moisture $s(t)$ are represented by the red spikes. (c) Temporal evolution of the overall, specific (i.e., for unit area) discharge. The soil, vegetation, and transport parameters employed for this simulations are $n = 0.55$, $Z_r = 30\,cm$, $s_w = 0.18$, $s_1 = 0.6$, $ET_{max} = 0.35\,cm/d$, and $k = 0.5\,d^{-1}$.

(Figure 2.2c):

$$\frac{dQ(t)}{dt} = -k\,Q(t) + k\,A\,\xi_t(\lambda;\gamma_P), \tag{2.6}$$

the stationary solution of which is [Botter et al., 2007c]:

$$p(Q, t \to \infty) = p(Q) \sim Q^{(\frac{\lambda}{k} - 1)} \exp(-\gamma_Q\,Q), \tag{2.7}$$

where $\gamma_Q = \gamma_P/(k\,A)$ represents the mean runoff increment due to incoming percolation events.

The probability distribution of Q is thus related to the underlying soil and vegetation properties (through the parameter λ) and to key rainfall properties (through both the parameters γ_p and λ), but it also depends on important geomorphic factors such as the mean residence time of subsurface flow ($1/k$) and the size of the basin (A). In particular, according to equation (2.7), the behavior of $p(Q)$ is chiefly controlled by the ratio between the percolation frequency, λ, and the inverse of the mean residence time in subsurface, k. When $\lambda/k > 1$ ('wet conditions') the pdf of the storage (runoff) is hump-shaped with $p(0) = 0$ (i.e., a zero storage (runoff) is characterized by zero probability), while for $\lambda/k < 1$ ('dry conditions') $p(Q)$ goes to infinity for $Q \to 0$, and it monotonically decreases for $Q > 0$.

Although the simplifications made, $p(Q)$ captures the observed behavior of the streamflow pdf reasonably well in many cases of practical interest [e.g. Botter et al., 2007a; Ceola et al., 2010].

2.3 Stochastic Storage-Discharge Relation

In order to take into account the stochasticity in the relation between storage and streamflow (due to transient connectivity, differential activation of preferential flowpaths, macroscopic effects of spatially heterogeneous and hysteretic subsurface properties), a Gaussian colored noise, $l(t)$, of the Ornstein−Uhlenbeck type [Gardiner, 2004] is added to the linear deterministic relation described by equation (2.5):

$$Q(t) = -k\, S(t) + l(t)\, S(t)\,. \tag{2.8}$$

The Ornstein−Uhlenbeck process is chosen because it is the simplest model of correlated noise. The property of non-zero correlation is physically meaningful, and is specifically chosen to describe the nature of fluctuations in hydrologic processes that have non-negligible correlation scales, both in space and time. Gaussian white noise is also recovered as the zero-correlation limit of the Ornstein−Uhlenbeck process. At stationarity, $l(t)$ has mean $\langle l(t) \rangle = 0$ and it is characterized by a correlation structure $\langle l(t) \cdot l(s) \rangle = (D\tau/2)\, e^{-|t-s|/\tau}$. The relaxation time τ indicates the characteristic time for which two fluctuations cease to be correlated, while D is the equivalent of a diffusion coefficient which represents the amplitude of stochastic fluctuations (note that $\langle \cdot \rangle$ is the ensemble average operator). A colored noise, differently from white noises that have infinite variance, does not suffer from this limitation [Gardiner, 2004] and thus $Q(t)$ in equation (2.8) is well defined for all t ensuing to this variable an appropriate hydrologic meaning. The Ornstein−Uhlenbeck process is described by the discrete Langevin equation:

$$l(t + \Delta t) = l(t) - \frac{1}{\tau} l(t)\Delta t + \sqrt{D} \cdot \Delta W(t), \tag{2.9}$$

where W is the Wiener process, i.e., $\Delta W(t) = W(t + \Delta t) - W(t)$ is a temporally uncorrelated normal random variable with mean zero and unit variance. Thus the stochastic storage-discharge relation can be studied by combining equations (2.4), (2.8) and (2.9). The evident hydrologic consequence of equation (2.8) is that now the streamflow corresponding to a given storage S can fluctuate around the deterministic value kS within a range that depends on the variance of the Ornstein−Uhlenbeck process.

Figure 2.3a illustrates the typical scatter produced by the colored noise in the storage−discharge relation (obtained by suitable MonteCarlo simulation), while Figure 2.3b shows the results produced by the [Brutsaert and Nieber, 1977] method that probes the nature of the $S - Q$ relation via numerical manipulation of the recession curves. This method can be obtained by differentiation of the axis of Figure 2.3a and it allows to estimate the mean residence time $1/k$ solely from the measured discharge time series. Sampling effects due to the time step used in the analysis, however, enhances the noise in the $S - Q$ relation (see comparison between Figures 2.3a and 2.3b). It is interesting to observe that the introduction of the artificial noise qualitatively resembles field and numerically simulated data [Kirchner, 2009; Brutsaert and Nieber, 1977].
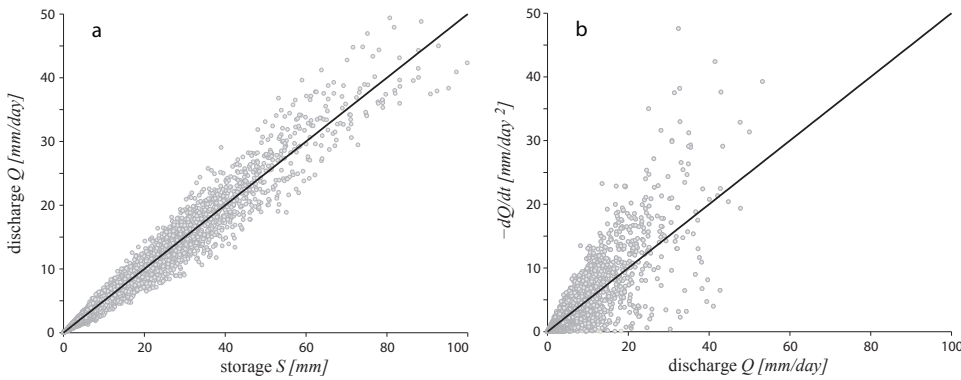
Figure 2.3: a)Storage−discharge relationship arising from addition of colored noise; b) Application of the Brutsaert and Nieber [1977] method for 6 hour time-step that probes the nature of the $S - Q$ relation via numerical simulation of the recession curves. The parameters employed are $k = 0.5\, d^{-1}$ and $Var(l) = 0.15\, k$.

Figure 2.4 summarizes the main results of the numerical simulation. For the integration of the Langevin equation (2.9) the Euler-Maruyama algorithm [Higham, 2001] is employed. Because negative values of the streamflow Q have no physical meaning, in the simulation a reflecting boundary conditions in zero for the process $(l(t) + k)$ is imposed. It can be noticed that the shape of the probability distribution of the storage $p(S)$ is not significantly affected in a broad range of noise conditions (Figure 2.5). On the contrary the noise has a major effect on the streamflow distribution $p(Q)$. In particular, the mode of $p(Q)$ shifts towards zero as the strength of the noise increases. The sensitivity of S and Q to noise in the Q-S relationship

depends on hydrologic measures, such as the Index of Dryness $\mathscr{I}_D = \langle ET \rangle / \langle R_t \rangle$. In fact, in arid or semi-arid climates where $\mathscr{I}_D \approx 1$ deterministic evapotranspiration drives the evolution of S and thus $p(S)$ and $p(Q)$ are less affected by intrinsic random fluctuations in the storage-discharge relation with respect to wet climates (see Figure 2.5). The limited effect of the noise on the $p(S)$ with respect to $p(Q)$ can be explained as follows. From equations (2.4) and (2.8), in fact, a realization of the trajectory of the storage as given by $S(t) \sim \exp(-\int_0^t [k + l(u)]\,du)$ is obtained. Thus for large t and for relaxation times τ short with respect to $1/k$, the stochastic contribution to the latter integral tends to zero ($\langle l(t) \rangle = 0$). This can be generalized for any colored noise with zero mean. In fact, because of the nature of the mass balance equation, S depends on the noise in an integral fashion thus regularizing its fluctuations.
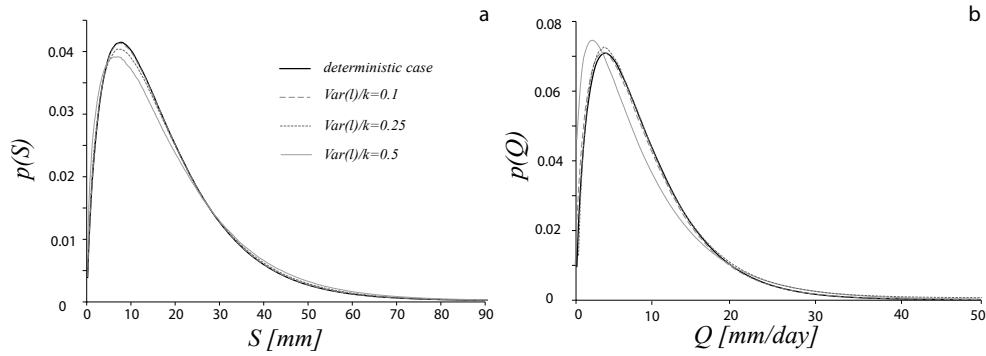
Figure 2.4: Probability distribution of a) the storage $p(S)$ and b) the streamflow $p(Q)$, calculated taking into account different variances of the colored noise in the storage-streamflow relation (equation (2.8)). The black curves correspond to the distributions derived from a deterministic storage-streamflow relation $Q = kS$. The parameters employed are $k = 0.5\,d^{-1}$, $\lambda = 0.5\,d^{-1}$ and $\tau = 1\,d$. The noise appreciably affects the shape of $p(Q)$ shifting the mode of the distributions toward zero even in the case of relatives low values of its variance.

The above results suggest that the noise in the storage-discharge relation significantly and interestingly influences the pdf of the streamflow. However this framework does not allow further analytical investigation. In order to achieve exact solutions that can improve the understanding of the effect of the noise on the streamflow distribution, the stochastic fluctuation in the storage-discharge relation by including a Gaussian multiplicative noise affecting the discharge equation of a linear reservoir is mimicked. The reliability of this modelling scheme is justified via numerical simulations (see Figure 2.5 and discussion in the following sections). In this way the stochastic relation between S and Q still holds and the effect of the noise on the discharge distribution can be deepened.

2.4   Noise Effects on the Streamflow Discharge Distribution

As discussed in the previous section, in order to mimic the stochasticity of the streamflow generation process, a Gaussian white noise, $\zeta(t)$, is added to the temporal variability of the
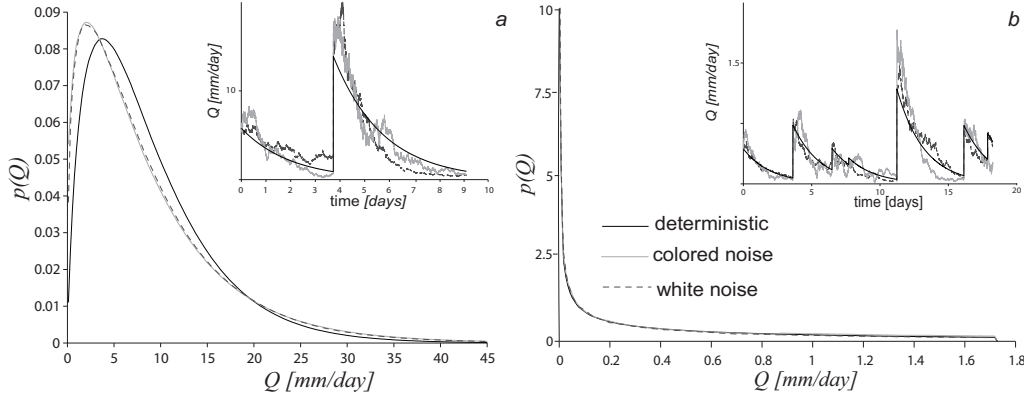
Figure 2.5: a) A comparison between different numerical streamflows probability density functions $p(Q)$ for wet climate ($\lambda = 0.875\,d^{-1}$, $1/\gamma_P = 20\,mm$ and $k = 0.5\,d^{-1}$). The effect of both colored and white noise is to shift the mode of the distribution toward zero. The parameters used for this calculation are respectively $\sigma^2 = 0.125\,d^{-1}$ for the white noise and $\tau = 1\,d$, $D = 0.0625\,d^{-1}$ for the colored noise. A numerical simulation of typical trajectories of streamflows discharge time series in a wet regime for the same cases is also reported as an inset. Note that the parameters of the white noise are tuned to reproduce the pdf of the colored noise simulation. This is deemed legitimate because it is simply claimed that white noise with appropriate parameters reproduces $p(Q)$ forced by colored noise. b) The same comparisons, but for dry climate ($\lambda = 0.2\,d^{-1}$, $1/\gamma_P = 10\,mm$ and $k = 0.5\,d^{-1}$) and same σ^2, τ and D used in a). It is evident how, for dry climate, Q (and S) are much less affected by the intrinsic random fluctuations.

outflow discharge equation:

$$\frac{dQ(t)}{dt} = -k\,Q(t) + k\,A\,\xi_t(\lambda;\gamma_P) + \zeta(t)\,Q(t), \tag{2.10}$$

where $\langle \zeta(t) \rangle = 0$ and $\langle \zeta(t)\zeta(s) \rangle = 2\sigma^2\delta(t-s)$; δ denotes the Dirac delta function. The physical interpretation of equation (2.10) is that the discharge responds to an instantaneous increase of the subsurface storage ΔS with an instantaneous increase of discharge $\Delta Q = k\Delta S$, as in the linear deterministic case, while between interstorms arrivals the linear decrease of the discharge is perturbed by the noise $\zeta(t)$. As a result, the generalized Langevin equation (2.10) describes the streamflow dynamics taking into account both the variability due to the rainfall filtered by the topsoil layer and the intrinsic stochasticity of the discharge generation processes. Equation (2.10) is valid as long as storage discharge relation may be thought to hold (say, not for overland-flow dominated regimes).

Note that equation (2.10) needs be to properly interpreted by a mathematical viewpoint (see Appendix A.2). The correct mathematical interpretation is obtained by imposing the condition of stationary storage reservoir. Taking the temporal mean of both sides of equation (2.4) the

following result is obtained:

$$\frac{1}{T}\int_0^T \frac{dS}{dt}dt = -\frac{1}{T}\int_0^T Q(t)dt + \frac{A}{T}\int_0^T \xi_t(\lambda;\gamma_P)dt \tag{2.11}$$

For $T \to \infty$ the system reaches statistical steady state and therefore the left hand side of equation (2.11) is equal to zero and $\frac{1}{T}\int_0^T x\,dt \equiv \langle x \rangle$, thus equation (2.11) reads as $\langle Q(t) \rangle = A\langle \xi_t(\lambda;\gamma_P) \rangle$. Analogously, from the generalized Langevin equation (2.10) the average discharge $\langle Q(t) \rangle = A\langle \xi_t(\lambda;\gamma_P) \rangle + \langle \zeta(t)Q(t) \rangle / k$ is obtained. Therefore the condition of stationary reservoir translate into $\langle \zeta(t)Q(t) \rangle = 0$. The latter condition is satisfied if the generalized Langevin equation is interpreted in Itô sense [Gardiner, 2004] (Appendix A.2).

Following the Itô interpretation, the steady state solution of equation (2.10) for the streamflow discharge pdf is (Appendix A.2):

$$p(Q) \sim p(Q, t \to \infty) = Ce^{-Q\gamma_Q}Q^{-\alpha+\frac{\beta}{2}-2}L^{\beta}_{\alpha-\frac{\beta}{2}}(Q\gamma_Q), \tag{2.12}$$

where C is the normalization constant; $\alpha = k/2\sigma^2 - 1/2$, $\beta = \sqrt{4\lambda\sigma^2 + (\sigma^2 + k)^2}/\sigma^2$; and $L^b_a(x)$ is the generalized Laguerre polynomial [Abramowitz and I.A., 1965]. A remarkable property of $p(Q)$ is that it has a power-law tail, i.e. for $Q \to \infty$

$$p(Q) \propto Q^{-(2+k/\sigma^2)}, \tag{2.13}$$

and thus that just adding a multiplicative Gaussian noise to the linear decay of the discharge, the tail of the streamflow distribution changes from the exponential (equation 2.7) to the power-law (equation 2.13). This also induces a change in the mode of the distribution, thus substantially altering the streamflow regimes with respect to the purely deterministic linear case.

2.5 Results and Discussion

A comparison between numerical streamflows probability density functions corresponding to three different cases (deterministic, white and colored noise) is shown in Figure 2.5. While for wet case (Figure 2.5a) both colored and white noise produce the effect to shift the mode of the distribution towards zero, in dry climates (Figure 2.5b) Q (and S) are much less affected by the noise. The results shown in Figure 2.5 support the assumption of considering a noisy storage-discharge relation as a byproduct of a memoryless stochasticity in the streamflow
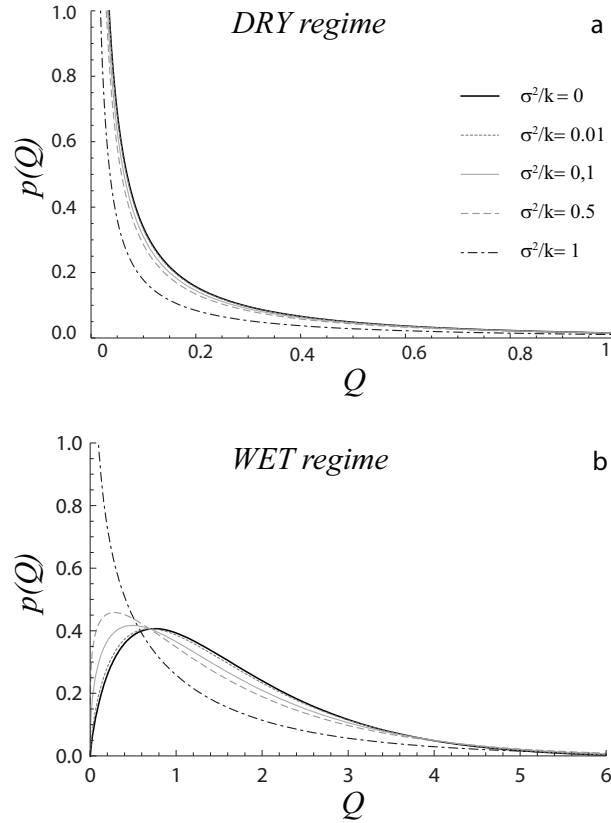
Figure 2.6: Analytical probability distribution of daily streamflow discharge (equation (2.12)) for: a) dry regime ($\lambda/k = 0.04$) and b) wet regime ($\lambda/k = 1.75$). Different lines represent different values of the dimensionless parameter ratio σ^2/k. Note that $\sigma^2/k = 0$ refers to the case of the deterministic reservoir (equation (2.7)).

generation processes. A numerical simulation of typical trajectories of streamflows discharge time series for the same cases is also reported as an inset of Figure 2.5.

Figure 2.6 shows the effects induced by the white multiplicative noise on a dry streamflow distribution (equation (2.12)). As the noise (σ^2/k) increases the probability of observing low values of streamflow increase, thus increasing the dryness on the streamflow regime. Notice, however, that $\langle Q \rangle$ remain constant, i.e. $\langle Q \rangle$ does not depend on σ^2, i.e. $\langle Q \rangle = \langle \xi_t(\lambda, \gamma_P) \rangle$. The higher probabilities for low Q are indeed balanced by increased probabilities of high discharge, owing to their algebraic decay (see equation(2.13)). In other terms, the presence of the noise does not modify the mean of the streamflow pdf, but it significantly increases its variance (see Figure 2.7a).

The noise induces similar effects on the wet streamflow distribution as shown in Figure 2.6 b. As the noise strengthens, an increased probability of high discharge is observed and it is balanced by a shift of the mode of $p(Q)$ towards zero. Interestingly, above a certain

(a) Variance in the streamflow discharge process.

(b) Noise induced phenomena in discharge distribution regime.
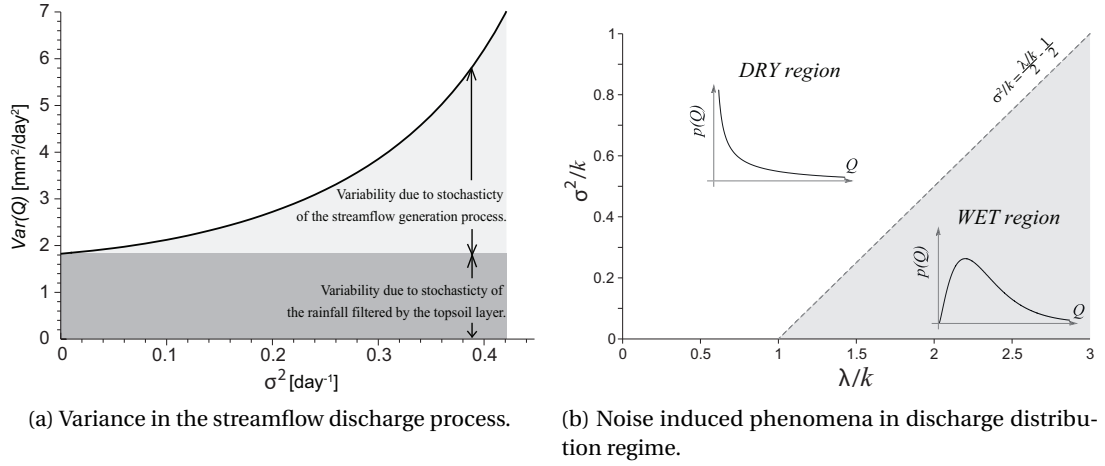
Figure 2.7: a) Numerical calculation of the variance $Var(Q) = \langle Q^2 \rangle - \langle Q \rangle^2$ of the streamflow discharge pdf (for parameters $\lambda = 0.875\,d^{-1}$ and $k = 0.5\,d^{-1}$), as a function of the intensity of the white noise σ^2. For $\sigma^2 = 0$, the variance of the Gamma distribution in equation (2.7) (with the same parameters λ and k that refers to deterministic linear reservoir) is exactly obtained. Different colors evidence the different contributions to $Var(Q)$.
b) Representation of the streamflow discharge distribution regime as function of the dimensionless parameter λ/k and σ^2/k. Notice that the noise induces a shift from wet to dry regime.

threshold of the noise, a shift from wet to dry regime occurs. According to equation (2.12) the shift between the two different regimes is controlled by the exponent of the power in equation (2.12). The function $f(\lambda/k, \sigma^2/k) = 3\sigma^2 + k - \sqrt{\sigma^4 + 2k\sigma^2 + 4\lambda\sigma^2 + k^2}/2\sigma^2$ of the two dimensionless parameters ratio λ/k and σ^2/k defines such shift. When $f(\lambda/k, \sigma^2/k) < 0$ the pdf of the runoff is hump-shaped with $p(Q = 0) = 0$ ('wet regime'), while for $f(\lambda/k, \sigma^2/k) > 0$ ('dry regime') $p(Q)$ goes to infinity for $Q \to 0$, and it monotonically decreases for $Q > 0$. The threshold between the two regimes is derived by imposing $f(\lambda/k, \sigma^2/k) = 0$, from which one finds (see Figure 2.7b):

$$\frac{\sigma^2}{k} = \frac{\lambda}{2k} - \frac{1}{2}. \tag{2.14}$$

Therefore taking a given value λ/k that would correspond to a wet regime when $\sigma^2 = 0$, may now correspond to dry regime if σ^2 is above $\lambda/2 - k/2$. This noise-induced shift, which is similar to those noted in other contexts [e.g. Horsthemke and Lefever, 2006; Ridolfi et al., 2011], occurs also for colored noise, as is has been verified numerically.

Notice that in the framework described in Section 2.4, for each realization of $Q(t)$ it is possible to obtain the corresponding time series of S by integration of equation (2.4). It is worth to underline the fact that, for a given time window, $\Delta Q / \Delta S$ is still free to fluctuate around its

corresponding deterministic value k.

Underestimating hydrologic variability may have ecological consequences. Noise induced wet−dry transitions (thought of as reproducing regime shifts from perennial to ephemeral streamflows) are important because low−stage discharges are crucial to establish natural flow conditions for the preservation of fish and riparian habitats. Note, in particular, that minimum flows are the most used targets in water resources management to sustain aquatic ecosystems.

The following conclusions are thus worth emphasizing:

- Numerical simulation has shown that a physically meaningful colored noise in the storage-discharge relation influences appreciably the discharge distribution. In particular, a shift of the mode of the pdf towards zero is observed. Noise in the storage-discharge relation is assumed to possibly be surrogated by considering stochasticity directly in the streamflow generation processes. In such a case, the analytical derivation of the probability distribution of streamflows is obtained, and its properties reflect the results obtained through numerical simulation by employing colored noise. The theoretical framework couples a stochastic description of soil moisture dynamics with a transport model that embeds the variability of the streamflow generation process through a multiplicative Gaussian noise;

- The effect of the multiplicative Gaussian noise on the streamflow distribution is significant. In particular, the tail of the streamflow distribution $p(Q)$ changes from exponential to a power-law type with heavy tail. As the noise increases, the probability of observing low values of streamflow also increases and the mode of $p(Q)$ shifts towards zero. Higher probabilities for low Q are balanced by an increased probability of high discharge, i.e. the power-law tail becomes fatter. Therefore the presence of the noise does not change the mean of Q, but significantly increases its variance. Thus neglecting these additional environmental fluctuations may produce underestimations of the variability of streamflows, with relevant ecohydrological consequences;

- Above a threshold in the noise magnitude, a shift from wet to dry regime occurs, implying a major ecological impact owing to a change from perennial to ephemeral streamflows. The transition between the two different regimes is controlled by parameters of clear physical meaning, and an analytical expression for the related threshold is proposed.

Chapter 3

Stochastic Modeling of Soil Salinity

The universe cannot be read until we have learnt the language and become familiar with the characters in which it is written.

G. Galilei

3.1 Introduction

Large areas of cultivated land worldwide are affected by soil salinity. Szabolcs [1989] estimates that 10% of arable land in over 100 countries, especially in arid and semi-arid regions [Tanji, 1989]. Salinity refers to large concentrations of easily soluble salts present in water and soil on a unit volume or weight basis (typically expressed as electrical conductivity (EC) of the soil moisture in dS/m, i.e. deciSiemens per meter at 25° C; for $NaCl$ 1 mg/l $\sim 15 \cdot 10^{-4}$ dS/m). High salinity causes both ion specific and osmotic stress effects, with important consequences for plant production and quality. Normally, yields of most crops are not significantly affected if EC ranges from 0 to 2 dS/m, while above levels of 8 dS/m most crops show severe yield reductions [Ayars et al., 1993; Hillel, 2000]. Prevention or remediation of soil salinity is usually done by leaching salts, and has resulted in the concept of leaching requirement [Richards, 1954; Hillel, 1998; Schleiff, 2008]. Alternative amelioration strategies by harvesting salt-accumulating plants appear to be less effective [Qadir et al., 2000].

Vertically-averaged soil moisture and salt balance equations have also been used [Allison et al., 1994; Hillel, 2000]. Despite their simplicity, these models have the advantage of parsimony, thus allowing a direct analysis of the interplay of the main processes, and provide an ideal starting point to include external, random hydroclimatic fluctuations in the analysis of long-

term salinization trends. The goal of this chapter is to offer a first step in this direction. With this purpose, here a minimalist model of soil primary salinization, describing analytically the long-term dynamics of salt in soils caused by wet (rain) and dry (aerosol) deposition is presented. The aim of this work is to quantify the salt mass and concentration probability density functions (pdfs) in the root zone, and the probability of crossing the crops salt tolerance threshold as a function of the main hydro-climatic parameters. The model framework is potentially extendible to systems including salt input from groundwater and irrigation.

3.2 Methods

The starting modelling scheme is a spatially lumped model [Bras and Seo, 1987] for the vertically averaged dynamics of soil moisture and salt in the root zone. As a first step input of salt due by irrigation or groundwater upflow will not be considered. As in the previous chapter, rainfall (R_t) is modeled as a marked point process with frequency λ_p and with daily rainfall depths exponentially distributed with mean $1/\gamma_p$. Also here, the averaged soil moisture dynamics are modeled assuming constant (spatially and temporally averaged) soil and ecohydrological parameters, i.e., root depth, Z_r, porosity, n, and maximum evapotranspiration rate, ET_{max}. Assuming a rain salt concentration C_R and a constant input $\mathcal{M}_d$ of salt mass per unit ground area and per unit time by dry deposition, the root-zone mass balance for soil moisture and salt mass m is given by:

$$nZ_r \frac{ds}{dt} = -ET(s) - L(s) + R_t, \tag{3.1}$$

$$\frac{dm}{dt} = C_R R_t + \mathcal{M}_d - CL(s), \tag{3.2}$$

where C is the salt concentration in the root zone; $L(s)$ represents deep percolation, while $ET(s)$ represents the losses resulting from plant transpiration and soil evaporation. All the rainfall input that cannot be accommodated is assumed to be lost as $L(s)$ at s_1 (an effective field-capacity threshold).

The next section will investigate how reduction in evapotranspiration affects salinization patterns (reduced evapotranspiration in turn increases the available soil moisture and thus reduces the concentration of salt in the soil and increases leaching frequencies).

3.2.1 Feedback Effect of Evapotranspiration on Soil Salinity

The evapotranspiration ET represents the sum of the losses resulting from the plant transpiration and evaporation from the soil. The effect of salinity on relative transpiration is called

osmotic adjustment. This means that salts increase the energy to which water is held in the soil. This can be taken in account by adding in the soil-water retention curve, given by the matric potential $\psi(s) = \psi(1)(\frac{s}{s_{sat}})^{-\tau}$ [Brooks and Corey, 1966], the effect of the osmotic potential $\pi(C)$. Assuming that the osmotic potential follows the Van't Hoff's law $\pi(C) = \kappa C$, then [Bras and Seo, 1987]:

$$ET(s,C) = \begin{cases} (s_{sat}/s^*)\psi^m(1)\{\psi(1)(s/s_{sat})^{-1/m} + \kappa C\}^{-m}ET_{max}, & s \leq s^*; \\ ET_{max}, & s > s^*, \end{cases} \quad (3.3)$$

where the soil moisture threshold s^* indicates the stomata's plants closure point (to prevent internal water losses), s_{sat} is the soil moisture saturation point and τ is a parameter related to the soil connectivity and tortuosity.

The effect of evapotranspiration on salinity is twofold. From one side, as evapotranspiration increases, the average soil moisture content decreases and therefore the salt concentration in the root zone increases. From the other side, there is a feedback effect of evapotranspiration on salinity. In fact, because of the the osmotic adjustment, the increased salt concentration of salt raises the actual content of water in the soil by decreasing the effective evapotranspiration. In this way, the probability of having leaching events is in turn increased generating a feedback effect on soil salinity that tends to diminish through percolation to deeper layer (groundwater). Simulations show that this feedback effect has an important impact on soil salinization (Figures 3.1a and 3.1b).

(a) Feedback effect of evapotranspiration on soil salinity.

(b) Effect of osmotic adjustment on salt concentration *pdf*s.
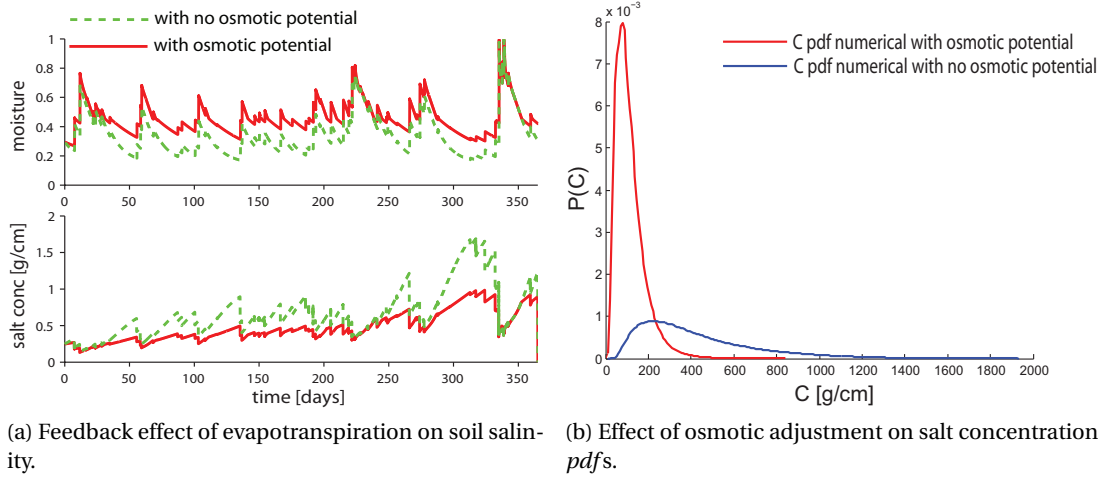
Figure 3.1: a) Numerical simulation of soil moisture and salt concentration in the root zone for given climate and soil parameters ($\lambda_P = 0.1\,d^{-1}, \gamma_P = 1\,cm^{-1}, n = 0.45, Z_r = 30\,cm$). The negative feedback of evapotranspiration drives down the salt concentration in the root zone. b)Numerical salt concentration *pdf*s in the root zone for a given climate and soil parameters ($\lambda_P = 0.1\,d^{-1}, \gamma_P = 1\,cm^{-1}, n = 0.45, Z_r = 30\,cm$). The negative feedback included by taking in account the osmotic adjustment shifts the mode of $p(C)$ towards zero.

In non salty soil, the minimalist model for soil moisture dynamics assumes evapotranspiration to be linear in the range of soil moisture comprised between the wilting point, s_w, and s_1, at which ET occurs at the maximum rate ET_{max} [Porporato et al., 2004]:

$$ET(s) = \begin{cases} ET_{max}\frac{s}{s_1}, & s_w \leq s \leq s_1; \\ ET_{max}, & s > s_1, \end{cases} \tag{3.4}$$

In spite of this strong linear simplification, the effect of the osmotic adjustment on salinization through an average reduction of ET_{max} can be taken in account. This is simply done here by keeping the same ET_{max} in Eqs. (3.3) and (3.4) (previous studies [Viola et al., 2008] have shown that, in the absence of osmotic effects, the minimalist model should have artificially higher ET_{max} to account for percolation losses below s_1).

A complete numerical model, in which the impact of osmotic stress in reducing ET is explicitly included through Eq. (3.3), has been also studied. Moreover, in the detailed model runoff takes place at saturation ($s_{sat} = 1$), while percolation occurs for $s > s_{fc}$ (the soil moisture field capacity), and it is proportional to the soil hydraulic conductivity $K_{sat}s^c$, where c is a soil-pore connectivity index and K_{sat} is the saturated hydraulic conductivity [Rodriguez-Iturbe I. and Porporato, 2004]. A comparison between the results of the two soil moisture models, presented in Figure 3.2a, suggests the viability of the simplified model. Simulations for wetter climates confirm this result.

3.2.2 Analytical Stochastic Model of Soil Salinity

The system (3.1) and (3.2) can be further simplified if one considers that the typical timescales for salt mass dynamics in the root zone are orders of magnitude larger than the ones characterizing rainfall (and thus wet deposition). Moreover, soil moisture typically reaches steady-state conditions within a growing season (e.g., $< 5 - 7$ months), while the salt mass balance only does so on much longer times scales (e.g., $>$ decades). Accordingly, at those long timescales, say T, the salt mass input flux can be assumed to take place at a constant rate, Υ, that is $\int_t^{t+T}(\mathcal{M}_d + C_R R_t)dt' \sim \mathcal{M}_d T + T C_R \lambda_P/\gamma_P = \Upsilon T$, and be interrupted by instantaneous and unfrequent leaching events induced by percolation. As a result, (3.2) can be rewritten as

$$\frac{dm}{dt} = \Upsilon - \frac{m}{n Z_r s} L(s). \tag{3.5}$$

The previous chapter has already shown how leakage may be modeled as a marked point process, with percolation depths exponentially distributed with parameter γ_P [Botter et al., 2007c]. For reasons of analytical tractability, the percolation events are assumed to occur according to a Poisson process with frequency λ given by the frequency of soil moisture crossing the threshold $s = s_1$. This can be expressed in terms of the soil moisture pdf as $\lambda = \rho(s_1)p(s_1)$, where the term $\rho(s) = (ET(s) + L(s))/nZ_r$ represents the normalized catchment-
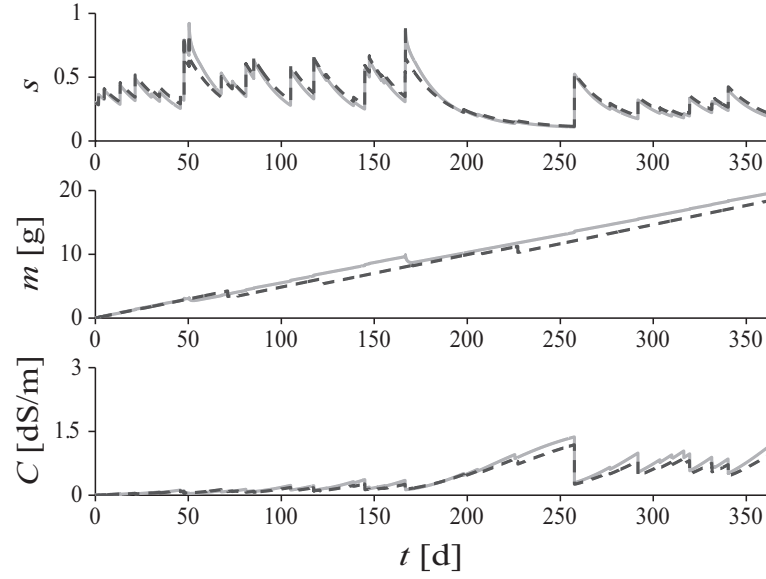
Figure 3.2: Comparison of soil moisture and salinity models: (a) Temporal evolution (equation (3.1)) of $s(t)$, forced by intermittent rainfall ($\lambda_P = 0.1$ d^{-1} and $1/\gamma_P = 1.79$ cm). The blue dashed line refers to the minimalist model, while the continuous red line is the complete numerical model (see text for details). (b) Temporal evolution of root-zone salt mass for the complete numerical model (red line) and the minimalist model (blue dashed line). (c) Temporal evolution of the corresponding specific salt concentration $C(t) = m(t)/nZ_r s(t)$ in the root zone for the same two cases of 3.2b. The unit of measure of C are transformed from mg/(cm m^2) to dS/m, by using mg/(cm m^2) = 10^{-1} mg/l. The soil and vegetation parameters employed for the simulation of the complete model are those typical for a sandy-loam soil, while the free parameters of the minimalist model are $s_1 = 0.8$, $b = 0.6$. In particular for both models $n = 0.45$, $Z_r = 30$ cm, $s_w = 0.1$, $ET_{max} = 0.35$ cm/d, $C_R \approx 3$ mg l^{-1} and $\mathcal{M}_d = 54$ mg d^{-1} m^{-2} (coastal area) have been used.

scale loss function (i.e. the total losses from the system due to evapotranspiration and leakage as a function of the soil moisture) [Rodriguez-Iturbe I. and Porporato, 2004]. Adopting the soil moisture minimalist model, for which the pdf is a truncated gamma distribution [e.g. Porporato et al., 2004], the leaching frequency is $\lambda = \eta \exp(-\gamma)\gamma^{\lambda_P/\eta}/\Gamma(\lambda_P/\eta, \gamma)$ [Botter et al., 2007c], where $\Gamma(x, y)$ is the lower incomplete gamma function, $\eta = ET_{max}/(nZ_r(s_1 - s_w))$ and $\gamma = \gamma_P nZ_r(s_1 - s_w)$. A leaching-efficiency parameter b is used to account for incomplete salt dissolution, further assuming that the typical value of soil moisture during leaching events can be approximated by the value s_1. With the above assumptions, the dynamics of the salt mass in the root zone can be described by a single equation

$$\frac{dm}{dt} = \Upsilon - mL'_t, \tag{3.6}$$

31

where L'_t is a marked Poisson noise [Van Den Broeck, 1983] with frequency λ, and (dimensionless) exponential marks with mean

$$\mu = \frac{b}{n\,Z_r\,s_1\,\gamma_P}. \tag{3.7}$$

Figures 3.2b and 3.2c compare the results of both salinity models. The free parameters s_1 and b are fitted with respect to the complete model of salt mass and concentration, respectively.

From a mathematical viewpoint, equation (3.6) is a stochastic differential equation with multiplicative white (jump) noise. In this case, since the soil solution can be considered in equilibrium during leaching events, one has to interpret (3.6) in the Stratonovich sense (see Van Den Broeck [1983] and Appendix A.3 for all the mathematical details). Accordingly, the normal rules of calculus are preserved, and equation (3.6) can be transformed into

$$\frac{dy}{dt} = \Upsilon e^{-y} - L'_t, \tag{3.8}$$

where $y(t) = \ln[m(t)]$.

3.3   Results and Discussion

The stationary solution of (3.8) can be obtain as in Rodriguez-Iturbe et al. [1999b]. Then using the derived distribution for m, i.e., $p(m) = p(y)dy/dm$, the probability distribution for the salt mass in the root zone is obtained

$$p(m) = \mathcal{N}\,\exp(-\frac{m\lambda}{\Upsilon})\,m^{1/\mu}, \tag{3.9}$$

where $\mathcal{N} = (\lambda/\Upsilon)^{\frac{1+\mu}{\mu}}/\Gamma(\frac{1+\mu}{\mu})$ and $\Gamma(x)$ is the Gamma function. Equation (3.9) summarizes the soil salinity statistics as a function of climate, soil and vegetation parameters.

Figure 3.3 is a graphical representation of the dependence of the mean salt concentration $\langle C \rangle = \langle m \rangle / nZ_r \langle s \rangle$ on the yearly rainfall and λ_P. The contour-lines connect equal values of the mean salt concentration in the soil, for a given input of salt Υ. The latter has been calculated for two different geographic regions. Typical salt inputs in coastal areas are $100-200$ kg/(ha yr) of salt, while values drop of an order of magnitude in continental regions [Hillel, 2000].

Between the black region and the light gray ones in Figure 3.3a, the behavior of $\langle C \rangle$ changes substantially. Above a certain total rainfall per year, the input of salt related to rainfall frequency becomes immaterial as leaching effectively washes out the salt mass from the root zone. For lower total rainfall values, however, the salt in the soil increases with increasing λ_P. For a

given annual precipitation depth, with low rainfall frequencies, rainfall events carry enough water to trigger leaching. Conversely, if λ_P is high, evapotranspiration dominates, leaching is largely reduced, thereby causing salt accumulation in the root zone. Therefore, $\langle m \rangle$ strongly increases with λ_P. Relatively small reductions of rainfall at the transition between these two regimes may entail a dramatic increase in long-term soil salinization. Figure 3.3 also shows the threshold of soil salinity below which vegetation is practically unaffected (e.g., $\langle C \rangle < 2$ dS/m) and the thresholds above which regular (e.g. non-halophytic) vegetation is damaged (e.g., $\langle C \rangle > 2$ dS/m). For coastal areas soil salinization may occur even in relatively more humid regions, especially when rainfall events are not very intense. On the contrary, in continental regions only arid climates may begin to develop soil salinization (in the absence of irrigation and groundwater input). Indeed, through the proposed model one can evaluate the risk of soil salinization in rain-fed agriculture just by estimating the typical salt inputs, total rainfall per year and the rainfall frequency. For example, a rain-fed crop in a semi-arid climate (e.g., rainfall depth of 70 cm/yr) in a continental region risks salinization only when rainfall events are not very intense (e.g., $\gamma_P^{-1} \leq 0.4$ cm or $\lambda_P \geq 0.48$ d^{-1}). If the same crop is located in a coastal area, salinization occurs for a wider range of rainfall parameters (e.g., $\gamma_P^{-1} \leq 1$ cm or $\lambda_P \geq 0.18$ d^{-1}).

The solution (3.9) may be used in conjunction with soil moisture statistics to obtain a full characterization of the salt concentration in the root zone. Because one may safely assume that equations (3.1) and (3.5) are decoupled over short time scales, the soil moisture $s(t)$ and the salt mass $m(t)$ may be treated as statistically independent random variables. By observing that the salt concentration in the root zone is equal to $C(t) = m(t)/nZ_r s(t)$ and assuming $s_w \sim 0$, the stationary probability distribution of the salt concentration $p(C)$ as the quotient distribution of two independent random variables [Curtiss, 1941] is achieved:

$$p(C) = \frac{\lambda(\frac{\Upsilon\gamma_P}{C\lambda}+1)^{-1/\mu}(\frac{\Upsilon\gamma_P}{C\lambda+\Upsilon\gamma_P})^{\frac{\lambda_P}{\eta}}(\Gamma(\frac{\lambda_P}{\eta}+\frac{1}{\mu}+1)-\Gamma(\frac{\lambda_P}{\eta}+\frac{1}{\mu}+1,nZ_r s_1(\frac{C\lambda}{\Upsilon}+\gamma_P)))}{\Gamma(1+\frac{1}{\mu})(C\lambda+\gamma_P\Upsilon)(\Gamma(\frac{\lambda_P}{\eta})-\Gamma(\frac{\lambda_P}{\eta},nZ_r s_1\gamma_P))}. \quad (3.10)$$

The comparison between analytical solutions and numerical simulations (Figure 3.4) shows that the analytical solution reproduces reasonably well the pdf of the complete model. By integrating equation (3.10) from a given concentration value C^* to infinity, one obtains the cumulative pdf of C, $P(C^*)$, which is the probability of having a salt concentration greater than a certain critical concentration value, C^*, as a function of the soil-plant-atmosphere parameters. The inset of Figure 3.4 confirms the impact that climate change may have on soil salinity. Note, in particular, that such an impact is marked only for semi-arid or drier climates (see Figure 3.3). For example with a reduction from $\lambda_P = 0.2$ to $\lambda_P = 0.15$ d^{-1}, the probability of crossing $C^* = 6$ dS/m is more than tripled. When coupled to a crossing analysis of concentration levels, the previous results may be used to evaluate the risk of plant salt stress. The analytical form of the results makes it suitable for computations of salinity risk at the
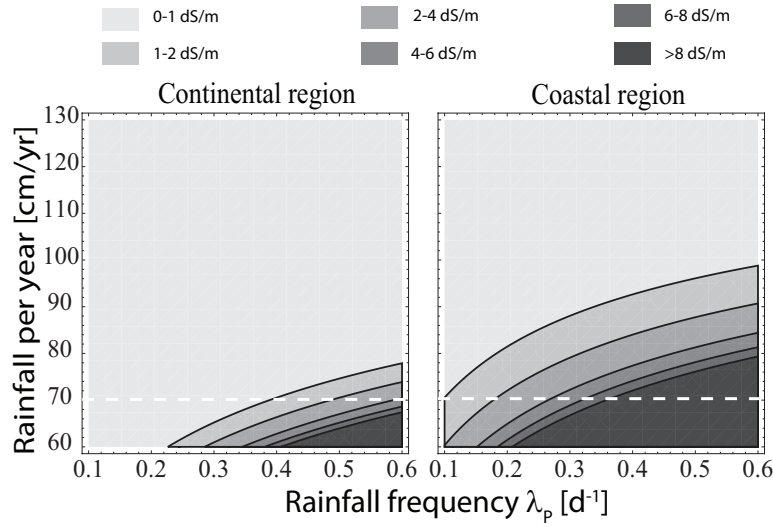
Figure 3.3: Contour plot of the asymptotic mean concentration of salt $\langle C \rangle$ from the exact solution of $\langle m \rangle$ as a function of yearly rainfall depth and frequency. The values reported in the legend refer to the corresponding salt concentration values with respect to the average soil moisture $\langle s \rangle$ (for its analytical expression see Porporato et al. [2004]). The contour lines represent significant soil salinity values (1,2,...,8 dS/m). The parameter μ has been calculated through equation (3.7); the others are as in figure 3.2 for the coastal region, while for continental areas $\Upsilon \approx 6 \, \text{mg} \, \text{d}^{-1} \, \text{m}^{-2}$.
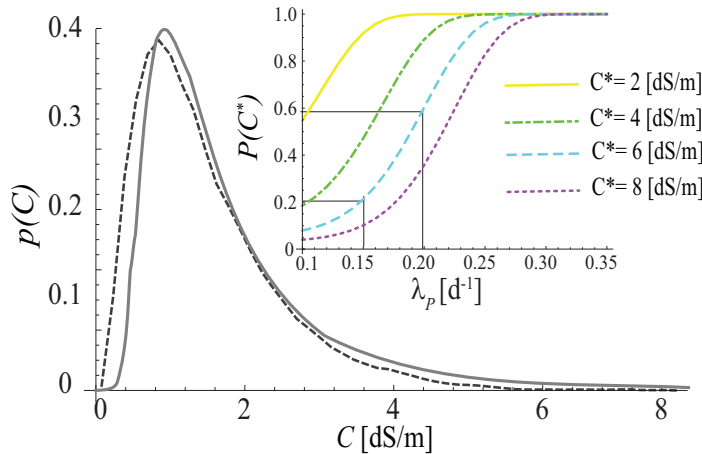
Figure 3.4: Comparison between the analytical form of $p(C)$ for the minimalist model obtained via equation (3.10) in which the free parameters (b and s_1) have been fitted (solid line) and the numerical simulations of the corresponding complete model (dashed line). The soil and hydro-climatic parameters are as in Figure 3.2. In the inset: probability of exceeding a soil salinization critical threshold C^* ($\int_{C^*}^{+\infty} p(C)dC$) as a function of the rainfall frequency λ_P.

global scale as a function of few measurable parameters, and facilitates their coupling with other models of long-term soil-plant biogeochemistry.

In future works, this modeling framework will be extended to investigate additional salt inputs from irrigation and groundwater by modifying accordingly the average salt input parameter Υ and calculating the corresponding soil moisture pdfs (e.g. see Vervoort and Van Der Zee [2008] for groundwater inputs and Vico and Porporato [2010] for irrigation).

Effects of the spatial structure of the environmental matrix on ecosystems dynamics

Overview

In this second part, it will be shown how hydrologic controls on the ecosystem dynamics are extended also through the connectivity structure of the space where the ecosystem lives.

The commonalities existing among all types of river basins and their drainage networks suggest the possible existence of general rules commanding the role that their structure plays in ecological patterns. Moreover, the ecological corridors provided by river networks induce anisotropy in the spreading of species, pathogenes or other agents along the waterways. Natural ecosystems are characterized by striking diversity of form and functions and yet exhibit deep symmetries emerging across scales of space, time and organizational complexity. Species-area relationships [Brown, 1995] and species-abundance distributions [Preston, 1948] are examples of emerging patterns irrespective of the details of the underlying ecosystem functions.

In this context, the so¯called neutral ecological theory [Hubbell, 2001; Bell, 2001; Vallade and Houchmandzadeh, 2003; Volkov et al., 2003], assuming that all species are competitively equivalent at a per capita level, offer a benchmark dynamics suggesting that many aspect of real biotic system may not require a more complicated model [Stanley, 2000]. In fact a general theory emerges on the effects of dendritic geometries on the ecological processes and dynamics operating on river basins and it explain the spatial pattern of species dispersal in riparian corridors [Muneepeerakul et al., 2008b; Rodriguez-Iturbe et al., 2009]. However, emerging patterns consistent with the neutrality assumptions do not necessarily imply that the underlying ecological processes are neutral [Harte, 2003].

The ability of the neutral theory to make testable predictions about a wide variety of ecosystem properties is fully explored in Chapter 4. The reader will be briefly introduced to neutral theory and a related lattice model that allow to calculate the local species persistence time (SPT) distributions. SPT is defined as the timespan between local colonizations and extinctions in a given geographic region and its distribution is proposed as a new macroecological pattern emerging from ecosystem dynamics. In this chapter it is also theoretically investigated how the scaling features of the SPT distribution depend on the structure of the spatial interaction network. In Chapter 5 empirical SPT distributions pertaining to four different ecosystems will be presented: North American breeding birds, Kansas grasslands, a forest in New Jersey and marine fishes of Bristol channel. Data will be analyzed in a new theoretical framework that accounts for length−biased sampling effects on the measurements. In spite of the differences between taxa and spatial scales of analysis, the SPT distributions all exhibit a power-law behavior limited by a cut-off determined by the rate of emergence of new species in that particular ecosystems. Empirical scaling exponents of SPT distributions are reproduced by the spatial neutral model. The framework developed in these chapters also allows to link the cut-off timescale with the spatial scale of analysis, and the persistence-time distribution to the species-area relationship.

Part II

Neutral Theory and Species Persistence Time Distributions

The enormous usefulness of mathematics in the natural sciences is something bordering on the mysterious and that there is no rational explanation for it

E. Wigner

4.1   Stochastic modelling of ecosystems dynamics

Neutral models are based on the assumption that, within a single trophic level, individual birth and death rates are species-independent. Neutral theory has been proposed as an unified context for understanding ecological patterns [Hubbell, 2001]. Since then, several studies have focused their attention on it [Houchmandzadeh and Vallade, 2003; Chave, 2004; Alonso and McKane, 2004; Azaele et al., 2006; Volkov et al., 2007; Chisholm and Lichstein, 2009]. Although the neutral hypothesis is easily falsifiable for specific cases, neutral models prove able to explain several universal macroecological patterns. This is the case, for instance, of the relative species abundance (RSA) distributions [McGill, 2003; Volkov et al., 2005], species-independent beta diversity patterns [Condit et al., 2002; Zillio et al., 2005] and species-area relationship [Brown, 1995; Zillio et al., 2008; O'Dwyer and Green, 2010]. Remarkably, this happens by invoking only basic ecological processes such as birth, death, migration and dispersal limitation. Although certain emergent ecological patterns are independent of the fine ecological details and often well predicted by neutral theory [Bell, 2001], it has been observed that this does not imply that the underling ecological processes are neutral [Harte, 2003; Purves and Pacala, 2005]. For these reasons, although the persistence biogeography

framework does not require neutral process as well, it will be kept as a reference frame for the theoretical speculations presented in this chapter.

According to the assumption of neutrality [Hubbell, 2001], the dynamics of a species in the ecosystem is fully specified by its effective birth and death rates $b(n)$ and $d(n)$ that depend exclusively on the population size n. Species abundance dynamics is described by the so called birth-death master equation (ME) [Volkov et al., 2003; Houchmandzadeh and Vallade, 2003]:

$$\frac{dP(n,t)}{dt} = b(n-1)P(n-1,t) + d(n+1)P(n+1,t) - (b(n)+d(n))P(n,t), \qquad (4.1)$$

where $P(n,t)$ is the probability, for a given species, of having a population of n individuals at time t. The birth and death rate, $b(n)$ and $d(n)$ respectively, take into account several ecological processes that may increase or decrease the number of individuals in a species over time as, for instance, immigration or emigration.

The continuum approximation of the dynamics described by the ME (4.1) exist. If the population of a species at time t is treated as a continuous variable, $x(t)$ (assumption which is valid when the population varies smoothly with time and is not too small), then it can be proved that the species abundance dynamics is described by the generalized Langevin equation [Pigolotti et al., 2004; Azaele et al., 2006]:

$$\frac{dx}{dt} = 2b(0) - x(t)(d(1)-b(1)) + \sqrt{(d(1)+b(1))x(t)/2}\,\xi(t), \qquad (4.2)$$

where $\xi(t)$ is a Gaussian white noise with zero mean value. From Eq. (4.2) one can derive the corresponding Fokker-Planck equation for the probability distribution function $p(x,t)$ of finding x individuals at time t in the community.

Although the dynamics generated by Eqs. (4.1) and (4.2) are equivalent (if the continuum assumption is valid), it is preferred to proceed using the discrete neutral model given by the ME (4.1).

4.2 Theoretical Persistence-Time Distributions

In this context, study of species persistence times (SPT) of trophically equivalent co-occurring species in relation to the spatial scale of observation is here adressed. The persistence time τ of a species within a geographic region is defined as the time incurred between its emergence and its local extinction (see Keitt and Stanley [1998]; Pigolotti et al. [2005] and Figure 4.1). At a local scale, persistence times are largely controlled by ecological processes operating at short timescales (e.g. population dynamics, dispersal, immigration, contraction/expansion of species geographic ranges) as local extinctions are dynamically balanced by colonizations [MacArthur and Wilson, 1967; Ricklefs, 1987]. At a global scale, originations and extinctions are controlled by mechanisms acting on macroevolutionary timescales.
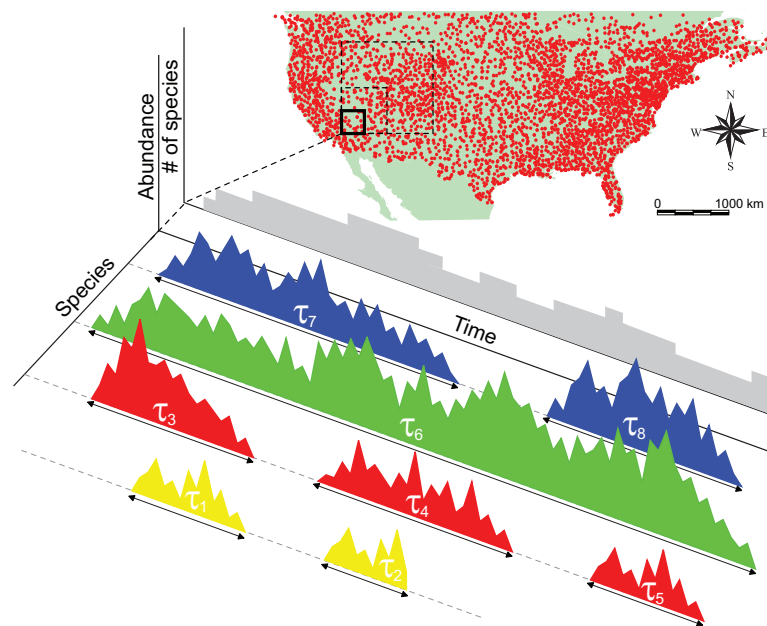
Figure 4.1: SPT τ within a geographic region is defined as the time incurred between a species' emergence and its local extinction. Recurrent colonizations of a species define different persistence times. The number of species in the ecosystem as a function of time (gray shaded area) crucially depends on species emergences and persistence times. The inset shows the observational routes of the Breeding Bird Survey, a long-term dataset that it will be analyzed in the next chapter. Aggregating local information comprised in a given geographic area, species presence-absence time-series that allow the estimation of persistence-time distributions is reconstructed.

From a theoretical viewpoint, the simplest baseline model for population dynamics is a random walk without drift, according to which the abundance of a species in a geographic region has the same probability of increasing or decreasing by one individual at every time step. According to this scheme, local extinction is equivalent to a random walker's first passage to zero, and thus the resulting persistence-time distribution has a power-law decay with exponent 3/2 [Chandrasekhar, 1943].

A more realistic description can be achieved by accounting for basic ecological processes through the neutral mean field schemes, as follows. Consider a community of N individuals belonging to different species. At every time step a randomly selected individual dies and space or resources are freed up for colonization. With probability ν the site is taken by an individual of a species not currently present in the system; ν is equivalent to a per-birth diversification rate and it accounts for both speciation and immigration from surrounding communities. With the residual probability $1 - \nu$ the died individual is replaced by one offspring of an individual randomly sampled within the community [Durrett and Levin, 1996; Chave et al., 2002]. As such the probability of colonization by a species depends solely on its relative abundance in the community. This model is known as voter model [Holley and Liggett, 1975].

Considering $b(n) = n(1 - \nu)$ and $d(n) \propto n$, the abundance dynamics in the voter model (in the limit of infinite number of nodes) is exactly described by Eq. (4.1). Assuming absorbing boundary conditions in $n = 0$, the probability density distribution (pdf) of the theoretical persistence time p_τ can be expressed as [Pigolotti et al., 2005]:

$$p_\tau(t) = \frac{dP(0, t)}{dt}. \tag{4.3}$$

The asymptotic behavior of the resulting persistence-time distribution (i.e. $p_\tau(t)$) exhibits a power-law scaling limited by an exponential cut-off:

$$p_\tau(t) \propto t^{-\alpha} f(\nu t), \tag{4.4}$$

with $f(z) = (z/(1 - e^{-z}))^2 e^{-z}$ and $\alpha = 2$ [Pigolotti et al., 2005]. In Eq. (4.4), time is expressed in generation time units [Hubbell, 2001], i.e. it has been rescaled such that the birth rate is equal to one. Notably, in the mean field scheme the probability distribution $p_\tau(t)$ depends solely on the diversification rate which accounts for speciation and migration processes and imposes a characteristic timescale $1/\nu$ for local extinctions. While per-birth speciation rates are not expected to vary with the spatial scale of analysis, per-birth immigration rates are argued to decrease as the spatial scale increases. In fact, the possible sources of migration (chiefly dependent on the geometrical properties of the boundary and the nature of dispersal processes) are argued to scale sub-linearly with the community size [Chisholm and Lichstein, 2009], which in turn is typically linearly proportional to geographic area [MacArthur and Wilson, 1967; Brown, 1995]. As continental scales are approached, migration processes (almost) vanish and the diversification rate ultimately reflects only the speciation rate.

Another issue of interest concerns the effects of the structure of the environmental matrix [Ricketts, 2001]. In particular, the role of the connectivity structure of environmental matrix on the SPT distribution is investigated.

4.3   Spatial Effects on Persistence-Time Distributions

In this section, a theoretical rationale for the universality of the scaling behavior of persistence-time distributions with respect to the topology of the interactions allowed by the environmental matrix is provided. In particular, evidences are presented on how non-trivial exponents ($\alpha \neq 1.5$ or $\alpha \neq 2$) can be reproduced by the neutral voter model once dispersal limitation and the actual network of spatial connections are taken into account.

The neutral game described above in regular one-, two- and three-dimensional lattices in which every site represents an individual [Durrett and Levin, 1996; Chave et al., 2002] have been implemented. Patterns emerging from the application of the model to dendritic structures mimicking riverine ecosystems where dispersal processes and ecological organization are constrained by the network structure have also been explored. Indeed, many features of riverine ecosystems have been shown to be affected by the connectivity of river networks

[Grant et al., 2007; Rodriguez-Iturbe et al., 2009]. In particular, river geometry has been studied in relation to extinction risk [Fagan, 2002], migration processes [Campos et al., 2006], persistence of amphibian populations [Grant et al., 2010], macroinvertebrate dispersal [Brown and Swan, 2010] and freshwater fish biodiversity [Muneepeerakul et al., 2008b; Bertuzzo et al., 2009]. For general calculations of the topological structure and metric properties relevant to dendritic ecological corridors, the features of Optimal Channel Networks (OCNs) [Rodriguez-Iturbe et al., 1992] are employed. They hold fractal characteristics known to closely conform to the scaling of real networks [Rinaldo et al., 1992]. Among the advantages of the use of OCNs, one recalls the possibility to fit one such construct into any assigned domain (e.g. a square, Figure 4.2), thus allowing exactly the same size and number of nodes of a two-dimensional lattice to be endowed with altered directionality of connections. To account for limited dispersal effects, only the offsprings of the nearest neighbors of the died individual to possibly colonize the empty site is allowed. In the networked landscape the neighborhood of a site is defined by the closest upstream and downstream sites. Limited dispersal promotes the clumping in space of species, which enhances their coexistence and survival probability [Chave et al., 2002; Kerr et al., 2002]. Indeed, it is found that in all the considered landscapes, persistence-time distributions is still compatible with a power-law now limited by an exponential cut-off

$$p_\tau(t) \propto t^{-\alpha} e^{-\nu t}, \tag{4.5}$$

with $\alpha = 1.92$ for the 3D, $\alpha = 1.82$ for the 2D, $\alpha = 1.62$ for the OCN, $\alpha = 1.50$ for the 1D landscape, Figure 4.2) limited by an exponential cut-off.

4.4 The Role of the Dispersal Range

The effects of dispersal ranges wider than nearest neighbors on persistence-time distributions are also investigated . To that end, the individual-based neutral [Hubbell, 2001; Chave et al., 2002] model with uniform-dispersal with varying radii r is implemented. The model in regular lattices in d spatial dimensions comprising N sites each or them occupied by one individual is thus simulated. To avoid edge effects, periodic boundary conditions have been prescribed. The model assumes neutral dynamics, therefore all the basic ecological processes reproduced in the model (birth, death, dispersal, colonization, and diversification) are equivalent for all the individuals of all species. The system is assumed to be always saturated, i.e., no available resources or sites are left unexploited. At each time step, a randomly selected individual dies and the resources are freed up and available for colonization by another individual. With probability ν, the diversification rate, the empty site is colonized by individual of a species not currently present in the system. The diversification is a rate per-birth that accounts for both speciation and immigration from surroundings communities. With the residual probability $1 - \nu$ the empty site is occupied by on offspring of an individual randomly sampled among all the sites that are less distant than r from the empty site, where r is measured in units of lattice spacing. For $r = 1$ the voter model reduces to the nearest-neighbor dispersal case
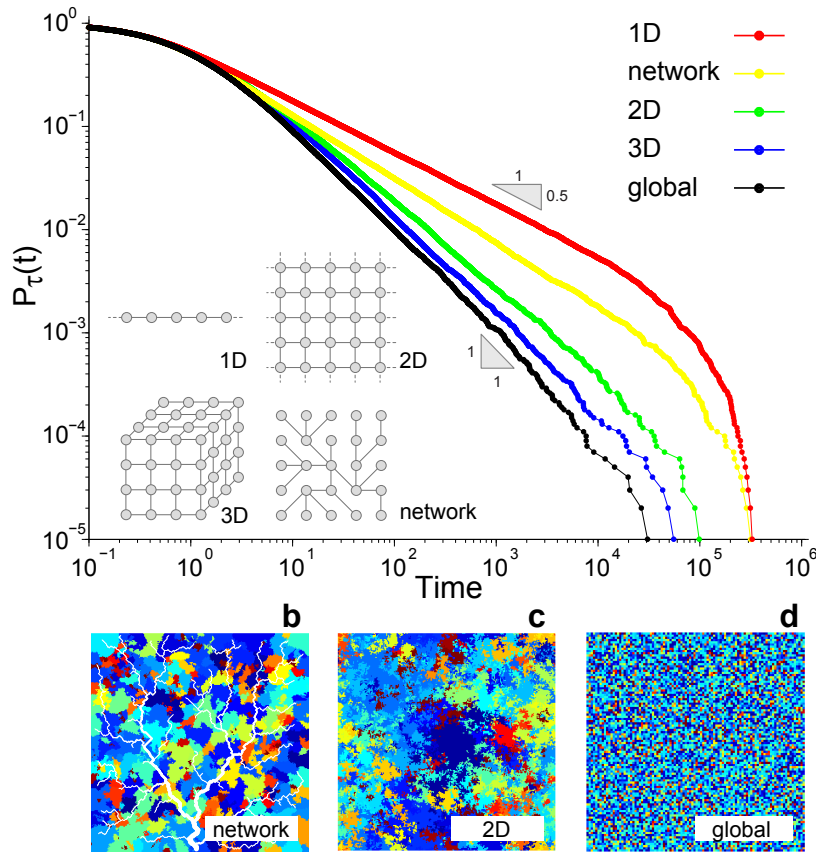
Figure 4.2: Persistence-time distributions are dependent on the structure of the spatial interaction networks. (a) Persistence-time exceedance probabilities $P_\tau(t)$ (probability that species' persistence times τ be $\geq t$) for the neutral individual-based model [Durrett and Levin, 1996; Chave et al., 2002] with nearest-neighbor dispersal implemented on the different topologies shown in the inset. Note that in the power-law regime if $p_\tau(t)$ scales as $t^{-\alpha}$, $P_\tau(t) \propto t^{-\alpha+1}$. The scaling exponent α is equal to 1.50 ± 0.01 for the one-dimensional lattice (red), $\alpha = 1.62 \pm 0.01$ for the networked landscape (yellow), 1.82 ± 0.01 and 1.92 ± 0.01 respectively for the 2D (green) and 3D (blue) lattices. Errors are estimated through the standard bootstrap method. The persistence-time distribution for the mean field model (global dispersal) reproduces the exact value $\alpha = 2$ (black curve). For all simulations $v = 10^{-5}$ and time is expressed in generation time units [Hubbell, 2001]. Bottom panels sketch the color-coded spatial arrangements of species in a networked landscape (b), in a two-dimensional lattice with nearest neighbor dispersal (c), and with global dispersal (d).

(von Neumann neighborhood). After the system has reached a statistically steady state, the persistence times of 10^5 species, are measured by tracking their abundances from emergence to extinction.

Figure 4.3 shows the persistence-time exceedance probability distribution for different dis-
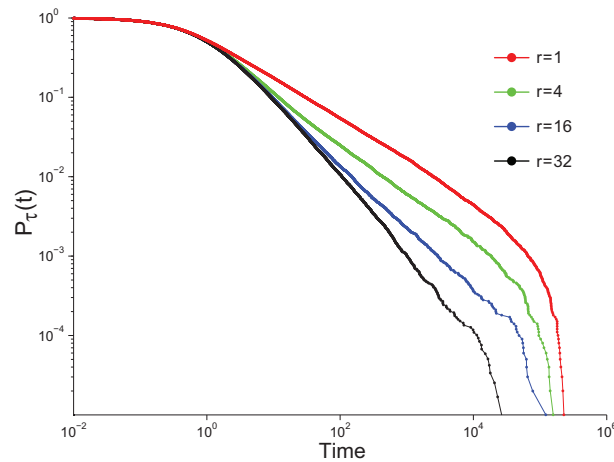
Figure 4.3: Persistence-time exceedance probability distributions $P_\tau(t)$ for the neutral model implemented in a one-dimensional lattice. Different colors refer to different uniform dispersal radii. For all simulations $\nu = 10^{-5}$ and time is expressed in generation time units [Hubbell, 2001].

persal radii. When the range of dispersal is small compared to the typical size of species' cluster, the persistence-time distributions tend to have, after a transient regime, the same scaling found in the nearest neighbor case. As expected, as the radius of dispersal increases, persistence-time distributions approach the scaling of the global dispersal.

4.5 Competition-Survival Trade-Off Model

The resulting persistence-time distributions once the neutral assumption is relaxed and differences among species are considered are finally analyzed . In this context, a competition-survival trade-off model [Tilman, 1994; Buttel et al., 2002; Chave, 2004], where species have different mortality rate μ, has been implemented on regular lattices. A trade-off between mortality and competitive advantage is assumed, so that species with higher mortality rates have an higher probability of outcompeting species with lower mortality rate in the engagement for the colonization of an empty site. Operationally, each species s is labeled by its mortality rate μ_s. At every time step an individual, randomly selected among all the individuals in the system, dies with a probability equal to the mortality rate of the species it belongs to. If a death occurs the empty site is colonized by one individual of a species not currently present in the system with probability ν, the diversification rate. The mortality rate of the new species is sampled from a uniform distribution between 0 and 1. With the remaining probability $1 - \nu$, the empty site is colonized by an offspring of an individual chosen among the nearest neighbors with probability proportional to their mortality rate.

Figure 4.4 shows the comparison between persistence-time distributions emerging from the neutral and the competition-survival trade-off model implemented in a two-dimensional

lattice. Also in the trade-off model the persistence-time distribution, after a transient regime, exhibits a power-law behavior with an exponent ($\alpha = 1.80 \pm 0.03$) close to that obtained under neutrality ($\alpha = 1.82 \pm 0.01$). Noticing that persistence time in the trade-off model has a larger mean value $\langle \tau \rangle$ and recalling that the mean number of species S in the system at a certain time is $S = \nu N \langle \tau \rangle$ [Rodriguez-Iturbe et al., 1987], the it can be concluded that trade-off mechanisms can facilitate the coexistence of species, a result already suggested in the literature [Chave et al., 2002; Tilman, 1994].

This confirms the expectation that a power-law distribution for SPT is the result of emergent behaviors independent of fine ecological details, thus supporting the neutral assumption that effective interaction strength among species is weak [Volkov et al., 2009] and does not significantly constrain the dynamics of ecosystems. It also worth to note that the above results are not seen as a test for the neutrality hypothesis for breeding birds or herbaceous plants dynamics, but rather as tools to reveal emerging universal and macroscopic patterns [Solé et al., 2002; Pueyo et al., 2007].
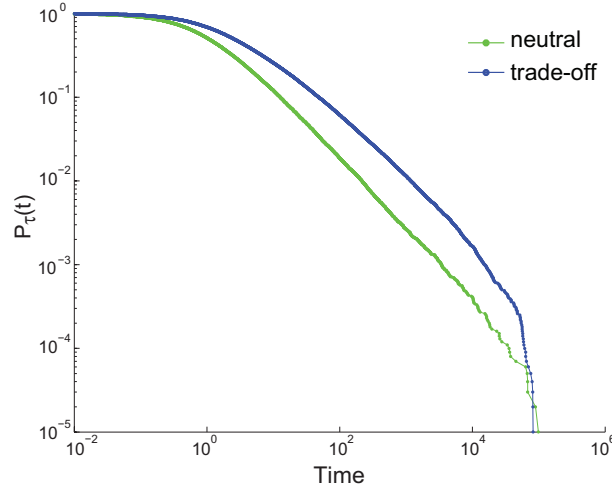
Figure 4.4: Comparison between the persistence-time exceedance probability distribution $P_\tau(t)$ for the neutral (green) and the competition-survival trade-off (blue) model implemented in a two-dimensional lattice. Note that in the power-law regime if $p_\tau(t)$ scales as $t^{-\alpha}$, $P_\tau(t) \propto t^{-\alpha+1}$. The scaling exponent α in the range $10^3 <$Time$< 10^4$ is equal to 1.82 ± 0.01 for the neutral model and 1.80 ± 0.03 for the trade-off model. Errors are estimated through bootstrap method (random sampling with replacement). For all simulations $\nu = 10^{-5}$ and time is expressed in generation time units [Hubbell, 2001].

4.6 Sampling Effects on SPT Distribution

Empirical real data on persistence times for different species are available, especially from succession studies and long term ecological databases [Pigolotti et al., 2005; Bertuzzo et al., 2011]. However, before testing the theory against the empirical data, length−bias sampling

effects on SPT are needed to be taken in account.

In fact, when dealing with observational data, the effect of the finiteness of the observed time window on the measured SPT must be properly taken into account. In this section, the analytical derivation of the distribution of two variables that can actually be measured from empirical data is presented: i) the persistence times τ' of species that emerge and go locally extinct within the observed time window ΔT_w; and ii) the variable τ'' that comprises the persistence times τ' and all the portions of SPT that are partially seen inside the observational time window but start or/and end outside (Figure 4.5). By matching analytical and observational distributions for $p_{\tau'}(t)$ and $p_{\tau''}(t)$ it is possible to infer the persistence-time distribution $p_\tau(t)$. In the theoretical framework described in the main text, the probability v of having a diversification event in a time step is constant, thus species emergence in the system due to migration or speciation is a uniform point Poisson process with rate $\lambda = v N$, where N is total number of individuals in the system and λ has the dimension of the inverse of a generation time. t_0 is the emergence time of a species in the system, and T_0 and $T_f = T_0 + \Delta T_w$ the beginning and the end of the observational time window, respectively. A species emerged at time t_0 will be continuously present in a geographic region for its persistence time τ until its local extinction at time $t_0 + \tau$. The distribution of τ'', the most complex case is first analyzed. The distribution of τ' will follow easily.
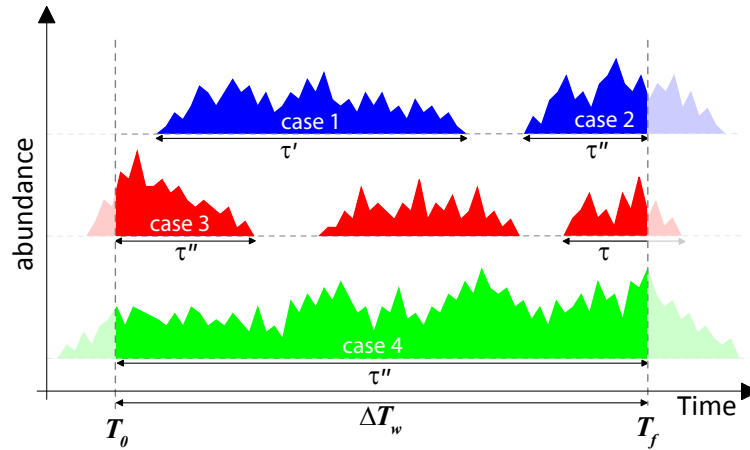
Figure 4.5: Schematic representation of the variables that can be measured from empirical data over a time window ΔT_w: τ' (persistence times that start and end inside the observational window) and τ'' (which comprises the persistence times τ' and all the portions of SPT that are partially seen inside the observational time window but start or/and end outside). Cases 1,2,3 and 4 refer to the description provided in the text.

The variable τ'' can be expressed as function of the random variables τ and t_0, which are probabilistically characterized. Four different cases (Figure 4.5) can be distinguished:

1. the species emerges and goes locally extinct within the time window;

2. the species emerges during the observations and it is still present at the end of the time window;

3. the species emerges before the beginning of the observations and goes locally extinct within the time window;

4. the species is always present for all the duration of the observations;

or, mathematically:

$$\tau'' = \begin{cases} \tau, & \text{if } T_0 \leq t_0 \leq T_f \text{ and } t_0 + \tau \leq T_f \\ T_f - t_0, & \text{if } T_0 \leq t_0 \leq T_f \text{ and } t_0 + \tau > T_f \\ t_0 + \tau - T_0, & \text{if } 0 < t_0 < T_0 \text{ and } T_0 \leq t_0 + \tau \leq T_f \\ T_f - T_0, & \text{if } 0 < t_0 < T_0 \text{ and } t_0 + \tau > T \end{cases} \tag{4.6}$$

The probability of observing τ'' conditional on a persistence time of duration τ is express as:

$$p_{\tau''}(t|\tau) =$$
$$\frac{1}{\mathcal{N}}\Big(\langle \delta(\tau - t)\Theta(t_0 - T_0)\Theta(T_f - (t_0 + \tau))\Theta(T_f - T_0 - \tau)\rangle +$$
$$+ \langle \delta(T_f - t_0 - t)\Theta(t_0 - T_0)\Theta(T_f - t_0)\Theta(t_0 - (T_f - \tau))\rangle +$$
$$+ \langle \delta(t_0 + \tau - T_0 - t)\Theta(t_0)\Theta(T_f - t_0 - \tau)\Theta(T_0 - t_0)\Theta(t_0 - T_0 + \tau)\rangle +$$
$$+ \langle \delta(T_f - T_0 - t)\Theta(t_0)\Theta(T_0 - t_0)\Theta(t_0 - (T_f - \tau))\Theta(\tau - (T_f - T_0))\rangle \Big), \tag{4.7}$$

where the operator $\langle \cdot \rangle$ is the ensemble average with respect to the random variable t_0, $\delta(x)$ and $\Theta(x)$ are the delta of Dirac distribution and the Heaviside function, respectively. $\mathcal{N}$ is the normalization. Solving the ensemble averages, the previous equation reads:

$$p_{\tau''}(t|\tau) = \frac{1}{\mathcal{N}}\Big(\delta(t - \tau)\int_{T_0}^{T_f - \tau}\Theta(T_f - T_0 - \tau)dt_0 +$$
$$+ \Theta(T_f - T_0 - t)\Theta(\tau - t) +$$
$$+ \Theta(T_f - T_0 - t)\Theta(T_0 + t - \tau)\Theta(\tau - t) +$$
$$+ \delta(t - (T_f - T_0))\min[T_0, T_0 - (T_f - \tau)]\Theta(\tau - (T_f - T_0))\Big). \tag{4.8}$$

Marginalizing with respect to τ, the probability distribution of τ'' is obtained:

$$p_{\tau''}(t) = \int_0^\infty p_{\tau''}(t|\tau)p_\tau(\tau)d\tau. \tag{4.9}$$

Eq. 4.8 combined with Eq. 4.9 yields:

$$
\begin{aligned}
p_{\tau''}(t) = \frac{1}{\mathcal{N}}\Big(& (T_f - T_0 - t)\, p_\tau(t)\,\Theta(T_f - T_0 - t) + \\
& + \Theta(T_f - T_0 - t)\int_{t>0}^{\infty} p_\tau(\tau)\,d\tau + \\
& + \Theta(T_f - T_0 - t)\int_{t>0}^{T_0+t} p_\tau(\tau)\,d\tau + \\
+ \delta(t-(T_f - T_0))\int_{T_f - T_0}^{\infty} & \min[T_0, T_0 - (T_f - \tau)]\,p_\tau(\tau)\,d\tau\Big).
\end{aligned}
\tag{4.10}
$$

The last term of Eq. 4.10 gives an atom probability in $t = \Delta T_w = T_f - T_0$ corresponding to the fraction of species that are always present during the observational window.

Recalling Eq. 4.7 and 4.9, the normalization constant $\mathcal{N}$ reads:

$$
\begin{aligned}
\mathcal{N} \;=\; \int_0^{\infty} \Big(& \langle \Theta(t_0 - T_0)\Theta(T_f - (t_0 + \tau))\Theta(T_f - T_0 - \tau)\rangle + \\
+ \; & \langle \Theta(t_0 - T_0)\Theta(t_0 - (T_f - \tau))\Theta(T_f - t_0)\rangle + \\
+ \; & \langle \Theta(t_0)\Theta(T_f - (t_0 + \tau))\Theta(T_0 - t_0)\Theta(t_0 - (T_0 - \tau))\rangle + \\
+ \; & \langle \Theta(t_0)\Theta(T_0 - t_0)\Theta(t_0 - (T_f - \tau))\rangle \Big) p_\tau(\tau)\,d\tau,
\end{aligned}
\tag{4.11}
$$

which simplifies to

$$
\begin{aligned}
\mathcal{N} \;=\; & \int_0^{T_f - T_0} (T_f - T_0 - \tau)\, p_\tau(\tau)\,d\tau \; + \\
+ \; & \int_0^{\infty} \min[T_f - T_0, \tau]\, p(\tau)\,d\tau \; + \\
+ \; & \int_0^{\infty} (\min[T_0, T_f - \tau] - \max[0, T_0 - \tau])\, p(\tau)\,d\tau \; + \\
+ \; & \int_{T_f - T_0}^{\infty} \min[T_0, T_0 - T_f + \tau)\, p_\tau(\tau)\,d\tau.
\end{aligned}
\tag{4.12}
$$

When comparing analytical and observational distributions, it is assumed that the system is at stationarity and unaffected by initial conditions, i.e. T_0 is far from the beginning of the process. Mathematically this is obtained taking the limit $T_0, T_f \to +\infty$ with $T_f - T_0 = \Delta T_w$ in

Eq.4.10, which finally takes the form:

$$
\begin{aligned}
p_{\tau''}(t) \;=\; & \frac{1}{\mathcal{N}}\Big((\Delta T_w - t)\,p_\tau(t)\Theta(\Delta T_w - t) \;+\; \\
& +\; \Theta(\Delta T_w - t)\int_{t>0}^{\infty} p_\tau(\tau)d\tau \;+\; \\
& +\; \Theta(\Delta T_w - t)\int_{t>0}^{\infty} p_\tau(\tau)d\tau \;+\; \\
& +\; \delta(t - \Delta T_w)\int_{\Delta T_w}^{\infty}(\tau - \Delta T_w)p_\tau(\tau)d\tau\Big),
\end{aligned}
\tag{4.13}
$$

where $\mathcal{N}$ simplifies to:

$$
\mathcal{N} = \Delta T_w + \langle \tau \rangle - 2\,\Delta T_w\, P_\tau(\Delta T_w) + 2\Big(\int_0^{\Delta T_w}(P_\tau(t) - t\,p_\tau(t))dt\Big),
\tag{4.14}
$$

with $P_\tau(t) = \int_t^{+\infty} p_\tau(\tau)d\tau$ being the exceedance cumulative distribution of the persistence-time probability density function.

The variable τ' comprises only the first of the four cases listed in Eq. 4.6. Thus the probability distribution $p_{\tau'}(t)$ follows directly from the first term of Eq. 4.13

$$
p_{\tau'}(t) \;=\; \frac{1}{\mathcal{N}'}(\Delta T_w - t)\,p_\tau(t)\Theta(\Delta T_w - t),
\tag{4.15}
$$

where the normalization constant $\mathcal{N}'$ is equal to

$$
\mathcal{N}' \;=\; \int_0^{\Delta T_w}(\Delta T_w - \tau)\,p_\tau(\tau)d\tau,
\tag{4.16}
$$

which completes the derivation.

Chapter 5

Empirical species persistence time distributions and implications for biodiversity

...he who does not admit to the unfathomable mystery cannot even be a scientist.

A. Einstein

5.1 Introduction

Empirical SPTs are defined as the number of consecutive years in which the measurements reveal the presence of the species in that geographic region (see figure 4.1) Assume that species abundance (or presence-absence) time series for a given single trophic ecosystem are available from field campaigns carried on in a time period of $\Delta T_w = T_f - T_0$ years. From the collected data, SPT can be measured and the empirical SPT obtained.

From an empirical viewpoint, species and genera persistence times deducted from fossil record data have been suggested to follow either power-law (with non-trivial exponents in the range $1.5 - 2$ [Sneppen et al., 1995; Sole and Bascompte, 1996; Newman and Sibani, 1999]) or exponential distributions [Van Valen, 1973; Sole and Bascompte, 1996]. It has been argued, however, that data quality, in particular for species, precludes a critical assessment [Pigolotti et al., 2005]. Also, local analyses of species persistence over ecological timescales suggest power-law distributions with non trivial exponents [Keitt and Stanley, 1998]. In what follows new empirical evidences for power-law behavior shall be provided. Results on empirical

SPT distribution will be compared with neutral model predictions on SPT distribution, and implications on emerging macroecological patterns will be examined, with special attention to possible biogeographical effects.

The problem of the scale of observation becomes a central one in this framework. In fact, at the local scale, say the observation site for presence/absence of breeding birds, persistence times of a species are controlled by ecological processes operating a short timescales, like population dynamics, dispersal, immigration, or contractions (expansions) of geographic ranges. At such scale local extinctions are dynamically balanced by colonizations [MacArthur and Wilson, 1967; Tilman, 1994; Ricklefs and Scheuerlein, 2003; Muneepeerakul et al., 2008b]. At the global scale, origination and extinctions are controlled by mechanisms acting on macroevolutionary timescales [Brown and Kodricbrown, 1977; Diamond, 1989]. For intermediate scales of observations, transitions from one regime to the other prove smooth. The scaling behavior proposed in the previous chapter to govern the transition from local to global scales is unsuitable to provide detailed descriptions about a specific species or a particular patch inside the ecosystem, and yet is capable of effectively describing the overall dynamical evolution of the ecosystem diversity. It also allows to predict SPT distributions for wide geographic areas from measures of persistence on smaller, way more tractable or observable areas. A biogeography of species persistence has thus been introduced.

5.2   Empirical SPT Distributions

SPT distributions are empirically characterized by analyzing four long-term datasets covering very different spatial scales: i) a 41-year survey of North American breeding birds [North American Breeding Bird Survey, 2008]; and ii) a 38-year inventory of herbaceous plants from Kansas prairies [Adler et al., 2007]; iii) a 44-year long study from the Buell-Small Succession (BSS) Study of plants in New Jersey [Institute of Ecosystem Studies, 2008] ; and iv) a 28-year long database of British marine fish collected at Hinkley Point (headland on the Bristol Channel coast of Somerset, England) [Henderson and Magurran, 2010].

North American Breeding Bird Survey Database

The North American Breeding Bird Survey consists of a record of annual abundance of more than 700 species over the 1966-present period along more than 5000 observational routes. The spatial location of the routes analyzed is shown in Figure 4.1. Only routes with a latitude less than 50° are considered, because density of routes with a long surveyed period drastically decreases above the 50th parallel. Noting that in many regions the survey started only in 1968, the first two years of observations are discarded in order to have simultaneous records for all the regions in the system. The spatial extent of the observational routes allows us to analyze species persistence times at different spatial scales. 20 different scales of analysis are considered, with linearly increasing values of the square root of the sampled area starting from $A = 10000 \, \text{km}^2$ to $A = 3.8 \cdot 10^6 \, \text{km}^2$. The whole system, which corresponds to an area of

$A = 7.8 \cdot 10^6$ km^2, is also analyzed. For every scale of analysis A several overlapping square cells of area A inside the system are considered. A three-dimensional presence-absence matrix P is thus built. Each element p_{stc} of the matrix is equal to 1 if species s is observed during year t in at least one of the observational routes comprised in cell c, otherwise $p_{stc} = 0$. For every scale of analysis the cells that (i) do not have a continuous record for the whole period (41 years) or (ii) have more than 5% of their area falling outside the system are discarded. For every cell and every species persistence times from presence-absence time series derived from the second dimension of matrix P are measured. Persistence time is defined as the length of a contiguous sequence of ones in the time series. For every scale of analysis all the measured persistence times, regardless of the species they belong to and the cell where they are measured, are considered.

Kansas Grass Prairies Database

The herbaceous plant dataset [Adler et al., 2007] comprises a series of 51 quadrats of 1 m^2 from mixed Kansas grass prairies where all individual plants were mapped every year from 1932 to 1972. In order to meet the data quality standard required for the analysis as discussed above for the breeding bird data, 10 quadrats are discarded and the first three years of observations. Due to the limited number of observational plots in the herbaceous plant dataset, the analysis is limited to quadrat spatial scale $A = 1$ m^2. Analogously to the previous case, the matrix P from presence-absence data for every species, year and quadrat, is reconstructed.

New Jersey BSS Field Study

The BSS study [Institute of Ecosystem Studies, 2008] includes ten fields that were released from agriculture and used to study succession dynamics. Permanent plots, measuring 48 m^2 each, were established in every field. The permanent plots were sampled every year from 1958 to 1979, after which they were sampled alternatively every other year to the present day. The fields differ in the year of release, thus only data collected after the latest field abandonment year (1968) are considered. In order to avoid missing data in alternate years, the minimum sampled area (afterwards named cell) considered in the calculation of the empirical SPT distributions comprises two adjacent fields (96 m^2). In this way each cell is populated with presence-absence records for each year. The same analysis is repeated for 3 other different scales: 4,6,8 adjacent fields (A=192, 288, 384 m^2, respectively). The whole system, which corresponds to an area of $A = 480$ m^2, is also analyzed. For every scale A of analysis, the presented results refer to the average over all possible combinations of adjacent plots within the system (moving average procedure). Again, the three-dimensional presence-absence matrix P is thus built and empirical SPT are calculated at every cell scale.

Bristol Channel Fishes Database

Fish samples were collected from the cooling-water filter screens at Hinkley Point B Power Station, situated on the southern bank of the Bristol Channel in Somerset, England. A full description of the intake configuration and sampling methodology is given in Henderson and Magurran [2010]. Only marine fish species (not crustaceous organisms) are considered. Empirical SPT pdf's can be computed as described for the previous datasets. Note, however, that this time there is not spatial scale variability, in fact the samples were collected in a single point in the space. Therefore the analysis is limited to a single spatial scale.

5.3   Comparison with Model Results

5.3.1   Fitting Results

Note that, when dealing with empirical survey data, the effect of the finiteness of the observational time window on the measured species persistence times must be properly taken into account, as explained in detail in the previous chapter. To this end, by matching analytical and observational distributions for $p_{\tau'}(t)$ and $p_{\tau''}(t)$ (described by Eqs (4.15) and (4.13)), it is possible to infer the persistence-time distribution $p_\tau(t)$.

The scaling exponent and the diversification rate have been determined with a simultaneous nonlinear fit of observational and analytical $p_{\tau'}(t)$ and $p_{\tau''}(t)$, giving a higher weight to the atom probability (the last point in each SPT distribution). Confidence intervals are equal to the standard error of the fit. For breeding birds and BSS forest plants, the nonlinear fit for different spatial scales of analysis is repeated. The reported scaling exponent and the confidence interval have been obtained by averaging results across spatial scales. Table 5.1 summarizes the results of the fit. Remarkably, in all cases the scaling exponents derived empirically are consistent with those predicted by the neutral voter model.

The SPT of breeding birds at different spatial scales of analysis and of herbaceous plants prove to be best fitted by a power-law distribution with an exponent $\alpha = 1.83 \pm 0.02$ and $\alpha = 1.78 \pm 0.08$, respectively (Figure 5.1), very close to that one predicted by neutral theory for ecosystem that live in 2D environment. On the other hand, for BSS forest plants and Bristol Channel marine fishes $\bar{\alpha} = 1.97 \pm 0.12$ and $\alpha = 1.97 \pm 0.06$ are obtained, which are compatible with the exponent $\alpha = 2$ predicted by the mean field voter model. In the case of the analyzed fishes (screened at Hinkley Point B Power Station), this result would be expected, not being any spatial information embedded in the empirical data. On the contrary, the negligible effects of the dispersal limitation on the ecosystem dynamic in the BSS plants community is quite unexpected and it suggests that the average plants dispersal radius R is most comparable with the linear dimension of the total sampled area, i.e. $R \approx \sqrt{A}$, and thus the net result is a mean field evolution not affected by spatial interaction as shown in the previous chapter by studying the effect of dispersal radius on the neutral Voter model. The fact that the total sampled area of the ecosystem is very small (total area of 480 m^2) strongly supports this interpretation of
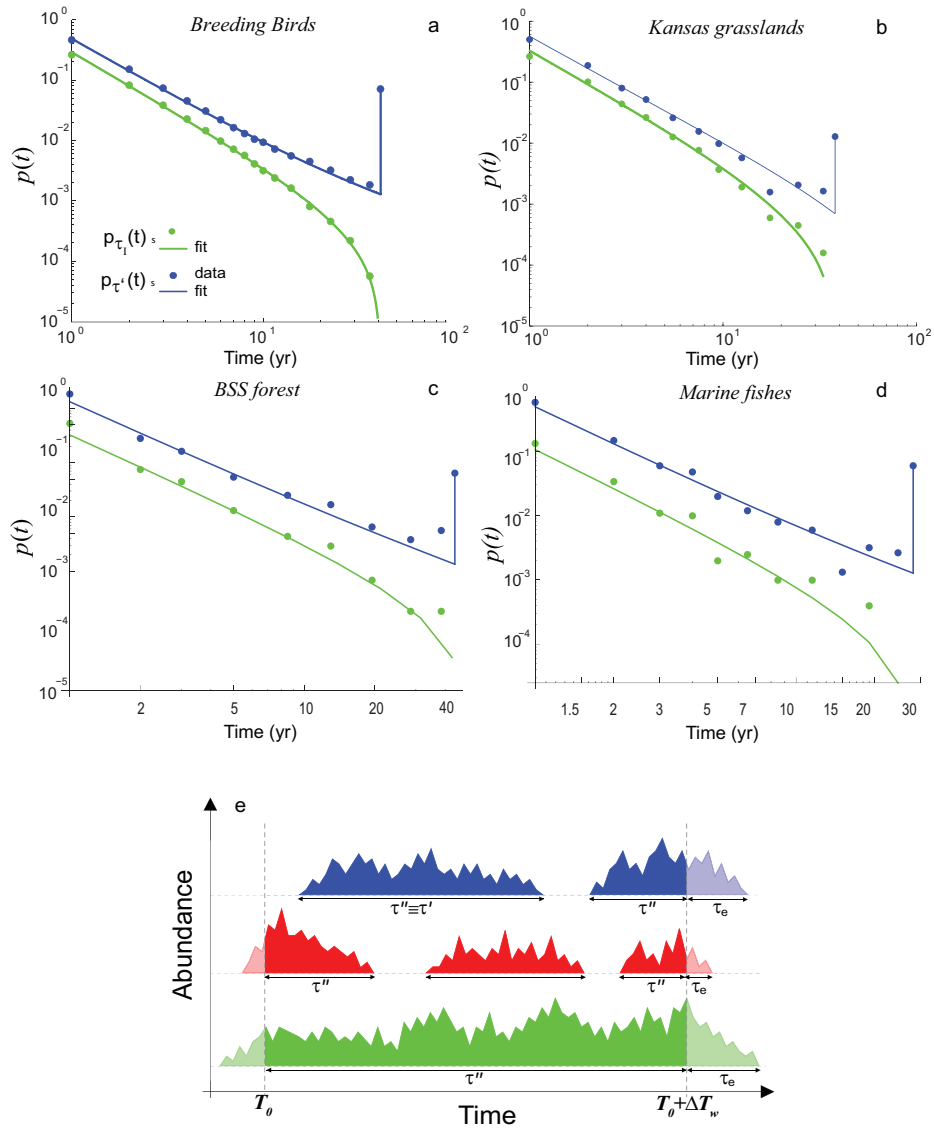
Figure 5.1: Comparison between empirical distributions for (a) North American Breeding birds, (b) Kansas grasslands, (c) New Jersey BSS forest, (d) marine fish community and the corresponding theoretical SPT probability density functions $p(t)$ of τ' (green), τ'' (blue). Filled circles and solid lines show observational distributions and fits, respectively. The spatial scale of analysis is $A = 10,000$ km^2 and $\Delta T_w = 41$ years for (a), A=1 m^2 and $\Delta T_w = 38$ years for (b), A=480 m^2 and $\Delta T_w = 44$ years for (c) and $\Delta T_w = 28$ years for (d) (for details on the analysis of the databases relevant to (a) and (b) see supplementary material in Bertuzzo et al., 2011). The finiteness of the time window imposes a cut-off to $p_{\tau'}(t)$ and an atom of probability in $t = \Delta T_w$ to $p_{\tau''}(t)$, which corresponds to the fraction of species that are always present during the observational time. $p_{\tau_I}(t)$ and $p_{\tau'}(t)$ have been shifted in the log-log plot for clarity. (d) A schematic representation of the variables that can be measured from empirical data over a time window ΔT_w: τ', persistence times that start and end inside the observational window, and τ'', which comprises τ' and all the portions of persistence times seen inside the time window that start or/and end outside. Times to local extinction τ_e are also presented.

the data.

ecosystem	$\bar{\alpha}$	AIC
Breeding Birds	1.83 ± 0.02	4
Kansas plant	1.78 ± 0.08	11
BSS forest	1.97 ± 0.012	15
Marine fishes	1.97 ± 0.06	16

Table 5.1: Scale invariant exponent from the simultaneous best fitting of Eqs (4.15) and (4.13) on empirical SPT distributions.

5.3.2 Persistence-Time Distributions for Breeding Bird Passeriformes Species

In order to test the robustness of the SPT pattern, the persistence-time distribution of a subset of the breeding bird data are characterized. Specifically, the passeriformes species, which comprise 282 out of the 644 species included in the dataset, are analyzed. Remarkably, also the persistence times of the passeriformes species prove to be best fitted by a power-law distribution with an exponent ($\alpha = 1.76 \pm 0.05$, Figure 5.2) close to the scaling exponent of the whole dataset ($\alpha = 1.83 \pm 0.02$)
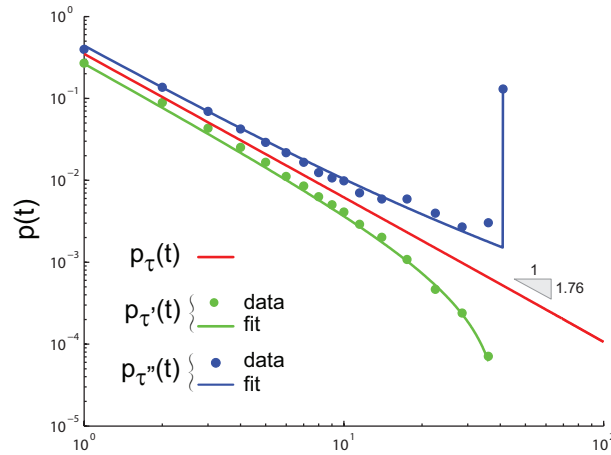
Figure 5.2: Persistence-time distributions of breeding bird passeriformes species. Probability density function of τ' (green), τ'' (blue) and persistence times τ (red) (see main text for explanation). Filled circles and solid lines show observational distributions and fits, respectively. The values of the best fit exponent is $\alpha = 1.76 \pm 0.05$.

5.3.3 Imperfect Detection

While studying animal communities on the basis of presence/absence (or count) data, imperfect detection of species is a source of concern, because animal species are routinely sampled with a detection probability < 1. This represents a well-known issue for the breeding bird dataset under analysis, as convincingly shown in the literature [Boulinier et al., 1998; Nichols

et al., 1998a,b; Cam et al., 2000; Boulinier et al., 2001; Cam et al., 2002; Alpizar-Jara et al., 2004]. The aggregation procedure of breeding bird data from route to cell level explained in the main text reduces the probability of imperfect detection. By terming p_r the route-level detection probability, i.e. the probability that a species is recorded given that it is present in the surveyed route. The probability of a pseudo-absence is therefore one minus the detection probability. Following [Alpizar-Jara et al., 2004], the detection probability at cell level p_c can be expressed as $p_c = 1 - (1 - p_r)^n$, where n is the number of routes comprised in the cell where the species is present. Therefore, the probability of detection at cell scale increases rapidly with n. Note that the in finest scale of the analysis ($A = 10000$ km^2), cells comprise an average of 6 routes. The
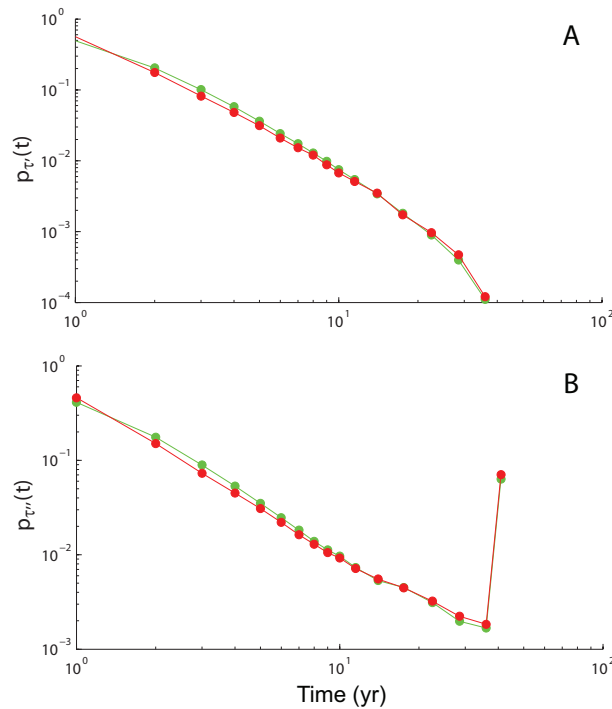
Figure 5.3: Imperfect detections in the breeding bird dataset. Comparison between the probability distributions of the variables τ' (A) and τ'' (B) (see main text for explanation) derived from the original data (red) and those derived after the addition of possibly undetected species (green).

sensitivity of the achieved results to imperfect detection have been explicitly tested. Starting from the data at route level, for every year and every route, species randomly chosen among the assemblage of species observed in that route in the whole observational window (41 years) have been added. The number of species added is chosen so that the resulting route-level detection probability has a constant value p_r. The data so modified have been then aggregated at cell level and analyzed to derive the persistence-time distribution. Given the randomness in the choice of the added species, the operation has been repeated 1000 times. An analysis on species imperfect detection in the Breeding Bird Survey dataset [Alpizar-Jara et al., 2004] has estimated an average route-level detection probability of about 0.8 for rare species and $p_r \simeq 0.9$ for more common species. The analysis have been run with the conservative value

$p_r = 0.8$ for all species. Figure 5.3 shows the comparison between the probability distribution of the variables τ' and τ'' derived from the original breeding bird data and those derived after the addition of possibly undetected species as explained above. The Figure refers to the finest scale of analysis, the more affected by imperfect detections. As expected, even with a low route-level detection probability ($p_r = 0.8$) imperfect detections have no significant impact on the resulting persistence-time distribution. The same procedure may be applied to the marine fishes databases. The problem of imperfect detection is far less relevant for the herbaceous plant dataset, given the sampling methodology and the small spatial scale of analysis [Adler et al., 2007].

5.4 Implication of the SPT pattern on Biodiversity

5.4.1 SPT and α diversity: the importance of being a river

In this section, SPT distribution, owing to its robustness and scale-invariant character, are suggested as a synthetic descriptor of ecosystem dynamics and of biodiversity. In fact, other key macroecological patterns are intimately related to the persistence-time distribution. A first clear example is the direct link with ecosystem diversity, as explained below. In the theoretical framework, species emerge as a point Poisson process with rate $\lambda = \nu N$ and last for a persistence time τ. The mean number of species S in the system at a given time is therefore $S = \lambda \langle \tau \rangle$ where $\langle \tau \rangle$ is the mean persistence time. Therefore, the smaller exponents obtained, say, for networked environments with respect to two-dimensional ones, imply longer mean persistence and, in turn, higher diversity. This echoes recent results suggesting a higher diversity of freshwater versus marine ray-finned fishes [de Aguiar et al., 2009; Moyle and Chech, 2003].

5.4.2 SPT and SAR

Another evidence of the effective way in which SPT distribution can characterize ecosystem diversity is the link with the species-area relationship (SAR), which characterizes the increase in the observed number of species with increasing sample area.

The scale invariant character of p_τ indicates that α depends only on the spatial connectivity of the environment (i.e. $1D, 2D,..$), and not on the area A where the species are considered. On the contrary ν, which accounts for immigration processes from species outside the local community, is argued to decreases with increasing sampled area A, thus being the only signature of the geographic scale in the analysis (see Figures 5.4a,b and 5.5a). The SAR characterizes how the average number of observed species increases with increasing sample area and it is usually characterized by a power law, i.e. $S \sim A^z$ [Brown, 1995]. The theoretical framework allows the linkage between the the cut-off timescale ν to the spatial scale of analysis, and the persistence-time distribution to the SAR.

In fact, as shown in the previous subsection, $\langle S \rangle = \nu N \langle \tau \rangle$, and thus yy deriving the scaling behavior of ν, N and $\langle \tau \rangle$ with respect to A, a connection between SPT and SAR can be made.

Scaling laws for breeding birds

The large spatial extent of the breeding bird datasets and the tools developed for the data analysis allow us to study with precision how the SPT distribution depends on the spatial scale of analysis. As expected, while the scaling exponent remains the same, the diversification rate ν decreases with the geographic area A and is found to closely follow a scaling relation of the type $\nu \propto A^{-\beta}$, with $\beta = 0.84 \pm 0.01$ in the case of breeding birds (Figure 5.4b) and , for a wide range of areas. This scaling form of the cut-off timescale $1/\nu$ can be related to the species-area relationship. Assuming that the number of individuals scales isometrically with the sampled geographic area [MacArthur and Wilson, 1967; Brown, 1995], i.e. $N \propto A$, and given that $\langle \tau \rangle = \int t p_\tau(t) dt \propto \nu^{\alpha-2}$ one gets:

$$S = \lambda \langle \tau \rangle \propto A^{1-\beta(\alpha-1)} = A^z. \tag{5.1}$$

The observational values $\beta = 0.84 \pm 0.01$ and $\alpha = 1.83 \pm 0.02$ give an exponent $z = 0.30 \pm 0.02$ which is close to the species-area relation measured directly on the data for the same range of areas ($z = 0.31 \pm 0.02$, Figure 5.4c). Conversely, one could have used the observed species-area exponent to infer the scaling properties of the diversification rate.

Scaling laws of BSS plants community

To infer with better precision the relation between ν and A from the few observational data, the information of the fraction of species always present in the ecosystem during ΔT_w are exploited. In fact the last term in Eq. (4.13) represents the analytical expression of this quantity, defined as $\mathscr{S}(\Delta T_w)$. To estimate with precision the diversification rate ν from the empirical SPT pdf, the value of ν that minimizes $\mathscr{S}(\Delta T_w)$ at every different spatial scales A (and using the scaling exponent $\bar{\alpha}$) is numerically obtained. The best fit of the scaling law between ν and A obtained in this way gives $\nu \sim A^{-\beta}$, with $\beta = 1 \pm 0.1$.

While for breeding birds the simple isometric relation $N \propto A$ [MacArthur and Wilson, 1967] assumption holds, the abundance data on the New Jersey plant community show that this is not the case in the BSS forest. In fact the average total species coverage percentage over all the plots is 165%. Samplers start with the topmost layer, usually the trees, and work their way down to the species closest to the ground [Institute of Ecosystem Studies, 2008]. Thus the exceedance in the coverage percentage is due to the overhanging among species. In order to evaluate the scaling relation between N and A, a minimalist model is then built, in which species are classified by their typical average size. Assuming biotic saturation, i.e. that each species covers completely the sampling region, in a plot of area A, at most $[A/a]$ individuals of characteristic size area $a \ll A$ can be found ($[x]$ denotes the smallest integer not greater that
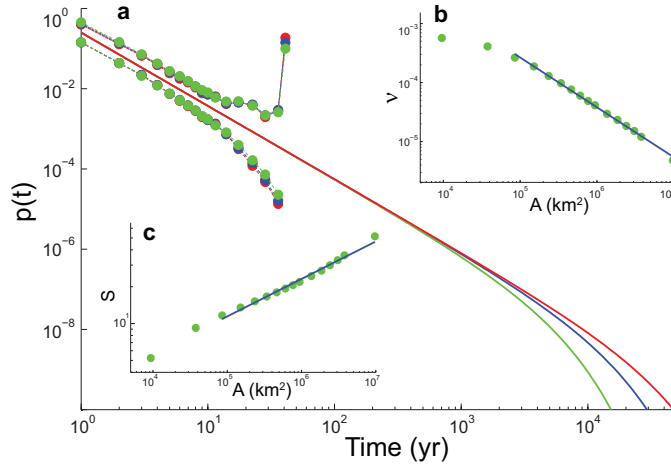
Figure 5.4: Biogeography of species persistence time. (a) Observational distributions $p_{\tau'}(t)$ and $p_{\tau''}(t)$ (interpolated solid circles) for the breeding bird dataset and corresponding fitted persistence-time distributions $p_\tau(t) \propto t^{-\alpha} e^{-\nu t}$ (solid lines) for different scales of analysis: Area $A = 8.5 \cdot 10^4$ km^2 (green), $A = 3.4 \cdot 10^5$ km^2 (blue), $A = 9.5 \cdot 10^5$ km^2 (red). $\nu(A)$ provides the cut-off for the distribution, whose scaling exponent is unaffected by geographic area. Note that the position of the cut-off of $p_\tau(t)$ is inferred from the estimate of the atom of probability of $p_{\tau''}(t)$ which is more sensitive to the scale of analysis; (b), Scaling of the diversification rate ν with the geographic area $\nu \propto A^{-\beta}$, $\beta = 0.84 \pm 0.01$; (c), Empirical species-area relationship (SAR). The plot shows the mean number of species S obtained in moving squares of size A. $S \propto A^z$, $z = 0.31 \pm 0.02$ has been obtained. Slope and confidence interval have been obtained averaging 41 SARs, one per year of observation.

x), $[A/2a]$ individuals of characteristic area $2a$ and so forth. In general, the total number of individuals N in a region of area A scales as

$$N(A) \propto \sum_{n=1}^{[A/a]} \frac{A}{na} \sim \frac{A}{a} H_{A/a}, \tag{5.2}$$

where $H_q = \sum_{n=1}^{q} 1/n$ is the harmonic number and $H_{A/a} \sim \ln(A/a)$. The constant of proportionality in Eq. (5.2) depends on the arbitrary choice of a. However since $H_{A/a} \sim \ln(A/a) +$ sub-leading terms, a change of a in λa leads to $H_{\lambda a} = \frac{1}{\lambda} H_{A/a} +$ sub-leading terms and so its precise value is irrelevant since it can be absorbed in the proportionality constant (see below) to be determined from the best fit of the data.

Putting all these scaling results together, the mean-field ($\alpha \approx 2$) neutral framework for this ecosystem predicts a power-law SAR with logarithmic correction:

$$\langle S(A) \rangle = K A^{1-\beta} \ln(A) H_A, \tag{5.3}$$

where K is the constant of proportionality. The constant K is determined in three different ways: 1) through the best fit of Eq. (5.3) on the empirical SAR data; 2) by imposing that in

the total area A^* the total number of species S^* is found; 3) by finding the value of K that gives the exact number of species in the smallest area ($A = 96$ m^2) and then predicting the complete SAR curve (upscaling). Remarkably, all three methods yield $K \approx 3.46$. The best fit of the power-law SAR $\left(S \sim A^z\right)$ on the empirical data has also been performed, giving $z = 0.34$. Therefore, in this range of areas, it is not distinguishable with the mean field solution given by Eq. (5.3).

In the special case of mean field interactions, the endemic-area curve (EAR) can also be calculated. EAR is defined as the relation between the number of species endemic to a region and area of that region [Green and Ostling, 2003]. The EAR has been recently proposed by [He and Hubbell, 2011] as the correct approach in order to estimate extinction from habitat losses. Results on SPT and SAR suggest that in the BSS ecosystem the phenomenon of spatial aggregation or clustering is noneligible and species can be considered randomly distributed in space. Therefore, the EAR curve can be can evaluated [He and Hubbell, 2011]:

$$\langle E(A) \rangle = S^* - \langle S(A^* - A) \rangle = S^* - K \ln(A^* - A) \, H_{A^* - A}. \tag{5.4}$$

Figure 5.5 summarizes all the above results and makes the comparison between the empirical and theoretical results of the SAR and EAR.

5.5 Conservation perspective and universality of the SPT pattern

Finally, from a conservation perspective, a meaningful assessment of species' local extinction rates is deemed valuable. The distribution of the times to local extinction τ_e (Figure 5.1A) are proposed as a tool to quantify the dynamical evolution of the species assembly currently observed within a given geographic area. Mathematically, τ_e is defined as the time to local extinction of a species randomly sampled from the system, regardless of its current abundance. When Eq. 4.4 holds for persistence times, the distribution of the times to local extinction $p_{\tau_e}(t)$ is shown to scale as $p_{\tau_e}(t) \propto t^{1-\alpha} e^{-\nu t}$. Therefore, not only do the developed theoretical and operational tools allow to infer the scaling behavior of persistence times, but also of the time to local extinction even from relatively short observational windows. Although these patterns cannot provide information about the behavior of a specific species or of a particular patch inside the ecosystem considered (e.g. a biodiversity hot-spot) they can effectively describe the overall dynamical evolution of the ecosystem diversity. In particular the scaling behavior allows to extrapolate SPT distributions for wide geographic areas, which are hard to estimate, from measures of persistence on smaller areas, which are, on the contrary, more practical and feasible. Empirical evidences that support the hypothesis of the existence and universality of the SPT pattern have been presented. Four different ecosystems have been analyzed: breeding birds, two plants community and marine fishes. In all cases the observed SPT distributions display a power-law shape with a cut-off due to the finiteness of the observational time window, which are well reproduced by the neutral theoretical model. The specific value of the exponents depends only a few key ingredients, namely the spatial dimension of the embedding
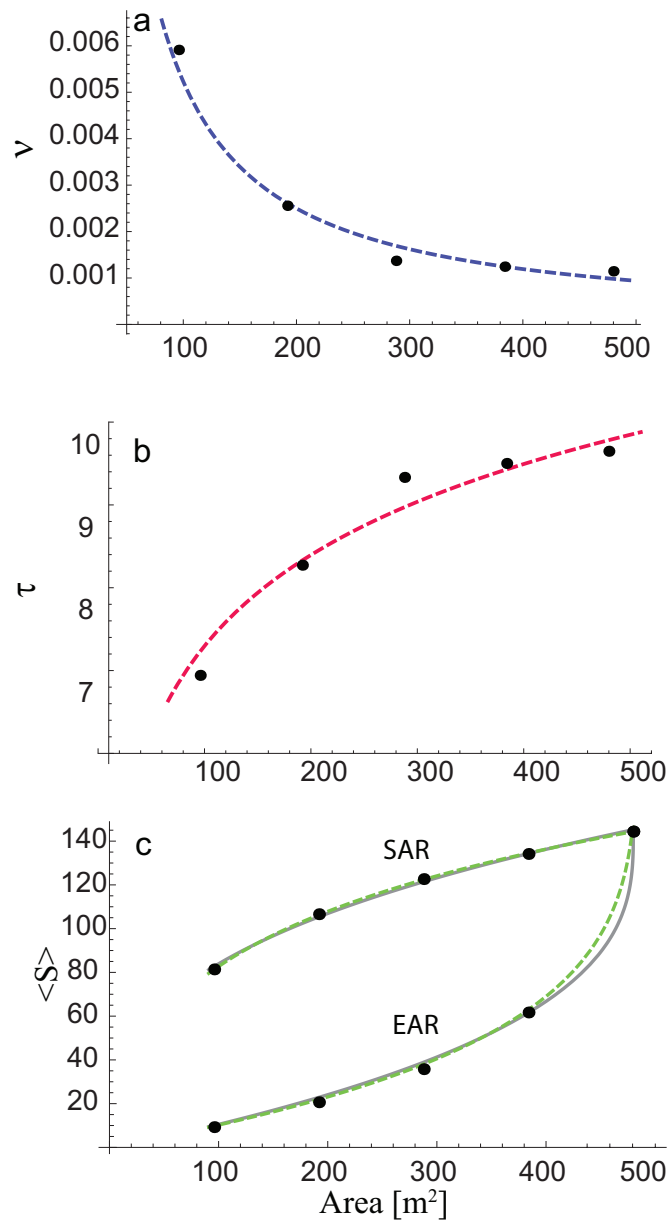
Figure 5.5: Comparison between a) Diversification rate for different areas (black dots) and scaling law $\nu \sim A^{-1}$ (dashed blue line); b) mean persistence time for each sample area (black dots) and logarithmic relation $\tau \sim \ln(A)$ (dashed red line) predicted by the neutral model for the mean field case ($\alpha = 2$); c) Empirical SAR and EAR relation (black points), power law fitting of the SAR and the corresponding EAR (gray continuous line) and the mean field SAR and EAR given by Eqs. (5.3) and (5.4) (green dashed).

ecosystem and the presence of dispersal limitation, or the lack of them. Because of the scale invariance character of SPT, only limited by biogeographical finite−size effects, and its relation with other important macro−ecological pattern, it is suggested that SPT distribution can be

considered as a syntectic descriptor of ecosystems biodiversity and dynamic.

For these reasons, field biologists and ecologists should perhaps invest renewed efforts in collecting improved long-term datasets of ecosystem dynamics at different spatial and temporal scales and under different dispersal conditions (for instance, archives beyond presence/absence of species within networked environments like riverine ecosystems for freshwater fish) as they prove crucial for a deeper understanding of universal patterns in macroecology. The present study shows that long-term datasets can be profitably used to highlight crucial properties and spatial effects of ecosystem dynamics, including the role played by the underlying connectivity structure in shaping SPT distributions.

New data would allow for the investigation of other implications of the theoretical predictions. For instance, the neutral voter model predicts that the scaling exponents of the SPT distributions of riparian ecosystems (i.e. networked environments where directional, anisotropic dispersal is forced by the structure of the fluvial environmental matrix) should be lower than those of 2D, 3D or mean-field ones. This result, if proven, would have remarkable consequences for conservation ecology, because it suggests that species that disperse isotropically have shorter average persistence times than species that are constrained to disperse along spatially constrained and anisotropic ecological corridors, like those provided by river networks. This, in turn, calls for long-term analysis of riparian ecosystems to test empirically the effects of river morphology on ecosystem dynamics.

The biogeographical characters of species persistence, stemming from the structure of the spatial interaction networks and from local constraints to species emergence rates, add a new ingredient to a rich literature bearing major implications for the inventory of life on Earth.

Water footprints of human society

Overview

The human society may be considered a very peculiar ecosystem. Human "species" is of course unique and yet its existence depends crucially on water availability; not only for drinking purposes, but in particular because of its impacts on food production.

Therefore, recurrent or ephemeral water shortages are a crucial worldwide challenge. The global character of this challenge is reflected in the trade among nations of virtual water, defined as the amount of water used to produce a given commodity. Virtual water may be seen as footprints of human society, as many key aspects of a given country are related to its population [Bettencourt et al., 2007], which sustainability depends in turn upon water.

In this part, complex network theory is utilized to analyze the structure of the global virtual water trade associated with the international food trade. In its complete representation, the global virtual water trade forms a weighted and directed network, in which the nodes correspond to nations and the links represent the flows of virtual water from the country of export to the country of import.

In Chapter 6, the topological and weighted properties of the undirected representation of this network are quantitatively characterized using statistical measures from complex network theory. In particular implications and consequences of the structure of the network will be discussed.

In Chapter 7, it is shown how all the key features of the directed and undirected network are well described by a stochastic model. It is able to reproduce both the topological and weighted properties of the global virtual water trade network, by assuming as sole controls each country's gross domestic product and yearly rainfall on agricultural areas. The high degree of globalization of water trade is captured and quantitatively described and it is shown how a small group of nations play a key role in the connectivity of the network and in the global redistribution of virtual water. Finally, examples of prediction of the structure of the network under future political, economic and climatic scenarios are illustrated, suggesting that the crucial importance of the countries that trade large volumes of water will be strengthened.

The above analysis provides the necessary framework for the development of a model of global virtual water trade aimed at applications ranging from network resilience analysis, to climate change impact evaluations.

Part III

The Structure of the Global Virtual Water Trade Network

Science is facts; just as houses are made of stones, so is science made of facts; but a pile of stones is not a house and a collection of facts is not necessarily science.

J.H. Poincaré

6.1 Trade, food and water: a network to share resources

Food production is by far the most freshwater-consuming process (80% of the total world water resources [Rost et al., 2008]). Due to population growth and economic development, water shortage is thus subject to increasing pressure at local and global scales. Several studies have recently focused on the issues of globalization of water [e.g. Hoekstra, 2002; D'Odorico et al., 2010b], using the concept of virtual water (VW) [Allan, 1993], defined as the volume of water used to produce a given commodity. They have highlighted the importance of tackling water management problems not only at the basin or country scales, but rather through a worldwide perspective [Hoekstra and Chapagain, 2008]. In fact, international trades link fortunes and resources of countries, providing potentially important conduits for geographically limited water resources to be transferred to food and water-stressed regions.

The virtual water trade between regions [Hoekstra and Hung, 2005; Chapagain et al., 2006; Yang et al., 2006] and the gross virtual water flow of nations [Hoekstra and Chapagain, 2008] have been quantified. These studies have primarily focused on agricultural commodities [Hoekstra and Hung, 2005; Rost et al., 2008], including those used for bio-fuel production

[Gerbens-Leenes et al., 2009], but the concept has also been extended to include industrial products [Hoekstra and Chapagain, 2008]. However, the global properties of virtual water trade have not yet been quantified or explored. In this chapter, complex network methods are utilized to characterize the global structure of the virtual water trade associated with the international food trade.
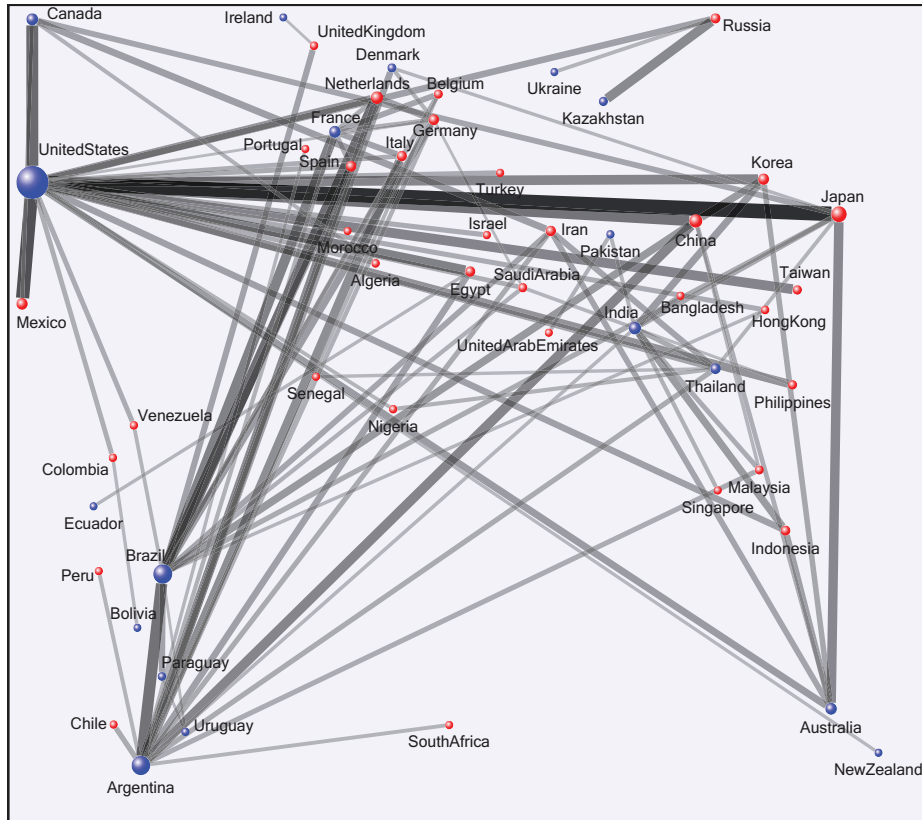
Figure 6.1: Backbone of the global Virtual Water Trade Network (GVWTN). Only 4% of the total number of links accounting for 80% of the total flow volume are shown. Resulting isolated nodes are consequently removed. The blue nodes represent the net exporter nations, while the red ones are the net importers. The weights of the links are color-coded by the grayscaling in the edge's colors (black is the link carrying the highest volume of VW.)

The origin of complex network theory can be traced back to the work of Erdös and Rényi [1961] on random graphs. Recently, much research has been devoted to the field of complex network analysis, both theoretically and as applied to real-world systems [Barabási and Albert, 1999; Newman et al., 2006]. This recent interest in complex networks is largely due to the discovery of organizing principles in networks, such as community structure [Watts and Strogatz, 1998] and scale-free properties [Barabási and Albert, 1999]. Additionally, network analysis has become increasingly popular due to its flexibility and generality for representing many natural structures [Barabasi, 2002; Newman, 2002], including street systems [Costa et al., 2007], the internet and world wide web [Park and Newman, 2003], international tourism [Miguens

and Mendes, 2008], international trade web [Fagiolo et al., 2008], financial transactions [Garlaschelli et al., 2005], Hollywood actors and scientific collaborations [Newman, 2002], among others.

6.2 H08 Model and Virtual Water Content Data

VW flows of 5 major crops (barley, corn, rice, soy, and wheat) and 3 livestock products (beef, poultry, and pork) has been calculated. The VW content of such commodities are calculated for each nation using a state-of-the-art global water resources model [Hanasaki et al., 2008a, 2010], at a spatial scale of $0.5^o \times 0.5^o$. A brief description of the H08 model [Hanasaki et al., 2008b,a, 2010] is provided here. The H08 model is a global water resources model composed of six modules: land surface hydrology, river routing, crop growth, reservoir operation, environmental flows requirements estimate, and anthropogenic water withdrawal. It simulates both natural and anthropogenic water flows globally on a daily basis at a spatial resolution of $0.5^o \times 0.5^o$ (longitude and latitude) with water and energy balance closure. In order to run the H08 model two types of input data are needed: meteorological forcing and land use. The former is provided by the near-surface meteorological database (NCC [Ngo-Duc et al., 2005]), while the latter is obtained by compiling published land-use data (for details see [Hanasaki et al., 2010]). Once the inputs are known simulations can be performed. First the meteorological forcing for the initial year of the simulation period (1980) was iteratively fed to the H08 model until the state variables of the model (such as soil moisture, river channel storage, and storage of large and medium-size reservoirs) reached equilibrium. Then twenty years were simulated. Finally, the amount of evapotranspiration from irrigated and non-irrigated cropland was obtained for each grid cell and averaged over decades. The virtual water content of raw crops is defined as the kg of water used to produce a kg of crop yield, while the virtual water content of livestock products is calculated as the water consumption per head of livestock divided by the livestock production per head. Thus combining the averaged outputs of the H08 model, the required livestock feed per head (estimated taking in account the life cycle of livestock) and the trade food data allowed us to calculate the virtual water flows between countries for each single food product.

6.3 Network Construction and Properties

Data on international food trade from the year 2000 and concerning the N=184 nations under study, have been obtained from FAO [2000a]. In there, 58 commodities from 8 major products (barley, corn, rice, soy, wheat, beef, chicken and pork) are considered. Eight 184×184 matrices (food trade matrices) $FTM^{(k)}$ with $k = 1, 2, .., 58$ are built, each one referring to the k-th commodity. The entry $FTM^k(i, j)$ is the volume of the k-th commodity exported from countries i to j. In order to avoid discrepancies in the trade volumes, the average of the values reported by the exporter and by the corresponding importer is taken. An exception was made where one of the two nations reported a trade while the other one does not: in this case the

reported values are used. The FAO reported trade data are considered to reflect the actual trade that is occurring among nations.

The weighted architecture [Barrat et al., 2004] of the undirected network (for a given food product) is described by the matrix W, whose elements ($w_{ij} = w_{ji}$) represent the total volume of VW exchange between nodes obtained by summing the corresponding import and export fluxes (edges). The topological structure of the network is fully specified by the adjacency matrix a with elements $a_{ij} = \Theta(w_{ij})$, where Θ is the Heaviside function.

The network is characterized by the following node's properties:

$$k_i \equiv \sum_{j=1}^{N} a_{ij}, \qquad \text{node degree} \tag{6.1}$$

$$s_i \equiv \sum_{j=1}^{N} w_{ij}, \qquad \text{node strength} \tag{6.2}$$

$$k_{nn,i} \equiv \frac{1}{k_i} \sum_{j=1}^{N} a_{ij} k_j, \qquad \text{average nearest neighbor degree} \tag{6.3}$$

$$s_{nn,i} \equiv \frac{1}{k_i} \sum_{j=1}^{N} a_{ij} s_j, \qquad \text{average nearest neighbor strength} \tag{6.4}$$

$$s_{nn,i}^{W} \equiv \frac{1}{s_i} \sum_{j} w_{ij} s_j, \qquad \text{strength-strength correlations} \tag{6.5}$$

$$C_i \equiv \frac{1}{k_i(k_i - 1)} \sum_{j,k=1}^{N} a_{ij} a_{jk} a_{ki}, \qquad \text{clustering coefficient.} \tag{6.6}$$

The generalization of Eqs. (6.1)-(6.6) for the directed network is straightforward.

In this work the countries are ranked based on the total volume of transfers s_i. However, some countries may have low water transfers because they have small populations. Some of these countries are more reliant on imports then the large volume countries. In order to study the virtual water transfers also related to population sizes, one may define, besides the total strength of the node s, also the pro-capite strength of node i

$$s_{proC} = \frac{s_i}{pop_i}, \tag{6.7}$$

where pop_i is the population of country i. Countries as United Arab Emirates, Qatar, Kuwait and Oman are found to be in the top ranks of highest s_{proC}, while they have very low total s. This show how also small volumes of VW but in water scarce low populated countries are very important to understand as these nations are likely to get into a bind with global water resources.

6.4 Undirected GVWTN

Key statistical characterizations of the undirected GVWTN are here presented and interpreted. The global properties of the directed network will be briefly presented in the next chapter.

The undirected network is characterized by the symmetric matrix W, whose elements ($w_{ij} = w_{ji}$) represent the total volume of VW exchanged between countries (nodes) and obtained by summing the corresponding import and export fluxes (Figure 6.1).

The global topology of a network is described by its degree probability density function (pdf) $p(k)$, i.e. $p(k)dk$ is the probability that the degree of a given node is k [Newman et al., 2006] (Figure 6.2a). It provides the number of edges connected to a given node regardless of the identity of the neighbors. To investigate how nodes are connected, the average nearest neighbor degree k_{nn} is studied. It shows a tendency of the nodes with high degree to provide connectivity to small degree nodes (Figure 6.2c). Such trend (known as disassortative behavior) denotes, differently from purely random networks, non-trivial nodal degree correlations [Newman, 2002]. Another interesting indicator is the local clustering coefficient C_i ($0 \leq C_i \leq 1$) which describes the ability of node i to form cliques, i.e. triangles of connected nodes. Figure 6.2d shows that poorly connected nations i tends to form connected trading food sub-markets ($C_i \approx 1$). On the contrary, high degree vertices j connect otherwise disconnected regions ($C_j \ll 1$). The average clustering coefficient is very high ($\bar{C} = 0.747$) and the graph has, in analogy to many real networks [Newman et al., 2006], an average node-to-node topological distance (d_{nn}) smaller than 5 ($d_{nn} = 4$). The GVWTN thus exhibits a small-world network behavior [Watts and Strogatz, 1998], providing a quantitative measure of the globalization of water resources [Hoekstra and Chapagain, 2008; D'Odorico et al., 2010b].

The hydrological features of the network are given by its weighted properties. The total volume imported and exported by nation i is quantified by its strength s_i, defined as the total VW volume exchanged by node i. The strength distribution shows a heavy-tailed pdf suggesting high heterogeneity of the volumes of traded VW (Figure 6.2b): only 4% of the total number of links accounts for 80% of the total flow volume, indicating established bonds among countries that rule the main fluxes in the GVWTN (Figure 6.1). Strengths between neighboring nodes are correlated. In fact, the average nearest neighbor strength s_{nn} [Serrano, 2008] displays a decreasing trend as a function of s (Figure 6.2e). Strength-strength correlations disentangled from degrees (s_{nn}^W) [Serrano, 2008] are not significant, i.e. s_{nn}^W does not depend on s. A power-law relation $s \sim k^b$ with exponent $b = 2.6 \pm 0.03$ ($R^2 = 0.982$) (see Figure 6.3) indicates a non-trivial correlation between node degrees and strengths [Barrat et al., 2004]. The above suggests that we live in a global water world where, on average, the export of VW from few water rich countries increases the food locally available to the connected nations. At the same time, there exist preferential VW routes, mainly driven by geographical, political and economical factors, through which most of the VW volume flows.

In the next Chapter a theoretical network model that robustly describes the topological and weighted properties of the global VW trade network (GVWTN) will be proposed.
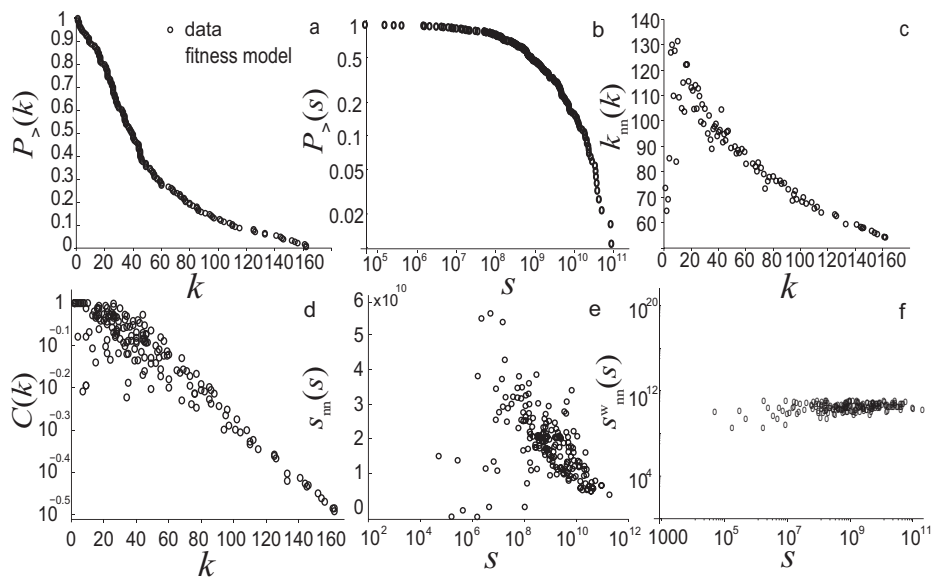
Figure 6.2: Topological and weighted properties of the GVWTN. a-b) Cumulative pdf of the node's degree $P_>(k)$ in linear scale and node strength $P_>(s)$ in log-log scale; c-d) average nearest neighbors degree k_{nn} and cluster coefficient C as a function of the nodes degree k; e-f) average nearest neighbors strength s_{nn} and strengths-strengths correlation $s_{nn}^W(s) \sim const$ in semilog-x and log-log scale, respectively.
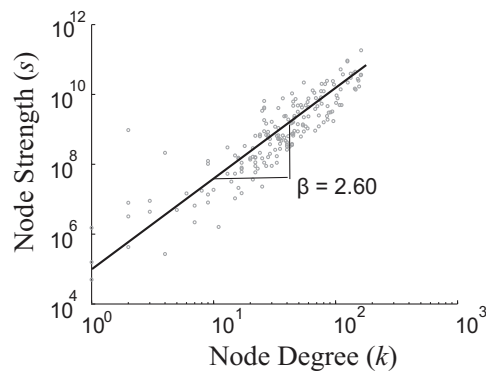
Figure 6.3: Empirical Relation between s and k: the best fit gives a power-law relation $s \sim k^{2.6}$

Chapter 7

Controls of the Global Virtual Water Trade Network and Future Scenarios

With four parameters I can fit an elephant, and with five I can make him wiggle his trunk .

J. von Neumann

7.1 The Fitness Model

The complexity of all factors (political, economical and environmental) involved in shaping the GVWTN structure is remarkable, and calls for investigating whether key variables and linkages exist through which the emerging structural properties of the network could be revealed. A model that can concisely describe all the above features of the GVWTN is developed. Specifically, it is assumed that the topological and weighted features of the network can be determined, respectively, by two external characteristics of each node: namely, the gross domestic product [World Bank, 2010] (GDP) and the (average) yearly rainfall [mm/yr] on agricultural area [km^2] (denoted by RAA [mm·km^2/yr]).

A 'fitness model' is a conventional definition for a general class of network models where vertices are labeled by an intrinsic quantity (i.e. the fitness, or hidden variable) that determines their connection and/or strength probability [Bianconi and Barabási, 2001; Caldarelli et al., 2002; Boguna and Pastor-Satorras, 2003; Garlaschelli and Loffredo, 2004]. Here a model of the global virtual water trade network (GVWTN) is described. It consists of: i) a fitness model describing the topological properties of the network, and ii) a fitness model, independent of i), characterizing its weights. Although both models are based on the individuation of the appropriate set of fitness variables, they have a different role and meaning.

The null hypothesis for the functional shape of the fitness function $p(x_i, x_j)$ (i.e. that giving the probability that node i and j – endowed respectively with fitness x_i and x_j – are connected) can be obtained through an entropy optimization principle [Park and Newman, 2004]. In particular, let $G \in \mathcal{G}$ represent a graph – chosen within a specified ensemble $\mathcal{G}$ – constrained to have a specified degree sequence $\{k_l\}_{l=1,2,..}$ (specifically, the sequence of node degrees from the real-life empirical GVWTN). If $P(G)$ is the probability of that graph within the ensemble, one may choose $P(G)$ such that the expected values of the degree sequence of G, $\{\langle k_l \rangle\}_{l=1,2,..}$, equal the observed empirical values $\{k_l\}_{l=1,2,..}$ [Park and Newman, 2004; Garlaschelli and Loffredo, 2009]. A choice of a $P(G)$ satisfying this requirement is obtained by maximizing the Gibbs entropy defined as $S \equiv -\sum_{G \in \mathcal{G}} P(G) \log[P(G)]$. The solution of the entropy optimization problem subject to the constraints:

$$\sum_{G \in \mathcal{G}} P(G) k_l(G) \;=\; \langle k_l \rangle \qquad l = 1, 2, \dots \tag{7.1}$$

$$\sum_{G \in \mathcal{G}} P(G) \;=\; 1, \tag{7.2}$$

is [Park and Newman, 2004]:

$$P(G) = \frac{e^{\mu L - H(G)}}{Z}, \tag{7.3}$$

where $H(G) \equiv \sum_l \alpha_l k_l$ is the graph Hamiltonian, $\{\alpha_l\}_{l=1,2,..}$ are the Lagrange multipliers, $Z = \sum_G e^{-H(G)}$ is the partition function, and μ is a parameter, analogous to the chemical potential, that controls the number of edges [Park and Newman, 2004; Garlaschelli and Loffredo, 2009]. More generally, the graph Hamiltonian can be defined as [Garlaschelli and Loffredo, 2006]

$$H = \sum_{i<j} \epsilon_{ij} a_{ij}, \tag{7.4}$$

where a_{ij} is the adjacency matrix and $\epsilon_{ij} = \alpha_i + \alpha_j$ is the coupling parameter between i and j that can be interpreted as the energy necessary to establish a link between i and j (that have fixed degree k_i and k_j, respectively). Consequently, the following partition function is obtained:

$$Z = \sum_{\{a_{ij}\}} e^{\mu L - \sum_{i \neq j} \epsilon_{ij} a_{ij}} = \prod_{i<j} Z_{ij}, \qquad \text{with } Z_{ij} = 1 + e^{\mu - \epsilon_{ij}} \tag{7.5}$$

where the first sum is over all possible configurations of $\{a_{ij}\} = 0, 1$. Note that the relation $L = \sum_{i<j} a_{ij}$ has been employed. Finally, the free energy of the network can be obtained as:

$$\Omega = -\log(Z) = -\sum_{i<j} \log((1 + e^{\mu - \epsilon_{ij}})), \tag{7.6}$$

and from Eq.(7.6) many properties of the network can be obtained exactly [Park and Newman, 2004; Garlaschelli and Loffredo, 2006]. For instance, the probability p_{ij} of having an undirected

link from i to j is

$$p_{ij} = \langle a_{ij} \rangle = -\frac{\partial \Omega_{ij}}{\partial \mu} = \frac{\sigma e^{-\epsilon_{ij}}}{1 + \sigma e^{-\epsilon_{ij}}}, \tag{7.7}$$

where $\sigma = e^{\mu}$.

The null hypothesis is that the fitness variable x_i (that is, the external quantity $x_i = GDP_i / (\sum_j GDP_j)$), determines the topological importance of node i by driving the number of its connections. Moreover, all graphs with degree sequence $\{k_l\}_{l=1,2,..}$ must appear in the ensemble with equal probability [Park and Newman, 2003]. For these reasons ϵ_{ij} are set as

$$e^{\epsilon_{ij}} = -x_i \cdot x_j, \tag{7.8}$$

and thus Eq. (7.7) becomes:

$$p_{ij} = p(x_i, x_j) = \frac{\sigma x_i x_j}{1 + \sigma x_i x_j}. \tag{7.9}$$

From Eq. (7.9), all topological properties can be computed: the node degree

$$\langle k_i \rangle = \sum_{j \neq i}^{N} p(x_i, x_j), \tag{7.10}$$

the average degree of the nearest neighbors

$$\langle k_{nn,i} \rangle = \frac{\sum_{j \neq i}^{N} \sum_{l \neq j} p(x_i, x_j) p(x_j, x_l)}{\langle k_i \rangle}, \tag{7.11}$$

and the cluster coefficient

$$\langle C_i \rangle = \frac{\sum_{j \neq i}^{N} \sum_{l \neq j,i}^{N} p(x_i, x_j) p(x_j, x_l) p(x_l, x_i)}{(\langle k_i \rangle - 1) \langle k_i \rangle}. \tag{7.12}$$

Although generalizations of the model exist [Garlaschelli and Loffredo, 2009], it is found that they are not suitable to describe the weights in the GVWTN. For instance, it is found that the null model for the weighted network in which the entropy is maximized by constraining both the degree and strength sequences, does not describe the properties of the empirical network. For this reason, a weighted fitness function [Caldarelli et al., 2002] $q(y_i, y_j)$ is built. It assigns a value (interpreted as the average weight $\langle w_{ij} \rangle$) to the link connecting i to j. Therefore, in this case, the weight $\langle w_{ij} \rangle$ is not interpreted as the number of edges between node i and j, but rather as a number associated with the assigned link between the two nodes ranking its importance.

The functional shape of $q(y_i, y_j)$ is in principle arbitrary. For sake of simplicity, and generalizing the concept of weighted configuration model [Serrano and Boguna, 2005; Garlaschelli and

Loffredo, 2009], the following expression for $q(y_i, y_j)$ is chosen:

$$q(y_i, y_j) = \langle w_{ij} \rangle = \eta y_i y_j, \tag{7.13}$$

where η is the parameter controlling the total flux of the network and y_i is the normalized fitness variable.

7.2 The Fitness Variables and Results

Each of the 184 countries (nodes) is assigned its corresponding normalized value of the *GDP* (*x*) and *RAA* (*y*) based on data from 2000 [United Nations, 2010; World Bank, 2010] (i.e., $x_i = GDP_i / \sum_{j=1}^{N} GDP_j$, $y_i = RAA_i / \sum_{j=1}^{N} RAA_j$). These variables are referred as fitness (or hidden) variables [e.g. Bianconi and Barabási, 2001; Caldarelli et al., 2002; Boguna and Pastor-Satorras, 2003; Garlaschelli and Loffredo, 2004; Park and Newman, 2004]. They measure the relative importance of the vertices in the GVWTN. *GDP* and *RAA* are assumed to be good candidates to explain the structure of the GVWTN. In fact the country *GDP* is closely related to its trade activity [Garlaschelli and Loffredo, 2004] and drive the numbers of trading partners that a country has. On the other hand, volumes of virtual water traded by a country in general does not crucially depend on its *GDP*, but rather on the amount of crops and meat produced in that country (even if they were correlated). For instance, Thailand has a *GDP* much lower than that of Japan, but it exports much more virtual water. Thailand is in fact a nation based on agriculture, and it is one of the major rice exporter in the world. Two important factors determine the virtual water productivity of a country. They are mean annual rainfall (in arid or semiarid climates agriculture and livestock holdings are unfavored) and the extent of agricultural areas. The latter plays a key role. In fact, without agricultural areas obviously no crop is produced. A good agreement between data and model results will support these assumptions.

Given the simple functional shape of Eq. (7.13), if an analytical approximation for the distribution of y exists, exact results can be obtained on the properties involving node strengths. The empirical cumulative distribution of y is well fitted ($R^2 \approx 0.998$) by a stretched exponential $\rho_>(y) = \exp\left(-(\frac{y}{\gamma})^\beta\right)$ (Figure 7.1). Then using the continuum approximation [Caldarelli et al., 2002], $\langle s(y) \rangle = N \int_0^\infty q(y,z)\rho(z)dz = N\langle y \rangle \eta y = NF(y)$ is obtained, and for large enough N one has: $p(s) = \rho[NF^{-1}(s/N)]\frac{d}{ds}F^{-1}(s/N) = \frac{1}{\eta}\rho(s/\eta)$, yielding:

$$P_>(s) = e^{-(\frac{s}{\eta\gamma})^\beta}. \tag{7.14}$$

Finally, the strength−strength correlation is obtained as:

$$\langle s_{nn}^W(y) \rangle = \frac{N \int_0^\infty q(y,z)\langle s(z) \rangle \rho(z)dz}{\langle s(y) \rangle} = N\eta\gamma^2\Gamma[1 + 2/\beta], \tag{7.15}$$

and it is found that it that does not depend on y. Although the same procedure cannot be
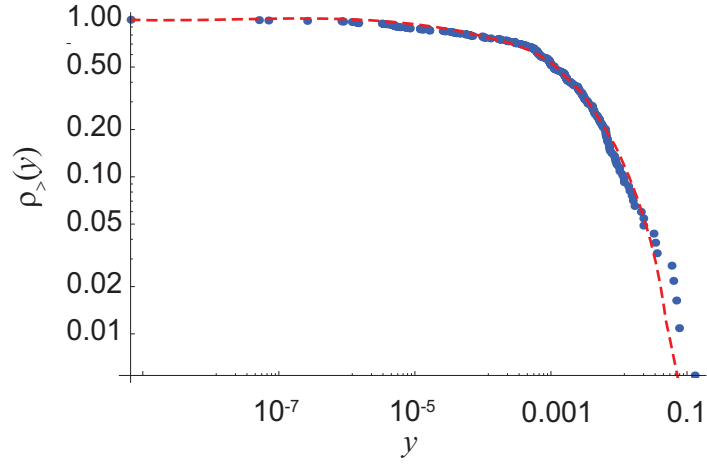
Figure 7.1: Best fit of the cumulative distribution function $\rho_>(y) = e^{-(\frac{y}{\gamma})^\beta}$ of the fitness variable y. The dots represent the empirical distribution constructed from the data $y_i = RAA_i/(\sum_j RAA_j)$, while the dashed line is $\rho_>(y)$. The best fit ($R^2 \approx 0.998$) gives $\beta \approx 0.482 \pm 0.005$ and $\gamma \approx 0.00236 \pm 0.00003$.
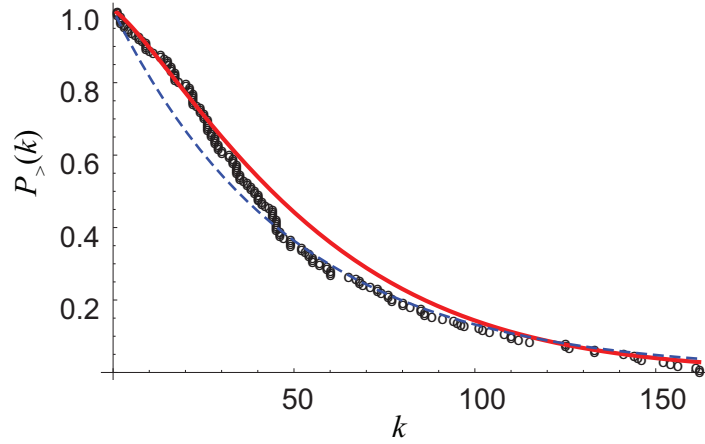
Figure 7.2: Comparison between empirical cumulative distribution function of the nodes degree (black dots), an exponential function $e^{-k/\langle k \rangle}$ (blue dashed line) used as null hypothesis $H0$ and the compressed stretched exponential $P_>(k)$ (null model H1) obtained from $P_>(s)$ using the derived approach distribution (in red). Computing the Kolmogorov Smirnov test, the null hypothesis $H0$ is rejected for any confidence interval, while $H1$ is accepted with $P - value = 0.05$.

repeated for the fitness variable x, a qualitative analytical behavior for the distribution of the node degree can be obtained by using the empirical relation $s = ak^b$ through a derived distribution approach, i.e. $p(k)dk = p(s)ds$. It is then obtained:

$$P_>(k) = e^{-(\frac{a}{\eta\gamma})^\beta k^{b\beta}},$$ (7.16)

which is a compressed exponential distribution, confirming the exponential−like tail observed

from the empirical analysis of the degree distribution (Figure 7.2).
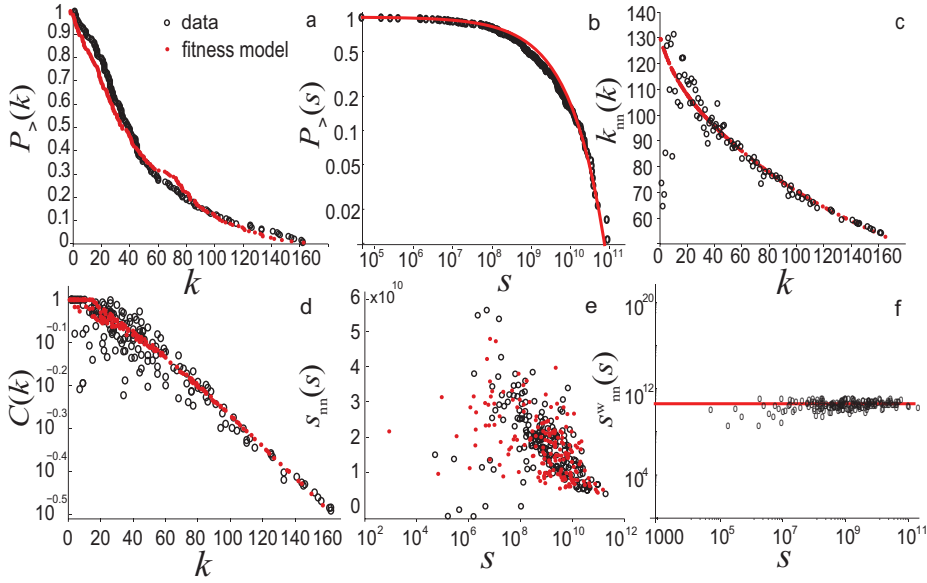
Figure 7.3: Topological and weighted properties of the GVWTN compared with the results of the fitness model (red line). a−b) Cumulative pdf of the node's degree $P_>(k)$ in linear scale and node strength $P_>(s)$ in log−log scale; c−d) average nearest neighbors degree k_{nn} and cluster coefficient C as a function of the nodes degree k; e−f) average nearest neighbors strength s_{nn} and strengths−strengths correlation $s_{nn}^W(s) \sim const$ in semilog x and log−log scale, respectively.

The parameters σ and η are determined through the normalization conditions (and not fitted):

$$\frac{1}{2}\sum_i\sum_{j\neq i} f(x_i; x_j) = L, \qquad \frac{1}{2}\sum_i\sum_{j\neq i} q(y_i, y_j) = \Phi, \tag{7.17}$$

where L and Φ are the total number of edges and the total flux in the network, respectively. $\sigma = 331892$ and $\eta = 1.306 \times 10^{12}$ are numerically obtained.

The agreement between Eqs. (7.10)−(7.12) and the topological data of the GVWTN, can be interpreted as the consequence of two independent factors, both necessary. First, the topology of the network is fully explained by its degree sequence: once it is fixed, all the topological network properties are established. Secondly, the degree of a nation in the GVWTN is determined by the hidden variable x. In particular, the normalized country's GDP ($x_i = GDP_i/(\sum_j GDP_j)$) is chosen. The good agreement between data and model (shown in the main text) confirms that the GDP plays indeed a major role in the connectivity of the nations in the GVWTN, as observed for the world trade web [Garlaschelli and Loffredo, 2004]. Analogously, by using the weighted fitness model a clear signature of the importance of rainfall on agricultural area (RAA) on the volume of virtual water traded between countries is found.

Figure 7.4: Comparison between empirical (black dots) and model results (red dots): a) The relationship between nodal degrees k and normalized GDP x shows how on average the number of connections is an increasing function of the nation's GDP; b) The relationship between node strength s and normalized RAA, y i.e. $s(y) = N\eta\langle y\rangle y$; c) The node independent relationship between strengths s^* and degrees k^* in the GVWTN ($s^* \sim k^{*q}$) implies a power−law correlation between GDP and RAA. The red line represents the best fit obtained from k and s $\langle s_i\rangle$ and $\langle k_i\rangle$, $i = 1, 2, ..., N$ generated by the fitness model. In this case $b = 2.69 \pm 0.03$ ($R^2 = 0.98$) is obtained. All the plots are in log−log scale.

7.3 Analysis and Modelling of the Complete Directed Virtual Water Network

The GVWTN can be also studied as a directed network, that is, a graph where the links incorporate the information on the direction of the traded virtual water. In brief, consider two nodes i and j. It is possible to have either only a link that goes from i to j (or from j to i), or two edges, one from i to j and one from j to i. If the network has only one link between each pair of nodes, then it is a pure unidirectional network. If instead it has always double edges between each pair of nodes, it is a pure bidirectional graph, that is, every country is both an exporter and an importer.

A usual way of quantifying where a real network lies between such extremes [Garlaschelli and Loffredo, 2006] lies in measuring its reciprocity coefficient r, defined as the ratio of the number of links pointing in both directions $\overleftrightarrow{L}$ to the total number of links $L = \overleftrightarrow{L} + \overrightarrow{L} + \overleftarrow{L}$

$$r = \frac{\overleftrightarrow{L}}{L}. \tag{7.18}$$

For the international virtual water trade network it is found $r = 0.49$.

In general the topological structure of a directed network is described by the adjacency matrix

$$a = \overleftrightarrow{a} + \overrightarrow{a} + \overleftarrow{a}, \tag{7.19}$$

where $\overleftrightarrow{a}_{ij} = 1$ means that there is a bidirectional link between node i and j. $\overleftrightarrow{a}$ is symmetric, while $\overrightarrow{a}_{ij}$ ($\overleftarrow{a}_{ij}$) is equal to 1 when there is a nonreciprocated edge from (to) i to (from) j.

83

Note that the possible configurations of a are $\{(0,0,0),(1,0,0),(0,1,0),(0,0,1)\}$, i.e. nonzero adjacency matrix elements are mutually excluding. The reciprocated and nonreciprocated degree of the network are defined as

$$\overleftrightarrow{k}_i \equiv \sum_j \vec{a}_{ji} \qquad \vec{k}_i \equiv \sum_j \vec{a}_{ij} \qquad \overleftarrow{k}_i \equiv \sum_j \overleftarrow{a}_{ij}. \tag{7.20}$$

In order to describe the non trivial reciprocity of the VWN, an existing general framework [Garlaschelli and Loffredo, 2006] is used. Reciprocated and nonreciprocated links are considered as different "species", each governed by the corresponding parameter. In particular, for each pair of vertices i, j (considered only once, i.e., $i < j$) a nonreciprocated link from i to j (j to i), $\rightarrow$ ($\leftarrow$), is regarded as a link, whose total number is controlled by the parameter $\vec{\sigma} = e^{\vec{\mu}}$ ($\overleftarrow{\sigma} = e^{\overleftarrow{\mu}}$), while two mutual links between i and j as a single edge of type $\leftrightarrow$ which total number depends on $\overleftrightarrow{\sigma} = e^{\overleftrightarrow{\mu}}$. The Hamiltonian defined in Eq.(7.4) can be thus be generalized to:

$$H_r = \sum_{i<j} (\vec{\epsilon}_{ij}\vec{a}_{ij} + \overleftarrow{\epsilon}_{ij}\overleftarrow{a}_{ij} + \overleftrightarrow{\epsilon}_{ij}\overleftrightarrow{a}_{ij}). \tag{7.21}$$

The generalization of the partition function Z and of the free energy Ω easily follow from Eqs. (7.4)−(7.5) using Eq. (7.21).

The reciprocity of the network can be calculated from the model from Eq. (7.18) as:

$$\vec{L} = -\frac{\partial \vec{\Omega}_{ij}}{\partial \vec{\mu}}; \qquad \overleftarrow{L} = -\frac{\partial \overleftarrow{\Omega}_{ij}}{\partial \overleftarrow{\mu}}; \qquad \overleftrightarrow{L} = -2\frac{\partial \overleftrightarrow{\Omega}_{ij}}{\partial \overleftrightarrow{\mu}}; \tag{7.22}$$

Note that because ($\rightarrow$) and ($\leftarrow$) lead to the same species which has been split in order to consider each pair of vertices only once, $\vec{Q}_{ij} = \overleftarrow{Q}_{ji}$ and $\overleftrightarrow{Q}_{ij} = \overleftrightarrow{Q}_{ji}$ for $i > j$ is set, where Q is a given property of the network [Garlaschelli and Loffredo, 2006].

In order to constrain the degree sequence $\{\overleftrightarrow{k}_l, \overleftarrow{k}_l, \vec{k}_l\}_{l=1,2,...}$ $\overleftarrow{\epsilon}_{ij} = \alpha_i + \beta_j$ and $\overleftrightarrow{\epsilon}_{ij} = \gamma_i + \gamma_j$ is set. Finally, the fitness variables x, y and z are introduced. x_i, y_i are the values determining the outgoing and incoming non−reciprocated links from (and to) node i, while z_i drives the reciprocated ones. To do so

$$e^{-(\alpha_i+\beta_j)} = x_i \cdot y_j \qquad e^{-(\gamma_i+\gamma_j)} = -z_i \cdot z_j, \tag{7.23}$$

is set, from which the following expression is achieved

$$\vec{p}_{ij} = \vec{p}(x_i, x_j) = \frac{\vec{\sigma} x_i y_j}{\mathcal{Z}_{ij}} \qquad \overleftrightarrow{p}_{ij} = \overleftrightarrow{p}(z_i, z_j) = \frac{\overleftrightarrow{\sigma} z_i z_j}{\mathcal{Z}_{ij}}, \tag{7.24}$$

where $\mathcal{Z}_{ij} = 1 + x_i y_j \overrightarrow{\sigma} + x_j y_i \overleftarrow{\sigma} + z_i z_j \overleftrightarrow{\sigma}$ is the normalization factor. Eventually from Eq. (7.24) the expected degrees of the directed network is computed:

$$\langle \overrightarrow{k}_i \rangle = \sum_{j \neq i} \overleftrightarrow{p}(z_i, z_j) \qquad \langle \overrightarrow{k}_i \rangle = \sum_{j \neq i} \overrightarrow{p}(x_i, y_j) \qquad \langle \overleftarrow{k}_i \rangle = \sum_{j \neq i} \overrightarrow{p}(x_j, y_i), \tag{7.25}$$

and using Eq. (7.25) the distribution of the expected degrees is built.

$z_i = GDP_i / (\sum_j GDP_j)$ is chosen to be the fitness variable for the reciprocated part of the VWN, as in the undirected case (the pure bidirectional part is equivalent to an undirected network). On the other hand, the non−reciprocated edges are determined by the normalized GDP and the normalized average yearly rainfall ($\bar{R}$) of the countries, i.e. $x = GDP_i / (\sum_j GDP_j)$ and $y = \bar{R}_i / (\sum_j \bar{R}_j)$. Again, the predicted results are in good agreement with the empirical results obtained from the virtual water directed network (see Fig. 2 a−c).

Similarly, the complete directed weights matrix is defined as:

$$W = \overleftrightarrow{W} + \overrightarrow{W} + \overleftarrow{W}, \tag{7.26}$$

describing the flux of the GVWTN through the corresponding reciprocated or non−reciprocated links. Accordingly, the definitions of the node strength are generalized as:

$$\overrightarrow{s}_i \equiv \sum_j \overrightarrow{W}_{ij} \qquad \overleftarrow{s}_i \equiv \sum_j \overleftarrow{W}_{ij} \qquad \overleftrightarrow{s}_i \equiv \sum_j \overleftrightarrow{W}_{ij} \tag{7.27}$$

The weighted fitness model for the complete network follows easily from Eq. (7.13). Assuming that the fitness variable controlling the flux on the reciprocated edges is t, while those determining the flux on the non−reciprocated links are u and v:

$$\overrightarrow{q}(u_i, v_j) = \overrightarrow{\eta} \, u_i v_j \qquad \overleftrightarrow{q}(t_i, t_j) = \overleftrightarrow{\eta} \, t_i t_j, \tag{7.28}$$

where $\overleftrightarrow{\eta}$ and $\overrightarrow{\eta}$ are the parameters controlling the total flux of the reciprocated and non−reciprocated network, respectively. Interestingly, the node strengths, both reciprocated and non−reciprocated and in and out (i.e. $\overleftarrow{s}, \overrightarrow{s}$ and $\overleftrightarrow{s}_{in}, \overleftrightarrow{s}_{out}$) are all well described by using only mean annual rainfall on agricultural area as fitness variable, i.e. $u_i = v_i = t_i = RAA_i / \sum_j RAA_j$.

The external data sets on GDP, $\bar{R}$ and RAA are sufficient to reproduce the global topological and weighted properties of both the reciprocated and non−reciprocated network described by the degree and strengths pdfs, respectively. The parameters $\overleftrightarrow{\sigma}, \overrightarrow{\sigma}, \overleftrightarrow{\eta}$ and $\overrightarrow{\eta}$ are evaluated by fixing respectively the number of links ($\overrightarrow{L}, \overleftrightarrow{L}$) and the total fluxes ($\overrightarrow{\Phi}, \overleftrightarrow{\Phi}$) of the GVWTN. All the results are summarized in Figure 7.5 corresponds to this single choice of parameters.
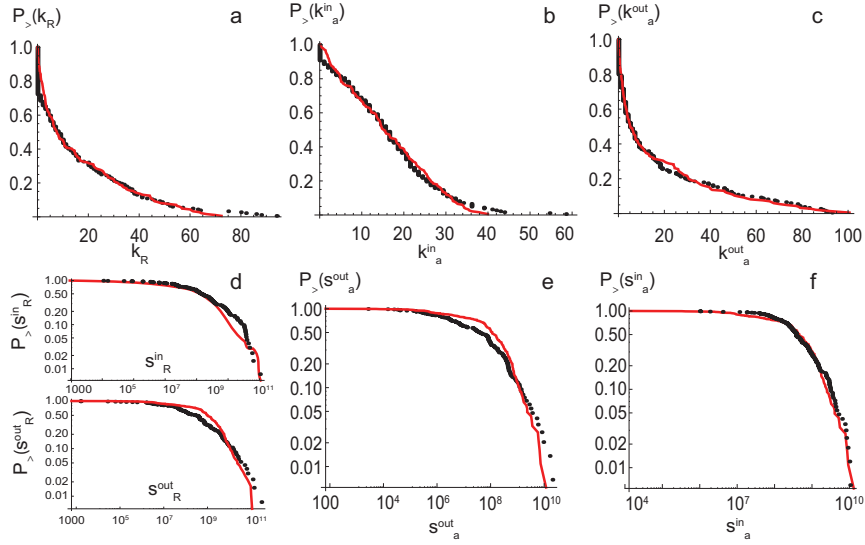
Figure 7.5: Comparison between empirical analysis (black points) and fitness model (in red) of a) reciprocal node degree (k_R) cumulative distribution; b−c)cumulative distribution of nonreciprocated (k_a) in and out degrees, respectively; d) Reciprocated in and out nodes strength (s_R); e−f) cumulative distribution of nonreciprocated (s_a) in and out nodes strength. Plots (a)−(b)−(c) are in natural scale, while plots (d)−(e)−(f) are in log−log scale.

7.4 Future Scenarios

The theoretical framework is suitable to investigate future scenarios of the GVWTN structure. To this aim, the annual rainfall for 2030−2050 from the A2 socio−economic scenario of the World Climate Research Programmes (WCRPs) Coupled Model Intercomparison Project Phase 3 (CMIP3) multi−model dataset [Meehl et al., 2007] is evaluated. 11 different climate change projections utilizing data from the CMIP3 archive[Meehl et al., 2007] are computed for the case of the most extremes scenarios (IPCC A2 scenarios) and in a $2^o \times 2^o$ spatial scale. The results presented here are obtained from the GISS−E−R scenario. It should be noted that the results for the other scenarios show the same general trends. The spatial mean is then calculated for each country in the network over this time horizon.

Then by using published projections of the GDP and agricultural area [FAO, 2000b; Fonseca et al., 2009] for 2030, the fitness functions $p(x_i^T, x_j^T) = \sigma' x_i^T x_j^T / (1 + \sigma' x_i^T x_j^T)$ and $q(y_i^T, y_j^T) = \eta' y_i^T y_j^T$ are built, where x^T and y^T are the projections of the fitness variables at year $T = 2030$. The parameters σ' and η' are to be determined by the future total number of connections L' and flux Φ'. In the simulation is assumed that $L' = L$ and $\Phi' = \Phi$, but in general they may be part of the scenarios under study. All A2 climate change scenarios [Meehl et al., 2007] yield a decrease in rainfall at a global scale, but the total arable land is predicted by [FAO, 2000b] to increase around 1%, thereby leading to an increase of the total RAA. Figure 7.6 summarizes the results of the structure of the GVWTN under the driest climate change scenario. The structure

of the GVWTN topology is robust with respect to these particular scenarios. A heavier tail in the strengths pdf is observed suggesting a rich−gets−richer phenomenon [Newman et al., 2006], where the nodes with large strengths benefit from the changes in RAA, becoming even stronger. The node independent relation $s^* \sim k^{*q}$, (where the vectors k^* and s^* are the sorted degrees and strengths in the GVWTN) is also observed to increases from $q = 2.69 \pm 0.03$ ($R^2 = 0.982$) to $q = 2.77 \pm 0.02$ ($R^2 = 0.986$) (see Figure 7.6). The disassortative behavior of the nearest neighbors average strength s_{nn} in the future scenario does not change. Through Eq. (7.15), it can be shown that s_{nn}^W increases of about 2%. Note that these results are valid for future total number of connections L' and total flux Φ' that are equal to the present values, i.e. $L' = 4550$ and $\Phi' = 6.24 \cdot 10^{11}$. For this case, $\sigma = 285505$ and $\eta = 1.3174 \times 10^{12}$ are obtained. Other constraints would yield different scenarios.

These results suggest that economic and climatic future scenarios will likely enhance the globalization of water resources, giving to water−rich countries even more inroad for reaching poorly connected nodes. At the same time, the observed rich−gets−richer phenomenon will intensify the reliance of most of the nations on the few VW hubs. As a consequence it will reduce the ability of the GVWTN to respond to disturbances whose impact may be dramatic when the VW trade supports carrying capacities beyond those supported by local resources [D'Odorico et al., 2010b]. Finally this study highlights how agricultural land management may indeed remarkably impact the future structure of the GVWTN.
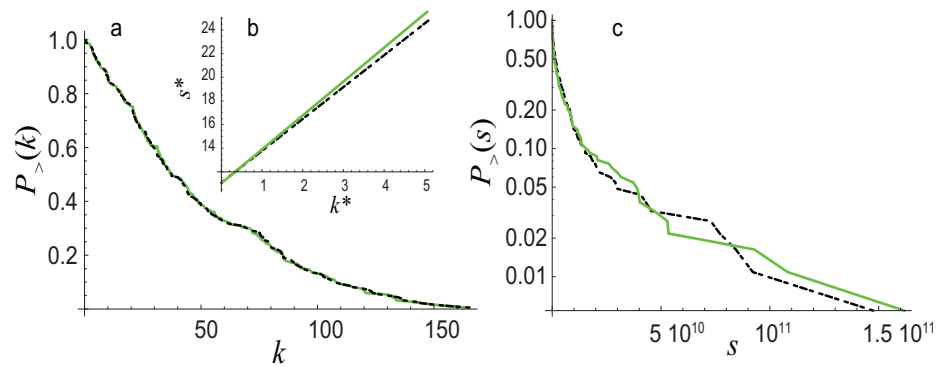
Figure 7.6: An example of predictive application of the fitness model of GVWTN for the driest case scenario: comparison between the global properties of the network for the year 2000 (dashed black) with those predicted by the fitness model for the year 2030 (green line): a) Cumulative degree pdf. Inset b): Relationships between sorted strengths and degrees); c) Cumulative pdf of the nodal strengths.

This work opens new quantitative and predictive perspectives in the study of stability and complexity of the GVWTN coupled to social, economic and political processes related to the international food trade. Ongoing research incorporates scenarios where L' and Φ' are different from values of the year 2000 and reflect the evolving and dynamic character of the global network.

8

Summary and Conclusions

The important thing is not to stop questioning. Curiosity has its own reason for existing. One cannot help but be in awe when he contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend a little of this mystery every day. Never lose a holy curiosity.

Albert Einstein

This thesis provided a set of coordinated attempts using methods of statistical mechanics to effectively study ecohydrological footprints. In particular, this work is focused on: the effect of stochasticity of green and blue water fluxes on water controlled ecosystems (Part I); the imprinting of the spatial structure of the environmental matrix on macroecological measures of ecosystems dynamics (Part II); water footprints of human societies through the analysis of the global virtual water trade network (Part III).

In Chapter 2 the impact of stochastic fluctuations in the storage-discharge relation defining the hydrologic response has been investigated. It is shown through numerical simulation that a physically meaningful, colored noise in the storage-discharge relation influences appreciably the discharge distributions. In particular, by assuming that noise in the storage-discharge relation may be surrogated by considering stochasticity directly in the streamflow generation processes, the probability distribution of streamflows, whose properties reflect the results obtained by numerical simulation, has been analytically derived. The theoretical framework couples a stochastic description of soil moisture dynamics with a transport model that embeds the variability of the streamflow generation process through a multiplicative Gaussian noise.

Its effect on the streamflow distribution is significant. In particular, the tail of $p(Q)$ changes from exponential to a power-law type with heavy tail. As the noise increases, the probability of observing low values of streamflow also increases and the mode of $p(Q)$ shifts towards zero. Higher probabilities for low Q are balanced by an increased probability of high discharge, i.e. the power-law tail becomes fatter. Therefore the presence of the noise does not change the mean of Q, but significantly increases its variance. Above a threshold in the noise magnitude, a shift from wet to dry regime occurs, implying a major ecological impact owing to a change from perennial to ephemeral streamflows. Thus neglecting these additional environmental fluctuations may produce underestimations of the variability of streamflows, with relevant ecohydrological consequences.

In Chapter 3, by means of the same mathematical approach, the problem of primary soil salinization has been studied. In particular, an analytical model for soil salinization has been presented, where the complexity of the problem is reduced by employing simplifying assumptions that permit us to describe high-dimensional components via random terms. By assuming time-averaged inputs of salt and instantaneous percolation processes, a decoupling from the soil moisture equation results in a simplified stochastic mass balance equation for the soil salt mass amenable to exact solution. Soil salinity statistics are obtained as a function of climate, soil and vegetation parameters. These can be combined with soil moisture statistics to obtain a full characterization of soil salt concentrations and the ensuing risk of primary salinization.

In Part II, theoretical and empirical evidences supporting the broad validity of a new macroecological pattern, namely the species persistence-time distributions, have been presented. Specifically, in Chapter 4 the neutral framework that allow us to model the relevant ecosystem dynamics and to obtain an analytical shape for the species persistence-time distributions, has been presented. A related length−bias sampling problem has also been proposed and solved. In Chapter 5 four rather diverse ecosystems (hosting respectively breeding birds, herbaceous plants, forest plant and marine fishes) have been statistically analyzed. In all cases the observed SPT distributions display a power-law shape with a cut-off due to the finiteness of the observational time window, which is reproduced by the theoretical model. These results confirm the expectation that power laws, observed for SPT distributions, are the result of emergent behaviors independent of fine details of the system dynamics.The specific values of the scaling exponents of the SPT distributions have been shown to depend only on a few key factors, namely the spatial dimension of the embedding ecosystem matrix and the nature of dispersal limitation. Because of the assessed, robust scale invariance character of the SPT distribution (limited only by biogeographical finite-size effects), and because of its relation with other macro-ecological patterns, the SPT distribution have been proposed as a powerful synthetic descriptor of ecosystem biodiversity and of its associated dynamics.

In Part III complex network theory has been utilized to analyze the structure of the global virtual water trade network associated with the international food exchanges. In Chapter 6, the topological and weighted properties of the undirected representation of this network have

been quantitatively characterized using statistical tools from complex network theory. The network is shown to be connected small-world-like, with compressed exponential degree distribution and displays disassortative behavior for correlation and clustering coefficients. Differently, node strengths follow heavy-tail stretched exponential distribution and a power law relation with non-trivial exponent holds between node degrees and strengths. All these properties suggest that few dominant nations connect otherwise disconnected portions of the network and form tight clusters with each other. The majority of the nations reside on the periphery of the trade network and participate in relatively small volume trade. In other words, the high degree of the globalization of water trade has been proven and evidences for the existence of the weighted rich-club phenomenon, i.e. a small group of nations that play a key role in the connectivity of the network and in the global redistribution of virtual water, has been provided. In Chapter 7 all the above key features of the network have been shown to be well described by a stochastic model that assumes as sole controls each country's gross domestic product and yearly rainfall on agricultural areas. The model has been presented in details and examples of prediction of the structure of the network under future political, economic and climatic scenarios have been illustrated. These scenarios suggest that the crucial importance of the countries that trade large volumes of water will be strengthened.

The results mentioned above suggest that a probabilistic approach to model ecohydrological systems and their footprints may be very effective in producing analytically manageable results. Exact results are important to lay down general principles to provide quantitative frameworks which, in turn, facilitate progress in our understanding of complex problems. Minimalist models, as simple as possible but not simpler – dubbing the famous say – address the simplest scheme of a given universality class of behaviors, and as such have received much attention is many fields. The present work attempts to handle ecohydrological complexity using minimalist– probabilistic approaches in a consolidated procedure derived from statistical mechanics, the science that translates such precepts into mathematical language. Several studies based on this approach have made, in fact, quite interesting progresses in the domain of ecohydrology, jointly with other fields of environmental science. Further developments lie ahead in several related areas. With regards to the topics presented in this thesis, the results that are planned for future research deal with, in particular:

- Comparing the theoretical probability distribution of streamflows with data from real basins, and finding empirical and theoretical methods in order to estimate the intensity of stochastic fluctuations in storage-discharge relations;

- Adding in the salinity model the effect of groundwater upflow and irrigation. In fact these two processes seems to be the main responsible of salinization in dry regions. Moreover, the problem of sodic soils could also be incorporated in the presented stochastic theoretical framework. If the total dissolved salts are not in a critical concentration in the soil, but Na levels are high or not balanced with the Ca and Mg, soil tilth can also be effected (alkali salinization or sodicity problem). In these conditions, the positively charged clay particles in the soil attach the negatively charged clay particles in the

soil, causing the soil to be sticky when wet, and hard and impermeable when dry with dramatic consequences on the vegetation;

- Incorporating an evolutionary framework in the neutral theory of biodiversity. Light will be thrown on how species and their fitness landscapes dynamically coevolve, through the study of evolutionary trajectories that causally emerge using a variational principle proposed in a spin-model-like setup of evolutionary systems. Game theory can also be used to estimate the effects of selection versus fluctuations in ecosystems dynamics;

- Developing a modeling framework and data analysis to study the effect of the trade of food products on the resilience of human societies with respect to drought and famine. Using existing data sets of water use, food consumption, population growth, and virtual water trade, this research will assess if and how virtual water is related to population growth in water poor country and how societal resilience varies with an increase in the interconnectedness of the global network of virtual water trade.

Appendix A

Itô–Stratonovich Dilemma in Generalized Langevin Equations

Everything should be made as simple as possible, but not simpler.

A. Einstien

A.1 Introduction

While dealing with the mathematical treatments of ecohydrologic footprints on soil salinity, it becomes necessary to face a particular class of mathematical problems generically referred to stochastic differential equations with multiplicative noise. Some mathematical methods, somewhat technically involved, have been developed to that end. Such results have been published independently in a technical journal of statistical physics. It has seen therefore deemed appropriate to include them in the thesis as an appendix: because they were originally developed as a part of the doctoral work, and because they can be seen as a methodological section of general nature.

In this section the reader is thus first briefly introduced to stochastic differential equations (SDE). As main references to this topics the books by Gardiner [2004] and Van Kampen [2007] are suggested.

In general a SDE is a differential equation in which one or more of the terms are a stochastic process. Therefore, the solution of a SDE is not a deterministic function or trajectory, but a stochastic process itself. SDEs can be used to model a large diverse class of phenomena ranging from fluctuating stock prices [Mantegna and Stanley, 1999], to biology [Azaele et al.,

2009], to ecology [McKane et al., 2000], and to several physical systems [Gardiner, 2004].

The most widespread type of noise used to model a wide class of random perturbations is the (continuous−time) Gaussian white noise (GWN), which can be thought as the derivative of Brownian motion (or the Wiener process) [Gardiner, 2004]. Another stochastic process that is well studied in the literature is the so called Ornstein−Uhlenbeck process [Uhlenbeck and Ornstein, 1930]. It is a continuous−time noise characterized by a non trivial correlation structure and a finite variance. It belongs to the class of noise known as Gaussian colored noise (GCN, in contrast to GWN, that lacks in correlation) and it describes the velocity of a massive Brownian particle under the influence of friction [Van Kampen, 2007]. SDEs characterized by these type of noises (both GWN and GCN) are also known as Langevin equations (LE).

Nevertheless, environmental variability is also often due to discrete−time random events. For instance rainfall, in first approximation, can be modeled as a compound Poisson process, i.e. a random sequence of discrete events (precipitations), characterized by random amplitude (rainfall depth). Intense and concentrated (state−dependent) forcing events may often be modeled as (multiplicative random) jumps, taking place according to an underlying point process (e.g. Poisson). In this case, the SDEs are termed generalized Langevin equations (GLE).

In a very general settings it can be assumed that, from a mathematic point of view, the evolution equation of a given system dynamics is governed by the deterministic differential equation

$$\dot{x}(t) = a(x, t). \tag{A.1}$$

If, in addition, the system is subjected to random perturbations, it is natural to use a Langevin kind of model in which a noise $\zeta(t)$ is added to the deterministic equation. The resulting SDE for white noise $\zeta(t)$, which can be either Gaussian or non Gaussian, can be written as

$$\dot{x}(t) = a(x, t) + b(x)\zeta(t), \tag{A.2}$$

where $b(x) = 1$ if the perturbations do not depend on the state of the system (additive noise).

The solution of Eq. (A.2) is *not* a given deterministic trajectory $x(t)$, but rather a probability distribution function $P(x, t)$ that provide the probability of finding the system in the state x at time t. In order to find this distribution the SDE (A.2) must to be transformed into a deterministic partial differential equation for $P(x, t)$:

$$\frac{\partial P(x, t)}{\partial t} = \mathscr{L}(x, \partial_x, \partial_x^2, \int_x) P(x, t), \tag{A.3}$$

where $\mathscr{L}(x, \partial_x, \partial_x^2, \int_x)$ is an appropriate intregro−differential operator that depends on Eq. (A.2). If x is a continuous variable, Eq. (A.3) is known as Fokker−Plank Equation (FPE). Sometimes it is more convenient to work with discrete state variables, in this case Eq. (A.3) is

called Master Equation (ME).

For instance, if in Eq. (A.2) $b(x) = 1$ and $\zeta(t)$ is a GWN, then the corresponding FPE (A.3) reads as

$$\frac{\partial}{\partial t} P(x,t) = -\frac{\partial}{\partial x}\big[a(x,t)P(x,t) \big] + D \frac{\partial^2}{\partial x^2} P(x,t). \tag{A.4}$$

Instead, if in Eq. (A.2) $b(x) = 1$, but $\zeta(t)$ is a compound Poisson process ($\zeta(t) = \sum_{i+1}^{N} z_i \delta(t - t_i)$, with $z_i \sim \rho(z)$ random variables drawn from a given pdf $\rho(z)$, and δ is the Dirac delta), then the corresponding ME (A.3) is

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\big[a(x,t)P(x,t) \big] + \nu \int_0^x \rho\,(x-z)\,P(z,t)dz - \nu\,P(x,t). \tag{A.5}$$

For generalization and mathematical details on how to reach Eqs. (A.4) and (A.5) from the SDE (A.2), see Gardiner [2004].

A.2   Analytical Solution of the GLE with Multiplicative Gaussian Noise and Additive Poisson Noise

In this section the derivation of the analytical solution of the generalized Langevin equation

$$\frac{dQ(t)}{dt} = -k\,Q(t) + \zeta(t)\,Q(t) + k\,A\,\xi_t(\lambda; \gamma_P) \tag{A.6}$$

that we used in Chapter 2 to describe the dynamics of the streamflows $Q(t)$ affected by intrinsic stochasticity $\zeta(t)$, and external noise (i.e. rainfall infiltration) ξ_t. In there, $\zeta(t)$ is a Gaussian white noise with mean $\langle\zeta(t)\rangle = 0$ and correlation $\langle\zeta(t)\zeta(s)\rangle = 2\sigma^2\delta(t - s)$; and

$$\xi_t(\lambda; \gamma_P) = \frac{1}{k\,A} \sum_{i=1}^{N(t)} \Delta Q_i\,\delta(t - t_i), \tag{A.7}$$

is a compound Poisson process [Snyder, 1975], where $\{N(t), t \geq 0\}$ is an homogeneous Poisson counting process of rate λ, and $\{Q_i\}$ is a sequence of mutually independent and also independent of $N(t)$, identically distributed random variables with a probability density function $b(Q)$. In this particular case $b(Q)$ is an exponential distribution with parameter $\gamma_Q = \gamma_P/A\,k$. From a mathematical point of view equation (A.6), as it stands, is meaningless [Gardiner, 2004]. In fact, according to equation (A.6), each pulse in $\zeta(t)$ gives rise to a pulse in $\dot{Q}$ and hence a jump in Q. Thus, the value of Q in the right hand side of equation (A.6) is undetermined. The

incremental form of equation (A.6) is:

$$dQ(t) = -kQ(t)\,dt + dW(t)\,Q(t) + kA\xi_t(\lambda;\gamma_P)dt, \tag{A.8}$$

where $W(t) = \int_0^t \zeta(t')dt'$ is the well known Wiener process. The solution of (A.8) can be written in the integral form:

$$Q(t) = Q(t_0) - k\int_{t_0}^t Q(t')\,dt' - \int_{t_0}^t dW(t')\,Q(t') + \int_{t_0}^t \xi_{t'}(\lambda;\gamma)dt'. \tag{A.9}$$

The second integral in the right hand side of equation (A.9) is a stochastic integral and it is defined as the limit of the partial sum [Gardiner, 2004]

$$S_n = \sum_{i=1}^n Q(\tau_i)[W(t_i) - W(t_{i-1})], \tag{A.10}$$

where $t_0 \leq t_1 \leq ... \leq t_{n-1} \leq t$ and $t_{i-1} \leq \tau_i \leq t_i$. It can be shown that S_n depends on the particular choice of τ_i [Gardiner, 2004] and thus its value depends on the particular interpretation chosen for (A.10). One of the most famous interpretations is that of Itô [Ito, 1951]:

$$S_n(It\hat{o}) = \sum_{i=1}^n Q(\tau_i = t_{i-1})[W(t_i) - W(t_{i-1})]. \tag{A.11}$$

Different interpretation lead to different results and so the correct and coherent interpretation of equations (A.6),(A.8), (A.9) is crucial both for analytical and computational calculations. In particular let us emphasize the fact that if equation (A.6) is interpreted in the Itô sense, then $\langle Q(t)\zeta(t)\rangle_{Ito} = \langle Q(t)\rangle\langle\zeta(t)\rangle = 0$. The former property straightforwardly derives from the observation that in the Itô interpretation $Q(t)$ at time t does not depend on the noise $\zeta(t)$ at the same time.

Accordingly, equation (A.6) corresponds to the forward Chapman-Kolmogorov equation [Gardiner, 2004]

$$\frac{\partial}{\partial Q}p(Q,t) = k\frac{\partial}{\partial Q}\left[Q\cdot p(Q,t)\right] + \sigma^2\frac{\partial^2}{\partial Q^2}\left[Q^2 p(Q,t)\right] - \lambda p(Q,t) + \lambda\int_0^Q \gamma_Q e^{-\gamma_Q(Q-z)}p(z,t)dz,$$

$$(A.12)$$

which in turn can be written in a form of a continuity equation

$$\frac{\partial}{\partial Q} p(Q,t) = -\frac{\partial}{\partial Q} J(Q,t),$$

(A.13)

where the probability current $J(Q,t)$ is [Daly and Porporato, 2006]:

$$J(Q,t) = -kp(Q,t) - \sigma^2 \frac{\partial}{\partial Q}\left[Q^2 p(Q,t)\right] + \lambda \int_0^Q e^{-\gamma_Q(Q-z)} p(z)\,dz.$$

(A.14)

Imposing natural boundary the probability current vanishes in steady-state conditions and so have that the steady-state solution $p(Q,t \to \infty)$ satisfies $J(Q) = 0 \quad \forall Q$. Multiplying both sides of the former equation by $e^{Q\gamma_Q}$ and differentiating with respect to Q, one gets the steady state equation:

$$Q^2 p''(Q) + \left(\left(\frac{k}{\sigma^2}+4\right)Q + \gamma_Q Q^2\right)p'(Q) + \left(\left(\frac{k\gamma_Q}{\sigma^2}+2\gamma_Q\right)Q + \left(2+\frac{k}{\sigma^2}-\frac{\lambda}{\sigma^2}\right)\right)p(Q) = 0. \quad (A.15)$$

The solution of equation (A.15) is [Polyanin and Zaitsev, 2003]:

$$p(Q) = e^{-Q\gamma_Q} Q^{-\alpha+\frac{\beta}{2}-2}\left(C_1 U\left[\frac{\beta}{2}-\alpha, \beta+1, Q\gamma_Q\right] + C_2 L^{\beta}_{\alpha-\frac{\beta}{2}}[Q\gamma_Q]\right),$$

(A.16)

where $\alpha = \frac{k}{2\sigma^2} - \frac{1}{2}$, $\beta = \frac{\sqrt{4\lambda\sigma^2+(\sigma^2+k)^2}}{\sigma^2}$, C_1, C_2 are integration constants and $U[a,b,z]$ and $L^{\nu}_n[z]$ are the confluent hypergeometric function and the generalized Laguerre polynomial respectively [Abramowitz and I.A., 1965]. For $C_1 \neq 0$ all moments of $p(Q)$ are infinite, then for physical reasons $C_1 = 0$ is taken. Thus the appropriate solution of (A.6) is:

$$p(Q) = C_2 e^{-Q\gamma_Q} Q^{-\alpha+\frac{\beta}{2}-2} L^{\beta}_{\alpha-\frac{\beta}{2}}[Q\gamma_Q],$$

(A.17)

where C_2 can be determined by the normalization condition, $\int_0^\infty p(Q)\,dQ = 1$. Finally notice

that in the limit $\sigma^2 \to 0$ the Chapman-Kolmogorov equation (A.12) becomes:

$$\frac{\partial}{\partial Q} p(Q, t) = k \frac{\partial}{\partial Q} \left[Q \cdot p(Q, t) \right] - \lambda p(Q, t) + \lambda \int_0^Q \gamma_Q e^{-\gamma_Q (Q-z)} p(z, t) dz, \tag{A.18}$$

that correspond exactly to the equation for the streamflow pdf previously derived [see Botter et al., 2007c] and with state solution $p(Q) \sim Q^{(\frac{\lambda}{k} - 1)} \exp(-\gamma_Q Q)$.

A.3 State−dependent discrete Poisson jumps processes

Unlike the cases of continuous−time (white or colored) Gaussian noise, which are described in all basic texts of stochastic processes [e.g. Gardiner, 2004; Van Kampen, 2007], state−dependent discrete jumps have been less investigated [Van Den Broeck, 1983; Denisov et al., 2009]. For instance, in general Eq. A.2 is ill−defined unless a prescription for the evaluation of the stochastic term $b(x)\zeta(t)$ is specified [Hanggi and Thomas, 1982]. While this issue is well understood for GWN [Kampen, 1981], a precise characterization of the noise prescriptions and a clear connection between the different interpretations are still missing for other kind of noises. The generalized Langevin equation (GLE) for white multiplicative noise $\zeta(t)$, which can be either Gaussian or non Gaussian,

$$\dot{x}(t) = a(x, t) + b(x)\zeta(t), \tag{A.19}$$

is ill-defined unless a prescription for the evaluation of the stochastic term $b(x)\zeta(t)$ is specified [Hanggi and Thomas, 1982].

In this appendix it is shown how different prescriptions corresponding to the Itô (I) and Stratonovich (S) interpretation of a stochastic differential equation (SDE) arise naturally for multiplicative jumps, depending on the relevant time scales of the process. The Master Equation (ME) for a GLE with multiplicative compound Poisson process in both the I and S prescriptions will be also presented. Moreover, it will be shown how, in the linear case $b(x) \propto x$, the difference between prescriptions is properly interpreted as a transformation of the jump size PDFs.

A.3.1 Connection between different prescriptions of a GLE and time scales of the process

Let us begin with a pedagogical example of a particle that experiences multiplicative impulsive forcing events, proportional to $\dot{\Theta}_\tau(t)$, of duration τ, in a field characterized by a friction coefficient ψ. The following analysis is inspired by the work in references Graham and Schenzle [1982]; Kupferman et al. [2004]. $\Theta_\tau(t) = \vartheta(t/\tau)$ with $\vartheta(z) \to 1 \, (0)$ in the limit $z \to \infty \, (-\infty)$ is chosen, so that $\dot{\Theta}_\tau(t) \to \delta(t)$ in the $\tau \to 0$ limit (in the distribution sense). Let us first consider

the case of a single jump event at $t = t_0 > 0$, where the dynamics is described by the Newton equation

$$m\,\ddot{x}(t) = -\psi\dot{x} + \psi b(x)w\dot{\Theta}_\tau(t - t_0),\tag{A.20}$$

where the random jump w is drawn from the jump size pdf $\rho(w)$. Thus in Eq. (A.20) we have two time scales, $\sigma = m/\psi$ and τ. The former is associated with the relaxation time toward stationarity, while the latter is related to the characteristic duration of the impulsive forcing. Different prescriptions of Eq. (A.19) arise depending on how the two emerging timescales σ and τ in Eq. (A.20) go to zero, i.e. $\sigma \to 0$ followed by $\tau \to 0$ or viceversa (see Figure A.1). For this reason, writing $\dot{x}(t) = b(x)w\delta(t - t_0)$ is ambiguous, being the result of two different limit procedures with different physical and mathematical meaning.

When $\sigma \ll \tau$ and then the zero limit of τ is taken in Eq. (A.20), the S prescription of the SDE (A.19), which preserves the usual rules of calculus, is obtained [Stratonovich, 1963; Van Kampen, 2007]. For example if $b(x) = x$, the resulting S-equation $d\ln(x)/dt = w\dot{\Theta}_\tau(t - t_0)$,
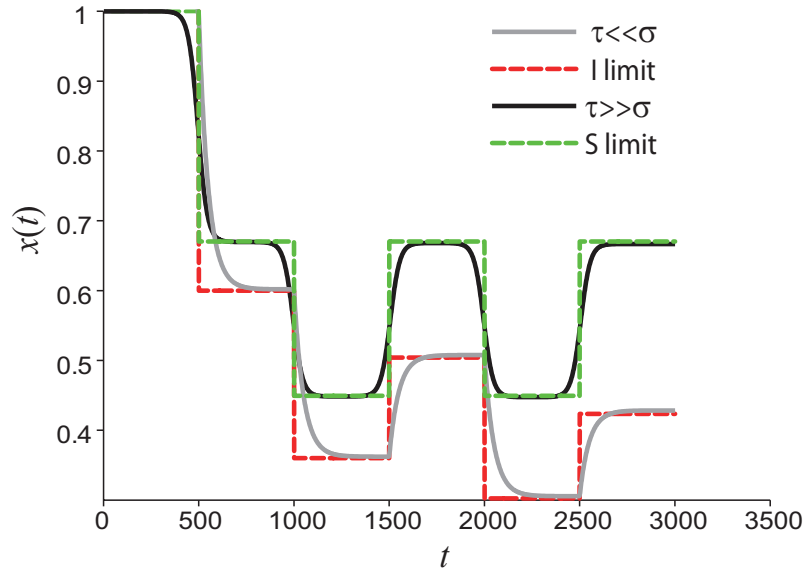
Figure A.1: Comparison between trajectories of a particle that undergoes impulsive multiplicative forcing in a viscosity field for different timescales (τ and σ), and the trajectories that result from the SDE $\dot{x}(t) = -x(t)\sum_{i=1}^{5} w_i \delta(t - t_i)$ interpreted in the I and S prescriptions. The jumps in this case are given by $w_i = \pm 0.4$.

after performing the limit $\tau \to 0$, has formal solution $x(t) = (1 + \Theta(t - t_0)(e^w - 1))\,x_0$, where $x_0 = x(0)$ and Θ is the Heaviside function. The corresponding pdf is

$$p^S(x, t) = \delta(x - x_0)(1 - \Theta(t - t_0)) + \Theta(t - t_0)\frac{\rho\left(\ln(\frac{x}{x_0})\right)}{x}\Theta(\frac{x}{x_0}),\tag{A.21}$$

with initial condition $\rho^S(x, 0) = \delta(x - x_0)$. If otherwise $\tau \ll \sigma$, then Eq. (A.20) becomes $\sigma \ddot{x} + \dot{x} = b(x) w \delta(t - t_0)$. Imposing the conditions of continuity and right and left differentiability in t_0, the initial conditions $x(t_0^-)$ and $\dot{x}(t_0^-)$, and taking the limit $\sigma \to 0$, the solution is (again for the case $b(x) = x$) $x(t) = x_0 + x_0 w \Theta(t - t_0)$. Note that the latter corresponds to the solution in the Itô prescription of the SDE (A.20). From the formal Itô solution of Eq. (A.20), the corresponding pdf in the I sense is obtained

$$p^I(x, t) = \delta(x - x_0)(1 - \Theta(t - t_0)) + \Theta(t - t_0) \frac{\rho\left(\frac{x - x_0}{x_0}\right)}{x_0}. \tag{A.22}$$

The latter equation can not be made to correspond to Eq. (A.21) for any choice of $\Theta(0)$. It is in fact interesting to observe that if $\Theta(0) = \alpha$ is set, then the parameter α defines where the $b(x)$ that multiplies the jump is evaluated: when $\alpha = 0$ $b(x)$ is evaluated before the jump, while $\alpha = 1/2$ corresponds to calculating $b(x)$ in the middle of the jump. In the literature on GWN, these choices are associated to the I and S prescriptions, respectively [Gardiner, 2004; Van Kampen, 2007]. Conversely, as just seen for a discrete jump process, the S interpretation of the SDE (A.20) does not correspond to any of the α prescriptions. In other words, there is not an immediate intuitive interpretation of the S prescription.

A.3.2 Multiplicative Compound Poisson noise

Let us generalize now the above analysis to a process described by the following SDE,

$$\dot{x}(t) = a(x, t) + b(x)\xi_\rho^\tau(v, t), \tag{A.23}$$

where $\xi_\rho^\tau(v, t) = \sum_{i=1}^{N(t)} w \dot{\Theta}_\tau(t - t_i)$ is a colored compound Poisson processes (CP), with jump heights w, each time drawn from a generic pdf $\rho(w)$, and $\{t_i\}$ are random times whose sequence is drawn from a homogeneous Poisson counting process $\{N(t), t \geq 0\}$ of rate v. The case in Section 2 corresponds to the special case of a finite deterministic number of jumps. As before, the I interpretation consists of taking $\tau = 0$ and, should a jump occur at time t, of evaluating $b(x)$ at the r.h.s of Eq. (A.23) before the jump occurrence, i.e. $x = x(t^-)$, while the S interpretation of Eq. (A.23) corresponds to performing the zero limit of the correlation time τ of the colored Poisson noise.

Derivation of the S Master Equation

The S ME associated with the GLE (A.23) can be derived through the generating function of $\xi_\rho^\tau(v, t)$

The stochastic process under study is described by the GLE (A.23) presented previously. For simplicity in the following $\xi_\rho^\tau(t, v) = \xi(t)$ has been set. The CP is characterized by the correla-

tion structure ($\langle \cdot \rangle$ denotes the ensemble average)

$$\langle \xi(t)\xi(t+\tau) \rangle \sim e^{-\frac{t}{\tau}}, \tag{A.24}$$

where τ is the characteristic time of the process and all the sub- and super- scripts have been omitted to simplify the notation. If Φ_t is the generating function of CP at time t, then

$$
\begin{aligned}
\Phi_t[v] &= \left\langle e^{i\int_0^t v(s)\xi(s)ds} \right\rangle = e^{\Psi_t[v]} \\
&= \sum_{n=0}^{+\infty} e^{-\nu t} \frac{(\nu t)^n}{n!} \int dw\rho(w) \int \prod_{j=1}^{n} \frac{dt_j}{t} e^{i\sum_{j=1}^{n} w_j \int_0^t v(s)\Theta_\tau(s-t_j)ds} \\
&= \sum_{n=0}^{+\infty} e^{-\nu t} \frac{\nu^n}{n!} \Big[\int_0^t dr \int dw\rho(w) \exp\Big[i\sum_{j=1}^{n} w_j \int_0^t v(s)\Theta_\tau(s-t_j)ds \Big] \Big].
\end{aligned}
\tag{A.25}
$$

Moreover if $\hat{\rho} = \int_{-\infty}^{+\infty} e^{ivw}\rho(w)dw$ is defined as the characteristic function of $\rho(w)$, then

$$\Phi_t[v] = \exp\Big[-\nu t + \nu \int_0^t dr\hat{\rho}\Big(\int_0^t v(s)\Theta_\tau(s-\tau)d\tau \Big) \Big], \tag{A.26}$$

and thus

$$\Psi_t[v] = \ln\Phi_t[v] = \nu \int_0^t dr\Big[\hat{\rho}\Big(\int_0^t v(s)\Theta_\tau(s-\tau)d\tau \Big) - 1 \Big]. \tag{A.27}$$

The Stratonovich interpretation of Eq. (A.67) arises when the limit $\Theta_\tau(t-\tau) \to \delta(t-\tau)$ is taken [Stratonovich, 1963; Kampen, 1981], that is considering a white Poisson process (WP) as the zero limit of the correlation time of the corresponding CP. For a WP the logarithm of the generating function thus reads

$$\Psi_t[v] = \nu \int_0^t dr[\hat{\rho}(v(r)) - 1]. \tag{A.28}$$

Finally because of the Kubo theorem [Kubo, 1962]

$$\Psi_t[v] = \sum_{n=1}^{\infty} \frac{i^n}{n!} \int_0^t ds_1 \cdots ds_n v(s_1) \cdots v(s_n) \langle\langle \xi(t_1) \cdots \xi(t_n) \rangle\rangle_n, \tag{A.29}$$

where $\langle\langle \cdots \rangle\rangle_j$ is the j-th cumulant, i.e., $\langle\langle \cdot \rangle\rangle_1 = \langle \cdot \rangle$, $\langle\langle \cdot\cdot \rangle\rangle_2 = \langle \cdot\cdot \rangle - \langle \cdot \rangle\langle \cdot \rangle$, etc...

From Eqs. (A.28) and (A.29) the explicit formula to calculate the cumulants is obtained

$$
\langle\langle \xi(t_1) \rangle\rangle_1 \;\; = \;\; \frac{\delta \Psi_t}{i \delta \nu(t_1)} = \frac{\nu}{i}\hat{\rho}'(\nu)_{|_{\nu=0}} \tag{A.30}
$$

$$
\langle\langle \xi(t_1)\xi(t_2) \rangle\rangle_2 \;\; = \;\; \frac{\delta^2 \Psi_t}{i^2 \delta \nu(t_1)\delta \nu(t_2)} = \frac{\nu}{i^2}\hat{\rho}''(\nu)_{|_{\nu=0}}\delta(t_2 - t_1) \tag{A.31}
$$

$$
\langle\langle \xi(t_1)\cdots\xi(t_n) \rangle\rangle_n \;\; = \;\; \frac{\delta^n \Psi_t}{i^n \delta \nu(t_1)\cdots\delta \nu(t_n)} = \frac{\nu}{i^n}\hat{\rho}^{(n)}(\nu)_{|_{\nu=0}}\delta(t_2 - t_1)\cdots\delta(t_n - t_{n-1}) \tag{A.32}
$$

In this way, once $\rho(w)$ is given, a complete description of the WP is achieved. For example in the case of exponential distributed jumps, i.e. $\rho(w) = \frac{1}{\langle w\rangle}e^{-\frac{w}{\langle w\rangle}}$, the WP is fully characterized by the moments

$$
\langle\langle \xi(t) \rangle\rangle_1 \;\; = \;\; \nu\langle w\rangle \tag{A.33}
$$

$$
\langle\langle \xi(t_1)\xi(t_2) \rangle\rangle_2 \;\; = \;\; \nu\langle w^2\rangle\delta(t_1 - t_2) \tag{A.34}
$$

$$
\langle\langle \xi(t_1)...\xi(t_n) \rangle\rangle_n \;\; = \;\; \nu\langle w^n\rangle\delta(t_1 - t_2)...\delta(t_{n-1} - t_n). \tag{A.35}
$$

Once all the moments of the WP process have been calculated, the ME corresponding to the GLE (A.67) can easily be achieved. For a given realization of ξ the solution of Eq. (A.23) is

$$
p^S(x,t|\xi) = \delta(x - x(t)). \tag{A.36}
$$

To obtain the general solution of Eq. (A.67), the ensemble average of different trajectories is taken

$$
\langle p^S(x,t|\xi) \rangle = P^S(x,t). \tag{A.37}
$$

Differentiating both sides of Eq. (A.36) and using Eq. (A.67) it follows that

$$
\partial_t p^S(x,t|\xi) \;\; = \;\; \partial_x\delta(x - x(t))[-\dot{x}(t)] \tag{A.38}
$$

$$
= \;\; -\partial_x\delta(x - x(t))[a(x(t),t) + b(x(t))\xi(t)] \tag{A.39}
$$

$$
= \;\; -\partial_x\delta(x - x(t))[a(x(t),t) + b(x(t))\xi(t)], \tag{A.40}
$$

and thus the forward ME for the pdf conditioned by a given realization of the WP is obtained

$$
\frac{\partial}{\partial t}p^S(x,t|\xi) = -\mathscr{O}(x,\partial_x,t)\,p^S(x,t|\xi), \tag{A.41}
$$

where $\mathscr{O}(x,\partial_x,t) = \partial_x[a(x(t),t) + b(x(t))\xi(t)]$ is the forward time evolution operator. The

solution of Eq. (A.41), for the initial condition $p^S(x(0), 0|\xi) = \delta(x − x(0))$ is

$$p^S(x, t|\xi) = T\left(\exp\left[-\int_0^t (\partial_x a(x, \tau) + \partial_x b(x(\tau))\xi(\tau))d\tau\right]\right)\delta(x − x(0)), \tag{A.42}$$

where T is the T-product operator. Using Eq. (A.37) and the Kubo relation (A.29), an explicit formula for the general formal solution of the GLE (A.67) in the Stratonovich prescription is obtained

$$P^S(x, t) = T\left(\exp\left[-\int_0^t \partial_x a(x, \tau)d\tau - \sum_{n=1}^\infty \int_0^t dt_1 \cdots \int_0^t dt_n \partial_x b(x(t_1)) \cdots \partial_x b(x(t_n))\right.\right. \tag{A.43}$$

$$\times \left.\left. \langle\langle\xi(t_1)\cdots\xi(t_n)\rangle\rangle\right]\right)\delta(x − x(0)). \tag{A.44}$$

Thanks to Eqs. (A.30), (A.31) and (A.32) complete characterization of the cumulants is obtained, and thus substituting Eq. (A.32) into Eq. (A.43) we have

$$P^S(x, t) = T\left(\exp\left[-\int_0^t \partial_x a(x, \tau)d\tau + \sum_{n=1}^\infty \nu \int_0^t (-\partial_x b(x(\tau)))^n \rho\hat{(0)}^{(n)} d\tau\right]\right)$$

$$= T\left(\exp\left[-\int_0^t \partial_x a(x, \tau)d\tau - \nu \int_0^t d\tau \langle e^{-\partial_x b(x(\tau))} − 1\rangle_{\rho(w)}\right]\right) \tag{A.45}$$

Eventually, differentiating Eq. (A.45) with respect to t, the ME corresponding to the GLE (A.67) in the Stratonovich interpretation is obtained:

$$\frac{\partial P^S(x, t)}{\partial t} = \left[-\frac{\partial}{\partial x} a(x, t) + \nu \langle e^{-w\frac{\partial}{\partial x} b(x)} − 1\rangle_{\rho(w)}\right] P^S(x, t), \tag{A.46}$$

where $\langle \cdot \rangle$ denotes the ensemble average operator. The same result can also be obtain through a more formal derivation [Hanggi, 1980; Sancho et al., 1987].

A simpler alternative derivation of the ME (A.46), can be obtained using the fact that in the S prescription the rules of calculus are preserved. In fact, the GLE (A.23) can be written as

$$\dot{x}(t) = \begin{cases} a(x, t), & \text{with probability } 1 − \nu dt; \\ b(x) w h_\tau(t), & \text{with probability } \nu dt, \end{cases} \tag{A.47}$$

where $h_\tau(t) = \dot{\Theta}_\tau(t)$. Let us now consider only the effect of the jumps on x. From Eq. (A.47) it follows that $dx/b(x(t)) = w h_\tau(t)dt$, and setting

$$\frac{d\eta(x)}{dx} = \frac{1}{b(x)} \quad \Rightarrow \quad \eta(x) = \int^x \frac{dx'}{b(x')}, \tag{A.48}$$

Eq. (A.47) becomes

$$d\eta(x(t)) = w h_\tau dt, \tag{A.49}$$

which integrated between t and $t + dt$ reads

$$\eta(x(t+dt)) = \eta(x(t)) + w\,\Delta\Theta_\tau(t) \quad \Rightarrow \quad x(t+dt) = \eta^{-1}\big[\eta(x(t)) + w\,\Delta\Theta_\tau(t)\big], \qquad \text{(A.50)}$$

where $\Delta\Theta_\tau(t) = \Theta_\tau(t+dt) - \Theta_\tau(t)$. Finally, discrete ME corresponding to the GLE (A.67) interpreted in the Stratonovich sense can be written

$$
\begin{aligned}
P^S(x, t+dt) \;=\; & (1-v)\,dt \int_0^\infty dx'\, P^S(x', t)\,\delta\big(x - (a(x')dt + x')\big) + \qquad \text{(A.51)} \\
& + v\,dt \int_0^\infty \int_0^\infty \rho(w) P^S(x', t)\delta\big(x - (\eta^{-1}[\eta(x') + w])\big)\, dw\, dx',
\end{aligned}
$$

where the limit $\tau \to 0$ of the GLE (A.23) has been performed and the fact that $\lim_{\tau \to 0} \Delta\Theta_\tau(t) = 1$ has been used. The integral in the r.h.s of Eq. (A.51) can be rewritten, inverting the Dirac Delta with respect to w and using the rule of the inverse function, as $\int_0^\infty \int_0^x \rho(w) P^S(x') \frac{\delta(w - (\eta(x) - \eta(x')))}{|1/\eta'(x)|}\, dw\, dx'$ and thus, after taking the continuum time limit, the Master Equation (A.51) becomes

$$\frac{\partial P^S(x, t)}{\partial t} = -\frac{\partial}{\partial x}\big[a(x, t)P^S(x, t)\big] + v \int_{-\infty}^\infty \frac{\rho(\eta(x) - \eta(x'))}{|b(x)|} P^S(x', t)dx' - v P^S(x, t). \quad \text{(A.52)}$$

Derivation of the I Master Equation

In the I prescription, $x(t)$ at time t does not depend on the noise $\xi_\rho^{\tau=0}(v, t) \equiv \xi_\rho(v, t)$ at the same time [Ito, 1951]. From this it follows that

$$\langle b(x)\xi_\rho(v, t)\rangle = \langle b(x)\rangle\langle \xi_\rho(v, t)\rangle. \qquad \text{(A.53)}$$

Therefore, if (A.23) with $\tau = 0$ is interpreted in the I sense, the size of the jumps can be changed from w to $b(x)w$, and the corresponding ME can be derived without ambiguity [Hanggi, 1980; Denisov et al., 2009]

$$\frac{\partial P^I(x, t)}{\partial t} = -\frac{\partial}{\partial x}\big[a(x, t)P^I(x, t)\big] + v \int_{-\infty}^\infty \rho\left(\frac{x - x'}{b(x')}\right)\frac{P^I(x', t)}{|b(x')|}dx' - v P^I(x, t). \qquad \text{(A.54)}$$

Alternatively, a different form of the I ME (A.54), that is the I analogous of the S ME (A.46), can be achieved

$$\frac{\partial P^I(x, t)}{\partial t} = -\frac{\partial}{\partial x}\big[a(x, t)P^I(x, t)\big] + v\langle : e^{-w\frac{\partial}{\partial x}b(x)} : -1\rangle_{\rho(w)} P^I(x, t), \qquad \text{(A.55)}$$

where $::$ is an operator (analogous to the normal order operator in quantum field theory) which indicates that all the derivatives must be placed on the left of the expression, i.e. : $e^{-w\frac{\partial}{\partial x}b(x)}F(x) := \sum_{n=0}^{+\infty} \frac{(-w)^n}{n!}(\frac{\partial}{\partial x})^n(b(x)^n F(x))$.

To derive Eq. (A.55), the integral in the ME (A.54) is rewritten as

$$\int_{-\infty}^{\infty} \rho\left(\frac{x-x'}{b(x')}\right) \frac{P^I(x',t)}{|b(x')|} dx' = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho(w)\delta(x-x'-wb(x'))P^I(x',t)dx'dw. \tag{A.56}$$

Formally expanding the Dirac delta

$$\delta(x-x'-wb(x')) = \sum_{n=0}^{+\infty} \frac{(-w)^n}{n!} \left(\frac{\partial}{\partial x}\right)^n b(x')^n \delta(x-x') \tag{A.57}$$

and substituting Eq. (A.57) in (A.56) it follows that

$$\int_{-\infty}^{\infty} \rho\left(\frac{x-x'}{b(x')}\right) \frac{P^I(x',t)}{|b(x')|} dx' = \langle(: e^{-w\frac{\partial}{\partial x}b(x)}:)P^I(x,t)\rangle_\rho, \tag{A.58}$$

where $(:e^{-w\frac{\partial}{\partial x}b(x)}:)P^I(x,t) \equiv \sum_{n=0}^{+\infty} \frac{(-w)^n}{n!}(\frac{\partial}{\partial x})^n b(x)^n P^I(x,t)$. Using the expression (A.58) in Eq. (A.54) the ME (A.55) is obtained.

It is worth to note that when $b(x) = b$ is constant, by using $e^{-bw\frac{\partial}{\partial x}}P^S(x,t) = P^S(x-bw,t)$ the I and S MEs become coincident, as expected.

A.3.3 FPE in the Limits of infinite frequency and infinitesimal amplitudes of the jumps

It is now shown that, taking the limit $\nu \to \infty$, $\langle w \rangle \to 0$, i.e. infinite frequency and infinitesimally small jumps, such that $\nu\langle w\rangle^2 = D$ remains constant, Eqs. (A.46) and (A.55) reduce to the well known I and S Fokker-Planck equation (FPE) for GWN, respectively [Gardiner, 2004].

In particular, the well known FPE corresponding to the GLE (A.23) is derived when $\xi(t)$ is a GWN with mean $\langle \xi(t) \rangle = 0$ and correlation $\langle \xi(t)\xi(s) \rangle = 2D\delta(t-s)$, from the MEs (A.46) and (A.55). The results are generalized to any jump size pdf of the form

$$\rho(w) = \gamma f(\gamma w), \tag{A.59}$$

with $\gamma > 0$ and $\int w^n \rho(w) dw = \langle w^n \rangle_\rho < \infty \; \forall n$. The latter condition implies $\gamma \int dw w^n f(\gamma w) = \gamma^{-n} \int dz z^n f(z) = \gamma^{-n} \langle z^n \rangle_f < \infty \; \forall n$.
Stratonovich Eq. The case for the Stratonovich prescription has been first presented in Van Den Broeck [1983]. The FPE corresponding to multiplicative GWN process interpreted in the Stratonovich sense is

$$\frac{\partial}{\partial t}P^S(x,t) = -\frac{\partial}{\partial x}\left[a(x,t)P^S(x,t)\right] + D\frac{\partial}{\partial x}b(x)\frac{\partial}{\partial x}b(x)P^S(x,t). \tag{A.60}$$

Once a zero mean WP process is considered, the ME (A.46) reads as [Van Den Broeck, 1983]

$$\frac{\partial P^S(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[\left[a(x,t) - \nu\langle w\rangle b(x)\right]P^S(x,t)\right] + \nu\langle e^{-w\frac{\partial}{\partial x}b(x)} - 1\rangle_\rho P^S(x,t) \quad (A.61)$$

$$= -\frac{\partial}{\partial x}\left[a(x,t)P^S(x,t)\right] + \nu\sum_{n=1}^{+\infty}(-\frac{1}{\gamma})^n\frac{\langle z^n\rangle_f}{n!}\left(\frac{\partial}{\partial x}b(x)\right)^n P^S(x) \quad (A.62)$$

where the integral in the r.h.s of Eq. (A.61) has been expanded as

$$\langle e^{-w\frac{\partial}{\partial x}b(x)}\rangle_\rho P^S(x) = \sum_{n=0}^{+\infty}(-1)^n\frac{\langle w^n\rangle}{n!}\left(\frac{\partial}{\partial x}b(x)\right)^n P^S(x) = \sum_{n=0}^{+\infty}(-\frac{1}{\gamma})^n\frac{\langle z^n\rangle_f}{n!}\left(\frac{\partial}{\partial x}b(x)\right)^n P^S(x). \quad (A.63)$$

Taking the limit $\nu, \gamma \to \infty$, such that $\frac{\nu}{\gamma^2} = D'$, then $\frac{\nu}{\gamma^n} \to 0$ for $n > 2$ and the latter ME (A.62) corresponds exactly to the FPE (A.60) with $D = D'\frac{\langle z\rangle_f}{2}$.

Itô Eq. The FPE corresponding to multiplicative GWN process interpreted with the Itô prescription is

$$\frac{\partial}{\partial t}P^I(x,t) = -\frac{\partial}{\partial x}\left[a(x,t)P(x,t)\right] + D\frac{\partial^2}{\partial x^2}\left[b(x)^2 P^I(x,t)\right]. \quad (A.64)$$

Let us now repeat the same procedure as before, starting from the zero mean I ME

$$\frac{\partial P^I(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[[a(x,t) - \nu\langle w\rangle b(x)]P^I(x,t)\right] + \nu\langle(: e^{-w\frac{\partial}{\partial x}b(x)} :) - 1\rangle_\rho P^I(x,t). \quad (A.65)$$

Expanding the r.h.s. and remembering that the operator :: means that all the derivatives must be placed on the left of the expression

$$\nu\langle(: e^{-w\frac{\partial}{\partial x}b(x)} :)-1\rangle_\rho P^I(x,t) = -\nu\langle w\rangle_\rho\frac{\partial}{\partial x}[b(x)P^I(x,t)]+\nu\sum_{n=2}^{+\infty}(-\frac{1}{\gamma})^n\frac{\langle z^n\rangle_f}{n!}\left(\frac{\partial}{\partial x}\right)^n[b(x)^n P^I(x,t)]$$

$$(A.66)$$

Eventually, inserting Eq. (A.66) in the I ME (A.65) and taking $\nu, \gamma \to \infty$ with $\frac{\nu}{\gamma^2} = D'$ and $D = D'\frac{\langle z\rangle_f}{2}$ the I FPE (A.64) is obtained.

A.3.4 Prescription-induced jump distributions

It is clear from the previous MEs (A.52) and (A.54) that the I and S prescriptions of the GLE

$$\dot{x}(t) = a(x,t) + b(x)\xi_\rho(\nu,t) \quad (A.67)$$

lead to different MEs. Let us now determine the connection between the two different interpretations. Specifically, the two jump PDFs in the I and S interpretation, ρ_I and ρ_S, which give rise to the same process are sought. Is also sought how to obtain one form when the other is

given. To this purpose it is sufficient to equate the two MEs, (A.52) and (A.54) for simplicity, from which

$$\frac{1}{|b(x')|}\rho_I(\frac{x-x'}{b(x')}) = \frac{1}{|b(x)|}\rho_S(\eta(x)-\eta(x')). \tag{A.68}$$

As a result, if Eq. (A.68) can be solved, given the jumps pdf and choosing the S (I) prescription for Eq. (A.67), the solutions ρ_I (ρ_S) of Eq. (A.68) give the equivalent corresponding I (S) GLE and ME. This is one of the main results of the paper and it provides the connection between the prescription-induced jump distributions ρ_I and ρ_S, allowing link the Itô ME and the Stratonovich ME corresponding to a GLE with multiplicative white Poisson noise.

The previous equation however has a solution only when $b(x)$ is a linear function of x. To show this Eq. (A.68) is rewritten as

$$\rho_I(y) = \frac{|b(x')|}{|b(x)|}\rho_S(\eta(x)-\eta(x')) \equiv F(x',y), \tag{A.69}$$

where $y = (x-x')/b(x')$. Because the l.h.s. of Eq. (A.69) does not depend on x', it must be $\frac{\partial F}{\partial x'} = 0$, that explicitly read as

$$
\begin{aligned}
0 &= \rho_S(\eta(x)-\eta(x'))\left[sgn[b(x')]\frac{b'(x')}{|b(x)|} - sgn[b(x)]\frac{b'(x)}{|b(x)|^2}|b(x')|(1+yb'(x'))\right] + \\
&+ \rho_S'(\eta(x)-\eta(x'))\frac{|b(x')|}{|b(x)|}\left[\eta'(x)(1+yb'(x)-\eta'(x'))\right].
\end{aligned} \tag{A.70}
$$

The latter, using Eq. (A.48), can be expressed as

$$\frac{\rho_S'(\eta(x)-\eta(x'))}{\rho_S(\eta(x)-\eta(x'))}\left(\frac{1}{b(x)}(1+yb'(x'))-\frac{1}{b(x')}\right) + \frac{b'(x')}{b(x')}-\frac{b'(x)}{b(x)}(1+yb'(x')) = 0. \tag{A.71}$$

Eq. (A.71) must hold for all ρ_S, then the solution of Eq. (A.71) is given by the function b that satisfies the conditions

$$
\begin{aligned}
b(x')(1+yb'(x')) &= b(x) \tag{A.72} \\
b(x')b'(x)(1+yb'(x')) &= b(x)b'(x'). \tag{A.73}
\end{aligned}
$$

Combining Eqs. (A.72) and (A.73) and using $x = b(x')y+x'$, the equation $b'(b(x')y+x') = b'(x')$ is obtained. If the derivative of both side with respect to the independent variable y is taken, then it follows that $b''(b(x')y+x')b(x') = 0$. This implies $b''(x) = 0 \ \forall x$, which solution is $b(x) = kx$ (with k any constant). For other functional shape of $b(x)$ the jumps pdf $\rho_I(w)$ depends also on the state of the system, i.e., the dependence on x of $\rho_I(w|x)$ cannot be factored

out. In this case, is not even clear to what a Stratonovich prescription would correspond to.

Finally, the distribution of the impulses that may be measured from the time series of the process is derived (see inset in Figure A.2). In fact, if a random jump (drawn from $\rho(w)$) occurs at time t, then the size of the impulse that the whole process experiences is $y_t = x(t + dt) - x(t)$. From the GLE (A.23) follows that with probability $v\,dt$, $\dot{x} = b(x)\,w\,\Theta_\tau(t)$. Taking the limit $\tau \to 0$, and using the definition of $\eta(x)$ it follows

$$
y(t) = \begin{cases} w\,b(x), & \text{(I)} \\ \eta^{-1}(\eta(x) + w) - x & \text{(S)} \end{cases} \tag{A.74}
$$

and thus

$$
\hat{P}^I(y, t) = \langle \delta(y - y_t) \rangle = \int_{-\infty}^{+\infty} dx\,dw \frac{1}{|b(x)|} P^I(x, t)\rho(w)\delta(w - y/b(x)) \tag{A.75}
$$

$$
\hat{P}^S(y, t) = \int_{-\infty}^{+\infty} dw\,dx P^S(x, t)\rho(w)\frac{\delta(w - [\eta(x + y) - \eta(x)])}{|b(x + y)|}, \tag{A.76}
$$

is obtained, i.e. the prescriptions characterize the pdf of the impulses of the whole process.

Summarizing, an approach to solve the Ito-Stratonovich (I-S) dilemma for GLE with multiplicative WP noise has been here proposed. It has been shown how different interpretations lead to different results and that choosing between the I and S prescriptions is crucial to describe correctly the dynamics of the model systems, and how this choice can be determined by physical information about the timescales involved in the process. Moreover, the related issue of finding a connection between the I and S interpretations in the case of linear WP noise has been addressed. Differently from the introduction of a drift previously proposed [Zygad\lo, 1993; Pirrotta, 2007], such connection has been found in a transformation of the jumps PDFs and tested these results numerically. The results are also consistent with the physics of the random forcing, which takes place at specific points in time, whereas a continuously-acting spurious drift would conceptually violate the causality of the process. In particular, once the GLE (A.67) is given, its I and S interpretations are shown to be equivalent if ρ_I and ρ_S satisfy the prescription-induced jumps pdf Eq. (A.68).

A.3.5   Application to Soil Salinization

The above mathematical problems naturally arise in the context of the process of soil salinization. This is an extremely relevant environmental problem as four million km^2 in arid and semi-arid lands are affected by soil salinization, causing vegetation dieoff and possible desertification [Hillel, 1998]. In natural salinization (unlike the anthropogenic one due to irrigation), salt may accumulate in surface soils by dry and wet deposition due to wind and rain. In this problem, state-dependent Poisson jumps arise naturally when writing the salt mass balance

equation at the daily-to-monthly time scale for soil root zone used as the control volume. Salt inputs due to rainfall and wind act almost continuously in time, while the state-dependent losses of salt occur through negative jumps due to the leaching caused by intense rainfall events. Schematically, the salt mass at time, $x(t)$, in the root zone is described by the GLE:

$$\frac{dx}{dt} = \Upsilon - x\xi_\rho(\nu, t), \tag{A.77}$$

where Υ is the time-averaged salt mass input flux, $\xi_\rho(\nu, t)$ is the leaching flux toward deeper layers, which can be approximated by a WP process with $\rho(w) = \mu \exp(-\mu w)\Theta(w)$. The leaching parameters ν (frequency of leaching events) and μ (mean jump) can be expressed in terms of the climatic, soil and vegetation properties. Because the typical duration of leaching events is on the order of a few hours, while the equilibration times of salt in the soil solution (proportional to the inverse of its dissolution rate) tend to be smaller (minutes to hours), this means that the inertia in the dynamics is small ($\sigma \ll \tau$) and the physically correct interpretation is likely to be the Stratonovich one.
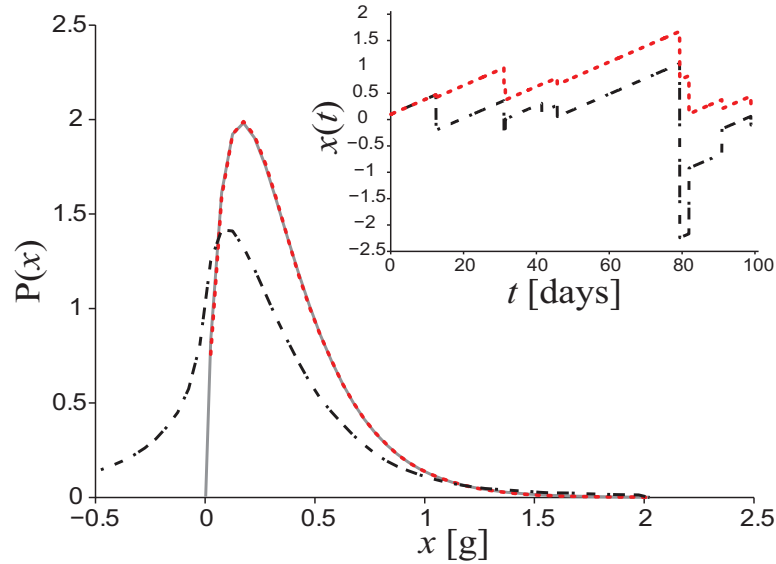
Figure A.2: Comparison of the steady state pdf of Eq. (A.77): S solution (solid line, obtained analytically from Eq. (A.78)), I solution (dash-dot, from numerical simulation), I solution using the jump distribution given by Eq. (A.80) (dotted line, from numerical simulation). The numerical simulations confirm the analytical results. Inset: simulated trajectory of the salt mass under the two different prescriptions. Note that if artificial reflecting barriers are not imposed, the salt mass given by the I prescription of Eq. (A.77) may assume unphysical negative values. The parameters used for the simulation are $\mu = 0.463$, $\nu = 0.15\,day^{-1}$ and $\Upsilon = 30\,mg/day$.

The stationary solution of Eq. (A.46) in the S prescription is a Gamma distribution (Figure A.2)

[Van Den Broeck, 1983]

$$P^S(x) = \mathcal{N} \, e^{-(x\nu/\Upsilon)} \, x^{1/\mu}, \tag{A.78}$$

for $x > 0$ and where $\mathcal{N} = (\frac{\nu}{\Upsilon})^{\frac{1+\mu}{\mu}} / \Gamma(\frac{1+\mu}{\mu})$ is the normalization constant and $\Gamma(z)$ the complete gamma function of argument z. Eq. (A.78) summarizes the soil salinity statistics as a function of climate, soil and vegetation parameters, which may in turn be used in conjunction with the soil moisture statistic to obtain a full characterization of the salt concentration in the root zone and the ensuing risk of salinization.
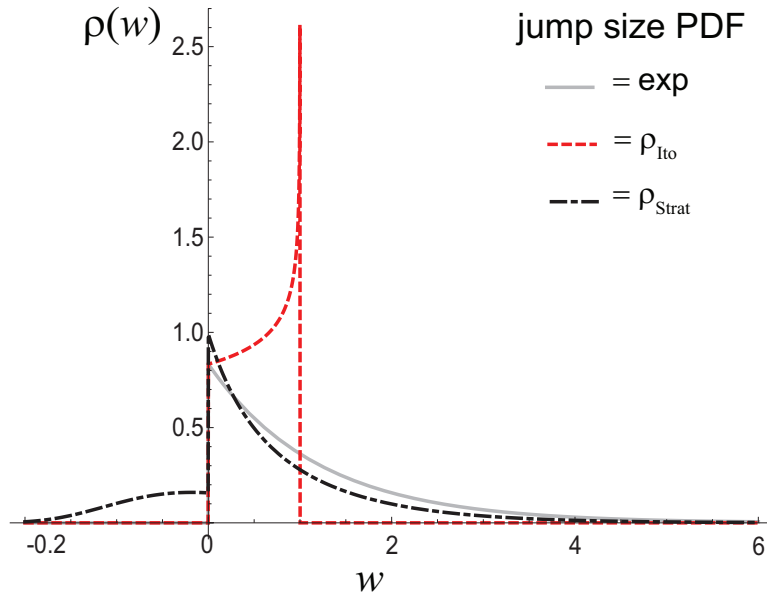
Figure A.3: Comparison between a jump exponential distribution $\rho(w)$ with mean $1/\gamma = 0.8$, and the solutions $\rho_I(z)$, $\rho_S(z)$ of the prescription induced jumps corresponding to Eqs. (A.80) and (A.81) respectively, corresponding to the given $\rho(w)$.

From Eq. (A.76) it is possible to derive the pdf of the impulses of the process for the S interpretation as

$$\hat{P}^S(y) = \epsilon e^{\epsilon y} \Theta(-y), \tag{A.79}$$

which is an exponential distribution controlled by the parameter $\epsilon = \nu/\Upsilon$, given by the ratio between the rate of leaching events and the average rate of salt input. Thus if time series of the process are available, the Stratonovich assumption can be checked by backtracking information on the physical timescales involved in the process, via a comparison with experimental data. A further support for the S interpretation of Eq. (A.77) is given by the fact that x must remain positive after a jump, a fact that is not ensured by the I interpretation unless a reflecting boundary in $x = 0$ is imposed (see Figure A.2). The prescription-induced jump distributions correspondence for this case ($b(x) = -x$) is computed, and corresponds to

$\rho_S(\ln|\frac{x'}{x}|) = |\frac{x}{x'}|\rho_I(1 - \frac{x}{x'})$, where x' and x are the variables before and after the jump, respectively. By taking into account that in the S prescription $x, x' > 0$, the I-jump pdf equivalent to $\rho_S(w) = \gamma e^{-\gamma w}\Theta(w)$ is

$$\rho_I(z) = \gamma(1 - z)^{\gamma - 1}, \qquad z \in [0, 1]. \tag{A.80}$$

This equivalence is indeed remarkable because it considerably facilitates the numerical simulation of the salinity equation in the S formulation (see Figure A.2). On the other hand, if the GLE (A.77) were interpreted in the I sense, the ratio x/x' could also be negative and the solution of Eq. (A.68), for $\rho_I = \gamma\Theta(w)e^{-\gamma w}$, would read

$$\rho_S(w) = \gamma e^{-\gamma - w}\left[\Theta(w)e^{\gamma e^{-w}} + e^{-\gamma e^{-w}}\right] \quad w \in]-\infty, +\infty[. \tag{A.81}$$

This implies that possible negative jumps (that occur for $x < 0$) in the I prescription for the given $\rho_I(w)$, would be explicitly present in the corresponding equivalent S-jump pdf $\rho_S(w)$ (see Figure (A.3)).

For further details on the soil salinization modelling scheme see Chapter 3.

Bibliography

Abramowitz, M. and I.A., S. (1965). *Handbook of Mathematical Functions*. New York Press, Dover.

Adler, P. B., Tyburczy, W. R., and Lauenroth, W. K. (2007). Long-term mapped quadrats from Kansas prairie: demographic information for herbaceous plants. *Ecology*, 88(10):2673.

Allan, T. (1993). Fortunately there are substitutes for water: otherwise our hydropolitical futures would be impossible. *Proceedings of the Conference on Priorities for Water Resources Allocation and Management*, 2:13–26.

Allison, G. B., Gee, G. W., and Tyler, S. W. (1994). Vadose-Zone Techniques for Estimating Groundwater Recharge in Arid and Semiarid Regions. *Soil Science Society of America*, 58:6–14.

Alonso, D. and McKane, J. (2004). Sampling Hubbell's neutral theory of biodiversity. *Ecology*, 7:901.

Alpizar-Jara, R., Nichols, J. D., Hines, J. E., Sauer, J. R., Pollock, K. H., and Rosenberry, C. S. (2004). The relationship between species detection probability and local extinction probability. *Oecologia*, 141(4):652–660.

Ayars, J. E., Hutmacher, R. B., Schoneman, R. A., Vail, S. S., and Pflaum, T. (1993). Long-Term Use Of Saline Water For Irrigation. *Irrigation Science*, 14(1):27–34.

Azaele, S., Banavar, J. R., and Maritan, A. (2009). Probing noise in gene expression and protein production. *Physical Review E*, 80:31916.

Azaele, S., Pigolotti, S., Banavar, J. R., and Maritan, A. (2006). Dynamical evolution of ecosystems. *Nature*, 444(7121):926–928.

Baird A. J. and Wilby, R. L. (1999). *Eco-hydrology: Plants and Water in Terrestrial and Aquatic Environments*. Routledge, London.

Bibliography

Balata, D., Piazzi, L., and Benedetti-Cecchi, L. (2007). Sediment disturbance and loss of beta diversity on subtidal rocky reefs. *Ecology*, 88(10):2455–2461.

Barabasi, A.-L. (2002). *Linked: The new science of networks.* Perseus Publishing.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286:509–512.

Barrat, a., Barthélemy, M., Pastor-Satorras, R., and Vespignani, a. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–52.

Bell, G. (2001). Ecology - Neutral macroecology. *Science*, 293(5539):2413–2418.

Berne, A., Uijlenhoet, R., and Troch, P. A. (2005). Similarity analysis of subsurface flow response of hillslopes with complex geometry. *Water Resources Research*, 41(9):W09410+.

Bertuzzo, E., Azaele, S., Maritan, A., Gatto, M., Rodriguez-Iturbe, I., and Rinaldo, A. (2008). On the space-time evolution of a cholera epidemic. *Water Resources Research*, 44(1).

Bertuzzo, E., Maritan, A., Gatto, M., Rodriguez-Iturbe, I., and Rinaldo, A. (2007). River networks and ecological corridors: Reactive transport on fractals, migration fronts, hydrochory. *Water Resources Research*, 43(4).

Bertuzzo, E., Muneepeerakul, R., Lynch, H. J., Fagan, W. F., Rodriguez-Iturbe, I., and Rinaldo, A. (2009). On the geographic range of freshwater fish in river basins. *Water Resources Research*, 45.

Bertuzzo, E., Suweis, S., Mari, L., Maritan, A., Rodriguez-Iturbe, I., and Rinaldo, A. (2011). Spatial Effects on Species Persistence and Implications for Biodiversity. *Proceeding of the National Academy of Science of the United States of America*, 108(11):4346–4351.

Bettencourt, L. M. a., Lobo, J., Helbing, D., Kühnert, C., and West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences of the United States of America*, 104(17):7301–6.

Beven, K. J. (2001). *Rainfall-Runoff Modeling: The Primer.* Wiley, Chichester.

Bianconi, G. and Barabási, A.-L. (2001). Bose-Einstein Condensation in Complex Networks. *Physical Review Letters*, 86(24):5632–5635.

Boguna, M. and Pastor-Satorras, R. (2003). Class of correlated random networks with hidden variables. *Physical Review E*, 68:36112.

Botter, G. (2010). Stochastic recession rates and the probabilistic structure of stream flows. *Water Resources Research*, 46.

Botter, G., Basu, N. B., Zanardo, S., Rao, P. S. C., and Rinaldo, A. (2010). Stochastic modeling of nutrient losses in streams: Interactions of climatic, hydrologic, and biogeochemical controls. *Water Resources Research*, 46.

Botter, G., Daly, E., Porporato, a., Rodriguez-Iturbe, I., and Rinaldo, a. (2008a). Probabilistic dynamics of soil nitrate: Coupling of ecohydrological and biogeochemical processes. *Water Resources Research*, 44(3):1–15.

Botter, G., Peratoner, F., Porporato, a., Rodriguez-Iturbe, I., and Rinaldo, a. (2007a). Signatures of large-scale soil moisture dynamics on streamflow statistics across U.S. climate regimes. *Water Resources Research*, 43(11):1–10.

Botter, G., Porporato, A., Daly, E., Rodriguez-Iturbe, I., and Rinaldo, A. (2007b). Probabilistic characterization of base flows in river basins: Roles of soil, vegetation, and geomorphology. *Water Resources Research*, 43(6).

Botter, G., Porporato, a., Rodriguez-Iturbe, I., and Rinaldo, a. (2007c). Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: Slow, leaching-prone components of the hydrologic response. *Water Resources Research*, 43(2):1–14.

Botter, G., Zanardo, S., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A. (2008b). Ecohydrological model of flow duration curves and annual minima. *Water Resources Research*, 44(8).

Boulinier, T., Nichols, J. D., Hines, J. E., Sauer, J. R., Flather, C. H., and Pollock, K. H. (2001). Forest fragmentation and bird community dynamics: Inference at regional scales. *Ecology*, 82(4):1159–1169.

Boulinier, T., Nichols, J. D., Sauer, J. R., Hines, J. E., and Pollock, K. H. (1998). Estimating species richness: The importance of heterogeneity in species detectability. *Ecology*, 79(3):1018–1028.

Bras, R. L. and Seo, D. (1987). Irrigation control in the presence of salinity: Extended linear quadratic approach. *Water Resources Research*, 23(7).

Brooks, R. H. and Corey, A. T. (1966). Properties of porous media affecting fluid flow. *Journal of Irrigation and Drainage Engineering*, IR(2):61.

Brown, B. L. and Swan, C. M. (2010). Dendritic network structure constrains metacommunity properties in riverine ecosystems. *Journal of Animal Ecology*, 79(3):571–580.

Brown, J. H. (1995). *Macroecology*. The University of Chicago Press, Chicago.

Brown, J. H. and Kodricbrown, A. (1977). Turnover Rates In Insular Biogeography - Effect Of Immigration On Extinction. *Ecology*, 58(2):445–449.

Brutsaert, W. (2005). *Hydrology. An introduction*. Cambridge Univ.Press, New York.

Bibliography

Brutsaert, W. and Nieber, J. L. (1977). Regionalized Drought Flow Hydrographs From a Mature Glaciated Plateau. *Water Resources Research*, 13(3).

Brutsaert, W. and Parlange, M. (1998). Hydrologic cycle explains the evaporation paradox. *Nature*, 396(6706):30.

Buttel, L. A., Durrett, R., and Levin, S. A. (2002). Competition and species packing in patchy environments. *Journal of Theoretical Biology*, 61(3):265–276.

Caldarelli, G., Capocci, A., De Los Rios, P., and Munoz, M. A. (2002). Scale-Free Networks from Varying Vertex Intrinsic Fitness. *Physical Review Letter*, 89(25):258702.

Cam, E., Nichols, J. D., Sauer, J. R., and Hines, J. E. (2002). On the estimation of species richness based on the accumulation of previously unrecorded species. *Ecography*, 25(1):102–108.

Cam, E., Nichols, J. D., Sauer, J. R., Hines, J. E., and Flather, C. H. (2000). Relative species richness and community completeness: Birds and urbanization in the Mid-Atlantic states. *Ecological Applications*, 10(4):1196–1210.

Campos, D., Fort, J., and Méndez, V. (2006). Transport on fractal river networks: application to migration fronts. *Theoretical population biology*, 69(1):88–93.

Ceola, S., Botter, G., Bertuzzo, E., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A. (2010). Comparative study of ecohydrological streamflow probability distributions. *Water Resources Research*, 46.

Chandrasekhar, S. (1943). Stochastic problems in physics and astronomy. *Review of Modern Physics*, 15(1):1–89.

Chapagain, A. K., Hoekstra, A. Y., and Savenije, H. H. G. (2006). Water saving through international trade of agricultural products. *Earth*, pages 455–468.

Chave, J. (2004). Neutral theory and community ecology. *Ecology Letters*, 7(3):241–253.

Chave, J., Muller-Landau, H. C., and Levin, S. A. (2002). Comparing classical community models: Theoretical consequences for patterns of diversity. *American Naturalist*, 159(1):1–23.

Chisholm, R. a. and Lichstein, J. W. (2009). Linking dispersal, immigration and scale in the neutral theory of biodiversity. *Ecology letters*, 12(12):1385–93.

Chow, V. T. (1988). *Applied Hydrology*. McGraw Hill, New York.

Condit, R., Pitman, N., Leigh, E. G., Chave, J., Terborgh, J., Foster, R. B., Nunez, P., Aguilar, S., Valencia, R., Villa, G., Muller-Landau, H. C., Losos, E., and Hubbell, S. P. (2002). Beta-diversity in tropical forest trees. *Science*, 295(5555):666–669.

Costa, L., Rodrigues, F. A., Travieso, G., and Boas, P. R. V. (2007). Characterization of complex networks: a survey of measurements. *Advances in Physics*, 56(1):167–242.

Curtiss, J. H. (1941). On the distribution of the quotient of two chance variables. *Annals Of Mathematical Statistics*, 12:409–421.

Daly, E. and Porporato, A. (2006). Probabilistic dynamics of some jump-diffusion systems. *Physical Review E*, 73(2):26108.

de Aguiar, M., Baranger, M., Baptestini, E. M., Kaufman, L., and Bar-Yam, Y. (2009). Global patterns of speciation and diversity. *Nature*, 460(7253):384–7.

Denisov, S. I., Horsthemke, W., and Hanggi, P. (2009). Generalized Fokker-Planck equation: Derivation and exact solutions. *European Physical Journal B*, 68:567–575.

Diamond, J. M. (1989). The present, past and future of human-caused exctinctions. *Philosophical Transactions of the Royal Society B*, 325(1228):469–477.

D'Odorico, P., Laio, F., Porporato, A., Ridolfi, L., Rinaldo, A., and Rodriguez-Iturbe, I. (2010a). Ecohydrology of the Terrestrial Ecosystems. *BioScience*, 60(11).

D'Odorico, P., Laio, F., and Ridolfi, L. (2010b). Does globalization of water reduce societal resilience to drought? *Geophysical Research Letter*, 37:L13403.

D'Odorico, P., Ridolfi, L., Porporato, A., and Rodriguez-Iturbe, I. (2000). Preferential states of seasonal soil moisture: The impact of climate fluctuations. *Water Resources Research*, 36(8):2209–2219.

Durrett, R. and Levin, S. (1996). Spatial models for species-area curves. *Journal of Theoretical Biology*, 179(2):119–127.

Eng, K. and Milly, P. C. D. (2007). Relating low-flow characteristics to the base flow recession time constant at partial record stream gauges. *Water Resources Research*, 43(1):W01201+.

Erdös, P. and Rényi, A. (1961). On random graphs. *Acta Mathematica Scientia Hungary*, 12:261–267.

Fagan, W. F. (2002). Connectivity, fragmentation, and extinction risk in dendritic metapopulations. *Ecology*, 83(12):3243–3249.

Fagiolo, G., Reyes, J., and Schiavo, S. (2008). On the topological properties of the world trade web: A weighted network analysis. *Physica A*, 387:3868–3873.

Falkenmark, M., Rockstrom, J., and H., S. (2004). Balancing Water for Humans ans Nature: the New Approach in Ecohydrology. *Earthscan*, 247.

FAO (2000a). Food Trade Data. *www.faostat.fao.org*.

FAO (2000b). World agriculture: towards 2030/2050. *http://ftp.fao.org/docrep/fao/009/a0607e/a0607e00.pdf*.

Fiori, A. and Russo, D. (2008). Travel time distribution in a hillslope: Insight from numerical simulations. *Water Resources Research*, 44(12).

Bibliography

Fonseca, J., Narrod, C., Rosegrant, M. W., Fernandez, M., Sinha, A., and Alder, J. (2009). Looking into the future for agriculture and AKST. *IAASTD*, 5:307–337.

Fraedrich, K., Kleidon, A., and Lunkeit, F. (1999). A green planet versus a desert world: Estimating the effect of vegetation extremes on the atmosphere. *Journal of Climate*, 12(10):3156–3163.

Gardiner, C. W. (2004). *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences*. Springer-Verlag, Berlin, 3rd edition.

Garlaschelli, D., Battiston, S., Castri, M., Servedio, D. P., and Caldarelli, G. (2005). The scale-free topology of market investments. *Physica A*, 350:491–499.

Garlaschelli, D. and Loffredo, M. I. (2004). Fitness-Dependent Topological Properties of the World Trade Web. *Physical Review Letter*, pages 1–4.

Garlaschelli, D. and Loffredo, M. I. (2006). Multispecies Grand-Canonical Models for Networks with Reciprocty. *Physical Review E.*, 73:015101(R).

Garlaschelli, D. and Loffredo, M. I. (2009). Generalized Bose-Fermi Statistics and Structural Correlations inWeighted Networks. *Physical Review Letter*, 102:38701.

Gerbens-Leenes, W., Hoekstra, A. Y., and van der Meer, T. H. (2009). The water footprint of bioenergy. *Proceeding of the National Academy of Science of the United States of America*, 106(25):10219–10223.

Gomi, T., Sidle, R. C., Miyata, S., Kosugi, K., and Onda, Y. (2008). Dynamic runoff connectivity of overland flow on steep forested hillslopes: Scale effects and runoff transfer. *Water Resources Research*, 44(8).

Gordon, L. and Folke, C. (2000). Ecohydrological landscape management for human well-being. *Water International*, 2(25):178–184.

Graham, R. and Schenzle, A. (1982). The validity of nonlinear Langevin equations. *Physical Review A*, 25:1731.

Grant, E. H. C., Lowe, W. H., and Fagan, W. F. (2007). Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecology Letters*, 10(2):165–175.

Grant, E. H. C., Nichols, J. D., Lowe, W. H., and Fagan, W. F. (2010). Use of multiple dispersal pathways facilitates amphibian persistence in stream networks. *Proceeding of the National Academy of Science of the United States of America*, 107(15):6936–6940.

Green, J. L. and Ostling, A. (2003). Endemics-area relationships: The influence of species dominance and spatial aggregation. *Ecology*, 84(11):3090–3097.

Hanasaki, N., Inuzuka, T., Kanae, S., and Oki, T. (2010). An estimation of global virtual water flow and sources of water withdrawal for major crops and livestock products using a global hydrological model. *Journal of Hydroly*, 384(3-4):232–244.

Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Motoya, K., Shirakawa, N., Shen, Y., and Tanaka, K. (2008a). An integrated model for the assessment of global water resources - part 2: Applications and assessments. *Hydrology and Earth System Sciences*, 12(3-4):1027–1037.

Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Shirakawa, N., Shen, Y., and Tanaka, K. (2008b). An integrated model for the assessment of global water resources - part 1: Model description and input meteorological forcing. *Hydrology and Earth System Sciences*, 12:1007–1025.

Hanggi, P. (1980). Langevin Description of Markovian Integro-Differential Master Equations. *Zeitschrift fur Physik B*, 36:271–282.

Hanggi, P. and Thomas, H. (1982). Stochastic processes : time evolution , symmetries and linear response. *Physics Reports*, 88(4):207–319.

Harman, C. J. and Silvapalan, M. (2009). Effects of hydraulic conductivity variability on hillslope-scale shallow subsurface flow response and storage-discharge relations. *Water Resources Research*.

Harte, J. (2003). Tail of death and resurrection. *Nature*, 424:1007.

He, F. and Hubbell, S. P. (2011). Species-area relationships always overestimate extinction rates from habitat loss. *Nature*, 473(7347):368–371.

Henderson, P. A. and Magurran, A. E. (2010). Linking species abundance distributions in numerical abundance and biomass through simple assumptions about community structure. *Proceedings of the Royal Society B-Biological Sciences*, 277(1687):1561–1570.

Higham, D. (2001). An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *SIAM Review*, 43(3):525–546.

Hillel, D. (1998). *Environmental Soil Physics*. Academic Press.

Hillel, D. (2000). *Salinity management for sustainable irrigation: Integrating science, environment, and economics*. The International Bank for Reconstruction and Development, the World Bank, Washington.

Hoekstra, A. Y. (2002). Virtual Water Trade: Proceedings of the International Expert Meeting on Virtual Water Trade. In *Value of Water Researh Report*, 12. UNESCO-IHE, Delft.

Hoekstra, A. Y. and Chapagain, A. K. (2008). *Globalization of Water*. Blackwell.

Hoekstra, A. Y. and Hung, P. Q. (2005). Globalisation of water resources: international virtual water flows in relation to crop trade. *Global Environmental Change*, 15:45–56.

Holley, R. and Liggett, T. (1975). Ergodic Theorems for Weakly Interacting Infinite Systems and Voter Model. *Annals of Probability*, 3(4):643–663.

Horsthemke, W. and Lefever, R. (2006). *Noise-induced Transitions: Theory and Applications in Physics, Chemistry, and Biology*. Springer-Verlag Berlin.

Bibliography

Houchmandzadeh, B. and Vallade, M. (2003). Clustering in neutral ecology. *Physical Review E*, 68(6):1–7.

Hubbell, S. (2001). *The Unified Theory of Biodiversity and Biogeography*. Princeton Univeristy Press.

Huntington, T. G. (2006). Evidence for intensification of the global water cycle: Review and synthesis. *Journal of Hydrology*, 319(1-4):83–95.

Institute of Ecosystem Studies (2008). Buell-Small Succession Study.

Ito, K. (1951). Multiple Wiener integral. *Memoirs of the American Mathematical Society*, 4:289.

Jackson, R. B., Carpenter, S. R., Dahm, C. N., McKnight, D. M., Naiman, R. J., Postel, S. L., and Running, S. W. (2001). Water in a changing world. *Ecological Applications*, 11(4):1027–1045.

Kampen, N. G. V. (1981). The validity of nonlinear Langevin equations. *Journal of Statistical Physics*, 24:175–187.

Keitt, T. H. and Stanley, H. E. (1998). Dynamics of North American breeding bird populations. *Nature*, 393(6682):257–260.

Kerr, B., Riley, M. A., Feldman, M. W., and Bohannan, B. J. M. (2002). Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature*, 418(6894):171–174.

Kirchner, J. W. (2009). Catchments as simple dynamical systems: catchment characterization, rainfall-runoff modeling, and doing hydrology backwards. *Water Resources Research*, 45:W02429.

Kottegoda N.T. and Horder, M. A. (1980). Daily flow model based on rainfall occurences using pulses and a transfer function. *Journal of Hydrology*, 47(3-4):215–234.

Kubo, R. (1962). Generalized cumulant expansion method. *Journal of the Physical Society of Japan*, 17:1100.

Kupferman, R., Pavliotis, G. A., and Stuart, A. M. (2004). Ito versus Stratonovich white-noise limits for systems with inertia and colored multiplicative noise. *Physical Review E*, 70:36120.

Laio, F., D'Odorico, P., and Ridolfi, L. (2006). An analytical model to relate the vertical root distribution to climate and soil properties. *Geophysical Research Letters*, 33(18):1–5.

Laio, F., Porporato, A., Fernandez-Illescas, C., and Rodriguez-Iturbe, I. (2001). Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress - IV. Discussion of real cases. *Advances in water resources*, 24(7):745–762.

Lamb R. and Beven, K. (1997). Using interactive recession curve analysis to specify a general catchment storage model. *Hydrology and Earth System Sciences*, 1:101–113.

122

Liu, J., Williams, J. R., Zehnder, A. J. B., and Yang, H. (2007). {GEPIC} - modelling wheat yield and crop water productivity with high resolution on a global scale. *Agricultural Systems*, 94:478–493.

Lyon, R. H. (1973). Propagation of Environmental Noise. *Science*, 179(4078):1083–1090.

MacArthur, R. H. and Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton Univeristy Press.

Manabe, S., Stouffer, R., Spelman, M., and Bryan, K. (1991). Transient Responses Of A Coupled Ocean Atmosphere Model To Gradual Changes Of Atmospheric Co2 .1. Annual Mean Response. *Journal of Climate*, 4(8):785–818.

Mantegna, R. N. and Stanley, H. E. (1999). *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press.

Mari, L., Bertuzzo, E., Casagrandi, R., Gatto, M., Levin, S. A., Rodriguez-Iturbe, I., and Rinaldo, A. (2011). Hydrologic controls and anthropogenic drivers of the zebra mussel invasion of the mississippi-missouri river system. *Water Resources Research*, 47.

McDonnell, J. J. (1990). A Rationale for Old Water Discharge Through Macropores in a Steep, Humid Catchment. *Water Resources Research*, 26(11).

McGill, B. J. (2003). A test of the unified neutral theory of biodiversity. *Nature*, 422(6934):881–885.

McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., Mcglynn, B. L., Welker, J. M., and Seibert, J. (2005). The role of topography on catchment-scale water residence time. *Water Resources Research*, 41(5).

McKane, A., Alonso, D., and Solé, R. V. (2000). Mean-field stochastic theory for species-rich assembled communities. *Physical Review E*, 62(6):8466–8484.

Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., Stouffer, R. J., and Taylor, K. E. (2007). The WCRP CMIP3 Multimodel Dataset: A New Era in Climate Change Research. *Bullettin of the American Meteorological Society*, 88:1383–1394.

Miguens, J. I. L. and Mendes, J. F. F. (2008). Weighted and directed network on traveling patterns. *Biowire*, 5151:145–154.

Milly, P. and Dunne, K. (1994). Sensitivity of the global water cycle to the water-holding capacity of land. *Journal of Climate*, 6(4):506–526.

Morbidelli, R., Corradini, C., and Govindaraju, R. S. (2006). A field-scale infiltration model accounting for spatial heterogeneity of rainfall and soil saturated hydraulic conductivity. *Hydrological Processes*, 20(7):1465–1481.

Bibliography

Moyle, P. B. and Chech, J. J. (2003). *An Introduction to Ichthyology 5th.* Benjamin Cummings, San Francisco.

Muneepeerakul, R., Bertuzzo, E., Lynch, H. J., Fagan, W. F., Rinaldo, A., and Rodriguez-Iturbe, I. (2008a). Neutral Metacommunity models predict fish diversity patterns in Mississipi-Missouri Basin. *Nature*, 453:220–222.

Muneepeerakul, R., Bertuzzo, E., Lynch, H. J., Fagan, W. F., Rinaldo, A., and Rodriguez-Iturbe, I. (2008b). Neutral metacommunity models predict fish diversity patterns in Mississippi-Missouri basin. *Nature*, 453(7192):220–2.

Muneepeerakul, R., Weitz, J. S., Levin, S. A., Rinaldo, A., and Rodriguez-Iturbe, I. (2007). A neutral metapopulation model of biodiversity in river networks. *Journal of Theoretical Biology*, 245(2):351–363.

Newman, M. (2002). Assortative Mixing in Networks. *Physical Review Letters*, 89(20):1–4.

Newman, M. E. J., Barabasi, A. L., and Watts, D. J. (2006). *The Structure and Dynamics of Networks.* Princeton University Press.

Newman, M. E. J. and Sibani, P. (1999). Extinction, diversity and survivorship of taxa in the fossil record. *Proceeding of the Royal Society B*, 266(1428):1593–1599.

Ngo-Duc, T., Polcher, J., and Laval, K. (2005). A 53-year forcing data set for land surface models. *Journal of Geophysical Research Atmospheres*, 110:D06116.

Nichols, J. D., Boulinier, T., Hines, J. E., Pollock, K. H., and Sauer, J. R. (1998a). Estimating rates of local species extinction, colonization, and turnover in animal communities. *Ecological Applications*, 8(4):1213–1225.

Nichols, J. D., Boulinier, T., Hines, J. E., Pollock, K. H., and Sauer, J. R. (1998b). Inference methods for spatial variation in species richness and community composition when not all species are detected. *Conservation Biology*, 12(6):1390–1398.

North American Breeding Bird Survey (2008). Department of the Interior, Geological Survey Patuxent Wildlife Research Center Laurel, MD of the United States of America.

Nour eldin, M. M. N., King, I. P., and Tanji, K. K. (1987). Salinity Management Model .1. Development. *Journal of Irrigation and Drainage Engineering-ASCE*, 113(4):440–453.

O'Dwyer, J. P. and Green, J. L. (2010). Field theory for biogeography: a spatially explicit model for predicting patterns of biodiversity. *Ecology letters*, 13(1):87–95.

Oki, T. and Kanae, S. (2006). Global hydrological cycles and world water resources. *Science*, 313(5790):1068–1072.

Pan, Y. and Nakagoshi, N. (2008). Effects of groundwater disturbance on vegetation implicated by ecohydrology. *Hikobia*, 15(2):177–184.

Park, J. and Newman, M. E. J. (2003). Origin of degree correlations in the Internet and other networks. *Physical Review E*, 68(2):1–7.

Park, J. and Newman, M. E. J. (2004). Statistical mechanics of networks. *Physical Review E*, 70:66117.

Pigolotti, S., Flammini, A., and Maritan, A. (2004). Stochastic model for the species abundance problem in an ecological community. *Physical Review E*, 70(1):11916.

Pigolotti, S., Flammini, A., Marsili, M., and Maritan, A. (2005). Species lifetime distribution for simple models of ecologies. *Proceeding of the National Academy of Science of the United States of America*, 102:15747.

Pirrotta, A. (2007). Multiplicative cases from additive cases: Extension of Kolmogorov-Feller equation to parametric Poisson white noise processes. *Probabilistic Engineering Mechanics*, 22(2):127–135.

Polyanin, A. and Zaitsev, V. F. (2003). *Handbook of Exact Solutions for Ordinary Differential Equations)*. Chapman and Hall, Crc, 2rd ed. edition.

Porporato, A., Daly, E., and Rodriguez-Iturbe, I. (2004). Soil water balance and ecosystem response to climate change. *The American naturalist*, 164(5):625–32.

Porporato, A., Laio, F., Ridolfi, L., and Rodriguez-Iturbe, I. (2001a). Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress - III. Vegetation water stress. *Advances in Water Resources*, 24(7):725–744.

Porporato, A., Laio, F., Ridolfi, L., and Rodriguez-Iturbe, I. (2001b). Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress: II. Probabilistic soil moisture dynamics. *Advances in Water Resources*, 24(7):707–723.

Porporato, A. and Rodriguez-Iturbe, I. (2002). Ecohydrology - a challenging multidisciplinary research perspective. *Hydrological Sciences Journal*, 47(5):811–821.

Preston, F. W. (1948). The commonness and rarity of species. *Ecology*, 29:254–283.

Pueyo, S., He, F., and Zillio, T. (2007). The maximum entropy formalism and the idiosyncratic theory of biodiversity. *Ecology Letters*, 10(11):1017–1028.

Purves, D. W. and Pacala, S. W. (2005). *Biotic Interactions in the Tropics*. Cambridge Univ. Press.

Qadir, M., Ghafoor, A., and Murtaza, G. (2000). Amelioration strategies for saline soils: A review. *Land Degradation and Development*, 11(6):501–521.

Raser, J. M. and O'Shea, E. K. (2005). Noise in gene expression: Origins, consequences, and control. *Science*, 309(5743):2010–2013.

Regonda, S. K., Rajagopalan, B., Clark, M., and Pitlick, J. (2005). Seasonal cycle shifts in hydroclimatology over the western United States. *Journal of Climate*, 18(2):372–384.

Bibliography

Richards, L. A. (1954). *Diagnosis and Improvement of Saline and Alkali Soils.* USDA Handbook No.60, Washington.

Ricketts, T. H. (2001). The matrix matters: Effective isolation in fragmented landscapes. *American Naturalist*, 158(1):87–99.

Ricklefs, R. E. (1987). Community diversity - relative roles of local and regional processes. *Science*, 235(4785):167–171.

Ricklefs, R. E. and Scheuerlein, A. (2003). Life span in the light of avian life histories. *Pop. Develop. Rev.*, 29(Suppl. S):71–98.

Ridolfi, L., D'Odorico, P., and Laio, F. (2011). *Noise-Induced Phenomena in the Environmental Sciences.* Cambridge University Press.

Rinaldo, A., Botter, G., Bertuzzo, E., Uccelli, A., Settin, T., and Marani, M. (2006). Transport at basin scales: 1. Theoretical framework. *Hydrology and Earth System Sciences*, 10(1):19–29.

Rinaldo, A., Rodriguez-Iturbe, I., Rigon, R., Bras, R. L., Vasquez, E., and Marani, A. (1992). Minimum Energy and Fractal Structures of Drainage Networks. *Water Resources Research*, 28(9):2183–2195.

Rodriguez-Iturbe, I. (2000). Ecohydrology: A hydrologic perspective of climate-soil-vegetation dynamies. *Water Resources Research*, 36:3–9.

Rodriguez-Iturbe, I., Cox, D. R., and Isham, V. (1987). Some Models For Rainfall Based On Stochastic Point-Processes. *Proceeding of the Royal Society A*, 410(1839):269–288.

Rodriguez-Iturbe, I., D'Odorico, P., Laio, F., Ridolfi, L., and Tamea, S. (2007). Challenges in humid land ecohydrology: Interactions of water table and unsaturated zone with climate, soil, and vegetation. *Water Resources Research*, 43(9):1–5.

Rodriguez-Iturbe, I., D'Odorico, P., Porporato, A., and Ridolfi, L. (1999a). On the spatial and temporal links between vegetation, climate, and soil moisture. *Water Resources Research*, 35(12):3709–3722.

Rodriguez-Iturbe, I., Muneepeerakul, R., Bertuzzo, E., Levin, S. a., and Rinaldo, A. (2009). River networks as ecological corridors: A complex systems perspective for integrating hydrologic, geomorphologic, and ecologic dynamics. *Water Resources Research*, 45(1):1–22.

Rodriguez-Iturbe, I., Porporato, A., Laio, F., and Ridolfi, L. (2001). Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress - I. Scope and general outline. *Advances in Water Resources*, 24(7):695–705.

Rodriguez-Iturbe, I., Porporato, A., Ridolfi, L., Isham, V., and Cox, D. R. (1999b). Probabilistic Modelling of Water Balance at a Point: The Role of Climate, Soil and Vegetation. *Proceedings of the Royal Society A*, 455(1990a):3789–3805.

126

Rodriguez-Iturbe, I. and Rinaldo, A. (1997). *Fractal River Basins: Chance and Self-Organization.* Cambridge University Press, New York.

Rodriguez-Iturbe, I., Rinaldo, A., Rigon, R., Bras, R., Ijjaszvasquez, E., and Marani, A. (1992). Fractal Structures As Least Energy Patterns - The Case Of River Networks. *Geophysical Research Letters*, 19(9):889–892.

Rodriguez-Iturbe I. and Porporato, A. (2004). *Ecohydrology of water controlled ecosystems: soil moisture and plant dynamics).* Cambridge Univ. Press, New York.

Rost, S., Gerten, D., Bondeau, A., Lucht, W., Rohwer, J., and Schaphoff, S. (2008). Agicultural green and blue water consumption and its influence on the global water system. *Water Resources Research*, 44:W09405, doi:10.1029/2007WR006331.

Sancho, J. M., San Miguel, M., Pesquera, L., and Rodriguez, M. A. (1987). Positivity requirements on fluctuating parameters. *Physica A*, 142(1-3):532–547.

Schleiff, U. (2008). Analysis of water supply of plants under saline soil conditions and conclusions for research on crop salt tolerance. *Journal of Agronomy and Crop Science*, 194(1):1–8.

Schoups, G., Hopmans, J., and Tanji, K. K. (2006). Evaluation of model complexity and space-time resolution on the prediction of long-term soil salinity dynamics, western San Joaquin Valley, California. *Hydrological Processes*, 20(13):2647–2668.

Serrano, M. A. (2008). Rich-club vs rich-multipolarization phenomena in weighted networks. *Physical Review E*, 78:26101.

Serrano, M. A. and Boguna, M. (2005). Weighted Configuration Model. *AIP Conference Proceedings*, 776:101.

Settin, T., Botter, G., Rodriguez-Iturbe, I., and Rinaldo, A. (2007). Numerical studies on soil moisture distributions in heterogeneous catchments. *Water Resources Research*, 43(5).

Shah, S., Vervoort, R., Suweis, S., Guswa, A., Rinaldo, A., and Van der Zee, S. (2011). Stochastic modeling of salt accumulation in the root zone due to capillary flux from brackish groundwater. *Water Resources Research*, 47:W09506.

Sneppen, K., Bak, P., Flyvbjerg, H., and Jensen, M. H. (1995). Evolution as a self-organized critical phenomenon. *Proceedings of the National Academy of Sciences of the United States of America*, 92(11):5209–13.

Snyder, D. L. (1975). *Random Point Process.* Wiley.

Solé, R. V., Alonso, D., and McKane, A. (2002). Self-organized instability in complex ecosystems. *Philosophical Transactions of the Royal Society B*, 357(1421):667–681.

Sole, R. V. and Bascompte, J. (1996). Are critical phenomena relevant to large-scale evolution? *Proceedings of The Royal Society of London Series B*, 263(1367):161–168.

Bibliography

Stanley, H. E. (2000). Scale Invariance and Universality: organizing principles in complex system. *Physica A*, 281:60–68.

Stratonovich, R. L. (1963). *Topics in the Theory of Random Noise*. Gordon and Breach, New York.

Suarez, D. L. and Simunek, J. (1991). Unsatchem Project, Version 1.1. *USSL Research Report*, 129:1–204.

Szabolcs, I. (1989). *Salt-affected soils*. Florida, CRC Press.

Tamea, S., Laio, F., Ridolfi, L., D'Odorico, P., and Rodriguez-Iturbe, I. (2009). Ecohydrology of groundwater-dependent ecosystems: 2. Stochastic soil moisture dynamics. *Water Resources Research*, 45(5):1–13.

Tanji, K. K. (1989). Agricultural Salinity - Nature, Extent and Concerns. In Austin, TA, editor, *National Water Conference*, pages 33–38.

Tilman, D. (1994). Competition And Biodiversity In Spatially Structured Habitats. *Ecology*, 75(1):2–16.

Uhlenbeck, G. E. and Ornstein, L. S. (1930). On the Theory of the Brownian Motion. *Physics Review*, 36(5):823–841.

United Nations (2010). United Nation (Statistic Division).

Vallade, M. and Houchmandzadeh, B. (2003). Analytical solution of a neutral model of biodiversity. *Physical Review E*, 68(6):061902.

Van Den Broeck, C. (1983). On the relation between white shot noise, Gaussian white noise, and the dichotomic Markov process. *Journal of Statistical Physics*, 31(3):467–483.

Van Kampen, N. G. (2007). *Stochastic Processes in Physics and Chemistry, Third Edition (North-Holland Personal Library)*. North Holland.

Van Valen, L. M. (1973). A new evolutionary law. *Evolutionary Theory*, 1:1:30.

Vervoort, W. R. and Van Der Zee, S. E. A. T. M. (2008). Simulating the effect of capillary flux on the soil water balance in a stochastic ecohydrological framework. *Water Resources Research*, 44(8):W08425+.

Vico, G. and Porporato, A. (2010). Traditional and micro-irrigation with stochastic soil moisture. *Water Resources Research*, 46:W03509.

Viola, F., Daly, E., Vico, G., Cannarozzo, M., and Porporato, A. (2008). Transient soil moisture dynamics and climate change in Mediterranean ecosystems. *Water Resources Research*, 44(11):W11412.

Volkov, I., Banavar, J., He, F., Hubbell, S., and Maritan, A. (2005). Density dependence explains tree species abundance and diversity in tropical forests. *Nature*, 438(7068):658–61.

Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. (2003). Neutral theory and relative species abundance in ecology. *Nature*, 424(6952):1035–7.

Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. (2007). Patterns of relative species abundance in rainforests and coral reefs. *Nature*, 450(7166):45–49.

Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. (2009). Inferring species interactions in tropical forests. *Proceedings of the National Academy of Sciences of the United States of America*, 106(33):13854–9.

Wackernagel, M., Onisto, L., Bello, P., Linares, A. C., Falfan, I., Garcia, J. M., Guerrero, A., and Guerrero, C. S. (1999). National natural capital accounting with the ecological footprint concept. *Ecological Economics*, 29(3):375–390.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–2.

World Bank (2010). World Bank Data. *http://data.worldbank.org/indicator*.

Xu, Z. X., Schultz, G. A., and Schumann, A. (2002). DA conceptually-based stochastic point process model for daily stream-flow generation. *Journal of Hydrology*, 16:3003–3017.

Yang, H., Wang, L., Abbaspour, K. C., and Zehnder, A. J. B. (2006). Virtual water trade: an assessment of water use efficiency in the international food trade. *Hydrology and Earth System Sciences*, 10:443–454.

Zalewski, M. (2000). Ecohydrology, the scientific background to use ecosystem properties as management tools toward sustainability of water resources. *Ecological Engineering*, 16(1):1–8.

Zillio, T., Banavar, J. R., Green, J. L., Harte, J., and Maritan, A. (2008). Incipient criticality in ecological communities. *Proceedings of the National Academy of Sciences of the United States of America*, 105(48):18714–7.

Zillio, T., Volkov, I., Banavar, J., Hubbell, S., and Maritan, A. (2005). Spatial Scaling in Model Plant Communities. *Physical Review Letters*, 95(9):1–4.

Zygad\lo, R. (1993). Relaxation and stationary properties of a nonlinear system driven by white shot noise: An exactly solvable model. *Physical Review E*, 47(6):4067–4075.

Curriculum Vitae of Samir Simon Suweis

Personal Information

Name: Samir Simon

Surname: Suweis

Date of Birth: 21 February 1984

Nationality: Italian

Address: Via G. Dupreé 33, 35134, Padova, Italy

Office phone number: +41216938026

Cell. number: +41787524041

Email: samir.suweis@epfl.ch

Web page: http://people.epfl.ch/188422

Education

Ph.D., Environmental Engineering, November 2011
Ecole Polytechnique Fédérale Lausanne (EPFL)

M.S., Physics, September 2008
University of Padova: GPA: 29.9/30;   110/110 cum laude

B.S., Physics, September 2006
University of Padova: GPA: 28/30;   108/110

International Experiences

- Participant at the 2007 Summer School University in Plasma Physics carried out by Max-Planck-Institut for Plasma Physics in Greifswald, Germany

- Participant in June 2009 at the Summer School of Environmental Dynamics, sponsored by the Istituto Veneto di Scienze Lettere ed Arti, Venice, Italy

- Visiting research collaboration project with professor I. Rodriguez-Iturbe at Princeton University, from May 2010 to January 2011.

Research interest: Ecohydrology, stochastic modeling of hydrologic processes, soil salinity, stochastic modelling of ecosystems dynamics, complex networks, virtual water, stochastic differential equations.

Publications

peer−reviewed journal articles:

- S. Suweis, E. Bertuzzo, G. Botter, A. Porporato, I. Rodriguez-Iturbe and A. Rinaldo, Impact of stochastic fluctuations in storage-discharge relations on streamflow distributions, Water R resources Research, Vol. 46, Nr. 3, 2010.

- A. Fiori, F. Boso, F.P.J. Barros, S. De Bartolo, A. Frampton, G. Severino, S. Suweis and G. Dagan, An indirect assessment on the impact of connectivity of conductivity classes upon longitudinal asymptotic macrodispersivity , Water Resources Research, Vol. 46, Nr. W08601, 2010.

- S. Suweis, A. Rinaldo, S. Van der Zee, E. Daly, A. Maritan and A. Porporato, Stochastic modeling of soil salinity, Geophysical Research Letters, Vol. 37, Nr. L07404, 2010.

- S. Shah, R. W. Vervoort, S. Suweis, A. Guswa, A. Rinaldo, and S. Van der Zee, Stochastic modeling of salt accumulation in the root zone due to capillary flux from brackish groundwater, Water Resources Research, Vol. 47, pp. W09506,2011

- S. Suweis, M. Konar, C. Dalin, N. Hanasaki, A. Rinaldo and I. Rodriguez-Iturbe, Structure and Controls of the Global Virtual Water Trade Network, Geophysical Research Letter, Vol. 38, pp. L10403, 2011.

- M. Konar, C. Dalin, S. Suweis, N. Hanasaki, A. Rinaldo and I. Rodriguez-Iturbe, Water for food: The global virtual water trade network, Water Resources Research, Vol. 47, pp. W005520, 2011.

- S. Suweis, A.M. Porporato, A. Rinaldo and A. Maritan, Prescription-induced jump distributions in multiplicative Poisson processes, Physical Review E, Vol. 83, Nr. 6, pp. 061119, 2011.

- E. Bertuzzo, S. Suweis, L. Mari, A. Maritan, I. Rodriguez-Iturbe and A. Rinaldo, Spatial effects on species persistence and implications for biodiversity, Proceedings of the National Academy of Sciences, Vol. 108, Nr. 11, pp. 4346-4351, 2011.

- S. Suweis, E. Bertuzzo, L. Mari, A. Maritan, I. Rodriguez-Iturbe and A. Rinaldo, On the Universality of the Species Persistence-Time Distribution, in review, Journal of Theoretical Biology, 2011.

Conference abstracts:

- S. Suweis, A. Rinaldo, S. Van der Zee, A. Maritan and A. Porporato, Stochastic Modeling of Soil Salinity, Geophysical Research, Vol. 12, EGU2010-6817, EGU General Assembly 2010, Vienna

- S. Suweis, E. Bertuzzo, L. Mari, A. Maritan, I. Rodriguez-Iturbe, and A. Rinaldo, Scaling and Universality of Species Lifetimes, Vol. 13, EGU2011-10386, EGU General Assembly 2011, Vienna

In the news:

- *The New Scientist*, "African land grab could lead to future water conflicts" (by Anil Ananthaswamy), Magazine issue 2814, May 2011 (http://www.newscientist.com/article/mg21028144.100)
- *Cyberpresse.ca*, "L'eau virtuelle domine le commerce de l'or bleu" (by Mathieu Perreault), La Presse, June 2011 (http://www.cyberpresse.ca/actualites/201106/30/01-4413857-leau-virtuelle-domine-le-commerce-de-lor-bleu.php)

Cultural skills and Synergistic Activities:

Languages:

- Italian: mother tongue
- English: proficient speaker (CEF level[1] Understanding C2; Speaking C1; Writing C1)
- French: independent speaker (CEF level, Understanding B2; Speaking B1; Writing B1)

Programming Languages and Softwares:

- Mathematica (expert), Matlab (good user)
- Windows, Microsoft Office, Adobe illustrator and LaTex (expert)

Teaching experiences:

- Teaching assistant for the class "Quantitative Methods II" given by professors Andrew Barry and Samuel Arey, during the second semester of the academic year 2008-2009, at EPFL University.

Awards:

- Ecohydrology symposium poster award, for outstanding poster. Latsis international symposium on ecohydrology, Lausanne, 2010.

Other Services

- Member of EGU and AGU since 2010
- Reviewer for three scientific journals (Water Resource Research, Advances in Water Resources, Journal of Theoretical Biology) and for the national science foundation (NSF) funding agency.

Advisors

- Amos Maritan, University of Padova;

---

[1]Common European Framework of Reference

- Andrea Rinaldo, Ecole Polytechnique Fédérale Lausanne (EPFL).

Personal profile and interests:

I find myself a curious person, interested about many things. Besides the interests related to my specific field of study, I'm also very fond of Philosophy, Literature and Cinema. As a ex co-leader of the student cultural association "Antonio Rosmini" I actually had the opportunity to handle with such subjects and to take active roles in organizing meetings and exhibitions on these topics.

I've been a professional player of the card game "Magic, the gathering". In 20001 I qualified and attended the "Pro Tour" international world tournament" in Tokyo. I've also been a professional table tennis player for seven years, from 1997 to 2004, participating in regional championships. I'm very interested in game theory, with particular focus on Poker and Backgammon.

My key strengths include strong analytical and logical skills, and communication and interpersonal skills. I enjoy working in a team and help others progress. At the same time I work well independently. As a highly motivated and driven individual I strive on taking up challenges.

References:

To get additional information about me please contact:

- Professor Amos Maritan (amos.maritan@pd.infn.it), Department of Physics, University of Padova, Italy;

- Professor Andrea Rinaldo (andrea.rinaldo@epfl.ch), Department of Civil and Environmental Engineering, Ecole Polytechnique Fédérale Lausanne, Switzerland.

- Professor Amilcare Porporato (amilcare.porporato@duke.edu), Department of Civil and Environmental Engineering, Duke University, North Carolina, USA

- Professor Ignacio Rodriguez−Iturbe (irodrigu@Princeton.EDU), Department of Civil and Environmental Engineering, Princeton University, New Jersey, USA

- Professor Paolo D'Odorico (paolo@virginia.edu), Department of Environmental Science, University of Virginia, Virginia, USA

November 14, 2011